

PELIGROSIDAD, ALGORITMOS Y *DUE PROCESS*: EL CASO *STATE v LOOMIS*¹

Lucía Martínez Garay

Universidad de Valencia

SUMARIO: I. La evolución de los métodos de determinación de la peligrosidad y el «*evidence-based sentencing*». II. El caso *State v Loomis*. 1. Razones para rechazar el recurso. 2. Limitaciones y advertencias. III. Comentario crítico. 1. Inseguridad jurídica. 2. Sesgos. 3. ¿Algoritmos secretos en la justicia penal? 4. La discriminación por género. IV. Consideraciones finales.

Resumen: El movimiento contemporáneo tendente a la sustitución de la noción de peligrosidad por el enfoque de la valoración del riesgo (*risk assessment*), que se inició en el ámbito del derecho penitenciario y de ejecución de las penas, está llegando a otras fases del proceso penal, como la de sentencia. Presentándose como una práctica «basada en la evidencia» (*evidence-based sentencing*), propugna tomar en consideración las valoraciones estructuradas del riesgo de reincidencia a la hora de determinar la clase y cuantía de la pena a imponer, para adecuarla mejor a las necesidades de prevención especial. El presente trabajo analiza críticamente una de las primeras sentencias de un tribunal supremo estatal de los EEUU que se ha enfrentado a los problemas que plantea esta propuesta, y en particular al siguiente: ¿es compatible con el derecho al debido proceso agravar la pena sobre la base de la valoración de riesgo hecha por un algoritmo, cuyo funcionamiento no se desvela al acusado porque está protegido como secreto de empresa?

Palabras clave: peligrosidad, valoración del riesgo de reincidencia, algoritmos secretos, determinación de la pena, derecho al debido proceso

¹ Trabajo elaborado en el marco del Proyecto de I+D+I DER2017-86336-R, financiado por el Ministerio de Economía, Industria y Competitividad.

Abstract: The contemporary tendency towards replacing the old notion of dangerousness by the risk assessment approach, that began in the correctional settings, is reaching other stages of the criminal justice, particularly the sentencing phase. Presenting itself as an evidence-based practice (evidence-based sentencing), it promotes introducing structured risk assessments in sentencing in order to help the judge impose the penalty that best prevents recidivism. This paper critically examines one of the first State Supreme Court decisions in the USA that has had to face the problems posed by this approach, and especially the following: is it compatible with due process rights to impose an enhanced sentence based on the risk assessment made by a secret algorithm, the details of which are not disclosed to the defendant due to its proprietary nature?

Keywords: dangerousness, violence risk assessment, secret algorithms, sentencing, due process

I. La evolución de los métodos de determinación de la peligrosidad y el «*evidence-based sentencing*»

El concepto de peligrosidad en su sentido moderno se introdujo en el discurso político criminal y en las legislaciones al calor del Positivismo criminológico a finales del siglo XIX y principios del XX. La peligrosidad se entendía entonces como una característica o disposición subjetiva que predisponía al sujeto a cometer delitos en el futuro, y el método predominante para determinarla fue hasta aproximadamente la década de 1990 el juicio clínico del experto —médico, preferentemente psiquiatra, o psicólogo—, combinado con la apreciación personal del juez².

Desde el mismo momento de su nacimiento el concepto de peligrosidad fue contemplado con escepticismo en la doctrina penal³, sobre todo desde las posturas críticas frente al positivismo criminológico y frente a las corrientes orientadas a la prevención especial. Sin embargo, en los años 70-80 del siglo XX esa desconfianza se generalizó, al hilo del examen muy crítico que las ciencias sociales y médicas hicieron en aquel momento sobre las posibilidades reales de detectar «científicamente» la peligrosidad: las dificultades teóricas y prácticas para prever la conducta futura de las personas, unidas a las altas tasas de error que fue posible constatar empíricamente en algunas ocasiones, restaban legitimidad a

² Sobre el entendimiento de la peligrosidad como rasgo personal, así como sobre la forma de determinarla en España en el contexto de la Ley de Peligrosidad y Rehabilitación Social de 1970, cfr. JORGE BARREIRO, A., *Las medidas de seguridad en el Derecho español (Estudio doctrinal y jurisprudencial de la Ley de Peligrosidad y Rehabilitación Social de 4-8-1970)*, Madrid (Civitas), 1976, p. 251.

³ Por todos, SANZ MORÁN, A., *Las medidas de corrección y de seguridad en el Derecho penal*, Valladolid (Lex Nova), 2003, p. 90.

las consecuencias jurídicas restrictivas de derechos que la legislación hacía depender de pronósticos tan inciertos, pues resultaban difícilmente compatibles con principios esenciales del Estado de Derecho como los de seguridad jurídica o proporcionalidad⁴.

Desde entonces, sin embargo, han cambiado muchas cosas. La idea de peligrosidad ha sido sustituida en Criminología por el enfoque de la «valoración del riesgo» de reincidencia o de violencia (*violence risk assessment*), según el cual de lo que se trata no es de averiguar si el individuo posee o no la cualidad subjetiva de peligroso (como si se le estuviera diagnosticando una enfermedad), sino de valorar un conjunto de factores, personales pero también ambientales, y cambiantes en el tiempo, que favorecen en mayor o menor grado la comisión de nuevos delitos y que permiten efectuar pronósticos sobre la reincidencia futura en términos de probabilidad⁵. Y el cambio en la conceptualización de aquello que se valora —de la peligrosidad al riesgo— ha ido acompañado de un cambio en la forma de medir ese riesgo. El tradicional método clínico tiende a considerarse poco fiable, sesgado, y carente de suficiente base empírica, y ha ido siendo sustituido por métodos estructurados de estimación del riesgo, que están basados en la observación empírica de la conducta de grandes grupos de sujetos, y en la cuantificación y combinación estadística de las variables que concurren en ellos y que demuestran estar significativamente asociadas a la aparición de conducta violenta y/o delictiva (factores de riesgo). Las herramientas estructuradas de valoración del riesgo (en adelante, HEVR) son listados contruidos con dichas variables (por ej., edad actual del sujeto, edad a la que cometió el primer delito, número de condenas previas, si padece o no adicciones, etc.), que cuando se administran a un individuo dan como resultado una estimación del riesgo de reincidencia que es esperable de dicha persona en un periodo de tiempo dado. En ocasiones esas estimaciones (por ej.: «un 25% de probabilidad de comisión de un nuevo delito sexual en un periodo de dos años») se obtienen automáticamente con un algoritmo que combina de determinada manera los valores numéricos asignados por la HEVR a cada factor de riesgo (HEVR

⁴ Cfr. VIVES ANTÓN, T.S., «Métodos de determinación de la peligrosidad», en VVAA, *Peligrosidad social y medidas de seguridad (La Ley de peligrosidad y rehabilitación social de 4 de agosto de 1970)*, Universidad de Valencia, 1974, pp. 389-417. Cfr. asimismo MARTÍNEZ GARAY, L. / MONTES SUAY, F., «El uso de valoraciones del riesgo de violencia en Derecho Penal: algunas cautelas necesarias», *InDret* 2/2018, pp. 4-5; MARTÍNEZ GARAY, L., «La incertidumbre de los pronósticos de peligrosidad: consecuencias para la dogmática de las medidas de seguridad», *InDret* 2/2014, pp. 19-20

⁵ ANDRÉS PUEYO, A. / REDONDO ILLESCAS, S., «Predicción de la violencia: entre la peligrosidad y la valoración del riesgo de violencia», *Papeles del Psicólogo*, vol. 28(3), 2007, y ANDRÉS PUEYO, A., «Peligrosidad criminal: análisis crítico de un concepto polisémico», en DEMETRIO CRESPO (dir.), *Neurociencias y Derecho Penal. Nuevas perspectivas en el ámbito de la culpabilidad y tratamiento jurídico-penal de la peligrosidad*, Madrid (Edisofer – BdF), 2013, pp. 483-503.

actuariales), y en otros casos incluyen en mayor o menor medida la valoración personal del experto que ha administrado la herramienta (HEVR de juicio clínico estructurado)⁶.

Con el cambio de enfoque y de metodología la investigación sobre valoración del riesgo de violencia ha experimentado un crecimiento espectacular, y también ha cambiado la opinión sobre sus resultados, pues frente al escepticismo al que nos hemos referido líneas arriba, en la Criminología actual se ha abierto paso un moderado (pero firme) optimismo sobre la capacidad predictiva de estos instrumentos, y sobre su utilidad en el ámbito forense⁷. Además, el éxito de estos nuevos métodos no se ha limitado al ámbito teórico o académico, sino que en muchos países y especialmente en el ámbito anglosajón se han introducido plenamente en la práctica de la administración de justicia penal: decretar prisión provisional o libertad bajo fianza, decidir sobre la suspensión de la pena (*probation*) o la libertad condicional (*parole*) así como determinar el nivel de supervisión que en su caso debe acompañarlas, imponer o no *civil commitment* a los *sexual violent predators*, y hasta imponer pena de muerte en los crímenes capitales, en todos estos estadios de un proceso penal intervienen las valoraciones de riesgo, y cada vez más éstas se realizan con herramientas estructuradas⁸.

Y desde hace algunos años esta tendencia está penetrando también en la fase de determinación de la pena. Bajo la atractiva denominación de «*evidence-based sentencing*» o incluso «*smart sentencing*», se trata de incorporar expresamente las valoraciones de riesgo como uno más de los criterios a tener en cuenta a la hora de establecer tanto el tipo como la cuantía de la pena, con el objetivo último de reducir la reincidencia⁹.

⁶ Aunque el empleo de tablas estadísticas y procedimientos actuariales en la valoración del riesgo en el ámbito penal tenía importantes precedentes desde al menos los años 30 del siglo pasado (HARCOURT, B. E., *Against prediction: profiling, policing, and punishing in an actuarial age*, Chicago (University of Chicago Press), 2007, pp. 39 y ss; LOINAZ, I., *Manual de evaluación del riesgo de violencia. Metodología y ámbitos de aplicación*, Madrid (Pirámide), 2017, pp. 41 y ss.), ha sido desde mitad de los años 80 cuando se ha multiplicado la investigación al respecto y se ha generalizado su uso en la práctica forense. Sobre las diferencias entre los instrumentos puramente actuariales y los de juicio clínico estructurado cfr. LOINAZ, I., 2017, pp. 42 y ss. y 101 y ss.

⁷ Cfr. por todos ANDRÉS PUEYO, A. / REDONDO ILLESCAS, S., *Papeles del Psicólogo*, vol. 28(3), 2007, p. 169.

⁸ STARR, S. B., «The Risk Assessment Era: An Overdue Debate», *Federal Sentencing Reporter*, 27:4, 2015; HARCOURT, B. E., 2007, pp. 77 y ss.

⁹ «*Evidence-based sentencing*» son «sentencing practices based on the best research evidence of what works to promote public safety by reducing recidivism» (WARREN, R. K., «Evidence-Based Sentencing in the State Courts», *Written statement of Judge (ret.) Roger K. Warren United States Sentencing Commission – Alternatives to incarceration panel*, Chicago, Illinois, September 10, 2009, accessible online en: https://www.uscc.gov/sites/default/files/pdf/amendment-process/public-hearings-and-meetings/20090909-10/Warren_testimony.pdf). Cfr. asimismo OLESON, J. C., «Risk in Sentencing: Constitutio-

Este planteamiento se fundamenta en varias premisas: la primera, que el objetivo de evitar la reiteración delictiva es uno de los fines legítimos del castigo penal; en segundo lugar, que la impresión del juez sobre el grado de peligro que presenta un sujeto es intuitiva y por ello al menos en parte arbitraria, subjetiva y sesgada; y por último, que las modernas HEVR proporcionan conocimientos rigurosos y empíricamente fundamentados acerca de dicho riesgo. En consecuencia, los resultados de la valoración estructurada del riesgo de reincidencia deben ser tenidos en cuenta en las sentencias porque constituyen evidencia científica de mejor calidad sobre la peligrosidad del sujeto que la impresión personal que pueda tener el juez.

El éxito que está teniendo este enfoque es notable: a fecha de 2014 más de 20 Estados en los EEUU habían implementado ya alguna forma de *evidence based sentencing* en la práctica¹⁰, y el *American Law Institute* lo ha incorporado también en una reciente revisión del *Model Penal Code*¹¹. Sin embargo, también está generando un intenso debate doctrinal, pues el uso de HEVR en Derecho penal (y no sólo en fase de sentencia) plantea importantes problemas y retos¹². En este trabajo me voy a referir a uno que ha empezado a ser objeto de atención en el ámbito estadounidense especialmente al hilo de un caso resuelto por el Tribunal Supremo de Wisconsin en 2016: ¿es compatible con el derecho al debido proceso (*due process of law*) usar valoraciones del riesgo de violencia para determinar la clase y la extensión de la pena?

nally Suspect Variables and Evidence-Based Sentencing», *SMU Law Review* 64, 2011, pp. 1330-1342, p. 1336; HANNAH-MOFFAT, K., «Actuarial Sentencing: An “Unsettled” Proposition», *Justice Quarterly*, 30:2, 2013, pp. 270-296, p. 271.

¹⁰ Cfr., con ulteriores referencias, STARR, S. B., «Evidence-Based Sentencing and the Scientific Rationalization of Discrimination», *Stanford Law Review* 66, 2014, pp. 805, 809 y ss. y nota 11 en esa página. En Kentucky, Ohio y Pennsylvania el juez debe disponer de una valoración de riesgos y necesidades criminógenas del sujeto en la fase de sentencia; en Arizona y Oklahoma es necesario tenerla en cuenta si el caso es susceptible de *probation* o de ser sancionado con un *community punishment*. Otros Estados no ordenan, pero permiten, que al establecer la pena el juez tenga en cuenta dichas valoraciones: Idaho, Indiana, Louisiana, Washington, West Virginia (cfr. las referencias de la normativa correspondiente en *State v. Loomis*, cit. infra en n. 13, ¶ 42, notas 23 y 24).

¹¹ Tras la última revisión éste establece que las *Sentencing Commissions* «deben» desarrollar instrumentos de valoración de riesgos y de necesidades criminógenas, y que cuando éstos demuestren ser suficientemente fiables, los «pueden» incorporar en las *sentencing guidelines* (THE AMERICAN LAW INSTITUTE, *Model Penal Code: Sentencing. Proposed Final Draft (April 10, 2017)*, Sección § 6B.09: «Evidence-Based Sentencing; Offender Treatment Needs and Risk of Reoffending»; el borrador fue aprobado en la asamblea de mayo de 2017 en los términos en que había sido propuesto).

¹² Cfr. muy crítica, STARR, S. B., *Stanford Law Review* 66, 2014, pp. 803-872, así como, con diferentes perspectivas, los trabajos cit. supra en nota 9. En castellano, me permito remitir a MARTÍNEZ GARAY, L., «La relación entre culpabilidad y peligrosidad», en VVAA, *La culpabilidad. I Seminario de Derecho penal de la UAM*, Editorial BdF, 2018 (en prensa).

II. El caso *State v Loomis*

El 13 de julio de 2016 el Tribunal Supremo del Estado de Wisconsin (en adelante, TSW) dictó sentencia en el caso *State v. Loomis*¹³. Eric Loomis había sido acusado de cinco delitos por su presunta intervención en un tiroteo efectuado desde un vehículo (dos cargos por posesión de armas de fuego, intento de evasión frente a un agente de tráfico, puesta en peligro de la seguridad pública, y conducción de un vehículo ajeno sin autorización). El acusado negó su participación en el tiroteo aunque admitió haber conducido el vehículo con posterioridad a los hechos, y llegó a un acuerdo con la Fiscalía para aceptar los dos últimos cargos y rechazar el resto de acusaciones.

El juez de instancia aceptó la conformidad y al concretar la pena tuvo en cuenta, entre otras cosas, el contenido de un «*presentence investigation report*»¹⁴ (en adelante, PSI), que incluía entre otros datos el resultado de la evaluación del sujeto mediante COMPAS, una herramienta estructurada que valora el riesgo de reincidencia y las necesidades criminógenas del sujeto. Según COMPAS Loomis presentaba un riesgo alto tanto de reincidencia en general como de reincidencia violenta. El juez condenó a Loomis a seis años de prisión y otros 5 de supervisión postpenitenciaria, y en la sentencia explicó que «A la hora de valorar los distintos factores, excluyo la suspensión de la condena [*probation*] debido a la gravedad del crimen y porque su historial, su historial en materia de supervisiones judiciales, y las herramientas de valoración del riesgo que se han aplicado, indican que usted presenta un riesgo extremadamente alto de reincidir»¹⁵.

Loomis solicitó al tribunal que revisara la cuantía de la pena, argumentando —además de otras razones¹⁶— que haberla fundamentado en el resultado del COMPAS había vulnerado su derecho a un proceso debido. Pero el tribunal no modificó la condena y explicó además que sólo había utilizado la valoración de riesgo para corroborar las conclusiones a las que ya había llegado a partir del resto de información disponible, añadiendo que habría impuesto exactamente la misma pena aunque no hubiera tomado en consideración el resultado del COMPAS¹⁷.

¹³ *State v. Loomis*, 881 N.W.2d 749 (Wis. 2016), accesible online en: <https://law.justia.com/cases/wisconsin/supreme-court/2016/2015ap000157-cr.html> (consultado por última vez el 17.7.2018).

¹⁴ Un informe previo a la sentencia que realiza el *probation officer* a petición del tribunal y que contiene información sobre el historial criminal del acusado convicto y también sobre aspectos educativos, familiares, laborales, socio-económicos, uso de sustancias, etc, información que es utilizada para determinar la clase y extensión de la pena.

¹⁵ *State v. Loomis*, ¶ 19.

¹⁶ Loomis alegó también que el tribunal había abusado de su discrecionalidad al tener en cuenta para fijar la cuantía de la pena la base fáctica de los delitos que él no había aceptado en su conformidad. Ese extremo no será analizado en este trabajo.

¹⁷ *State v. Loomis*, ¶ 23-28.

Loomis recurrió entonces ante el TSW alegando que el uso de COMPAS para determinar la pena había vulnerado su derecho a un proceso con todas las garantías (*right to due process*) por tres razones: (1) se había violado su derecho a ser condenado sobre la base de información fiable y precisa (*to be sentenced upon accurate information*), ya que la naturaleza de secreto comercial (*proprietary nature*) del algoritmo con el que funciona COMPAS le había impedido conocer cómo se habían calculado las estimaciones de riesgo y por tanto refutar la validez científica del nivel de riesgo que dicha herramienta le había asignado; (2) se había vulnerado su derecho a obtener una sentencia individualizada, y (3) se había tenido en cuenta indebidamente su género para determinar la pena.

El TSW no estimó ninguno de los motivos del recurso. Sin embargo, estableció algunas limitaciones sobre cómo y para qué pueden ser usadas las evaluaciones de riesgo en la fase de determinación de la pena.

1. Razones para rechazar el recurso

La empresa propietaria de COMPAS publica información sobre esta herramienta¹⁸, pero no desvela cómo funciona el algoritmo que calcula los niveles de riesgo (cómo se miden, ponderan y combinan las distintas variables) porque entiende que es un secreto de empresa, esencial para su negocio. Según el TSW, no haber podido acceder a esa información no vulneró el derecho que tiene todo acusado a ser condenado sobre la base de información fiable y exacta, que incluye el derecho a revisar y verificar la información contenida en el PSI¹⁹, porque Loomis sí tuvo acceso tanto a los resultados (los niveles de riesgo que COMPAS le asignó) como a la información con la que estos se calcularon, que consiste básicamente en datos objetivos sobre el historial criminal del acusado. Como la valoración del riesgo se basó en las respuestas del propio acusado al cuestionario que se le pasó, así como en información objetiva y públicamente accesible sobre sus antecedentes, el acusado tuvo oportunidad de verificar que dichos datos eran correctos. Y podría haber cuestionado los resultados alegando otros factores o datos diferentes que demostrasen su inexactitud²⁰.

El TSW tampoco consideró que se hubiera vulnerado el derecho de Loomis a una sentencia individualizada. El tribunal explicó que el recurrente habría tenido razón si la sentencia se hubiera basado únicamente

¹⁸ Cfr. la *Practitioner's Guide to COMPAS Core* (fechada en marzo de 2015), accesible en la web de la compañía (http://www.northpointeinc.com/downloads/compas/Practitioners-Guide-COMPAS-Core_031915.pdf), consultada por última vez el 1.9.2018), un documento de más de 60 páginas que explica en qué consiste, qué riesgos y necesidades criminógenas mide, en qué estudios se basa, etc.

¹⁹ *State v. Loomis*, ¶ 47 y ss.

²⁰ *State v. Loomis*, ¶ 54 y ss.

en la valoración estructurada de riesgo o si ésta hubiera sido el factor decisivo para establecer la pena, porque las HEVR sirven para identificar *grupos* de personas en riesgo de reincidencia, pero no a *individuos* concretos de alto riesgo²¹. Pero ello no significa que las valoraciones de riesgo no se puedan utilizar en la fase de sentencia, más bien al contrario: aportan al juez información que, unida al resto de circunstancias concurrentes, contribuye a que la condena se ajuste lo más posible al caso concreto. Lo que los jueces deben hacer cuando manejen estas herramientas es aplicar a sus resultados un margen de discrecionalidad, poniéndolos en el contexto del resto de circunstancias particulares propias del caso y del individuo concretos²².

Por último, el TSW tampoco aceptó que el uso de COMPAS supusiera una utilización indebida del género —masculino— de Loomis en la determinación de la sentencia. Destacó que tanto la Fiscalía como la defensa habían aceptado en el juicio que toda la evidencia científica disponible indica que los varones tienen unas tasas de reincidencia y de delincuencia violenta superiores a las de las mujeres. Siendo esto así, afirmó el TSW, prescindir del factor género en la valoración del riesgo disminuiría sensiblemente el acierto de las estimaciones, porque aumentarían las tasas de error tanto para los hombres como para las mujeres. En consecuencia, si incluir el género aumenta la precisión en las estimaciones de riesgo, ello no persigue un objetivo discriminatorio sino que va en beneficio tanto de los acusados como de la administración de justicia en general²³.

2. Limitaciones y advertencias

A pesar de rechazar todos los motivos del recurso y mostrarse partidario en principio del *evidence-based sentencing*, el TSW hace en la sentencia un repaso bastante pormenorizado de los problemas que la literatura científica ha puesto de manifiesto en relación con el uso de HEVR: desde dudas sobre el nivel de acierto de estos instrumentos, hasta problemas de posible discriminación de minorías, pasando por la falta de suficientes estudios de validación²⁴. Y sienta una doctrina general sobre las cautelas que deben observarse cuando se utilicen valoraciones estructuradas de riesgo al establecer las penas en las sentencias.

En primer lugar, afirma que las valoraciones de riesgo sólo pueden utilizarse para: (a) imponer a delincuentes de bajo riesgo medidas alternativas a la prisión; (b) valorar si un delincuente puede ser supervi-

²¹ Ibid., ¶ 68-69.

²² Ibid., ¶ 71-72.

²³ Ibid., ¶ 83, 86.

²⁴ *State v. Loomis*, ¶ 43, 59 y ss.,

sado de manera segura y efectiva en la comunidad, y (c) determinar los plazos y condiciones de la suspensión de la pena y de la libertad condicional y las consecuencias de infringirlas²⁵. Pero no pueden usarse para decidir el encarcelamiento del sujeto, ni para determinar la severidad de la sentencia²⁶.

Por otro lado, establece que las valoraciones de riesgo no pueden ser nunca el elemento determinante a la hora de decidir si el sujeto puede ser supervisado de manera segura y efectiva en la comunidad (y por tanto eludir una condena a prisión). Los tribunales deben explicar siempre qué otros factores han tenido en cuenta para fijar la condena, además de la valoración de riesgo.

Por último, la sentencia ordena que cualquier PSI que incluya una valoración de riesgo hecha con COMPAS debe no sólo advertir expresamente sobre estas dos limitaciones, sino también informar al tribunal de los siguientes extremos:

- de que se ha invocado la naturaleza de secreto de empresa del algoritmo con el que funciona COMPAS para no explicar cómo se ponderan los factores de riesgo ni cómo se calculan las puntuaciones;
- que las puntuaciones de COMPAS se elaboran con datos grupales, y por ello identifican grupos de delincuentes de alto riesgo, pero no individuos con alto riesgo de reincidencia;
- que algunos estudios han suscitado dudas sobre si COMPAS asigna a sujetos pertenecientes a minorías puntuaciones de riesgo desproporcionadamente altas;
- que no existe ningún estudio de validación para COMPAS realizado sobre población del Estado de Wisconsin;
- que para mantener el nivel de precisión de las estimaciones, las herramientas de valoración de riesgo deben ser constantemente revisadas para adaptarlas a los cambios que se producen en la población,
- y de que COMPAS no se diseñó para ser utilizado en la fase de sentencia sino en el ámbito penitenciario, en decisiones relativas al tratamiento, supervisión y libertad condicional²⁷.

Una vez establecidas estas cautelas, el TSW las aplica al caso concreto de Loomis y concluye que el tribunal de instancia las respetó ade-

²⁵ *State v. Loomis*, ¶ 88.

²⁶ *Ibid.*, ¶ 98. En este punto el Tribunal hace suyas advertencias que ya contenía un informe del *Department of Corrections* del Estado de Wisconsin, cfr. *State v. Loomis*, (cit. supra n. 13), ¶ 44.

²⁷ *State v. Loomis*, ¶ 98-101.

cuadramente, porque fue consciente de las limitaciones inherentes a las HEVR y no dio al resultado del COMPAS un peso excesivo: de hecho, le reconoció escasa o incluso nula relevancia porque indicó que habría impuesto exactamente la misma sentencia si no hubiera dispuesto de esa información²⁸.

III. Comentario crítico²⁹

State v Loomis ha suscitado considerable interés en el ámbito estadounidense. Es la segunda sentencia de una corte suprema estatal que acepta expresamente la entrada de cálculos de riesgo matemáticos en un reducto hasta ahora tan ajeno a ellos como el dictado de una condena criminal³⁰, y la primera que además permite que esos cálculos se hagan con un algoritmo cuyo funcionamiento es secreto.

El caso pone sobre la mesa la contradicción que plantea el *evidence-based sentencing*: por un lado permite incorporar información que a juicio del Tribunal es rigurosa y relevante para fijar la condena, pero por otro lado presenta riesgos para los derechos del acusado y para el correcto funcionamiento del sistema judicial. Dicho dilema debería resolverse como en Derecho se resuelven estas cosas: si los riesgos para los derechos fundamentales del acusado son reales y no pueden ser evitados por otros medios, para protegerlos habrá que prescindir de las ventajas que pudiera suponer la incorporación de valoraciones de riesgo en la fijación de las condenas (de la misma forma que por ejemplo se prohíbe utilizar como medio de prueba una confesión obtenida en una intervención telefónica realizada sin autorización judicial, por mucho que pudiera ser decisiva para revelar la verdad de lo acontecido). Si por el contrario fuera posible fijar una serie de restricciones o requisitos que permitan introducir las valoraciones de riesgo en el juicio sin merma para los derechos del acusado, hay que establecer con claridad dichos requisitos para que en el futuro los tribunales y los acusados sepan a qué atenerse. La sentencia del caso Loomis, sin embargo, no hace ni lo uno ni lo otro.

²⁸ *Ibid.*, ¶ 102-110.

²⁹ Cuando este trabajo se encontraba ya en proceso de edición he accedido a una publicación en castellano que comenta también esta sentencia: ROMEO CASABONA, C. M., «Riesgo, procedimientos actuariales basados en inteligencia artificial y medidas de seguridad», *Revista penal* n.º 42, 2018, pp. 165-179. Lamentablemente las restricciones de tiempo y espacio del proceso editorial impiden que pueda comentarla en mi exposición, por lo que me limito a dejar constancia de su existencia para el lector interesado.

³⁰ La primera fue el caso *Malenchik v. State of Indiana*, 9 de febrero de 2010 (accesible online en <https://www.in.gov/judiciary/opinions/pdf/06091001bd.pdf>, consultada por última vez el 24.8.2018)

1. Inseguridad jurídica

En contra de lo que pudiera parecer, ni siquiera deja claro para qué decisiones es legítimo utilizar valoraciones de riesgo y para cuáles no. La sentencia dice que pueden utilizarse para imponer a delincuentes de bajo riesgo medidas alternativas a la prisión y para valorar si un delincuente puede ser supervisado de manera segura y efectiva en la comunidad, pero no para decidir el encarcelamiento del sujeto. Sin embargo, ello encierra a mi juicio una contradicción. Porque si la valoración de riesgo indica que el delincuente no podrá ser supervisado eficazmente en la comunidad, ello será un argumento muy relevante para decretar la entrada en prisión. Luego autorizar las valoraciones de riesgo para lo primero las convierte también en decisivas para lo segundo. Y entonces, ¿hay algo para lo que realmente no pueden ser utilizadas?

Por otro lado, la sentencia no establece unos requisitos claros cuya concurrencia haga legítimo el uso de valoraciones de riesgo, sino un listado de «advertencias» sobre los peligros que éstas entrañan. No es fácil deducir qué se supone que deben entender los jueces a partir de ese listado: ¿se les está recomendando que utilicen valoraciones de riesgo, porque aportan información de interés, o se les está desaconsejando que lo hagan, por la cantidad de riesgos que implican?

Podría pensarse que quizá lo segundo, habida cuenta de que el TSW expresamente señala que la finalidad que persigue con ello es evitar que los jueces al imponer las penas den una importancia indebida (por excesiva) a las valoraciones estructuradas del riesgo de reincidencia³¹. Sin embargo, cabe dudar de que las advertencias lleguen a provocar en los destinatarios ese deseado escepticismo. En primer lugar, porque no se dice a los jueces si esos problemas de los que se les advierte llegan a poner seriamente en cuestión el uso de las valoraciones de riesgo —y en su caso por qué— o si son meras objeciones de detalle³². Y sin indicaciones más precisas no es fácil que los jueces sepan valorarlo por sí mismos, porque probablemente la mayoría simplemente no entiende cómo funcionan las HEVR, ni qué significa la información que existe sobre los niveles de acierto de sus estimaciones. En la propia sentencia del caso Loomis hay un voto particular que subraya este problema, afirmando que «en el caso presente un problema significativo ha sido que el tribunal no entiende el COMPAS. En los informes orales el tribunal preguntó repetidamente tanto al Estado como al abogado de la defensa cómo funciona COMPAS. Y obtuvo pocas respuestas»³³.

³¹ *State v. Loomis*, ¶ 110: «those cautions are required in part to ensure that undue weight is not given to the COMPAS risk scores».

³² En este sentido, RECENT CASES, «Criminal Law – Sentencing Guidelines – *State v. Loomis*. Wisconsin Supreme Court Requires Warning Before Use of Algorithmic Risk Assessments in Sentencing», *Harvard Law Review* 130, 2017, pp. 1531, 1534.

³³ *State v. Loomis*, ¶ 133, voto particular (concurrente) de la juez Shirley S. Abrahamson.

2. Sesgos

Por otra parte, esta llamada indeterminada a la cautela se produce en un ambiente de presión creciente a favor de la incorporación de *risk assessments* en la fase de sentencia: el *Model Penal Code*, muchas las voces en la judicatura y en la doctrina, y las legislaciones en diversos Estados, recomiendan o incluso ordenan utilizarlos. Algunos ya han advertido de que en este contexto la doctrina establecida en *Loomis*, en lugar de promover el escepticismo frente al uso de algoritmos para valorar la peligrosidad al imponer las penas, podría tener el efecto contrario. De un lado, porque la siempre difícil tarea de decidir sobre la gravedad de la pena a imponer —y especialmente si conlleva privación de la libertad— puede hacerse con la conciencia un poco más tranquila si los jueces pueden confiar en que su intuición está corroborada por una valoración de riesgo científicamente fundada, y por tanto querrán disponer de estas valoraciones siempre que sea posible. Y una vez dispongan de ellas, hay que contar con el probable sesgo de anclaje que provocarán. Porque aunque *State v. Loomis* ordena que el resultado de la valoración de riesgo sólo pueda servir como corroboración de otros indicios, eso no necesariamente impedirá que en la práctica pueda funcionar justo al contrario: que una vez conocido el nivel de riesgo estimado por la herramienta, el juez elabore su razonamiento condicionado —inconscientemente— por dicha cifra, a la que amoldará su valoración del resto de circunstancias del caso concreto para que encajen con ella³⁴.

El de anclaje no es por otra parte el único sesgo que resulta preocupante en esta materia. Cuando se afirma que la información que proporcionan las HEVR es objetiva y es por ello preferible a la valoración personal del juez, que puede estar sesgada por prejuicios inconscientes (raciales, de clase, de género, etc.), se olvida que los algoritmos son un mero reflejo de los datos con los que se los alimenta, de modo que si esos datos incorporan sesgos el algoritmo los reproducirá, o peor aún, los exacerbará. Por ejemplo, en EEUU se ha denunciado que la valoración estructurada del riesgo de violencia tiene a incrementar la ya desproporcionada sobrerrepresentación de las minorías en el sistema penitenciario, y en particular la de la afroamericana. Aunque la raza no aparezca como factor a valorar en ninguna de las modernas HEVR³⁵, en todas ellas tienen un peso muy importante los antecedentes penales del

³⁴ RECENT CASES, *Harvard Law Review* 130, 2017, p. 1536; ISRANI, E.: «Algorithmic Due Process: Mistaken Accountability and Attribution in *State v. Loomis*», *Harvard Journal of Law and Technology. Digest*, 31.8.2017 (<https://jolt.law.harvard.edu/digest/algorithmic-due-process-mistaken-accountability-and-attribution-in-state-v-loomis-1>).

³⁵ Cosa que sería inconstitucional: como afirmó el Tribunal Supremo de los EEUU en *Buck v. Davis* (22.2.2017): «It would be patently unconstitutional for a state to argue that a defendant is liable to be a future danger because of his race.»

sujeto, que están fuertemente correlacionados con la raza porque los miembros de las minorías étnicas tienen mayor número de condenas, reciben penas más largas y están más tiempo en prisión. Si el historial criminal actúa como factor decisivo para considerar que el riesgo de reincidencia es mayor, y conlleva por ello un tratamiento más duro en caso de una nueva condena, aumentará todavía más la proporción de minorías en el sistema penitenciario porque recibirán menos suspensiones de condena y penas aún más largas. Lo que a su vez empeorará su puntuación cuando su riesgo de reincidencia vuelva a ser estimado, y así sucesivamente³⁶.

Es por ello peligroso presentar las estimaciones realizadas con HEVR como objetivas y carentes de sesgos: ciertamente, carecen de los sesgos personales que pueden aquejar a cada juez como individuo, pero tienen «sesgos sistémicos»: porque incorporan, perpetúan y exacerbaban las discriminaciones estructurales del sistema en el que actúan y del que se alimentan³⁷.

3. *¿Algoritmos secretos en la justicia penal?*

A mi juicio el aspecto más sorprendente del caso Loomis es que el TSW aceptase que las sentencias puedan basarse en parte (siquiera sea en una mínima parte) en valoraciones de riesgo realizadas mediante herramientas comercializadas por compañías privadas, cuyo funcionamiento no se revela al público ni al acusado porque se considera un secreto de empresa. Si como afirmó el TSW el acusado tiene derecho a ser condenado sobre la base de información fiable y exacta, y eso incluye el derecho a revisar y verificar la información contenida en el PSI (cfr. supra, II.1.), no puede seriamente pretenderse que para satisfacer estas exigencias baste con darle al acusado acceso a los resultados de su valoración de riesgo y a los factores con los que se calcularon esas puntuaciones, pero sin permitirle saber cuánto peso se ha dado a cada factor, si todos ellos o sólo algunos han influido al final en su puntuación, ni cómo se han combinado. Basta pensar que dos sujetos con exactamente los mismos factores de riesgo podrían obtener dos puntuaciones diversas en dos HEVR distintas si cada una de ellas pondera y combina esos mismos factores de manera diferente. Para poder

³⁶ HARCOURT, 2007, pp. 145 y ss.; TONRY, M., (2014): «Legal and Ethical Issues in the Prediction of Recidivism», *Federal Sentencing Reporter* 26:3, 2014, p. 173. Sobre estas cuestiones, con más detalle, cfr. MARTÍNEZ GARAY, L., 2018 (en prensa).

³⁷ Para más información sobre estas características de los algoritmos en general, y la dificultad de corregir estos efectos en el ámbito de la justicia penal, cfr. O'NEIL, C.: *Weapons of Math Destruction*, New York (Crown Publishers), 2016, especialmente caps. 1 y 5.

cuestionar con una mínima eficacia el resultado, el acusado tiene que poder saber cómo se ha llegado a él³⁸.

La decisión del TSW llama además especialmente la atención porque el Tribunal conoce y cita en la sentencia estudios que cuestionan tanto los niveles de acierto de COMPAS como los posibles efectos discriminatorios entre minorías que puede producir —sobreestimando el riesgo de reincidencia en los afroamericanos, e infravalorándolo en los blancos—³⁹, lo que haría aún más necesario conocer en detalle su funcionamiento antes de aceptar que una sentencia pueda fundamentar la gravedad de la condena en los resultados que arroja esta herramienta. Como han destacado diversos autores, el TSW parece ignorar lo obvio: que el legítimo ánimo de lucro de la compañía propietaria de la herramienta⁴⁰ puede chocar con lo deseable desde el punto de vista del rigor científico, y hacer que la información que la empresa ofrece sobre su producto subraye sus bondades pero a la vez omita datos relevantes (por ejemplo, no informe sobre estudios en los que los niveles de acierto no alcanzaron lo esperado). En este contexto, resulta difícilmente comprensible que el TSW no reconociera al acusado las más amplias posibilidades para cuestionar una valoración de riesgo que le resultaba desfavorable⁴¹.

Sobre todo si tenemos en cuenta que además de las herramientas comercializadas por empresas privadas existen también HEVR públicas, o de acceso gratuito: algunas jurisdicciones han optado por desarrollar sus propias herramientas en lugar de comprarlas a empresas privadas, y en otros casos son fundaciones o investigadores quienes han creado herramientas de código abierto y disponibles para su utilización por cualquiera⁴². Es más, no existe evidencia científica de que las herramientas

³⁸ Cfr., críticos con la sentencia en este punto, DE MIGUEL BERIAIN, I.: «Does the use of risk assessment in sentences respect the right to due process? A critical analysis of the *Wisconsin v. Loomis* ruling», *Law, Probability and Risk* 17, 2018, pp. 48 y ss.; FREEMAN, K.: «Algorithmic injustice: how the Wisconsin Supreme Court failed to protect due process rights in *State v. Loomis*», *North Carolina Journal of Law & Technology*, vol. 18, 2016, pp. 91 y ss.; LIGHTBOURNE, J.: «Damned lies & criminal sentencing using evidence-based tools», *Duke Law & Technology Review* 15, 2017, pp. 334-337.

³⁹ La sentencia cita un estudio publicado por ProPublica en 2016 que denunciaba esta situación; cfr. <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing> (consultado por última vez el 1.9.2018)

⁴⁰ Que alcanza cifras millonarias: el importe del contrato entre Northpointe, compañía propietaria de COMPAS, y el *Department of Corrections* del Estado de Wisconsin para el periodo 2015-2017 ascendió a 1.765.334 dólares (cfr. FREEMAN, K., *North Carolina Journal of Law & Technology*, vol. 18, 2016, p. 92; puede accederse al contrato en: <https://epic.org/algorithmic-transparency/crim-justice/>).

⁴¹ DE MIGUEL BERIAIN, I., *Law, Probability and Risk* 17, 2018, pp. 49-53; FREEMAN, K., *North Carolina Journal of Law & Technology*, vol. 18, 2016, pp. 92 y ss.

⁴² Por ejemplo el Public Safety Assessment (PSA) creado por la Laura and John Arnold Foundation (cfr. un resumen de su funcionamiento en <https://www.arnoldfoundation.org/wp-content/uploads/PSA-Risk-Factors-and-Formula.pdf>, consultada por última vez el 24.8.2018).

privadas tengan mejor rendimiento que las desarrolladas por las Administraciones públicas o las de código abierto, ni tampoco de que las más sofisticadas (porque incorporan más factores de riesgo) realicen mejores estimaciones de riesgo que las más simples⁴³.

En mi opinión, los tribunales no deberían admitir valoraciones de riesgo hechas con algoritmos secretos. De ningún modo si son propiedad de empresas privadas que los consideran secretos de empresa: no debe permitirse que se haga negocio a costa de las garantías procesales de los acusados. Pero tampoco si la confidencialidad deriva de algún otro tipo de protección legal: si la valoración de riesgo influye en alguna medida en la gravedad de la condena, los derechos reconocidos en el art. 24.2 CE implican a mi juicio que el acusado tiene que poder cuestionar siempre no sólo sus resultados, sino también su diseño y funcionamiento.

4. La discriminación por género

Loomis alegaba que para COMPAS ser varón es un factor que incrementa el riesgo de reincidencia, y que por ello el juez, cuando le impuso una condena severa atendiendo entre otras cosas al resultado de la valoración de riesgo, le discriminó por ser varón. Ya hemos visto (supra, II. 1) que el TSW resolvió el problema con el argumento de la tasa de acierto: como prescindir del sexo provocaría más errores en las estimaciones de riesgo (tanto de hombres como de mujeres), incluirlo no resulta discriminatorio, sino beneficioso para el sistema de justicia y para ambos grupos.

Este argumento invita a formularse la siguiente pregunta: ¿tan acertadas son las estimaciones de riesgo que merece la pena manejar un criterio tan controvertido como el del sexo/género en aras de mantener unos resultados óptimos? Dicho de otra forma: ¿tanto se perdería si por prescindir de este factor perdiéramos algo de precisión en las estimaciones? Cuando para valorar la calidad de las estimaciones hechas con HEVR se dice que suelen alcanzar un valor de 0.70 de área bajo la curva ROC⁴⁴, esto no significa que en un 70% de los casos las estimaciones de

⁴³ Estudios recientes apuntan que modelos basados en unos pocos factores de riesgo (como edad y número de condenas previas), e incluso la opinión de público lego en Derecho con una información básica sobre cada acusado, alcanzan valores predictivos casi idénticos a los de herramientas complicadas y protegidas por normas de copyright como COMPAS: cfr. DRESSEL, J. / FARID, H., «The accuracy, fairness and limits of predicting recidivism», *Science Advances* 4, 2018; ANGELINO, E. / LARUS-STONE, N. / ALABI, D. / SELTZER, M. / RUDIN, C., «Learning Certifiably Optimal Rule Lists for Categorical Data», *Journal of Machine Learning Research* 18, 2018. Y ello sin poder entrar ahora, por razones de espacio, en la cuestión de hasta qué punto las HEVR realizan realmente estimaciones mucho mejores que el juicio clínico de los expertos.

⁴⁴ *State v. Loomis*, ¶ 61; *Practitioner's Guide to COMPAS Core*, cit., p. 13.

riesgo se vean confirmadas por la conducta posterior del sujeto, sino que hay un 70% de posibilidades de que un individuo reincidente seleccionado al azar haya recibido con esa HEVR una clasificación de riesgo superior a la de un individuo no reincidente seleccionado al azar. Por regla general, del grupo de sujetos que una HEVR clasifica como de riesgo alto, menos de la mitad (a veces, muchos menos) vuelven a reincidir en el periodo de tiempo estipulado⁴⁵. Quizá estos porcentajes de acierto puedan considerarse suficientes cuando los utiliza un funcionario penitenciario para asignar a una persona a un programa de tratamiento (que, no se olvide, es el contexto inicial para el cual se diseñaron las HEVR). Pero en un juicio penal, donde se presume que toda persona es inocente hasta que no se prueba lo contrario más allá de toda duda razonable, y donde tradicionalmente ha regido la máxima de que es preferible que haya 10 culpables absueltos antes que un solo inocente condenado injustamente, resulta a mi juicio muy discutible que esos porcentajes de acierto puedan servir como base legítima para optar por una pena más grave dentro de los márgenes legales, o para tomar decisiones tan aflictivas para los derechos fundamentales como la suspensión o no de la condena⁴⁶. Si además para conseguir esos modestos índices de acierto es necesario valorar como factor de riesgo el sexo/género, fuentes especialmente sospechosas de trato discriminatorio, ello puede ser un argumento más en contra del uso de HEVR en la determinación de la pena.

Porque además, y en contra de lo que pudiera parecer, que los varones como grupo tengan mayor riesgo de reincidencia y de conducta violenta no es algo que beneficie necesariamente a las mujeres. Sí las beneficia cuando las HEVR consideran el sexo/género como un factor de riesgo, al que el algoritmo asigna mayor puntuación en caso de ser el sujeto un varón, y menor si es una mujer. Pero hay otras maneras de incluir el sexo/género en la valoración de riesgo, y parece que precisamente es el caso de COMPAS. Según la Guía que publica la compañía propietaria, las escalas de riesgo de COMPAS se construyeron aplicando la herramienta a una muestra inicial (grupo de referencia), y ordenando todas las puntuaciones que se obtuvieron de menor a mayor. Luego ese conjunto se dividió en 10 segmentos iguales, cada uno de los cuales contiene el 10% de los casos evaluados, ordenados de menor a mayor nivel de riesgo. La puntuación que se obtiene cuando se aplica COMPAS a un

⁴⁵ Me he ocupado de estas cuestiones *in extenso* en MARTÍNEZ GARAY, L., *InDret* 2/2014; MARTÍNEZ GARAY, L., «Errores conceptuales en la estimación de riesgo de reincidencia», *Revista Española de Investigación Criminológica*, n.º 14, 2016; MARTÍNEZ GARAY, L. / MONTES SUAY, F., *InDret* 2/2018. Sobre el significado del área bajo la curva ROC, cfr. con detalle la última de las publicaciones citadas.

⁴⁶ En el mismo sentido ISRANI, E., *Harvard Journal of Law and Technology. Digest*, 2017, y LIGHTBOURNE, J., *Duke Law & Technology Review* 15, 2017, p. 336: «If the criminal justice system prefers type II errors over type I, then sentences should reflect the same sentiment that objectively less risky defendants should not be subject to overly severe punishment.»

nuevo sujeto lo clasifica dentro de uno de estos 10 deciles en que se divide cada escala⁴⁷. Un resultado que coloque al sujeto en el decil más alto indica que su puntuación entra en el mismo rango que la del 10% de los sujetos más peligrosos en el grupo de referencia con el que se construyó la escala. Como advierte la propia Guía, el significado de estos deciles está ligado a las características de ese grupo de referencia: si el grupo estaba formado por sujetos con un alto riesgo de delincuencia violenta, obtener una puntuación en el decil más bajo no significa que el riesgo de esta persona sea bajo, sólo que está entre los más bajos de un grupo de alto riesgo; por el contrario, obtener una puntuación situada en el decil superior si el grupo de referencia estaba constituido por delincuentes de bajo riesgo significa que el riesgo del sujeto está entre los más altos de un grupo de bajo riesgo, pero no necesariamente que sea muy alto en términos absolutos⁴⁸.

Pues bien, COMPAS tiene escalas de riesgo diferentes para varones y para mujeres, construidas con grupos de referencia separados por sexo⁴⁹. Como los varones tienen un mayor riesgo de reincidencia y de reincidencia violenta que las mujeres, un varón con una puntuación X puede ser clasificado en un decil bajo de la escala correspondiente a su grupo normativo de varones, mientras que una mujer con la misma puntuación bruta puede entrar en uno de los deciles altos de su escala referenciada a mujeres, en la que como la media de riesgo es más baja, la misma puntuación X corresponde a un decil (nivel de riesgo) superior. De modo que la consideración del sexo en este caso más bien perjudicaría a la mujer⁵⁰.

En definitiva, la pregunta de hasta qué punto es admisible incluir el sexo/género en las valoraciones de riesgo no admite una respuesta tan sencilla como la que ofreció el TSW en el caso Loomis, pues ni el argumento de que mejora los niveles de acierto de las estimaciones es a mi juicio por sí solo decisivo, ni resulta sencillo determinar si produce o no efectos discriminatorios entre varones y mujeres.

IV. Consideraciones finales

Desde las décadas de 1970-80 han cambiado mucho tanto el concepto de peligrosidad como sus métodos de determinación, y también las normas que lo consideran presupuesto para la imposición de consecuencias jurídico-penales afflictivas. Sin embargo, continúa plenamente justificada la preocupación que en aquella época expresaran muchos autores sobre su compatibilidad con las garantías penales y procesales básicas en un

⁴⁷ *Practitioner's Guide to COMPAS Core*, cit., p. 8.

⁴⁸ *Ibid.*, p. 11.

⁴⁹ *Ibid.*, p. 11 in fine.

⁵⁰ LIGHTBOURNE, J., *Duke Law & Technology Review* 15, 2017, pp. 339 y ss.

Estado de Derecho. El caso Loomis ejemplifica las nuevas formas en las que hoy en día se presentan algunos de estos —viejos— problemas de compatibilidad, así como las dificultades que encuentran los tribunales al enfrentarse a estas cuestiones.

Loomis recurrió su condena ante el Tribunal Supremo de los EEUU, pero éste rechazó en junio de 2017 revisar el caso⁵¹. Antes de tomar la decisión, no obstante, el Tribunal invitó al *Solicitor General* del Gobierno de los EEUU a exponer su opinión sobre el asunto, y al hacerlo el Gobierno reconoció que planteaba importantes cuestiones de relevancia constitucional, e incluso admitió que seguramente antes o después deberán ser resueltas por la Corte Suprema⁵². No obstante, aconsejó al Tribunal no admitir el recurso de Loomis, entre otras razones porque la utilización de valoraciones de riesgo en las sentencias era aún muy reciente y había dado lugar todavía a muy pocos pronunciamientos de los tribunales de distrito y estatales, no existiendo tampoco aún recomendaciones oficiales de la *United States Sentencing Commission*, lo que aconsejaba esperar a que el tema hubiera madurado previamente en los tribunales y en la doctrina científica antes de sentar jurisprudencia.

Estas líneas pretenden contribuir, en una modestísima medida, a dicho proceso de «maduración». La valoración estructurada de riesgo supone un cambio en la forma de tomar decisiones en Derecho penal y penitenciario que afecta de forma novedosa a importantes derechos fundamentales, no sólo los cuestionados en el caso Loomis y no sólo en el momento de dictar sentencia⁵³. Es necesario conocer mejor estos nuevos riesgos para las garantías penales y construir los criterios jurídicos a los que deba someterse el uso de esta nueva metodología, en especial si continúa extendiéndose, cosa que parece probable.

⁵¹ Cfr. <http://www.scotusblog.com/case-files/cases/loomis-v-wisconsin/>, consultada por última vez el 23.7.2018.

⁵² WALL, J. B. (Acting Solicitor General) *et al.*, *Brief for The United States as amicus curiae* (mayo 2017), online en www.scotusblog.com, loc. cit. en nota anterior, pp. 12 y 21.

⁵³ Por ejemplo, ¿puede el reo —que sabe que con sus respuestas se hará una valoración de riesgo que podría perjudicarlo— alegar el derecho a guardar silencio frente a todas o algunas de las preguntas que se le formulan? ¿Está justificada y resulta proporcional a la utilidad que de ellas se extrae la extensa intromisión en la intimidad que supone la administración de algunas HEVR? En relación con la fase de ejecución de la pena de prisión, algunos abogados han expresado preocupaciones de este tipo en relación con el RisCanvi, una HEVR implantada en el sistema penitenciario de Cataluña: <http://www.micapinforma.com/2015/11/los-sistemas-informaticos-de-valoracion.html> (consultada por última vez el 1.9.2018)