

eticas

08.03.2022


Auditoría Externa del Sistema VioGén



FUNDACIÓN | RED DE MUJERES
Ana Bella | SUPERVIVIENTES

CONTENIDO

1- Introducción	4
2- Evaluación de riesgos en el contexto de violencia de género	9
3- El sistema VioGén	12
¿Qué es VioGén?	12
¿Cómo funciona la evaluación de riesgo de VioGén?	14
4- Marco Metodológico	20
5- Auditoría de VioGén	22
Acceso al Sistema VioGén	22
La entrevista y el formulario de VioGén	26
Percepción del Algoritmo de VioGén	32
Resultados y recomendaciones	34
6- Conclusiones	38
7- Agradecimientos	40
8- Bibliografía	40
9- Apéndice - El Formulario VPR5.0	43



**Auditoría Externa del Sistema
VioGén**

1- Introducción

La violencia de género contra las mujeres es un fenómeno mundial. La ONU calcula que, en todo el mundo, 736 millones de mujeres -casi una de cada tres- han sido objeto de violencia de pareja y de violencia sexual fuera de la pareja al menos una vez en su vida¹. Las zonas más desarrolladas del mundo no son una excepción a esta tendencia. La violencia contra las mujeres también ha sido un problema clave en España, donde 1.126 mujeres fueron asesinadas por sus (ex)parejas entre 2003 y 2021². El 32,4% de las mujeres en España de 16 años ó más (aproximadamente 6,6 millones) han sufrido a lo largo de su vida violencia física, sexual y/o psicológica (emocional, de control, económica y de miedo) por parte de sus (ex)parejas (Delegación del Gobierno contra la Violencia de Género, 2019b). Ante la necesidad de proporcionar una protección adecuada a las mujeres que denuncian casos de violencia de género, muchos países han desarrollado programas especializados. En algunos de estos países, entre los cuales se encuentra España, dichos programas incluyen una herramienta de evaluación de riesgos que proporciona una puntuación algorítmica que evalúa el riesgo y que se emplea para tomar decisiones o recomendaciones sobre qué acciones policiales y legales deben promulgarse para proteger a las mujeres.

El uso de sistemas automatizados para predecir el riesgo ha ido en aumento en los últimos años, a menudo suscitando preocupaciones sobre su equidad y transparencia. En nuestro propio trabajo en Eticas, hemos destacado cómo estos sistemas suelen desplegarse inicialmente en contextos que afectan a las poblaciones vulnerables, lo que también suscita preocupaciones sobre su posible impacto y posterior reparación. Para examinar estas cuestiones más de cerca, en 2021 pusimos en marcha nuestro Proyecto de Auditoría Externa, en el que actualmente estamos haciendo ingeniería inversa de sistemas algorítmicos de diferentes ámbitos (justicia penal, empleo, redes sociales y banca)³, en colaboración con las partes afectadas interesadas y las organizaciones de la sociedad civil. VioGén es nuestro primer informe de lo que esperamos sea una larga serie.

En Eticas somos conocidos por nuestras Auditorías Internas, que han sido realizadas para clientes públicos y privados desde hace cinco años. En estos casos, somos contratados por quienes desarrollan y/o implementan un sistema algorítmico para que identifiquemos y corrijamos faltas de equidad e ineficiencias que presenta. Nuestra metodología de auditoría interna es una solución integral, tanto técnica como social, que examina los inputs, modelos y procesos, así como también los resultados y el impacto del sistema. Una de las cosas que hemos aprendido a través de nuestra experiencia de auditoría práctica es que la ingeniería inversa de los sistemas es un buen ejercicio incluso cuando se puede acceder al código y a los datos relevantes, ya que algunos tipos de sesgo pueden no ser evidentes antes de que se conviertan en impacto. Este entendimiento nos llevó a considerar la posibilidad de desarrollar una metodología de

¹ <https://www.unwomen.org/en/what-we-do/ending-violence-against-women/facts-and-figures>

² https://violenciagenero.igualdad.gob.es/violenciaEnCifras/victimasmortales/fichaMujeres/pdf/VMortales_2022_01_25_2.pdf

³ <https://eticasfoundation.org/auditorias-externas-algoritmos/>

Auditoría Externa para aquellos casos en los que no es posible acceder al código o a los datos.

Una auditoría externa es un proceso por el cual un tercero independiente examina el impacto y, en la medida de lo posible, el funcionamiento de un sistema algorítmico con el fin de detectar posibles anomalías o prácticas que pudieran ser injustas o perjudiciales para grupos protegidos o para la sociedad en su conjunto. La principal particularidad de las auditorías externas es que el acceso al algoritmo y a las bases de datos utilizadas para diseñar, desarrollar, probar y validarlo suele ser restringido. Por ello, el impacto suele evaluarse mediante análisis sistémicos de las poblaciones afectadas, fuentes secundarias y bases de datos que contienen datos recogidos a través de diferentes mecanismos de recogida. Cuando buscamos bibliografía sobre cómo auditar externamente los algoritmos, las fuentes que encontramos hacían referencia en su mayoría a la auditoría externa de las redes sociales o de los servicios digitales (Ada Lovelace Institute, 2021; Sandvig et al., 2014), por lo que nos propusimos desarrollar y probar diferentes metodologías en distintos ámbitos y áreas, así como en diferentes sistemas técnicos.

En el caso de VioGén, la herramienta de evaluación de riesgos de violencia de género utilizada por el Ministerio del Interior español, se cuenta con una trayectoria. Preocupados por su funcionamiento e impacto, en 2018 nos dirigimos al Ministerio para solicitar información y una reunión. Dicho encuentro sólo se celebró después de que solicitáramos la asistencia de un diputado del Congreso. Desde entonces, hemos mantenido varias reuniones con los responsables del sistema, en las que Éticas ha ofrecido una auditoría interna confidencial pro-bono del sistema VioGén, así como el considerar la incorporación de herramientas de aprendizaje automático supervisado para recabar información de las grandes cantidades de datos históricos que produce el sistema. Aunque estas sugerencias fueron bien recibidas, no se tomó ninguna medida y nuestra oferta de realizar una auditoría interna nunca llegó a materializarse. Por lo tanto, en 2021 nos propusimos iniciar nuestro **proyecto de Auditoría Externa de VioGén**, con la colaboración de la Fundación Ana Bella, una entidad líder perteneciente a la sociedad civil que trabaja con mujeres que han sobrevivido a la violencia doméstica y que, por lo tanto, en algún momento, han recibido una puntuación de riesgo producida por el propio sistema VioGén.

Las cuestiones que hemos abordado a través de este ejercicio de Auditoría Externa sobre VioGén son múltiples. Por un lado, nos preocupa la **transparencia** del sistema y las obligaciones del Ministerio del Interior al respecto. En 2015 se revisó el régimen jurídico del sector público en España para incluir un precepto según el cual **las actuaciones automatizadas desplegadas por la administración pública** (por ejemplo, a través de sistemas algorítmicos) **están sujetas a un conjunto de órganos y procesos** de gobierno para garantizar su adecuado funcionamiento, incluida la auditoría.⁴ Aunque el ámbito de aplicación de este precepto abarca exclusivamente a los sistemas totalmente automatizados (en los que no interviene ningún humano), nuestros datos muestran que el riesgo algorítmico asignado por VioGén permanece inalterado el 95% de las veces (Zurita

⁴ Ley 40/2015 art. 41

Bayona, 2014). A la luz de esto, argumentamos **que para los sistemas de toma de decisiones altamente automatizados y financiados con fondos públicos, de enorme impacto social como VioGén** (que, en algunos casos, toma decisiones de vida o muerte), **deberían exigirse por ley auditorías independientes.**

También hemos observado que **la mayor parte 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), **y por personas que trabajan o que tienen intereses particulares en el ministerio y las fuerzas policiales.** Esto refuerza nuestro argumento sobre la necesidad de una supervisión independiente del sistema, y esperamos que esta auditoría externa impulse a los responsables de VioGén a encargar una auditoría independiente y publicar sus resultados.

El papel de los agentes de policía en la validación o modificación de la puntuación de riesgo VioGén también plantea cuestiones de **responsabilidad.** Tener una puntuación de riesgo VioGén más alta o más baja significa que una mujer recibirá diferentes niveles de protección policial. Sin embargo, **no está claro quién es el responsable de esa decisión.** Dado que el 95% de los agentes de policía decidieron no modificar la puntuación de riesgo sugerida, está claro que consideran dicha puntuación algo más que una simple recomendación, y que le están delegando su agencia al sistema. Además, si su disposición a cuestionar el sistema disminuye cuando aumenta su carga de trabajo, como sugieren algunos autores (Estévez Mendoza, 2020), está claro que hay que prestar más atención al impacto del "humano en el circuito" en el rendimiento, la equidad y la responsabilidad del sistema.

Otro aspecto preocupante es la **falta de participación de las poblaciones afectadas** en el diseño y seguimiento del sistema. Gran parte de la literatura existente sobre VioGén se centra en sus aspectos técnicos -especialmente en lo relativo a la validez predictiva del algoritmo- y no en su impacto social, el papel o las experiencias de las mujeres afectadas por el mismo. Aunque ha habido intentos de realizar encuestas sobre la satisfacción de las usuarias (González-Álvarez & Garrido, 2015), es preocupante que para un sistema destinado a ser utilizado con poblaciones muy vulnerables, no se haya tenido en cuenta ni consultado ni a las usuarias finales ni a los grupos de usuarios. Este aspecto es algo que abordamos ampliamente en esta Auditoría Externa, centrándonos no sólo en cuestiones técnicas, sino también sobre el impacto de VioGén en las víctimas⁵ de la violencia de género, **siendo nuestros resultados muy preocupantes.**

Este aspecto es especialmente relevante en un momento en el que el **Ministerio del Interior parece estar considerando la incorporación de Machine Learning (ML)** al sistema. Aunque hasta la fecha de realización de este informe no se ha confirmado oficialmente, ha habido iniciativas recientes para incorporar ML y técnicas avanzadas de

⁵ Siguiendo la terminología oficial de las instituciones legales y judiciales en España, este informe se refiere a las mujeres que son objeto de agresión masculina por parte de sus ex o actuales parejas con el término "víctimas" y a los hombres que son autores de esta agresión con el término "agresores". Esta elección terminológica se adapta por razones prácticas y sirve para ser compatible con los relatos oficiales. Por otro lado, reconocemos que el término "víctima" es un concepto muy discutido y criticado por victimizar aún más a las mujeres. Por ello, muchas activistas de los derechos de las mujeres y académicas optan por el término "superviviente". Véase, por ejemplo: Fundación Ana Bella, Red de Mujeres Supervivientes. <https://www.fundacionanabella.org/>

análisis de datos al sistema VioGén (Pinedo, 2021). En diciembre de 2020, la empresa de software SAS anunció que el Ministerio del Interior y la Unidad de Violencia de Género habían llegado a un acuerdo con SAS para incorporar la analítica de datos y lo que se ha denominado el "agente digital" para automatizar y agilizar ciertos procesos con el fin de aumentar la protección⁶. Eticas se puso en contacto con SAS para conocer mejor la naturaleza de este acuerdo, pero se negaron a revelar cualquier información dado que el propietario del sistema es el Ministerio del Interior. Como describimos a continuación, el actual VioGén es un sistema actuarial que utiliza modelos estadísticos para inferir el riesgo futuro. Como tal, es un sistema algorítmico bastante simple en el que la información se introduce en un formato específico (un formulario), y se evalúa utilizando diferentes ponderaciones. Si bien podríamos argumentar que **las posibilidades de datos de VioGén están infrautilizadas**, no está claro que el ML -la creación de bucles de retroalimentación automatizados en el sistema, convirtiéndolo en un modelo de Inteligencia Artificial- sea deseable desde el punto de vista de la responsabilidad y la transparencia. En cualquier caso, **si el debate surge, no debería producirse a puerta cerrada** y sin tener en consideración, al menos, el punto de vista de las mujeres afectadas.

Por lo tanto, cuando a mediados de 2021 creamos el equipo para auditar externamente a VioGén, **teníamos preocupaciones en relación a la transparencia, la supervisión independiente, la rendición de cuentas, la participación de los usuarios finales y la transición al ML**. El proceso de auditoría nos ha permitido convertir las preocupaciones generales en conflictos específicos, construyendo nuestro caso para una mayor transparencia y supervisión, una mejor rendición de cuentas y una evaluación del impacto social en torno a datos concretos. Aunque las herramientas externas no nos permiten ser concluyentes, nos han proporcionado datos para plantear y justificar nuestras preguntas. Los resultados que presentamos a continuación son preocupantes, pero el proceso nos ha convencido de la utilidad de nuestro enfoque. **Si hemos conseguido llegar hasta aquí sin tener acceso a los datos pertinentes, imaginemos lo que podríamos haber hecho nosotros, y la sociedad en su conjunto, con acceso a ellos**. Esperamos que este informe impulse un cambio en la forma de trabajar y evolucionar de VioGén.

En las secciones siguientes, presentamos **los resultados de 7 meses de trabajo de Eticas y de la Fundación Ana Bella**, empleando los datos disponibles e involucrando a las mujeres afectadas y otras partes interesadas. Como se ha mencionado anteriormente, forma parte de un proyecto más amplio de auditoría externa en el que Eticas, en colaboración con otras organizaciones de la sociedad civil, realiza una ingeniería inversa y evalúa el impacto de los algoritmos en diferentes ámbitos. Con este proyecto pretendemos desarrollar herramientas metodológicas para auditar externamente sistemas automatizados de evaluación de riesgo en ausencia de acceso al código, a la entrada, a la salida y a los datos administrativos. Con ello se busca proporcionar herramientas metodológicas a las organizaciones comunitarias para evaluar y corregir algoritmos con impacto social, abogando así por el cambio de políticas. De este modo,

⁶https://www.sas.com/es_es/news/press-releases/locales/2020/viogen-secretaria-estado-seguridad-y-sas-unidos-lucha-contra-violencia-genero-analitica-avanzada-ia.html

pretendemos apoyar a los movimientos de auditoría algorítmica ascendente realizados por organizaciones y por grupos de usuarios finales.

Este informe **se estructura** en torno al proceso que vive una víctima de violencia de género al presentar una denuncia policial, desde el principio hasta el final, con el objetivo de cerrar la brecha entre la literatura existente sobre la validez técnica de VioGén y las experiencias vividas por aquellas mujeres cuya vida se ha visto afectada, o incluso determinada, por él. En primer lugar, ofrecemos una descripción técnica del sistema VioGén. Luego, discutimos críticamente el sistema y exploramos sus fortalezas y trampas internas mediante el estudio inductivo de las percepciones y experiencias de algunas de las partes interesadas, como son las supervivientes de violencia doméstica que han pasado por el sistema VioGén, sus abogados y las organizaciones de la sociedad civil que trabajan en este campo. De este modo, sentamos las bases para comparar el diseño y la evaluación del sistema con su funcionamiento real y el modo en que lo experimentan las principales partes interesadas.

2- Evaluación de riesgos en el contexto de violencia de género

Tanto los científicos como los responsables políticos se han esforzado por desarrollar soluciones rápidas, eficaces, precisas y prácticas para la prevención de la violencia. En un mundo hipotético con recursos infinitos, la mejor estrategia sería proporcionar un alto nivel de protección y vigilancia a todas las mujeres que se sientan en riesgo de sufrir violencia de género. Lamentablemente, los recursos policiales son limitados y, por lo tanto, deben asignarse cuidadosamente a quienes se prevé que corren un mayor riesgo. En este sentido, la evaluación de riesgos desempeña un papel crucial en la prevención de la violencia de género.

Las herramientas de evaluación de riesgos están diseñadas para clasificar los casos de violencia de género según el nivel de riesgo que se puede prever. Por lo tanto, pretenden ofrecer una predicción precisa de qué víctimas de violencia de género tienen más probabilidades de volver a ser agredidas y, por lo tanto, necesitan protección. Una puntuación de riesgo, en este contexto, no evalúa la gravedad de los incidentes pasados o actuales, sino que predice la probabilidad de tener un futuro episodio de violencia de género -lo que se evalúa es el riesgo de reincidencia del agresor.

Aunque las herramientas de evaluación de riesgos que predicen la violencia de género no son nuevas, en las últimas tres décadas se han producido importantes avances en cuanto a su precisión y status científico. En primer lugar, los estudios clínicos y sociopsicológicos han dado pasos importantes en la identificación de los principales factores de riesgo de la violencia de género. Aunque todavía hay poco consenso en la literatura sobre lo que se entiende por riesgo en el contexto de la violencia de género (Kropp, 2004)⁷, hay un acuerdo considerable sobre lo que constituye un factor de riesgo (Campbell et al., 2001; Riggs et al., 2000). Estas listas de factores de riesgo han proporcionado la base científica sobre la que se construyen las herramientas de evaluación de riesgo. En segundo lugar, los avances en las tecnologías de la información y la comunicación han permitido a las instituciones públicas con competencias en la prevención de la violencia de género compartir información y sincronizar sus acciones. Por lo tanto, las herramientas de evaluación de riesgo pueden basarse en múltiples bases de datos que reúnen diferentes tipos de información. En tercer lugar, los avances en la ciencia de los datos y la analítica han proporcionado mejores métodos de extracción/descubrimiento de conocimiento a partir de los datos, el análisis de datos/patronos, y han hecho posible los modelos predictivos (Sarker, 2021).

A pesar de los avances en este campo, todavía existen controversias sobre cómo debe realizarse la evaluación de riesgo, por quién, qué papel tienen los profesionales y las víctimas en este proceso, y cómo una evaluación de riesgos debe informar del proceso

⁷ Esto se debe a que no existe el "no riesgo" en el contexto de la violencia en la (ex)pareja, así como a que hay diferentes tipos de riesgos que varían en función de la inminencia, naturaleza (como emocional, física, sexual, etc.), frecuencia y gravedad (Kropp, 2004). Esto también muestra la dificultad de evaluar el riesgo basándose en una escala unidimensional (por ejemplo, bajo-alto), ya que éste tiene múltiples dimensiones que deben ser consideradas.

de gestión de riesgos. La literatura menciona **tres enfoques principales** para la evaluación de riesgo (Heilbrun et al., 2011; Kropp, 2004).

- **La evaluación clínica no estructurada** implica la evaluación profesional de cada caso específico y la adaptación individualizada de la gestión de riesgos. La ventaja de este enfoque es su capacidad para tener en cuenta condiciones únicas, inusuales y específicas del contexto que deben ser evaluadas caso por caso por los profesionales. Por otro lado, esta discreción profesional puede tener como consecuencia el coste de disponer de un sistema fiable y válido, ya que la evaluación depende en gran medida de la formación, las preferencias y los prejuicios del profesional.
- **La evaluación actuarial** se basa en la evaluación estadística de factores de riesgo predeterminados y definidos científicamente. Este enfoque está diseñado para producir evaluaciones de riesgo objetivas y estandarizadas con rigor científico sin depender de lo cualificado que esté el evaluador. Los modelos actuariales son criticados por ser "mecánicos y algorítmicos" (Grove y Meehl, 1996), por basarse en suposiciones lineales y por no ser lo suficientemente buenos para captar la información específica del contexto.
- **El juicio profesional estructurado** tiende un puente entre los dos primeros métodos. Aunque algunos pasos están estandarizados como en el modelo actuarial, el último paso no se hace de forma algorítmica y tiene en consideración la responsabilidad y la discreción profesional del evaluador (Kropp, 2004).

Además de estos tres modelos clásicos de evaluación de riesgo, los avances en las metodologías computacionales, así como la disponibilidad de grandes datos digitales, permiten aplicar técnicas de **Machine Learning** (ML) para predecir los delitos de género. Los sistemas de ML utilizan algoritmos inferenciales, basados en datos para extraer patrones de datos históricos (Tolan et al., 2019). Son capaces de hacer predicciones en un contexto de alta incertidumbre, pudiendo tratar con un gran número de características como input, y permitiendo la escalabilidad. Como aspecto negativo, cabe destacar que estos sistemas dependen en gran medida de datos históricos, lo que puede contribuir a perpetuar los sesgos estructurales y las desigualdades, mientras que su complejidad interna puede conducir a una falta de transparencia y al llamado efecto caja negra.

A día de hoy, VioGén es un sistema actuarial que utiliza modelos estadísticos para inferir el riesgo que puede correr una víctima (tanto de agresión como de homicidio) así como su evolución en base a un conjunto de indicadores que han sido determinados y posteriormente evaluados por un grupo de expertos. Como se comenta posteriormente, se ha discutido la posibilidad de incorporar al sistema VioGén un algoritmo de aprendizaje automático, basado en la técnica de Nearest Centroid para la clasificación -o un modelo híbrido que implemente una mezcla estocástica del sistema actual y de Nearest Centroid- que aparentemente superaría a VioGén (González-Prieto et al., 2021, p. 6), pero no hay evidencia de que el Ministerio esté considerando esta opción.

Cada enfoque de la evaluación de riesgos tiene sus propias ventajas y limitaciones, y ofrece diferentes perspectivas para evaluar el riesgo de reincidencia (p.ej. predecir la probabilidad de nuevos actos de violencia) y gestionarlo (p.ej. proporcionar información para la planificación de la prevención del riesgo). Sin embargo, se realizan esfuerzos continuos para mejorar los sistemas automatizados de evaluación de riesgos y para ofrecer una mejor protección a las víctimas, a la vez que se gestionan eficazmente los recursos. En este contexto, este informe presenta un nuevo enfoque para estudiar externamente los sistemas automatizados de evaluación de riesgos en general, y una evaluación crítica para mejorar VioGén en particular, analizando tanto la naturaleza del sistema como el impacto que tiene sobre las víctimas de violencia de género

3- El sistema VioGén

¿Qué es VioGén?

El Sistema de Seguimiento Integral de los casos de Violencia de Género (Sistema VioGén) es una aplicación web, integrada en la Red SARA española (Sistemas y Redes de Aplicaciones para Administradores). Está diseñada para coordinar las actuaciones de los profesionales públicos españoles que se encargan del seguimiento, asistencia y protección de las mujeres denunciadas de violencia de género y de sus hijos. De este modo, VioGén pretende establecer una densa red de instituciones con competencias en el ámbito de la prevención de la violencia de género y dar respuestas rápidas, integrales, eficaces y de alto nivel a la violencia de género en todo el país⁸. Tiene su origen legal en los mandatos del artículo 31 y 32 de la Ley Orgánica 1/2004 sobre las "Medidas de Protección Integral contra la Violencia de Género"⁹

El sistema fue creado por la Secretaría de Estado de Seguridad (SES) del Ministerio del Interior y puesto en marcha en todo el país (excepto Cataluña y el País Vasco) en 2007. Hasta la fecha ha realizado más de **3 millones de evaluaciones de riesgo** (López-Ossorio et al., 2019). En enero de 2022, **hay 673.912 casos en el sistema VioGén, de los cuales 69.391 son casos activos que requieren supervisión policial**, lo que convierte al sistema español de evaluación de riesgos **en el primero del mundo por volumen de casos** (González-Álvarez et al., 2018, p. 37).

El sistema VioGén está oficialmente diseñado para cumplir los siguientes objetivos:¹⁰

- Aglutinar a todas las instituciones públicas que tienen competencias en materia de violencia de género;
- Integrar toda la información relevante;
- Realizar una predicción del riesgo;
- Vigilar y proteger a las víctimas de la violencia de género realizando una labor preventiva, emitiendo avisos y alertas y adoptando otras medidas necesarias en función del nivel de riesgo.

El sistema busca integrar diferentes servicios públicos, como las fuerzas de seguridad (Guardia Civil y Policía Nacional), justicia, sanidad, servicios sociales, igualdad y los sistemas penitenciarios para facilitar el intercambio de información. Cuenta con más de 30.000 usuarios con diferentes niveles de permiso y acceso (González-Álvarez et al., 2018). Mientras que todos estos usuarios mencionados pueden acceder al sistema -e incluso algunos de ellos pueden aportar información relevante-, sólo los agentes de la ley (agentes de la Policía y de la Guardia Civil) pueden registrar casos. En cumplimiento del RGPD europeo (Reglamento General de Protección de Datos) y de la Ley Orgánica de

⁸ <http://www.interior.gob.es/web/servicios-al-ciudadano/violencia-contra-la-mujer/sistema-viogen>

⁹ https://www.boe.es/diario_boe/txt.php?id=BOE-A-2004-21760

¹⁰ <http://www.interior.gob.es/web/servicios-al-ciudadano/violencia-contra-la-mujer/sistema-viogen>

Protección de Datos (15/1999), los usuarios acceden al sistema con un nombre de usuario y una contraseña intransferible.

Guardia Civil	16.239
Policia Nacional	5.429
Policia Local	1.806
Mossos d'Esquadra, Policía Foral, Unidades Adscritas de Galicia y Comunidad Valenciana	573
Instituciones Penitenciarias	755
Unidades de Coordinación y Violencia	128
Servicios Sociales y de Igualdad	542
Ministerio de Justicia y Poder Judicial	8.433
Usuarios totales	33.905

Tabla 1: Usuarios habilitados del Sistema VioGén a 31/05/2020. Fuente: López-Ossorio (2020)

La Ley Orgánica española 1/2004 define **violencia de género** como "manifestación de la discriminación, la situación de desigualdad y las relaciones de poder de los hombres sobre las mujeres, se ejerce sobre éstas por parte de quienes sean o hayan sido sus cónyuges o de quienes estén o hayan estado ligados a ellas por relaciones similares de afectividad, aun sin convivencia".¹¹ En otros países, este tipo de violencia suele denominarse "violencia de pareja" (IPV) del hombre agresor a la mujer víctima.

En el sistema VioGén, **un caso** contiene una única víctima femenina y un único agresor masculino. Esto significa que cuando una mujer es víctima de múltiples agresores, habrá un caso diferente para cada uno de sus agresores. Del mismo modo, cuando un agresor masculino se dirige a diferentes mujeres, tendrá múltiples casos. Por lo tanto, el número de casos supera a las personas (González-Álvarez et al., 2018, p. 33). Un caso de violencia de género se registra en el sistema durante la denuncia oficial de la víctima ante la policía. Un **caso activo** significa que es seguido y supervisado activamente por las fuerzas policiales. Un caso se convierte en **inactivo** cuando deja de necesitar atención policial. Un caso se da de **baja** del sistema (caso de baja) cuando no hay previsión de reincidencia. Existen tres condiciones para que un caso sea dado de baja (González-Álvarez et al., 2018, p. 33):

- Absolución firme del acusado.
- Sobreseimiento de las actuaciones del investigado o procesado
- Condena firme ejecutada en la que ha transcurrido el plazo legal de anulación.

¹¹ Ley Orgánica 1/2004, Artículo 1: "la violencia que, como manifestación de la discriminación, la situación de desigualdad y las relaciones de poder de los hombres sobre las mujeres, se ejerce sobre éstas por parte de quienes sean o hayan sido sus cónyuges o de quienes estén o hayan estado ligados a ellas por relaciones similares de afectividad, aun sin convivencia."

¿Cómo funciona la evaluación de riesgo de VioGén?

El sistema VioGén pretende homogeneizar la valoración policial del riesgo de violencia de género y las medidas de protección y prevención asociadas en toda España. El sistema funciona a través de dos formularios (Protocolo Dual): Valoración Policial del Riesgo (VPR) y Valoración Policial de la Evolución del Riesgo (VPER). El formulario VPR realiza la primera valoración del riesgo en el momento de la denuncia de la agresión a la policía, mientras que el formulario VPER realiza el seguimiento de la evolución del riesgo de violencia de género. Estos protocolos de valoración son revisados y corregidos por un equipo multidisciplinar de expertos. La quinta versión, la más actualizada, se publicó en marzo de 2019. Desde entonces, la valoración del riesgo se realiza a través del VPR_{5,0}-H y VPER_{4,1}.

Cuando una mujer presenta una denuncia oficial contra su agresor, los agentes de policía rellenan el formulario VPR_{5,0}-H. Este formulario incluye **5 dominios** con **35 indicadores de riesgo** (véase el apéndice). Cada ítem se valora como "presente" y "no presente". De este modo, la recogida de información está estandarizada en todo el país. Una vez cumplimentado el formulario, el sistema asigna una puntuación de riesgo de violencia de género. Los niveles de esta puntuación de riesgo son **"no apreciado", "bajo", "medio", "alto" y "extremo"**. Los agentes de policía sólo pueden modificar la puntuación a un nivel de riesgo más alto, y no al revés: es decir, no se puede bajar la puntuación de riesgo calculada por el algoritmo VioGén. Sin embargo, y a pesar de que los agentes pueden aumentar la puntuación de riesgo asignada automáticamente, se informa de que rara vez lo hacen. En el 95% de los casos, los agentes mantienen la puntuación de riesgo asignada automáticamente (Zurita Bayona, 2014). Además, a medida que la pandemia de Covid-19 aumentó la carga de trabajo de los agentes de la ley, se ha observado en los últimos meses que los agentes presentan una mayor tendencia a confiar en las decisiones automatizadas (Estévez Mendoza, 2020). Lamentablemente, no disponemos de datos sobre cómo la carga de trabajo puede estar repercutiendo en los derechos, las posibilidades y la protección de las mujeres.

Como puede verse en la siguiente tabla, la distribución de los casos según las puntuaciones de riesgo asignadas parece ser estable a lo largo del tiempo. **La inmensa mayoría de los casos activos se consideran situaciones de riesgo no apreciado o de bajo riesgo**, y sólo una minoría de los casos entra en la categoría de niveles de riesgo medio/alto/extremo que requieren medidas de protección específicas por parte de la policía. También hay que señalar que el número de casos activos ha ido aumentando cada año. Es decir, aunque la distribución de las categorías de riesgo se ha mantenido estable a lo largo del tiempo, cada año hay más casos con puntuaciones de riesgo más altas que requieren una atención policial especial.

	Casos Totales	Casos Activos	No apreciado	Bajo	Medio	Alto	Extremo
2021	670,061	69,469	45.30%	41.82%	12.02%	0.83%	0.02%
2020	621,907	63,656	48.78%	40.62%	9.93%	0.66%	0.01%
2019	577,907	61,355	49.65%	39.49%	10.18%	0.65%	0.02%
2018	529,762	58,498	43.48%	45.95%	10.09%	0.45%	0.04%
2017	485,439	54,793	49.62%	41.58%	8.38%	0.39%	0.03%
2016	439,307	52,635	56.37%	36.11%	7.19%	0.32%	0.02%
2015	396,552	52,005	68.05%	26.11%	5.63%	0.19%	0.01%

Tabla 2: Número de casos totales y activos de VioGén y distribución de las puntuaciones de riesgo. Fuente: Boletín Estadístico Mensual de la Delegación del Gobierno contra la Violencia de Género (Ministerio de Igualdad).¹²

El seguimiento de la evolución del riesgo se realiza mediante el $VPER_{4,1}$. Si la evaluación continua tiene por objeto un control periódico sin incidentes, se denomina $VPER_{4,1}-S$ ("Sin incidente"). El periodo de esta evaluación viene determinado por el nivel de riesgo: nivel extremo: antes de 72 horas, nivel alto: antes de 7 días, nivel medio: cada 30 días, y nivel bajo: cada 60 días. Tras la aplicación de la $VPR_{5,0}-H$, si se produce un nuevo incidente, se realiza la $VPER_{4,1}-C$ ("Con incidente"). Entre 2007 y 2019 se han realizado más de 3 millones de evaluaciones (VPR y VPER). Según los desarrolladores de VioGén, este es uno de los números más altos de evaluaciones de riesgo en el mundo (López-Ossorio, Muñoz Vicente, et al. 2020).

En 2019, el sistema VioGén se ajustó para detectar los casos con riesgo de agresión letal, así como los casos en los que los niños están expuestos a la violencia¹³. El sistema VioGén actualizado ejecuta dos evaluaciones en paralelo ($VPR_{5,0}$ más la Escala H): una es la de reincidencia (la probabilidad de una nueva agresión por parte del mismo agresor) y la otra es del riesgo de homicidio. Los casos que conllevan riesgo de homicidio se reportan como casos de especial relevancia. En López-Ossorio et. al. (2020), los autores identificaron 13 indicadores que tienen una significación positiva en cuanto al riesgo de homicidio en la pareja íntima (IPH). Estos indicadores recogen determinadas variables de antecedentes penales del agresor, trastornos mentales y psiquiátricos, determinados patrones de conducta, problemas vitales o laborales, así como la salud mental de la víctima y la adicción a sustancias (López-Ossorio, González-Álvarez, et al., 2020, p. 50). En este sentido, los coeficientes OR reflejan la importancia de cada variable como predictora del riesgo de homicidio, siendo las amenazas suicidas ($OR=8,087$, $p<.001$), los problemas económicos y laborales en los últimos seis meses ($OR=6,324$, $p<.001$), y cualquier tipo de adicción o abuso de sustancias en la víctima ($OR=5,101$, $p<.001$) las más destacadas (Ibid.).

¹² <https://violenciagenero.igualdad.gob.es/violenciaEnCifras/boletines/boletinMensual/home.htm>

¹³ <https://www.rtve.es/noticias/20190313/entra-vigor-nuevo-protocolo-policial-para-valorar-riesgo-victimas-violencia-genero/1900900.shtml>

Los autores también aportan valores de sensibilidad (TPR=0,084) y especificidad (TNR=0,60), con el fin de atestiguar la capacidad de clasificación de la Escala H (López-Ossorio, González-Álvarez et al., 2020, p. 51). Con respecto a los niños, el nuevo sistema VioGén identifica dos categorías: "niños en situación de vulnerabilidad" y "niños en situación de riesgo" (instrucción 4/2019).

En estos tres casos -de especial relevancia, con niños en situación de vulnerabilidad y con niños en situación de riesgo- se adjunta a la Evaluación Policial de Riesgo, una Diligencia Automatizada para llamar la atención de jueces y fiscales y recomendar una evaluación pericial adicional del caso.

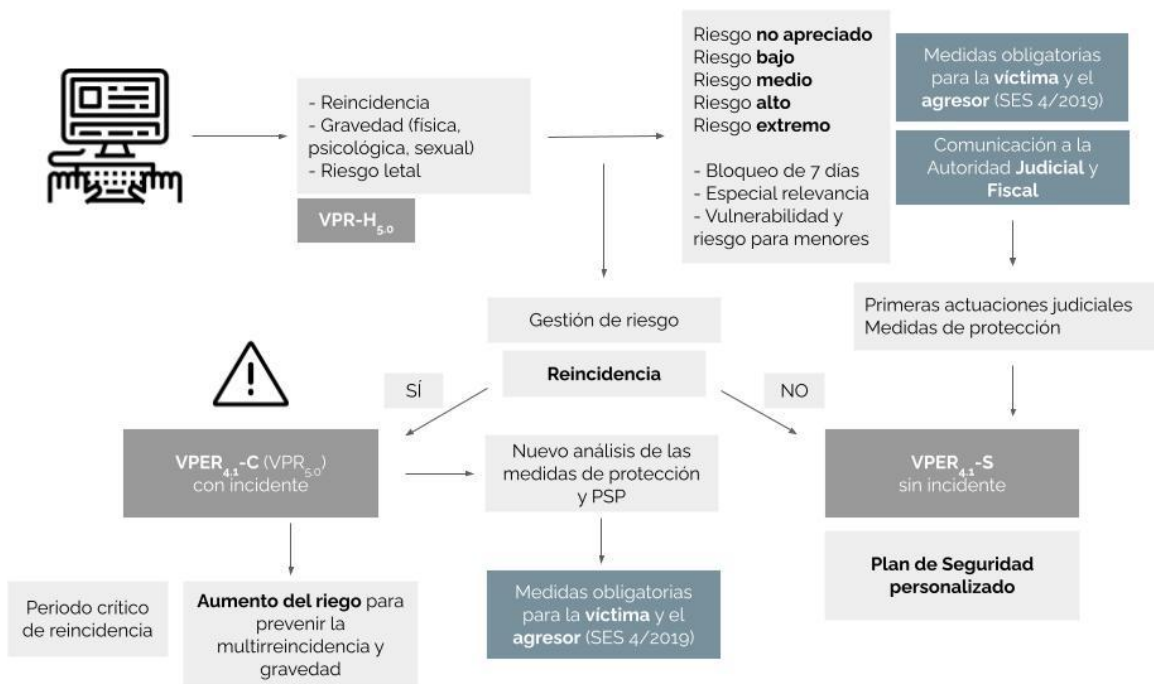


Figura 1: El proceso del sistema VioGén (adaptado de López-Ossorio, Muñoz Vicente, et al., 2020, p. 9).

El sistema VioGén activa automáticamente las medidas de protección policial para cada víctima en función de su puntuación de riesgo. La actualización de 2019 también ofrece un Plan de Seguridad Personalizado (PSP). El PSP tiene en cuenta las condiciones específicas de cada víctima, incluyendo si la víctima tiene hijos, trabaja fuera de su casa, vive con su agresor, etc. Además, cada vez se intenta más adaptar el PSP a las tecnologías disponibles, e incluye medidas como el cambio de número de teléfono, el bloqueo de llamadas del agresor, la instalación de la app AlertCop¹⁴, el registro de números de emergencia, la activación de la geolocalización, etc. Desde el punto de vista técnico, **VioGén utiliza modelos estadísticos clásicos para realizar una evaluación de**

¹⁴AlertCop es un servicio de aplicación móvil proporcionado por el Ministerio del Interior español que está diseñado para enviar alertas de seguridad con imágenes o videos al centro de emergencias más cercano, chatear directamente con un agente de asistencia o recibir noticias y notificaciones de seguridad enviadas por los servicios públicos de seguridad. Fuente: <https://alertcops.ses.mires/mialertcops/>

riesgos basada en la suma ponderada de todas las respuestas según unas ponderaciones preestablecidas para cada variable (González-Prieto et al., 2021, p. 11).

Con el fin de validar el $VPR_{4,0}$, se realizó un estudio longitudinal prospectivo estratificado para luego desarrollar el modelo predictivo. La muestra utilizada para ello estuvo constituida por víctimas de violencia de género entre el 24 de septiembre y el 1 de diciembre de 2015, con una ventana de seguimiento que abarcaba hasta el 29 de abril de 2016 para permitir la detección de reincidencia. La muestra ($n=3907$) se dividió entonces en dos grupos: 60% para construir el modelo y 40% para validarlo. Los indicadores de riesgo se extrajeron de $VPR_{4,0}$ y se utilizó el test de chi cuadrado de Pearson para calcular la significación estadística, siendo la Odds Ratio (OR) la principal variable dependiente (López-Ossorio, 2017, p. 191).

Además, se incluyeron 13 indicadores como identificadores críticos -con ponderación experta- para cribar el riesgo de homicidio. La ponderación de estos indicadores se multiplicó por dos, excepto en los casos de trastorno mental, a los que directamente se les dio una puntuación de 3 (López-Ossorio, 2017, p. 192). Estos indicadores son:

- Violencia física grave o muy grave
- Violencia sexual grave o muy grave
- Uso de armas (excepto armas de fuego)
- Amenazas de muerte por parte del agresor
- Acumulación de amenazas o agresiones durante los últimos seis meses
- Signos de celos extremos por parte del agresor en los últimos seis meses
- Comportamientos de acoso por parte del agresor en los últimos seis meses
- Agresiones a otras personas o animales por parte del agresor durante el último año
- Trastorno mental o psiquiátrico en el agresor
- Presencia de ideas o intentos de suicidio por parte del agresor
- Adicción o abuso de sustancias (alcohol, drogas o medicamentos) por parte del agresor
- La víctima manifestó su intención de terminar la relación en los últimos seis meses
- La víctima piensa que el agresor puede hacerle mucho daño o incluso matarla

El valor global se obtiene sumando la presencia ponderada de un indicador. Esto significa que todos los indicadores significativos suman en el resultado final, mientras que el peso de cada indicador varía en función de los datos empíricos obtenidos. Así, la escala teórica del $VPR_{4,}$ calculada mediante la suma de todos los indicadores, fue de (0 a 77,019) y la empírica de (0 a 68,062). Los autores estimaron que una suma ponderada aumentaba la capacidad predictiva del modelo en 2 puntos porcentuales de AUC (López-Ossorio, 2017, pp. 193-194)

Para clasificar cada caso, se fijaron los umbrales de cada categoría:

Clase de Riesgo	VPR _{4.0} Intervalos de Riesgo
Riesgo no apreciado	$0 \leq x \leq 9.353$
Riesgo bajo	$9.353 < x \leq 21.886$
Riesgo Medio	$21.886 < x \leq 34.715$
Riesgo Alto	$34.715 < x \leq 45.284$
Riesgo Extremo	$x > 45.284$

Tabla 3. Umbrales de riesgo para VPR_{4.0}. Fuente: López-Ossorio (2017, p. 194).

Para establecer los umbrales, se utilizó la escala empírica teniendo en consideración tres criterios principales: el primer umbral debía fijarse de modo que la tasa de falsos negativos fuera baja -independientemente del aumento de la tasa de falsos positivos-. La mayoría de los casos graves debía clasificarse por encima del "riesgo medio". Por último, la agrupación global debería ser proporcional a los recursos disponibles para aplicar las medidas de protección (López-Ossorio, 2017, p. 194). Sin embargo, todas las víctimas que pasan por el sistema VioGén consideran que están bajo una amenaza suficiente como para presentar una denuncia policial, lo que deja a las que tienen puntuaciones de riesgo más bajas en una posición vulnerable, como se analiza más adelante en este informe.

El desarrollo del VPER-4.0, en cambio, se realizó con importantes diferencias metodológicas, utilizando un método de diseño retrospectivo, con las primeras evaluaciones reincidentes como casos, y las primeras evaluaciones periódicas como controles (López-Ossorio, 2017, p. 197). En este caso, la escala teórica del VPER4.0 fue (-11,257 a 73,018) y la escala empírica fue (-11,257 a 25,080). Los umbrales para cada categoría se fijaron en:

Clase de Riesgo	VPER _{4.0} Intervalos de Riesgo
Riesgo no apreciado	$-11.257 \leq x \leq -3.087$
Riesgo Bajo	$-3.087 < x \leq -2.185$
Riesgo Medio	$-2.185 < x \leq 12.069$
Riesgo Alto	$12.069 < x \leq 19.751$
Riesgo Extremo	$x > 19.751$

Tabla 4: Umbrales de riesgo para VPER_{4.0}. Fuente: López-Ossorio (2017, p. 198).

Sin embargo, también se ha discutido un modelo alternativo de **Machine Learning** basado en la técnica de Nearest Centroid para la clasificación -o un modelo híbrido que implemente una mezcla estocástica del sistema actual y de Nearest Centroid- que aparentemente superaría a VioGén (González-Prieto et al., 2021, p. 6). Además, González Prieto et al. discuten el uso de una nueva métrica de evaluación denominada Protección Policial, que puede obtenerse como la suma de la precisión para la clase de riesgo "inexistente / no", la puntuación F1 para la clase de riesgo "bajo" y la sensibilidad para la clase de riesgo "alto" (González-Prieto et al., 2021, p. 6). Esta medida pretende asegurar la correcta identificación del riesgo para los modelos que, por ejemplo, tienen una buena precisión en el riesgo de reincidencia de la clase "alta" pero una mala sensibilidad. En esos casos, es posible que la puntuación F1 (i.e. la media armónica de la precisión y la sensibilidad) se encuentre dentro del rango de valores admisibles, aunque ello signifique dejar sin protección a la mayoría de los peores casos, lo que sería inadmisibles. Sin embargo, los autores también identifican dos **dificultades principales con la incorporación del Machine Learning en el sistema VioGén** que son intrínsecas al problema y se interponen en el camino de una solución basada en el Machine Learning. La primera es la imposibilidad -por razones legales, éticas y sociales- de establecer un grupo de control para medir el impacto del modelo. Se puede observar la evolución general de este tipo de delitos a lo largo de los años, pero también hay que tener en cuenta otros componentes como los cambios demográficos, las iniciativas políticas y policiales, y todas las demás variables que pueden afectar a la incidencia de la violencia de género. La segunda se deriva del intento de predecir el riesgo de sufrir violencia de género en sí mismo. En principio, parece que el sistema mide las recaídas, y no una valoración directa del riesgo (González-Prieto et al., 2021, p. 9). Sin embargo, los datos históricos no reflejan la reincidencia en sí misma, sino la nueva detención y la nueva condena (Christian, 2020, p. 75). Esto significa que el modelo ML -de forma similar a como lo hacen los expertos- debería aprender a clasificar el riesgo de victimización según los datos históricos de nueva detención y nueva condena, y no según la reincidencia o el riesgo real. Sin embargo, este problema es intrínseco a la evaluación de riesgos, y el enfoque híbrido podría facilitar una transición fluida del enfoque actuarial al basado en el ML.

4- Marco Metodológico

Esta auditoría externa se ha llevado a cabo mediante una investigación **multimétodo**. Gracias a nuestra experiencia con las auditorías internas, hemos aprendido que el mejor enfoque para entender cómo funcionan los algoritmos y su impacto en los diferentes grupos es mediante la combinación de enfoques cuantitativos y cualitativos. Los sistemas de alto impacto son siempre sistemas sociotécnicos, ya que los datos proceden de procesos sociales y sociológicos y repercuten en las dinámicas personales y sociales. Por lo tanto, y como se ve en la Guía de Auditoría Algorítmica de Éticas (Éticas, 2021), cualquier método o proceso diseñado para abrir la caja negra de los algoritmos y la Inteligencia Artificial requerirá ir más allá, por encima de un análisis puramente técnico, que no sólo sería incompleto sino también engañoso.

La metodología elegida reúne un análisis estadístico de los casos de IPH (Homicidio en la Pareja Íntima) para evaluar la validez predictiva de la evaluación de riesgos de homicidio de VioGén en diferentes grupos sociales de mujeres, así como métodos cualitativos que exploran diferentes percepciones y experiencias con el sistema VioGén. Este enfoque nos ha permitido maximizar el uso de los datos disponibles, tanto los que provienen de las bases de datos como la información que hemos recabado de la realización de entrevistas con diferentes actores implicados.

Dado que la petición de Éticas para acceder al conjunto de datos originales utilizados para construir y validar VioGén no fue concedida, **el análisis técnico se ha basado en el registro público de 1.000 víctimas de IPH** proporcionado por el Consejo General del Poder Judicial (CGPJ), ya que es la única fuente de datos que está disponible públicamente y nos permite abordar las cuestiones que estamos planteando en esta Auditoría Externa. El análisis cuantitativo ha sido diseñado para identificar las tasas de falsos negativos y las disparidades en la sensibilidad entre los diferentes estratos (es decir, grupos protegidos), con el fin de contextualizar la precisión predictiva de VioGén cuando se trata de casos de riesgo extremo. La sensibilidad es una métrica que refleja el número de casos positivos que se clasifican correctamente. Por lo tanto, las disparidades entre los grupos protegidos indicarían un sesgo subyacente en el algoritmo.

El conjunto de datos se restringió primero para reflejar exclusivamente los casos ocurridos entre 2009 y 2019 (585 casos de IPH) y, posteriormente, se restringió a los casos de toda España excepto Cataluña y el País Vasco, donde VioGén no está activo, lo que dio como resultado un total de 475 casos de IPH. Dado que VioGén sólo actúa sobre los casos denunciados de IPH, y los datos originales reflejan tanto los casos denunciados como los no denunciados de IPH, la muestra final está constituida por 126 casos denunciados de IPH que resultaron en homicidio. Cada caso de IPH utilizado incluye información sobre el lugar en el que se presentó la denuncia, la nacionalidad y la edad de la víctima y del agresor, su relación, la causa de la muerte, el número de hijos y menores de edad que tenía la víctima, la respuesta del agresor y si había una denuncia policial previa y una orden de protección posterior. Por último, los falsos negativos se definieron como cualquier caso de IPH con una denuncia policial previa pero sin

protección policial, mientras que la protección insuficiente se definió como cualquier caso de IPH con una denuncia policial previa que derivó en algún tipo de protección. Los resultados de este análisis se describen en la siguiente sección.

El **trabajo de campo cualitativo** de esta auditoría incluyó entrevistas telefónicas semiestructuradas y una investigación mediante encuestas con preguntas cerradas y abiertas. **Entrevistamos a 31 mujeres que sufrieron violencia de género y pasaron por el sistema VioGén.** Todas estas entrevistas fueron realizadas por un experto en el campo de la violencia de género con experiencia en métodos cualitativos de entrevista. Se obtuvo el consentimiento informado de las participantes al inicio de cada entrevista. Las preguntas de la entrevista fueron diseñadas para indagar sobre las experiencias y percepciones de las mujeres durante su paso por el sistema VioGén. Se evitaron las preguntas que pudieran crear angustia y una carga emocional para las participantes y se les recordó que podían omitir cualquier pregunta y/o retirarse del estudio en cualquier momento. Las entrevistas se grabaron (sólo en audio) y se transcribieron, pero toda la información personal, incluidos los nombres de las personas y los lugares, se anonimizaron para garantizar la confidencialidad. Utilizamos el código W (por ejemplo, W1, W2, etc.) para identificar a cada víctima en nuestro análisis.

Pudimos llegar a las supervivientes de violencia de género a través de la red de la Fundación Ana Bella, que incluye a más de 27.000 supervivientes en su Red de Mujeres Supervivientes.¹⁵ Como criterio de muestreo, incluimos a las mujeres que pasaron por el sistema VioGén entre 2019 y 2021 y que denunciaron la agresión en Andalucía, Valencia, Madrid o Galicia, que se encuentran entre las cinco primeras regiones de España con más casos activos de VioGén.

También **se contactó con abogados especializados en violencia de género** a través de una encuesta online que incluía preguntas abiertas y cerradas. **7 abogados respondieron a nuestra encuesta. También entrevistamos a representantes de la Fundación Ana Bella,** para conocer la perspectiva de la sociedad civil sobre este tema. Todos los datos recogidos durante el trabajo de campo se analizaron mediante un análisis temático para identificar los patrones y temas emergentes. Para estas entrevistas a expertos seguimos el procedimiento de consentimiento mencionado anteriormente, y codificamos a los entrevistados con "L" para los abogados y "C" para los representantes de la sociedad civil.

	Número de participantes	Código de entrevista
Mujer superviviente	31	W
Abogados	7	L
Sociedad Civil	2	C

Tabla 5: Distribución de los participantes en el trabajo de campo.

¹⁵ <https://www.fundacionanabella.org/>

5- Auditoría de VioGén

Acceso al Sistema VioGén

El sistema de valoración de riesgo VioGén sólo puede activarse si la víctima denuncia oficialmente a su agresor ante la policía. Aunque por ley todas las mujeres en España tienen derecho a denunciar una agresión de género, en la práctica solo un pequeño porcentaje lo hace. Según la Macroencuesta de Violencia sobre la Mujer (2019a), sólo el 21,7% de las mujeres mayores de 16 años que sufrieron violencia de género denunciaron a su agresor. El 78,3% restante de las mujeres no denunció y, por tanto, no sus casos no fueron evaluados por el sistema VioGén. Nuestro análisis del informe del CGPJ muestra que, de las 347 víctimas mortales de violencia de género en España entre 2009-2019, sólo 126 de ellas presentaron previamente una denuncia policial. Esto significa que **el 73% de las mujeres que fueron asesinadas por sus (ex)parejas no denunciaron previamente a su agresor y no recibieron una puntuación de riesgo VioGén.**

Conociendo estos datos, nos propusimos averiguar cuáles eran los factores que podrían estar impidiendo a las mujeres denunciar a sus parejas violentas, ya que esto influye mucho en la eficacia del sistema VioGén en general. Existe una amplia gama de obstáculos que pueden disuadir a las mujeres de denunciar oficialmente a su agresor. A partir de los resultados de las entrevistas realizadas, se pueden agrupar estos obstáculos en tres categorías para analizarlos más sistemáticamente:

- Barreras emocionales individuales
- Barreras estructurales de tipo grupal
- Barreras institucionales

Las barreras emocionales individuales contra la denuncia de las agresiones de género están relacionadas con los factores personales de la víctima. Estos pueden existir independientemente de la situación socioeconómica, la educación, la edad, la raza o la etnia de las víctimas. Algunos de estos obstáculos emocionales individuales encontrados durante nuestro trabajo de campo son: no ser consciente de la violencia de género en curso (negación), reticencia o miedo a cambiar el statu quo, evitar la carga emocional de presentar una denuncia policial, miedo al agresor, miedo al juicio de la sociedad, sentimientos de vergüenza, sentimientos de culpa o esperanzas de que el agresor cambie su comportamiento (C1 y C2). Estas barreras son muy comunes y casi todas las víctimas se enfrentan a alguna de ellas.

Las barreras estructurales de tipo grupal tienen lugar cuando una víctima se encuentra en una posición estructuralmente desventajosa para denunciar la agresión debido al grupo al que pertenece. Aunque denunciar la agresión de la (ex)pareja no es fácil para ninguna mujer, hay algunos grupos de mujeres que encuentran más dificultades estructurales que otros. Estos grupos vulnerables incluyen, pero no se limitan exclusivamente, a las mujeres con hijos pequeños, con pocos recursos, con discapacidades, que viven en zonas rurales, las mujeres migrantes y los miembros de la

comunidad LGBT+, así como las diversas intersecciones entre estas categorías. En este informe, nos hemos centrado en tres categorías de vulnerabilidad con el fin de examinar más de cerca los obstáculos a los que se enfrentan en su esfuerzo por denunciar oficialmente a sus agresores. Estos grupos son las mujeres con hijos pequeños, las mujeres socioeconómicamente desfavorecidas y las mujeres inmigrantes indocumentadas.

- **Mujeres con hijos menores de edad:** No sólo las mujeres, sino también sus hijos sufren la violencia de género. Las actualizaciones de 2019 en el protocolo VioGén pretenden captar mejor a los niños en riesgo de sufrir los efectos de la violencia de género. Entre 2019 y 2021 se han activado 7.008 casos VioGén con niños en situación de vulnerabilidad y 2.376 casos con niños en situación de riesgo. Las víctimas de violencia de género con hijos menores de edad se encuentran en una situación difícil cuando el agresor es también el padre de sus hijos. Nuestro trabajo de campo ha demostrado que las mujeres en esta situación están muy preocupadas por la seguridad y el bienestar de sus hijos menores cuando el agresor mantiene su custodia y derecho de visitas. Como dijo W17:

"El agresor no es un buen padre. Se droga, ni siquiera se preocupa por sí mismo. ... Ejerce una violencia indirecta sobre mí y pide la custodia compartida. ... No debería tener los mismos derechos que un buen padre. ... Me trató mal incluso cuando estaba embarazada y con mi hija en brazos. El sistema falla."

En algunos casos, **el miedo a dejar a los niños a solas con el agresor ha retrasado la decisión de presentar una denuncia policial** contra el agresor y pedir una orden de protección. Se ha subrayado el temor a que una posible orden de alejamiento decidida por el juzgado para proteger a la mujer contra su agresor, suponga también que la mujer no pudiera estar junto a sus hijos durante el horario de visitas del padre, que también es el agresor. Como dijo C1:

"Muchas mujeres no denuncian a su agresor por miedo a que después el padre visite a los niños sin que ella esté presente. Cuando presentan la denuncia, se separan del agresor, pero ¿qué pasa con los niños?"

Desde enero de 2022, se ha producido un importante cambio en la ley que ha dificultado que los agresores mantengan sus derechos de custodia compartida¹⁶. La nueva ley establece que " *No procederá la Guarda conjunta cuando cualquiera de los padres esté incurso en un Proceso Penal iniciado por intentar atentarse contra la vida, la integridad física, la libertad, la integridad moral o la libertad e indemnidad sexual del otro cónyuge o de los hijos que convivan con ambos*¹⁷". Aunque es pronto para ver los resultados de este cambio legal, la sociedad civil lo considera un paso importante en la protección de las mujeres y los niños contra la violencia de

¹⁶ Ley 17/2021, de 15 de diciembre, por la que se modifican el Código Civil, la Ley Hipotecaria y el Código Procesal Civil, en materia de régimen jurídico de los animales ("B.O.E." 16 diciembre); del art. 92.7 entre otros.

¹⁷ <https://www.boe.es/eli/es/l/2021/12/15/17>

género. Sin embargo, las mujeres que entrevistamos no conocían este cambio legal.

- **Mujeres socioeconómicamente desfavorecidas:** Aunque las víctimas de la violencia de género pueden provenir de todas las clases socioeconómicas, las mujeres con pocos recursos encuentran especialmente difícil presentar una denuncia policial contra su agresor y buscar un recurso legal. La dependencia económica del agresor es un importante obstáculo que impide a las mujeres denunciar a su agresor, a pesar de que en España existen ayudas sociales y económicas destinadas a las víctimas de violencia de género. Por ejemplo, y en función de sus posibilidades, las víctimas pueden recibir diversas ayudas como la Renta Activa de Inserción (RAI), la Renta Mínima de Inserción (RMI) y, en algunos casos, el Ingreso Mínimo Vital (IMV). Pueden tener acceso prioritario a la vivienda social y pueden recibir una ayuda especial para encontrar un empleo. Pero todas estas medidas requieren cierto tipo de trámites administrativos y pueden estar disponibles siempre y cuando la víctima reciba la orden de protección por parte del tribunal. Por lo tanto, no suelen considerarse adecuadas para "dar el salto fuera de una relación abusiva" (W 16).

La educación es otro factor importante que debe considerarse al evaluar el acceso de las víctimas al sistema VioGén. El papel de la educación debe considerarse junto con otros factores interrelacionados, como las oportunidades de empleo, las fuentes económicas, el momento de formación de la familia, la selección de la pareja y las actitudes de la familia hacia la igualdad de género (Weitzman, 2018). El nivel de educación tiene un papel importante en el apoyo a la alfabetización burocrática de las mujeres y su comprensión de cómo/dónde pedir ayuda formal contra la violencia de género. En este sentido, las mujeres con un nivel educativo más bajo tienen más dificultades para presentar una denuncia contra su agresor. Como explicó L7: "el nivel de estudios hace que sea más fácil entender lo que se pide y explicar cómo se siente y lo que sufre". Dicho esto, los expertos destacan que las mujeres profesionales con mayor nivel de estudios también evitan denunciar a su agresor, porque pretenden proteger su reputación social y su carrera profesional. En palabras de una representante de la sociedad civil que entrevistamos, "las mujeres sin estudios esperan 8 años y las mujeres con doctorado esperan 13 años para denunciar a su agresor". Es decir, "a las mujeres profesionales con estudios les resulta muy difícil denunciar a su agresor, porque creen que eso cuestionaría su carrera profesional" (C 1).

- **Mujeres inmigrantes:** Las mujeres inmigrantes encuentran desafíos específicos en el proceso de denuncia de la violencia de género. Estos retos incluyen: haber sido criadas en culturas que carecen de la noción de igualdad de género; estar encerradas en el círculo del agresor sin tener su propia red de apoyo familiar y de amigos y/o autonomía económica; no dominar los idiomas del país de acogida; carecer del conocimiento de las leyes, reglamentos y procesos oficiales en España; y estar en situación irregular en España y temer la deportación. Como nos explicó W3, su situación irregular le impidió buscar ayuda. Su agresor le decía constantemente que no podía hacer nada, porque si no la deportaban. Ella dijo:

"No me puse en contacto con ninguna organización, no tuve ningún apoyo y estuve completamente sola".

La legislación española (Ley Orgánica 4/2000, de 11 de enero, sobre derechos y libertades de los extranjeros en España y su integración social) proporciona una protección especial a las mujeres inmigrantes irregulares que son víctimas de violencia de género¹⁸. Según esta ley, si la presencia de la situación de violencia de género es confirmada por el tribunal y se asignan medidas de protección, esto abre el camino para la regularización de la estancia de las mujeres indocumentadas en España. Pero si el tribunal no confirma la situación de violencia de género y no asigna la orden de protección, entonces se inicia el proceso de deportación. Si bien la ley proporciona una importante protección a las víctimas de violencia de género indocumentadas, también crea un gran dilema si se tiene en cuenta que probar un caso de violencia de género no siempre es sencillo y a menudo requiere de pruebas debidamente recopiladas y una buena asistencia legal para preparar y presentar el caso ante el tribunal. El hecho de que VioGén clasifique aproximadamente el 45% de los casos como casos sin riesgo ("no apreciado"), valida el dilema que las mujeres indocumentadas pueden tener para buscar protección.

Barreras institucionales: Cuando una mujer supera los obstáculos anteriormente mencionados y decide denunciar a su agresor, la actitud de los agentes de policía desempeña un papel importante en su experiencia global. Las actitudes poco acogedoras, la falta de empatía y comprensión, y el comportamiento con prejuicios por parte de la policía pueden incluso hacer que las víctimas abandonen la comisaría sin poder denunciar oficialmente la agresión. Nuestro trabajo de campo muestra que las mujeres tienen experiencias y percepciones muy diferentes de los agentes de policía al presentar la denuncia de la agresión. En un caso, la víctima describió su experiencia con los policías como "estoy muy satisfecha, me atendieron y me calmaron" (W 16). En otros casos, las víctimas expresaron su experiencia con la policía como "desagradable" (W 4), "pasiva y sin empatía" (W 10) y "sin preparación" (W 21) para tratar los casos de violencia de género. Tres de los casos de nuestro trabajo de campo mostraron cómo un mismo caso de violencia de género recibió un trato diferente por parte de distintos policías. W 13, víctima de violencia de género de origen inmigrante, tuvo que viajar a otra ciudad para poder denunciar la agresión de su pareja. Sólo después de viajar a otra ciudad, pudo ser evaluada por el sistema VioGén, y recibió una puntuación de alto riesgo.

"En mi primer intento, no aceptaron mi queja y me dijeron que me volviera a mi país y que todo era culpa mía. Después, en otra ciudad, me trataron muy bien". (W 13)

En otro caso, W14 tuvo que esperar 12 horas en la comisaría por el turno de oficio, para poder evitar al policía que se dirigió a ella con poca amabilidad y con el que se encontró inicialmente y presentar la denuncia con otro agente.

¹⁸ <https://violenciagenero.igualdad.gob.es/informacionUtil/extranjeras/proteccion/home.htm>

"Me fui sin presentar la denuncia y tuve que esperar mucho tiempo. La actitud del agente de policía estuvo llena de prejuicios. Me sentí incómoda y culpable. ... Un trato horrible y me llamó 'mentirosa'. ... Me quedé allí hasta el turno de oficio. El otro policía me atendió muy bien y pude presentar la denuncia. Tardé 12 horas, desde las 4 de la tarde hasta las 4 de la mañana, en presentar la denuncia". (W 14)

El trabajo de campo realizado para esta Auditoría Externa señala la importancia de evaluar los sistemas técnicos también desde la perspectiva de sus usuarios e incluso antes de que el sistema técnico intervenga, ya que lo que ocurre antes y después de la solución técnica puede tener un gran impacto en los índices de participación, la representatividad y la evaluación general de un problema concreto.

La entrevista y el formulario de VioGén

Cuando una mujer denuncia oficialmente un caso de violencia de género ante la policía (o ante la guardia civil), **su caso se activa y se evalúa mediante el formulario VPR**. Este formulario incluye 35 indicadores de riesgo que se evalúan como "presente" o "no presente". Una vez que la víctima responde a las preguntas que cubren estos indicadores de riesgo, **el sistema asigna una puntuación de riesgo que evalúa su probabilidad de encontrarse con futuras agresiones por parte del mismo agresor**. De este modo, el sistema VioGén busca estandarizar el proceso de evaluación de riesgo en todo el país realizando la misma VPR en todos los casos de violencia de género, supuestamente de la misma manera, y proporcionando una puntuación de riesgo objetiva que, presumiblemente, está libre de los prejuicios de los agentes y de su nivel de experiencia. **El sistema VioGén se basa en el supuesto de que las mujeres que sufren violencia de género entienden y responden con claridad a los 35 puntos del formulario VPR y los agentes de policía transforman objetivamente las declaraciones de las mujeres en respuestas binarias (presente/no presente)** en el formulario VPR. Pero en la realidad, el proceso rara vez funciona de este modo idealizado. En nuestro trabajo de campo, **más del 80% de las mujeres entrevistadas señalaron diferentes problemas con el formulario VioGén**. Esto significa que la calidad de los datos introducidos en el sistema algorítmico podría verse comprometida durante la fase de generación de datos, lo que daría lugar a posibles fuentes de sesgo y tergiversación dentro del sistema.

A continuación se enumeran algunos de los problemas más destacados identificados durante el trabajo de campo:

- **Falta de información:** Para que las mujeres entiendan y contesten correctamente el formulario VioGén, deben tener información adecuada sobre qué es el sistema VioGén, cómo funciona, qué hace y qué se espera de ellas. Los agentes de policía/guardia civil tienen un papel fundamental en la información de las víctimas sobre el sistema antes de empezar a realizar el formulario. Las supervivientes de la violencia de género que entrevistamos, especialmente en el caso de presentación de denuncia por primera vez, destacaron que no estaban completamente informadas sobre el sistema VioGén. Una de las entrevistadas

afirmó que "Ellos (los policías) utilizaban la palabra VioGén, pero yo no sabía lo que significaba esa palabra" (W 29). Otra entrevistada (W 28) pensaba que la puntuación de riesgo VioGén es algo que calcula los metros de una orden de alejamiento. **El 35% de las mujeres que entrevistamos no estaban informadas sobre su puntuación de riesgo VioGén y por lo tanto no sabían qué nivel de riesgo les asignaba el sistema.**

- **Momento en el que se realiza el formulario VioGén:** El formulario VioGén se realiza en el momento de presentar la denuncia policial por violencia de género. Dado que muchas mujeres que sufren violencia de género llegan a la comisaría y presentan la denuncia justo después de un incidente violento, se encuentran en un estado de shock y no están preparadas ni física ni emocionalmente para dar respuestas precisas. Nuestras entrevistadas mencionaron cómo les resultaba difícil recordar toda la información pasada, organizar sus pensamientos y dar respuestas exhaustivas a las preguntas de VioGén. En uno de los casos, **W 25 tuvo un ataque de ansiedad durante el formulario VioGén y fue llevada al hospital. Otra entrevistada (W 23) describió el momento de responder a las preguntas de VioGén como "surrealista"** y explicó que en ese momento no sabía lo que estaba haciendo. Otra entrevistada lo recuerda como "un momento nublado con preguntas absurdas en las que se cometen errores al rellenar el cuestionario". (W 23)
- **Preguntas de la VPR:** Nuestras entrevistadas destacaron que aunque algunas preguntas de VioGén eran claras y directas, otras eran **"ambiguas"** (W15) y no tenían mucho sentido en ese contexto. Algunos abogados consideran que las preguntas de la VPR son **"rígidas"**, que no permiten explicaciones (L2), y **"genéricas"** (L4) y que carecen de capacidad para referirse adecuadamente a situaciones individuales. Según uno de los abogados entrevistados, citado anteriormente (L7), la comprensión y la respuesta a las preguntas del VPR dependen en gran medida del nivel de educación de la víctima.
- **Falta de apoyo legal:** Las mujeres que sufren violencia de género tienen derecho a solicitar un abogado para presentar su denuncia, si bien, sólo unas pocas conocen este derecho (C1). A menudo se les proporciona un abogado justo antes de que su caso sea visto por un tribunal. Por lo tanto, responden a las preguntas de VioGén sin poder hablar con un abogado que les informe sobre el proceso y sobre lo que se espera de ellas.
- **Falta de apoyo psicológico:** Una de las cuestiones importantes que se plantea durante la entrevista VioGén es el malestar emocional que genera por parte de las víctimas. No contar con ayuda y apoyo psicológico, antes y durante este proceso, agrava esta carga emocional, sobre todo cuando las preguntas de la VPR son realizadas por policías que no están especializados en violencia de género y carecen de la formación adecuada sobre el sistema VioGén.

Como en cualquier sistema de datos, la calidad de los datos introducidos es clave para la calidad, representatividad y equidad de los resultados. Los factores identificados se

traducen en **entradas pobres o sesgadas en el 80% de los casos**. Aunque no podemos generalizar a partir de nuestra pequeña muestra, los resultados obtenidos apuntan a la necesidad urgente de revisar las condiciones de acceso de las mujeres al sistema y al formulario, ya que actuar en esta fase es clave para garantizar la calidad de los datos. Como se ha mencionado anteriormente, **la calidad de los datos que se introducen en el sistema algorítmico puede verse comprometida durante la etapa de generación de los datos** (es decir, cuando las mujeres responden a los indicadores de riesgo de la VPR), **lo que da lugar a posibles fuentes de sesgo y tergiversación dentro del sistema**.

Hemos podido complementar esta evaluación cualitativa con los **resultados del análisis de datos realizado sobre la base de datos de 1.000 casos de homicidio (IPH)** descrito anteriormente, lo que nos ha permitido conocer los aspectos clave desde una perspectiva técnica. El formulario VioGén VPR_{5,0} ha sido dotado de un protocolo adicional (la escala H) diseñado para identificar a las víctimas potenciales de IPH. Según sus desarrolladores, tiene una alta capacidad predictiva.¹⁹ Sin embargo, y como sugieren los hallazgos cuantitativos que se presentan a continuación, la IPH es un mecanismo singular de violencia de género que, en la mayoría de los casos, no es el resultado final de una escalada de maltrato histórico, sino un evento puntual de violencia extrema, lo que hace más difícil su predicción. Por ello, y a pesar de la denegación de la petición de Éticas de acceder a la base de datos original de VPI (Violencia de Pareja Íntima) en España por parte del ministerio, se ha realizado una valoración crítica de la escala H a partir de los datos recogidos en el registro público de víctimas de homicidio (IPH) difundido por el CGPJ (Consejo General del Poder Judicial).²⁰

- **La escala H y la predictibilidad del homicidio:** El objetivo inicial del análisis era identificar posibles fuentes de sesgo y varianza en la prevalencia de homicidios (IPH) a lo largo de los diferentes grupos protegidos. Como Éticas no ha podido tener acceso a la base de datos original, hemos utilizado los datos disponibles para identificar cuestiones que apuntan a posibles problemas que deberían abordarse en el sistema VioGén. Los datos del CGPJ sólo permiten ver lo que hemos definido como "zona de muerte" (véase la figura 1). **El objetivo principal era identificar las variaciones de sensibilidad entre las clases protegidas, ya que esto indicaría problemas de sesgo y representatividad en las predicciones realizadas por VioGén**, con el fin de identificar y discutir mecanismos de corrección. Los datos utilizados para llevar a cabo el estudio captan los efectos del sistema global (incluida la decisión del juez) y no los del algoritmo exclusivamente. Por lo tanto, y para que los resultados sean generalizables, este estudio debería ampliarse a la base de datos original para identificar y abordar posibles patrones de discriminación en VioGén.

La figura 2 es una representación abstracta de la población de casos de violencia de género denunciados (es decir, los casos que son procesados por el sistema VioGén). La

¹⁹ The main performance metrics presented in (López-Ossorio, González-Álvarez, et al., 2020) claim a sensitivity of 84%, a specificity of 60%, an OR = 8.130, an AUC = .80, a PPV = .19 and a NPV = .97

²⁰ Disponible en: <https://www.poderjudicial.es/cgpj/es/Temas/Violencia-domestica-y-de-genero/Actividad-del-Observatorio/Datos-estadisticos>

mitad izquierda de la imagen representa los elementos relevantes (es decir, los casos que corren el riesgo de sufrir más violencia), siendo los falsos negativos (**FN**) el subconjunto de casos predichos como negativos pero realmente positivos, y los verdaderos positivos (**TP**) el subconjunto de casos positivos clasificados como tales. La mitad derecha del cuadro representa los elementos negativos reales (es decir, los casos que no corren el riesgo de sufrir más violencia), siendo los falsos positivos (**FP**) el subconjunto de casos predichos como positivos pero realmente negativos, y los verdaderos negativos (**TN**) el subconjunto de casos negativos clasificados como tales. Este análisis se centra en la "zona de muerte", un subconjunto que incluye exclusivamente los casos TP y FN que resultaron en homicidio.

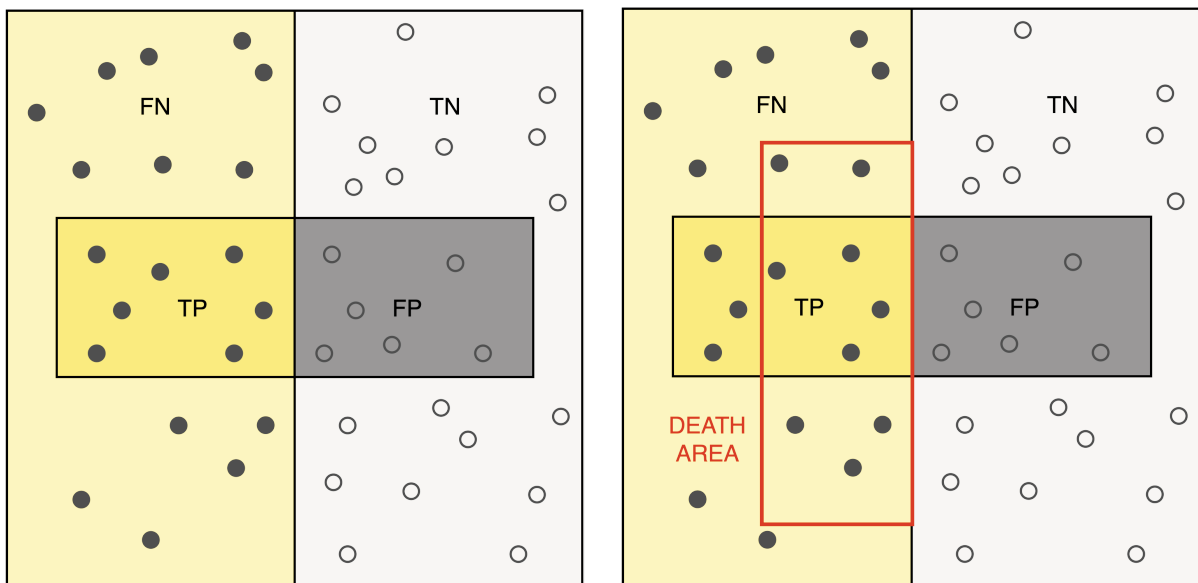


Figura 2: Representación abstracta de la población (izquierda) vs al área de visibilidad debido a la disponibilidad de datos (derecha) . Las áreas no son proporcionales a la población real. Fuente: Los autores²¹.

Para realizar el análisis, tras limpiar los datos disponibles, los segmentamos en los siguientes grupos:

- Falsos negativos: se refiere a los casos en los que las mujeres asesinadas habían presentado previamente una denuncia policial pero no recibieron una orden de protección porque VioGén no apreció el riesgo ("riesgo no apreciado") o predijo un riesgo bajo. El 56% de las víctimas (N=126) entran en esta categoría. Su distribución según los grupos aportados por el CGPJ en el conjunto de datos es la siguiente:

²¹ Nótese que hay casos verdaderamente positivos que quedan fuera del área de muerte. Este conjunto está constituido por casos de víctimas de IPV que fueron identificados como de riesgo, pero la reincidencia no fue en forma de IPH.

Nacionalidad de las víctimas	Total FN (homicidio)	
	ESP = 46	nESP = 25
Nacionalidad del agresor	ESP = 45	nESP = 26
Hijos (con hijos/sin hijos) *	C = 52	nC = 15
Hijos menores (con hijos menores/ sin hijos menores)*	mC = 38	nmC = 29
Edad (-54,+55) **	x≤54 = 61	x>54 = 9

* En cuatro casos no estaba especificado si la víctima tenía hijos o no

** Una víctima era menor de edad y no ha sido incluida en el grupo x≤54

Tabla 6: Recuento de casos de Falsos Negativos (es decir, denuncia policial previa sin medidas de protección que resultó en homicidio) calculado para cada atributo protegido a partir de los datos publicados por el CGPJ.

Fuente: Los autores.

- Protección insuficiente: se refiere a los casos en los que las mujeres asesinadas habían presentado previamente una denuncia policial y recibieron una orden de protección. El 44% de las víctimas (N=126) entran en esta categoría. Esta clase es un subconjunto de los verdaderos positivos en los que las medidas de protección no fueron suficientes para evitar el homicidio. Su distribución según los grupos proporcionados por el CGPJ en el conjunto de datos es la siguiente:

Nacionalidad de la Víctima	Total IP (homicidio)	
	ESP = 34	nESP = 21
Nacionalidad del agresor	ESP = 36	nESP = 19
Hijos (con hijos/sin hijos) *	H = 48	nH = 6
Hijos menores (con hijos menores/ sin hijos menores)*	Hm = 30	nHm = 24
Edad (-54,+55)	x≤54 = 50	x>54 = 5

* En un caso no se especificaba si la víctima tenía hijos o no

Tabla 7: Recuento de casos de Protección Insuficiente (es decir, informe policial previo con medidas de protección que ha resultado en homicidio) calculado para cada atributo protegido, a partir de los datos publicados por el CGPJ. Fuente: Los autores

Los recuentos estratificados de los casos de protección negativa e insuficiente permiten calcular la sensibilidad en función de los distintos atributos protegidos. De este modo, la comparación de estos con el valor de la sensibilidad en toda la población permite identificar el sesgo y la discriminación. La sensibilidad puede formularse como:

$$\text{Sensibilidad} = \frac{TP}{TP + FN}$$

A partir de los datos disponibles, observamos una diferencia significativa en la sensibilidad respecto a una dimensión concreta: tener hijos frente a no tenerlos. Esto significa que las víctimas que no tenían hijos fueron percibidas significativamente como de menor riesgo. A primera vista, esto es incoherente con la configuración de la escala H en VPR_{5,0}, donde ninguno de los 13 predictores incluidos refleja la presencia o ausencia de hijos.

	Sensibilidad (homicidio)		$ x_1-x_2 /rc$
Sensibilidad (TOTAL)	$rc = 55 / (55+71) = 0.437$		-
Nacionalidad de la Víctima	ESP = 0.425 (34/80)	nESP = 0.457 (21/46)	7%
Nacionalidad del Agresor	ESP = 0.444 (36/81)	nESP = 0.422 (19/45)	5%
Hijos (con hijos/sin hijos)	C = 0.48 (48/100)	nC = 0.285 (6/21)	44.62%
Hijos menores (con hijos menores/ sin hijos menores)*	mC = 0.442 (30/68)	nmC = 0.453 (24/53)	2.5%
Edad (-54,+55)	$x \leq 54 = 0.450 (50/111)$	$x > 54 = 0.357 (5/14)$	21.28%

Table 8: Sensibilidad calculated for each protected attribute, from the data released by CGPJ. Source: The Authors

El homicidio de la pareja íntima (IPH) es un crimen atroz, pero constituye un pequeño porcentaje de la población total de víctimas de la violencia de pareja íntima (IPV). Tomando, por ejemplo, el primer semestre de 2021, el número total de informes policiales por IPV fue de 70.723, mientras que el número de casos de IPH en el mismo período ascendió a 21 (aproximadamente el 0,03%). **El principal problema relacionado con la IPH, sin embargo, no es el tamaño de este grupo, sino el hecho de que la mayoría de los casos de IPV no son casos previamente denunciados.** Los datos aportados por el CGPJ muestran que el número de casos de IPH sin denuncia policial previa entre 2009 y 2019 (acotados para las CCAA. de España en las que actúa VioGén) ascendió a 347 (73%), mientras que el número de víctimas de IPH que habían presentado previamente una denuncia policial ascendió a 126 (27%). **Esto significa que el 73% de las víctimas de homicidio (IPH) durante ese periodo no habían presentado previamente una denuncia policial, y por tanto no se les asignó una puntuación VPR o VPER.** Además, si consideramos a todas las víctimas de IPV de ese periodo, **sólo 55 (el 11,6% sobre el total) recibieron algún tipo de protección policial.** Estos hallazgos son consistentes con la Macroencuesta de Violencia contra la Mujer, como se ha señalado al principio de este apartado, lo que apunta a un problema estructural en cuanto a la visibilidad de la IPV que VioGén debe tener en cuenta. Esto refuerza los hallazgos de nuestro enfoque cualitativo, descrito anteriormente en "Acceso a VioGén", y señala la necesidad urgente de trabajar para eliminar o disminuir las barreras existentes que impiden a las mujeres buscar protección.

La capacidad predictiva del VPR_{5,0}-H no se ve cuestionada por estos resultados: estudios anteriores muestran cómo la validez predictiva de la herramienta VioGén ha ido

mejorando, alcanzando un valor de AUC (Área Bajo la Curva) de 0,80 en su nueva escala de homicidios (escala H).²² Sin embargo, **las investigaciones existentes han evaluado el rendimiento del sistema VioGén a nivel agregado y no han estudiado si el sistema se comporta de forma diferente para los distintos grupos de mujeres según su edad, origen o si tienen hijos o no. En este estudio, hemos llevado a cabo un análisis estratificado de los casos de IPH entre diferentes categorías de mujeres con el propósito de indagar sobre los tipos de vulnerabilidad que podría producir y aplicar el sistema.** Al hacerlo, hemos comprobado que a las mujeres sin hijos se les asigna sistemáticamente una puntuación de riesgo menor que a las mujeres con hijos. Y lo mismo debería hacerse con VioGén, para comprender mejor cómo se calibra y si se tienen en cuenta las disparidades entre los grupos protegidos.

Además, el hecho de que sólo 1 de cada 4 víctimas de homicidio (IPH) entre en el sistema VioGén a través de informes previos de IPV plantea cuestiones difíciles sobre la sensibilidad de los casos de riesgo extremo por parte de VioGén. Una de las razones de la aparentemente alta capacidad predictiva de la escala H podría ser que los indicadores de riesgo extremo (especialmente dentro de la escala H) han sido adaptados para identificar un subconjunto particular de casos de IPH, es decir, aquellos que derivan de una acumulación de violencia de una historia más larga de VPI, donde el peso de otros factores es suficiente como para discriminar entre las etiquetas de riesgo extremo y las de riesgo menor. Pero éstos constituyen 1 de cada 4 casos. Para el otro 75% de los casos, debería realizarse un análisis más exhaustivo del contexto y de los perfiles para identificar las causas de esta falta de visibilidad, y se debería subrayar la limitación de la validez de la escala H sobre la población total de víctimas de IPH: la alta precisión predictiva de la escala H sólo cubre el 25% de los casos de IPH. **Esto significa que la mayoría de los casos de IPH (75%) siguen sin ser tratados por el nuevo protocolo VPR_{5,0}-H.**

Percepción del Algoritmo de VioGén

Mientras que la precisión predictiva del sistema VioGén está técnicamente estudiada por sus desarrolladores (López-Ossorio et al., 2017, 2019), se ha prestado menos atención a la percepción y la confianza en el sistema por parte de los principales interesados. El estudio de 2015 de Gonzales et. al. (2015) se centra principalmente en la satisfacción de las víctimas con la actuación policial, pero no en su percepción del algoritmo VioGén y su puntuación de riesgo asignada. Nuestro trabajo de campo, **las entrevistas con supervivientes de violencia de género y sus abogados indican cierta desconfianza general del funcionamiento del algoritmo VioGén.**

- **El 48% de las mujeres que entrevistamos evaluaron negativamente su experiencia con el sistema; el 32% de ellas destacaron tanto aspectos negativos como positivos, y sólo el 19% de ellas evaluaron positivamente su experiencia general con el sistema VioGén.**

²² El área bajo la curva (AUC) es uno de los modos más comunes de evaluar la validez predictiva de los instrumentos de evaluación de riesgos. Su valor puede oscilar entre 0 (predicción negativa perfecta) y 1 (predicción positiva perfecta), y 0,50 indica una predicción casual (Douglas et al., 2005)

- **Todos los abogados que respondieron a nuestra encuesta mostraron también poca confianza en el sistema VioGén.**

Aunque nuestra muestra no es representativa de la población más amplia de supervivientes y abogados y, por lo tanto, nuestras conclusiones no son generalizables, nuestro trabajo de campo plantea cuestiones importantes que deben estudiarse de forma más sistemática, e idealmente abordarse a nivel institucional.

Una de las principales preocupaciones con el algoritmo VioGén es que aproximadamente el 45% de los casos reciben la puntuación de "sin riesgo". En el contexto de la violencia de género, la categoría de "sin riesgo" es ya una cuestión muy discutida (Kropp, 2004). Tal y como mencionan nuestras encuestadas, el acto de presentar una denuncia policial contra el agresor es en sí mismo una conducta de riesgo que puede provocar reacciones. Cuando una víctima recibe una puntuación de riesgo "no apreciado" o "bajo riesgo", se crea una gran brecha entre cómo se autoevalúa y percibe su riesgo de revictimización y lo que predice el sistema. Como explicó un abogado entrevistado "Las víctimas se sienten engañadas una vez que presentan la denuncia y que no se les cree". (L5)

La otra cuestión que se puso de manifiesto durante nuestras entrevistas fue la percepción de cómo el algoritmo VioGén infravalora la violencia psicológica y quizás las formas más nuevas de violencia no física (como el acoso a través de las redes sociales), poniendo el énfasis en la violencia física. En este sentido, existe la creencia generalizada, tanto por parte de las víctimas como de sus abogados, de que los parámetros de VioGén no tienen en cuenta adecuadamente la violencia psicológica. Como afirma L4: "No hace falta una paliza, ni una agresión, para que exista el riesgo. Parece que los parámetros olvidaron el maltrato psicológico".

Las cuestiones de percepción son importantes ya que, en última instancia, la confianza en un sistema como VioGén tendrá un efecto importante en la calidad de los datos que son recibidos de las mujeres. Como todos los sistemas que hacen un uso intensivo de datos, el sistema VioGén se basa en datos humanos, por lo que si las fuentes de los datos se encuentran asustadas, son reticentes o dudan, la calidad de las entradas de datos se verá afectada. Nuestros medios y los datos disponibles no nos permiten ir más allá en nuestro análisis, pero proporcionan, en nuestra opinión, un fuerte argumento para una mejora sistemática de los puntos débiles que nuestra Auditoría Externa ha identificado.

Resultados y recomendaciones

Cuando iniciamos la Auditoría Externa de VioGén, nos preocupaba la transparencia, la supervisión independiente, la responsabilidad, el compromiso de los usuarios finales y la transición al ML. Tras realizar la auditoría, podemos confirmar que:

- **VioGén no es transparente.** No pudimos acceder a ningún dato o información del sistema más allá de lo producido por los expertos que participaron en la definición del sistema. Ni los auditores externos ni los grupos de mujeres tienen ningún tipo de acceso a los datos de VioGén. Para un sistema financiado con fondos públicos y de gran impacto como VioGén, esto resulta inaceptable.
- **VioGén no ha sido evaluado ni auditado de forma independiente.** Los recursos y las encuestas al alcance público sobre la validez y la conveniencia de VioGén han sido realizados por personas que trabajan o tienen intereses en el ministerio y las fuerzas policiales. Los auditores o investigadores externos no tienen ninguna vía oficial o pública para acceder a los datos, y el acceso parece ser proporcionado por el Ministerio a su discreción.
- **VioGén no rinde cuentas.** Aunque el Ministerio del Interior considera que VioGén es un sistema de recomendación, los altos índices de aceptación de los resultados algorítmicos (95%) apuntan a un sistema automatizado, que debería someterse a un mayor escrutinio según el Régimen Jurídico de la Función Pública.
- **VioGén no involucra a los usuarios finales.** En nuestro trabajo de campo hemos constatado que las mujeres y las organizaciones de mujeres nunca han sido consultadas sobre el sistema, ni en su fase de diseño ni posteriormente durante las diferentes decisiones sobre cómo modificar el sistema VioGén. Además, hemos comprobado que el 80% de las mujeres que han utilizado el sistema tienen comentarios negativos sobre el mismo. No están informadas de lo que hace ni de cómo funciona, lo que genera desconfianza.
- **La transición de VioGén a ML plantea nuevos interrogantes.** Aunque en la literatura académica se explora el proceso de transición a una versión de VioGén basada en el ML, la naturaleza y el alcance de la colaboración entre la empresa de software SAS y el ministerio no se han divulgado públicamente. Si bien la falta de un debate público y abierto sobre este proceso es en sí mismo preocupante, el hecho de que la evolución técnica del sistema se desvincule de la investigación y la supervisión técnica está destinado a generar más problemas.

El proceso de auditoría también nos ha permitido ir más allá de nuestras preocupaciones iniciales para identificar nuevos problemas que merecen especial atención.

En primer lugar, queremos destacar que, a través de esta auditoría, hemos comprobado que el sistema VioGén adapta la agrupación de las valoraciones de riesgo a los recursos disponibles. Esto significa que el sistema sólo da el número de puntuaciones de riesgo "extremo" que puede permitirse, por lo que **los posibles recortes en la financiación tendrían un impacto directo y cuantificable en las posibilidades de que las mujeres reciban protección efectiva después de solicitar protección policial.** Como el número de casos de VioGén crece cada año, hay más mujeres que reciben protección policial.

Mientras que en 2015 unas 3.000 mujeres recibieron protección policial -con puntuaciones de riesgo medio, alto y extremo-, en 2021 esta cifra se elevó hasta casi 9.000 mujeres. Sin embargo, sigue existiendo una gran brecha entre las mujeres que reciben protección policial respecto a las que no la reciben, a pesar de denunciar el caso de violencia de género a la policía. En cuanto a la calibración,²³ **nos preocupa el número de casos que el sistema VioGén "descarta" al otorgarles una puntuación de riesgo "inapreciable"**. Tal y como está diseñado actualmente, la puntuación de riesgo que da VioGén no sólo viene determinada por los hechos objetivos que pretende desvelar el formulario, sino también por la distribución global de los casos de violencia de género, que viene determinada por los recursos disponibles. Por ello, en 2021, **sólo 1 de cada 7 mujeres que acudieron a la policía en busca de protección la recibió realmente.**²⁴

Esto es aún más grave si tenemos en consideración las barreras que hemos identificado para acceder a VioGén, que es una de las razones por las que sólo el 21,7% de las mujeres víctimas de violencia de género solicitan protección. Estas cifras suponen que **sólo el 3% de las mujeres víctimas de violencia de género reciben una puntuación de riesgo "media" o superior y, por tanto, protección policial efectiva.**

En segundo lugar, hemos identificado que el hecho de no tener hijos tiene un impacto significativamente negativo en la percepción de los casos de riesgo extremo. Nuestro análisis de datos pone de manifiesto que a **las mujeres que fueron asesinadas por sus parejas y no tenían hijos se les asignó sistemáticamente una puntuación de riesgo más baja que a las que sí los tenían, con una diferencia de sensibilidad entre grupos del 44%.**

También queremos poner en duda la representatividad del valor AUC de la escala H que afirman los investigadores principales de VioGén. Si bien es cierto que con los datos disponibles la escala H es capaz de identificar los casos de riesgo extremo que pueden desembocar en un homicidio, **el hecho de que sólo 1 de cada 4 casos de homicidio se produzca después de un informe policial previo indica que la mayoría de las víctimas de homicidio seguirán estando desprotegidas, incluso con el despliegue de VPR_{5,0}-H.** Esto significa que, aunque VioGén está ahora mejor preparado para identificar ciertos casos de riesgo extremo, la mayoría de los casos de homicidio seguirán sin ser atendidos.

En cuarto lugar, hemos observado que **VioGén es, en la práctica, un sistema automatizado con una supervisión humana mínima e inconsistente.** Los agentes de policía sólo aumentan el riesgo observado en el 5% de los casos, cifra que disminuye ante la sobrecarga de trabajo. Esto es muy problemático, ya que una implementación no responsable de la supervisión humana puede provocar problemas de explicabilidad y transparencia. Si los agentes de policía no tienen instrucciones claras sobre cuándo y cómo intervenir, su papel puede introducir sesgo en el sistema, y las mujeres pueden recibir diferentes puntuaciones dependiendo de quién atienda su caso. Evaluar el papel

²³ El error de calibración puede entenderse aquí como la diferencia entre las probabilidades previstas de los resultados y las probabilidades reales de esos resultados

²⁴ Todas las cifras proceden del Boletín Estadístico Mensual elaborado por la Delegación del Gobierno contra la Violencia de Género.

de la supervisión humana a lo largo del tiempo debería formar parte de cualquier esfuerzo de auditoría y transparencia.

Aunque nuestra muestra no es representativa de la población más amplia de víctimas y abogados y, por lo tanto, nuestras conclusiones no son generalizables, nuestro trabajo de campo plantea cuestiones importantes que deben estudiarse de forma más sistemática, e idealmente abordarse a nivel institucional.

A la luz de lo anterior, **recomendamos:**

- Promulgar políticas destinadas a **eliminar las barreras de acceso a nivel individual, grupal e institucional**. Las mujeres que sufren violencia de género deben poder denunciar esta agresión libremente y buscar ayuda oficial. Dado que existe un amplio abanico de barreras que impiden el acceso de las mujeres al sistema VioGén, las medidas para paliar estas barreras deben abordar las cuestiones específicas de cada nivel y aportar soluciones efectivas.
- **Aumentar el número de agentes especializados en violencia de género**. Los agentes de policía tienen un enorme impacto en el funcionamiento del sistema VioGén en la práctica, en cómo se vive y en cómo lo perciben las víctimas. Actualmente, 27.000 agentes participan en el seguimiento de VioGén, pero sólo 2.000 están especializados en violencia de género²⁵. Es urgente aumentar el número de policías especializados en violencia de género.
- Cuando las mujeres acceden efectivamente al formulario VioGén, deben recibir **apoyo jurídico y psicológico**. Las mujeres que sufren violencia de género suelen carecer de información sobre sus derechos y deberes legales. Aunque tienen derecho a pedir un abogado, esta ayuda legal suele llegar en la fase del juicio. Sin embargo, un apoyo jurídico más temprano, antes y durante la entrevista de VioGén, ayudaría a las mujeres a evaluar y presentar mejor su situación. Además, no es fácil para las mujeres denunciar a sus (ex)compañeros sentimentales y responder al formulario VioGén. Como el proceso tiene una gran carga emocional y es angustiante, las víctimas deben recibir apoyo psicológico desde el principio, en algunos casos en la fase de presentación de la denuncia.
- Para abordar los problemas de **responsabilidad**, recomendamos **acompañar la puntuación VioGén con la justificación de los agentes de policía**. Como los agentes de policía llevan a cabo el protocolo VioGén, conocen bien el caso y sus detalles contextuales, y son ellos quienes aprueban la puntuación VioGén. Su punto de vista profesional es importante y no puede expresarse únicamente en un valor de puntuación de riesgo. Por lo tanto, la puntuación de riesgo VioGén debería ir acompañada de un informe policial que justifique la puntuación y aporte una opinión profesional adicional si es necesario, tanto cuando deciden no modificar la puntuación como cuando utilizan su opción discrecional para aumentarla. Esto apoyaría la responsabilidad de la Policía en la evaluación de

²⁵<https://english.elpais.com/society/2021-11-26/what-life-is-like-for-the-women-and-children-in-spain-under-police-protection-due-to-gender-violence.html>

riesgos. También ayudaría a los jueces en su esfuerzo por comprender e interpretar las puntuaciones de riesgo de VioGén.

- Como demuestra este informe, es muy valioso **auditar los sistemas automatizados**, especialmente cuando se financian con fondos públicos y tienen un impacto en comunidades vulnerables como las mujeres víctimas de la violencia de género. El impacto que un sistema automatizado de valoración de riesgos tiene sobre las víctimas de violencia de género, pero también sobre la sociedad en su conjunto, justifica que se habiliten auditorías de terceros. Hacerlo no sólo confirmaría el buen funcionamiento del sistema, sino que permitiría a investigadores independientes identificar posibles consecuencias perjudiciales de su uso. La auditoría de algoritmos se perfila como una de las prácticas que pueden garantizar la responsabilidad de los sistemas técnicos y promover la confianza. **Las instituciones públicas pueden y deben liderar la promoción de prácticas de datos responsables.**
- Utilizar datos históricos para inferir patrones de violencia de género. Dado que VioGén se apoya en una base de datos considerablemente grande para realizar sus predicciones, sería recomendable contrastar la configuración real con técnicas avanzadas de análisis de datos, con el fin de validar los factores de riesgo y los identificadores utilizados por el sistema para evaluar el riesgo. Además, estas evaluaciones deberían ponerse a disposición del público en general para fomentar la transparencia y la confianza en el sistema.
- Si VioGén pretende **incorporar técnicas de ML, esto debería ir acompañado de un debate público sobre los beneficios y riesgos** de ello. La investigación académica ha demostrado que algunas técnicas de ML superan el diseño actual de VioGén (González-Prieto et al., 2021). A la luz de esto, debería promoverse desde el ministerio un espacio de deliberación pública sobre los beneficios y riesgos de migrar hacia un enfoque de ML -en sus diferentes formas-, donde todos los interesados y expertos pudieran discutir su conveniencia.
- Por último, es urgente **recabar periódicamente la opinión de las víctimas y otras partes interesadas**. La evaluación del rendimiento del sistema no debe limitarse a su análisis técnico, sino que también debe tener en cuenta las experiencias y percepciones de las partes interesadas que pasan por el sistema o trabajan con él. El sistema VioGén no puede mejorarse únicamente a través de la contemplación científica desde una butaca, sino que también requiere una investigación activa en el terreno, metodologías de co-diseño y mecanismos de retroalimentación con las organizaciones de mujeres como una forma de explorar las vulnerabilidades actuales y emergentes, y cómo se aplica y experimenta el sistema en la práctica. **La desconfianza en el sistema mostrada por las mujeres es alarmante.**

6- Conclusiones

Las conclusiones y recomendaciones derivadas de la Auditoría Externa del sistema VioGén han sido destacadas anteriormente. En esta sección de conclusiones, queremos centrarnos en las lecciones aprendidas sobre cómo auditar externamente algoritmos, sus posibilidades, límites y riesgos.

Como se mencionó en la Introducción, este es el primer informe de nuestra serie de proyectos de Auditoría Externa, que finalmente dará lugar a una Guía de Auditoría Externa con recomendaciones prácticas sobre cómo realizar una ingeniería inversa de datos y sistemas algorítmicos. El caso de VioGén es un caso de grave falta de transparencia, en el que no sólo el acceso al sistema técnico es difícil (necesitaríamos ser mujeres víctimas para acceder a él), sino que también el acceso para las personas que han sido impactados por el sistema es laborioso. Por lo tanto, los métodos de auditoría sugeridos por la mayoría de los autores, que se centran en la ingeniería inversa de las redes sociales o los servicios de Internet, no han sido útiles. Tras una revisión exhaustiva de las fuentes disponibles para este caso, que son principalmente artículos académicos elaborados por expertos que participaron en el diseño del formulario VioGén y el conjunto de datos del CGPJ, decidimos proceder con un enfoque de métodos mixtos que sabíamos desde el inicio que no nos permitiría llegar a conclusiones representativas, pero esperábamos que nos permitiera formular las preguntas adecuadas. **Los resultados descritos han superado nuestras expectativas en cuanto a las posibilidades de un enfoque externo de la transparencia algorítmica.**

Las lecciones aprendidas, que incluiremos en una próxima Guía de Auditoría Externa, incluyen:

- **Ajustar las expectativas a los datos disponibles:** ninguna evaluación externa podrá alcanzar la representatividad de una conclusión alcanzada por un proceso de auditoría interna. Sin embargo, las auditorías externas deberían ser capaces de informar sobre las cuestiones correctas, e impulsar a los desarrolladores de sistemas a abordar los problemas identificados y a tratar las cuestiones de transparencia. Pero aunque los ejercicios externos nunca serán concluyentes, pueden ayudar a identificar dinámicas de sesgo que pueden no ser obvias durante la fase de diseño (en el caso de que una o más de las variables utilizadas en el algoritmo sea un indicador de una cualidad o grupo protegido), proporcionar una oportunidad para involucrar a los usuarios finales y a las comunidades afectadas, y capacitar a las Organizaciones de la Sociedad Civil tradicionales es una cuestión importante de abordar.
- **Analizar el sistema de principio a fin:** Durante las fases de diseño, desarrollo y validación de VioGén, no se evaluaron a fondo ni la experiencia del usuario ni los procesos de recogida de datos. Sin embargo, mediante esta auditoría externa se han identificado y discutido muchos de los factores que alteran la calidad de los datos de entrada, así como la experiencia general de las víctimas con el sistema.

Por lo tanto, y con el fin de comprender plenamente el impacto de un sistema algorítmico, los análisis deben trascender las métricas de evaluación, centrándose también en las limitaciones de diseño y su implementación que juegan un papel crucial en relación con la calidad de las predicciones que tales sistemas realizan - más aún con asuntos sensibles como en el caso de VioGén.

- **Evaluar el sistema mediante un enfoque multimétodo:** Una evaluación exhaustiva de los sistemas algorítmicos requiere el uso de múltiples métodos, dado el amplio abanico de cuestiones que es necesario abordar. Mientras que la evaluación técnica del sistema requiere el cálculo y la discusión crítica de diversas métricas de evaluación, que pueden verse perjudicadas por la falta de disponibilidad de datos, la exploración de cómo el sistema es experimentado y percibido por las diferentes partes interesadas puede hacerse a través de técnicas de investigación cualitativas, incluyendo el trabajo de campo etnográfico, entrevistas y cuestionarios. La combinación de métodos cuantitativos y cualitativos facilita un enfoque más completo del problema en cuestión.
- **Buscar datos alternativos para evitar las barreras de derechos de la propiedad.** La mayoría de los sistemas algorítmicos no pueden evaluarse directamente debido a limitaciones de privacidad. Sin embargo, en este caso, el modelo estadístico estaba disponible, pero el conjunto de datos original del que se infirieron los indicadores y sus ponderaciones no lo estaba. En este sentido, recurrir a fuentes de datos alternativas como el CGPJ permitió realizar una evaluación indirecta del impacto, identificando patrones que tendrían que ser evaluados más a fondo dentro del conjunto de datos originales.
- **Evaluar la brecha entre el diseño y el contexto de la experiencia.** Una de las principales limitaciones de los sistemas algorítmicos para evaluar el riesgo penal que hemos identificado es la existencia de un desfase entre el diseño y el contexto de su aplicación. No tener en cuenta las barreras lingüísticas, culturales o socioeconómicas a la hora de diseñar el formulario puede dificultar la calidad de los datos que se recogen, poniendo en duda la validez de las predicciones realizadas por el sistema. En este sentido, evaluar si el proceso de diseño incluyó a las partes interesadas para aumentar su sensibilidad al contexto puede desencadenar un conjunto de preguntas de investigación valiosas para el auditor.

7- Agradecimientos

Nos gustaría agradecer a la Fundación Ana Bella por habernos puesto en contacto con las principales partes interesadas. En segundo lugar, también queremos agradecer a las víctimas, abogados y miembros de la sociedad civil cuyas inestimables respuestas han contribuido a las conclusiones y debates presentados en este informe. Por último, queremos agradecer a Amparo Quiroga Albino por la realización de las entrevistas a las supervivientes de la violencia de género.

Equipo de proyecto: Auditorías Externas

Lider de proyecto & Directora de Investigación: Dr. Gemma Galdon-Clavell, Fundadora de Eticas

Investigadores:

- Dr. Evren Yalaz, Investigadora de Ética y Tecnología en Eticas
- Toni Lorente, Investigador de Ética y Tecnología en Eticas

Recomendación de cita bibliográfica: "Eticas. (2022). The External Audit of the VioGén System. Association Eticas Research and Innovation."

8- Bibliografía

Ada Lovelace Institute. (2021). *Technical methods for regulatory inspection of algorithmic systems*. Ada Lovelace Institute.
<https://www.adalovelaceinstitute.org/wp-content/uploads/2020/11/Inspecting-algorithms-in-social-media-platforms.pdf>

Campbell, J. C., Sharps, P., & Glass, N. (2001). Risk assessment for intimate partner homicide. In *Clinical assessment of dangerousness: Empirical contributions* (pp. 136–157). Cambridge University Press.
<https://doi.org/10.1017/CBO9780511500015.009>

Christian, B. (2020). *The Alignment Problem: Machine Learning and Human Values*. Norton & Company.

Delegación del Gobierno contra la Violencia de Género. (2019a). *Delegación del Gobierno contra la Violencia de Género*. Ministerio de Igualdad.
<https://violenciagenero.igualdad.gob.es/violenciaEnCifras/macroencuesta2015/Macroencuesta2019/home.htm>

Delegación del Gobierno contra la Violencia de Género. (2019b). *Macroencuesta de Violencia Contra la Mujer*. Ministerio de Igualdad.
https://violenciagenero.igualdad.gob.es/violenciaEnCifras/macroencuesta2015/pdf/Macroencuesta_2019_estudio_investigacion.pdf

Douglas, K. S., Guy, L. S., Reeves, K. A., & Weir, J. (2005). *HCR-20 Violence Risk Assessment*

- Scheme: Overview and Annotated Bibliography*. Implementation Science and Practice Advances Research Center Publications.
https://escholarship.umassmed.edu/psych_cmhsr/335
- Estévez Mendoza, L. M. (2020). Inteligencia artificial y violencia contra las mujeres: ¿funcionan los sistemas automatizados de evaluación del riesgo? *Perspectivas*, 3, Article 3.
<https://revistas.ucalp.edu.ar/index.php/Perspectivas/article/view/148>
- Eticas. (2021). *Guide to Algorithmic Auditing*. Eticas Research and Consulting.
<https://www.eticasconsulting.com/resources/>
- González-Álvarez, J. L., & Garrido, M. J. (2015). Satisfacción de las víctimas de violencia de género con la actuación policial en España. Validación del Sistema VioGen. *Anuario de Psicología Jurídica*, 25(1), 29–38.
<https://doi.org/10.1016/j.apj.2015.02.003>
- González-Álvarez, J. L., López-Ossorio, J., Urruela, C., & Díaz, M. (2018). Integral Monitoring System in Cases of Gender Violence-VioGén System. *Behavior & Law Journal*, 4, 29–40. <https://doi.org/10.47442/blj.v4.i1.56>
- González-Prieto, Á., Brú, A., Nuño, J. C., & González-Álvarez, J. L. (2021). Machine learning for risk assessment in gender-based crime. *ArXiv:2106.11847 [Cs, Stat]*.
<http://arxiv.org/abs/2106.11847>
- Grove, W. M., & Meehl, P. E. (1996). Comparative efficiency of informal (subjective, impressionistic) and formal (mechanical, algorithmic) prediction procedures: The clinical–statistical controversy. *Psychology, Public Policy, and Law*, 2(2), 293–323.
<https://doi.org/10.1037/1076-8971.2.2.293>
- Heilbrun, K., Yasuhara, K., & Shah, S. (2011). Violence Risk Assessment Tools: Overview and Critical Analysis. In R. K. Otto & K. S. Douglas (Eds.), *Handbook of Violence Risk Assessment* (pp. 1–18). Routledge.
- Kropp, P. R. (2004). Some Questions Regarding Spousal Assault Risk Assessment. *Violence Against Women*, 10(6), 676–697. <https://doi.org/10.1177/1077801204265019>
- López-Ossorio, J. J. (2017). *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* [Universidad Autónoma de Madrid, Facultad de Psicología, Departamento de Psicología Biológica y de la Salud]. <http://hdl.handle.net/10486/678510>
- López-Ossorio, J. J. (2020, June 19). *The VioGen System* [Espai Societat Oberta]. Predecir el futuro: herramientas predictivas y sociedades segmentadas.
<https://www.youtube.com/watch?v=b7ytXGfKAUg>
- López-Ossorio, J. J., González, J., Buquerín, S., Garcia, L., & Buela-Casal, G. (2017). Risk factors related to intimate partner violence police recidivism in Spain. *International Journal of Clinical and Health Psychology*, 17.

<https://doi.org/10.1016/j.ijchp.2016.12.001>

- López-Ossorio, J. J., González, J., Muñoz Vicente, J., Urruela, C., & andres-pueyo, A. (2019). Validation and Calibration of the Spanish Police Intimate Partner Violence Risk Assessment System (VioGén). *Journal of Police and Criminal Psychology*, 34, 1–11. <https://doi.org/10.1007/s11896-019-09322-9>
- López-Ossorio, J. J., González-Álvarez, J. L., Loinaz, I., Martínez-Martínez, A., & Pineda, D. (2020). Intimate partner homicide risk assessment by police in Spain: The dual protocol VPR5.0-H. *Psychosocial Intervention*, 30(1), 47–55. <https://doi.org/10.5093/pi2020a16>
- López-Ossorio, J. J., Muñoz Vicente, J., Andrés-Pueyo, A., & Pastor Bravo, M. (2020). *Guía de Aplicación del Formulario VFR5.0-H en la Valoración Forense del Riesgo*. Gobierno de España, Ministerio del Interior.
- Pinedo, M. (2021, September 2). Matemáticas e inteligencia artificial contra el maltrato machista. *El País*. <https://elpais.com/sociedad/2021-09-02/matematicas-e-inteligencia-artificial-contra-el-maltrato-machista.html>
- Riggs, D. S., Caulfield, M. B., & Street, A. E. (2000). Risk for domestic violence: Factors associated with perpetration and victimization. *Journal of Clinical Psychology*, 56(10), 1289–1316. [https://doi.org/10.1002/1097-4679\(200010\)56:10<1289::AID-JCLP4>3.0.CO;2-Z](https://doi.org/10.1002/1097-4679(200010)56:10<1289::AID-JCLP4>3.0.CO;2-Z)
- Sandvig, C., Hamilton, K., Karahalios, K., & Langbort, C. (2014). Auditing Algorithms: Research Methods for Detecting Discrimination on Internet Platforms. *Data and Discrimination: Converting Critical Concerns into Productive Inquiry*, 23. <http://www-personal.umich.edu/~csandvig/research/Auditing%20Algorithms%20--%20Sandvig%20--%20ICA%202014%20Data%20and%20Discrimination%20Preconference.pdf>
- Sarker, I. H. (2021). Data Science and Analytics: An Overview from Data-Driven Smart Computing, Decision-Making and Applications Perspective. *SN Computer Science*, 2(5), 377. <https://doi.org/10.1007/s42979-021-00765-8>
- Tolan, S., Miron, M., Gómez, E., & Castillo, C. (2019). Why Machine Learning May Lead to Unfairness: Evidence from Risk Assessment for Juvenile Justice in Catalonia. *Proceedings of the Seventeenth International Conference on Artificial Intelligence and Law*, 83–92. <https://doi.org/10.1145/3322640.3326705>
- Weitzman, A. (2018). Does Increasing Women's Education Reduce Their Risk of Intimate Partner Violence? Evidence from an Education Policy Reform. *Criminology: An Interdisciplinary Journal*, 56(3), 574–607. <https://doi.org/10.1111/1745-9125.12181>
- Zurita Bayona, J. (2014). *Violencia contra la mujer: Marco histórico evolutivo y predicción del nivel de riesgo* [Universidad Autónoma de Madrid, Facultad de Psicología, Departamento de Psicología Biológica y de la Salud]. <https://repositorio.uam.es/handle/10486/661810>



info@eticasfoundation.org

+34 936 005 400

Mir Geribert, 8, 3rd
08014, Barcelona