

La incertidumbre de los pronósticos de peligrosidad: consecuencias para la dogmática de las medidas de seguridad

Lucía Martínez Garay

Universidad de Valencia

Abstract^{*/**}

La legitimidad de las medidas de seguridad depende de que puedan demostrar su capacidad para alcanzar el fin que las justifica: evitar la comisión de delitos futuros. Esta racionalidad instrumental presupone a su vez que ha de ser posible identificar con claridad su presupuesto de aplicación: la probabilidad de comisión de futuros delitos por un sujeto concreto. Si la peligrosidad no puede establecerse con un suficiente grado de certeza, la aplicación de las medidas sería arbitraria y, en consecuencia, ilegítima. Este trabajo analiza los porcentajes de acierto que evidencian los distintos métodos de predicción de la peligrosidad criminal, y da cuenta de las limitaciones estructurales que presentan a partir de los datos ofrecidos por recientes investigaciones en relación con los instrumentos de predicción estadísticos o actuariales y de juicio clínico estructurado. A partir de dicho análisis se concluye no sólo que la capacidad para identificar correctamente a los sujetos que volverán a delinquir es muy limitada sea cual sea el método empleado para predecir la peligrosidad, sino también que dicha capacidad predictiva no parece probable que vaya a mejorar de manera significativa en un futuro cercano. Ello hace que resulten cuestionables diversos aspectos de las medidas de seguridad, y en particular la legitimidad de las medidas imponibles a sujetos responsables de forma acumulada a la pena correspondiente al injusto culpable.

Die Legitimität der Maßnahmen der Besserung und Sicherung hängt davon ab, ob sie das Ziel erreichen können, das sie rechtfertigt: die Vermeidung von zukünftigen Straftaten. Diese instrumentelle Rationalität setzt wiederum voraus, dass es möglich sein muss, ihre Anwendungsvoraussetzung eindeutig zu identifizieren: die Wahrscheinlichkeit, dass eine bestimmte Person in der Zukunft weitere Straftaten begehen wird. Wenn die Gefährlichkeit nicht mit hinreichender Sicherheit festgestellt werden kann, wäre die Anwendung der Maßnahmen willkürlich und daher ungerechtfertigt. Dieser Aufsatz analysiert die Erfolgsraten, die die verschiedenen Methoden der Vorhersage kriminellen Verhaltens aufweisen, und stellt ihre strukturellen Unzulänglichkeiten dar, anhand von den Ergebnissen, die neuere Forschungen über nomothetische Prognoseinstrumente ergeben. Aus dieser Analyse folgt erstens, dass die Möglichkeiten, zukünftige Straftäter zu identifizieren, unabhängig von der verwendeten Methode sehr beschränkt sind, und zweitens, dass es unwahrscheinlich ist, dass die Vorhersagekraft der Gefährlichkeitsprognosen in absehbarer Zukunft sich deutlich verbessern wird. Das macht verschiedene Aspekte der Sicherheitsmaßnahmen fragwürdig, insbesondere die Legitimität derer, die gegen verantwortliche Täter kumulativ zur Strafe für das schuldhaft Unrecht verhängt werden.

The legitimacy of the security measures depends upon whether they are able to achieve the purpose that justifies them: to avoid the commission of further crimes. This instrumental rationality presupposes in turn that it must be possible to clearly identify the cases to which they are applicable: the likelihood of committing future crimes by a particular subject. If dangerousness cannot be established with a sufficient degree of certainty, to apply the

* Este trabajo se ha realizado en el marco de los Proyectos de I+D DER2009-13295, dirigido por el Prof. Dr. E. Orts Berenguer, y DER2010-18825, dirigido por el Prof. Dr. J.C. Carbonell Mateu, ambos financiados por el Ministerio de Ciencia e Innovación (ahora Ministerio de Economía y Competitividad).

** El epígrafe 5 ha sido redactado en coautoría con el Prof. Dr. Francisco Montes Suay, Catedrático de Estadística e Investigación Operativa de la Universidad de Valencia.

measures would be arbitrary and therefore illegitimate. This paper analyses the success rates of the various methods of predicting criminal dangerousness, and describes the structural limitations they present, based on recent investigations on actuarial risk assessment instruments and those of structured clinical judgment. From this analysis it follows that not only the ability to correctly identify subjects who will re-offend is very limited regardless of the method used to predict dangerousness, but also that this predictive ability does not seem likely to be improving significantly in the near future. This makes different aspects of the security measures questionable, and particularly the legitimacy of measures imposed on responsible offenders in addition to the penalty they deserve for the wrongdoing they already committed.

Titel: Die Unsicherheit der Gefährlichkeitsprognose: Folgerungen für die Dogmatik der Maßregeln der Besserung und Sicherung.

Title: The uncertainty of risk assessment in criminal law: consequences for the concept, legitimacy and limits of security measures.

Palabras clave: pronósticos de peligrosidad, peligrosidad criminal, *risk assessment*, falsos positivos, medidas de seguridad, proporcionalidad, culpabilidad.

Stichwörter: Gefährlichkeitsprognose, falsch Positiven, Maßnahmen der Besserung und Sicherung, Verhältnismäßigkeit, Schuld.

Keywords: prediction of dangerousness, risk assessment, false positives, security measures, proportionality, culpability.

Sumario

1. *Introducción*
2. *Peligrosidad – valoración del riesgo (risk assessment) – falsos positivos*
3. *Cómo se predice la peligrosidad*
4. *Algunos datos empíricos*
 - 4.1 *En otros países*
 - 4.2 *En España*
5. *Limitaciones inherentes a los métodos estadísticos de predicción de la peligrosidad*
 - 5.1 *Sensibilidad – especificidad – valor predictivo*
 - 5.2 *Tasa de prevalencia del fenómeno (base rate) y errores que estamos dispuestos a asumir*
 - 5.3 *Valores predictivos de los instrumentos de valoración del riesgo*
 - 5.4 *Los márgenes de error y la traslación de los resultados grupales a la predicción del comportamiento individual*
 - 5.5 *El problema del «riesgo moderado»*
6. *La cuestión en los tribunales*
7. *Conclusiones*
8. *Consecuencias para la dogmática de las medidas de seguridad y para algunas penas*
 - 8.1 *Fundamento de las medidas de seguridad*
 - 8.2 *Proporcionalidad de las medidas de seguridad*
 - 8.3 *Pronósticos en diferentes contextos. Relación entre culpabilidad y peligrosidad*

8.4 Prisión permanente revisable

9. Consideraciones finales

10. Tabla de jurisprudencia

11. Bibliografía

«Prediction is very difficult, especially if it's about the future»

Niels Bohr

1. Introducción

La tendencia a la intensificación de la respuesta penal frente a los delincuentes considerados peligrosos es un fenómeno que va mucho más allá de nuestras fronteras, y que hace ya décadas que impregna la política criminal en muchos países.¹ Hasta 2010 en España este endurecimiento se había venido haciendo a través de la agravación de las penas y de su régimen de cumplimiento, y no mediante la previsión de medidas de seguridad; sin embargo, esto cambió con la LO 5/2010, que introdujo la libertad vigilada para delincuentes plenamente responsables en relación con los delitos contra libertad sexual y de terrorismo. Y todo apunta a que esta tendencia se va a consolidar, como así lo evidencia la ampliación que experimenta la libertad vigilada en el Proyecto de reforma del Código Penal recientemente presentado por el Gobierno ante las Cortes,² así como el intento – si bien de momento (¿temporalmente?) abandonado – de introducir en nuestro ordenamiento la custodia de seguridad, tal como lo contemplaba el Anteproyecto de reforma del Código Penal que precedió al Proyecto de ley que hoy está en tramitación parlamentaria.

Estos cambios legales han ido paralelos a una evolución en la doctrina penal española, que según creo también parece confiar cada vez más en lo que pueda dar de sí la peligrosidad, habida cuenta que hace ya años vienen aumentando las voces que abogan por la necesidad de ampliar el limitado margen que dejó el legislador en 1995 para las medidas de seguridad, insistiendo en la conveniencia

¹ Cfr. por todos el conocido trabajo de GARLAND, *The culture of control*, 2001, cuyo análisis se refiere fundamentalmente a la evolución de la política criminal en el ámbito estadounidense y británico. En relación con nuestro país, cabe distinguir también perfectamente una creciente importancia de la inocuización del delincuente considerado peligroso dentro de la tendencia general al aumento del rigor punitivo que ha caracterizado todas las reformas penales de los últimos años; entre los primeros autores que destacaron este aspecto, si bien desde perspectivas diferentes, cabe citar a Díez RIPOLLÉS, «De la sociedad del riesgo a la seguridad ciudadana: un debate desenfocado», *RECPC*, (07-01), 2005, pp. 17 y ss., 26 y ss., y SILVA SÁNCHEZ, «El retorno de la inocuización» en REDONDO (coord.), *Delincuencia sexual y sociedad*, 2002, pp. 156 y ss. Entre la numerosa bibliografía que desde entonces se ha ido publicando en nuestro país sobre el tratamiento penal del delincuente peligroso, cfr., por citar sólo dos monografías recientes, ACALE SÁNCHEZ, *Medición de la respuesta punitiva. Tratamiento penal del delincuente imputable peligroso*, 2010, y ARMAZA ARMAZA, *El tratamiento penal del delincuente imputable peligroso*, 2013.

² Cfr. Proyecto de Ley Orgánica por la que se modifica la Ley Orgánica 10/1995, de 23 de noviembre, del Código Penal, BOCG. Congreso de los Diputados Núm. A-66-1 de 04/10/2013.

de introducir en nuestro Derecho penal medidas de seguridad para imputables peligrosos.³ No cabe duda de que se trata además de un tema de indudable actualidad y relevancia mediática: buena muestra de ello es el revuelo que ha producido la reciente sentencia del TEDH desautorizando la «doctrina Parot»,⁴ y sobre todo la cobertura mediática de las excarcelaciones que la han seguido, pues si bien en los primeros días la atención estuvo centrada en la salida de prisión de internos condenados por delitos de terrorismo, así como en las reacciones de las asociaciones de las víctimas de estos delitos, el foco se ha ampliado después a las excarcelaciones de asesinos comunes y violadores múltiples considerados aún peligrosos por la Administración penitenciaria. Mientras que en el caso de los terroristas la crítica a las excarcelaciones ha recurrido a argumentos de tipo retributivo («no puede valer igual un asesinato que veinte»), en el caso de los delincuentes sexuales la preocupación se centra en el temor a la reincidencia, y en la inexistencia de medidas penales susceptibles de controlar dicho peligro.⁵

La peligrosidad es, sin embargo, un concepto al que tradicionalmente la doctrina penal se había acercado con mucha cautela. Como es sabido se ha advertido repetidamente de que se trata de un concepto muy «peligroso», adjetivo que se debe entre otras cosas al reconocimiento generalizado de que no disponemos de conocimientos empíricos suficientes para garantizar que el pronóstico de

³ Ha sido siempre partidario de esta vía CEREZO MIR, véase por ejemplo su trabajo «Consideraciones político-criminales sobre el nuevo Código Penal de 1995» en *Estudios jurídicos en memoria del profesor Dr. D. José Ramón Casabó Ruiz*, v. 1, 1998, pp. 390 y ss., donde calificaba la ausencia de medidas de seguridad para delincuentes imputables peligrosos como «uno de los mayores defectos del Código penal, desde el punto de vista político-criminal» (ob. cit., p. 390). Pero también son de esta opinión otros autores, si bien con diferencias importantes en las propuestas que formula cada uno de ellos, pues no necesariamente abogan por la acumulación sucesiva de pena y medida de seguridad. Véase por ejemplo JORGE BARREIRO, «Reflexiones sobre la compatibilidad de la regulación de las medidas de seguridad en el CP de 1995 con las exigencias del Estado de Derecho» en *Homenaje al Profesor Dr. Gonzalo Rodríguez Mourullo*, 2005, pp. 581 y ss., 585, o SANZ MORÁN, *Las medidas de corrección y de seguridad en el Derecho Penal*, 2003, pp. 215 y ss. y especialmente p. 218. Cfr. asimismo GRACIA MARTÍN, «Las medidas de seguridad y de reinserción social» en EL MISMO (coord.), *Tratado de las consecuencias jurídicas del delito*, 2006, pp. 435 a 437, 457, 469; ZUGALDÍA ESPINAR, «Medidas de seguridad complementarias y acumulativas para autores peligrosos tras el cumplimiento de la pena», *RDPC*, 2009, pp. 199 y ss. (que aboga por abrir el debate al respecto, y considera que cabría admitir medidas post-condena tanto no privativas de libertad como de internamiento, incluyendo la custodia de seguridad para casos excepcionales, pp. 209 y s.); REQUEJO RODRÍGUEZ, «Peligrosidad criminal y Constitución», *InDret*, (3), 2008, p. 20 (considerando asumible añadir una medida de seguridad tras el cumplimiento de la pena para delincuentes imputables, como vía para superar «las deficiencias de prevención especial» que evidencia nuestro Código penal), o GARCÍA ALBERO, «Ejecución de penas en el proyecto de reforma. Estudio de un problema concreto: ¿qué hacer con los reos habituales o reincidentes en los que subsiste la peligrosidad criminal tras el licenciamiento definitivo?» en ÁLVAREZ GARCÍA (dir.), *La adecuación del Derecho penal español al ordenamiento de la Unión Europea. La política criminal europea*, 2009, especialmente pp. 135 y ss. (quien propone introducir una libertad vigilada que tenga formalmente la naturaleza de pena accesoria - aunque materialmente sea una medida de seguridad - porque entiende que en el marco constitucional español existen dificultades para justificar medidas de seguridad añadidas a la pena, pp. 137 y s.). Hay que citar también en este contexto la evolución de SILVA SÁNCHEZ, partidario ahora de prever medidas de internamiento incluso permanente para sujetos especialmente peligrosos («El contexto del Anteproyecto de Reforma del Código Penal de 2008» en AAVV, *El Anteproyecto de modificación del Código Penal de 2008. Algunos aspectos (Cuadernos penales José María Lidón, Núm. 6)*, 2009, pp. 24 y s.), frente a la opinión en principio contraria a las medidas de seguridad adicionales a la pena para delincuentes imputables peligrosos que manifestó en trabajos anteriores (como en «La regulación de las medidas de seguridad (artículo 6)» en EL MISMO, *El nuevo Código penal: cinco cuestiones fundamentales*, 1997, p. 45).

⁴ Sentencia del TEDH 21.10.2013, caso Del Río Prada contra España.

⁵ Cfr. por ej. el amplio reportaje aparecido en el diario *El País*, de 26 de noviembre de 2013, con el expresivo título «Poco que hacer frente a los violadores libres».

comisión de futuros delitos por parte de un sujeto concreto se realice con un suficiente grado de certeza.⁶ Y si no estamos en condiciones de identificar correctamente la peligrosidad, entonces la aplicación de consecuencias jurídicas desfavorables para la persona sobre esta base tan incierta resulta altamente cuestionable. También se ha señalado que los estudios empíricos demostrarían que existe una tendencia a sobrevalorar la peligrosidad, es decir, a clasificar como peligrosos a sujetos que no lo son (falsos positivos), y que esta sobreestimación del peligro vendría además alentada por la presión que ejercen los medios de comunicación sobre los operadores jurídicos, pues el enfoque alarmista que aplican aquéllos cuando se pone de manifiesto un falso negativo (persona a la que no se consideró peligrosa, y a la que por ello se le concedió un permiso de salida, o la libertad condicional, o no se le aplicó medida de seguridad, y que sin embargo luego delinque) acentúa la tendencia en los últimos a ser excesivamente cautos e inclinarse demasiadas veces por apreciar la existencia de peligrosidad.⁷

Partiendo de estos antecedentes, y si se tiene en cuenta el clima de permanente alarma social en relación con el delito en el que vivimos desde hace años, así como el entusiasmo – por emplear un término suave – con el que el legislador responde o incluso se anticipa a cualquier demanda real o presunta de endurecimiento de la legislación penal que proceda de la opinión pública (o publicada), no parece que corran precisamente los mejores tiempos para adentrarse en este complicado jardín y ampliar el ámbito de la peligrosidad y de las medidas de seguridad. Sin embargo, es precisamente este mismo clima de alarma y temor a la delincuencia el que hace que, como hemos visto, el concepto de peligrosidad (y con él la necesidad de su determinación a través de unos u otros procedimientos) esté cada día más presente en el Derecho penal, especialmente por lo que respecta a la imposición de medidas de seguridad, y más en concreto medidas de seguridad adicionales a la pena para delinquentes imputables peligrosos.

Creo que a la vista de este estado de cosas es necesario detenerse a hacer una reflexión, y preguntarse cuál es el estado actual de la cuestión por lo que respecta al pronóstico de peligrosidad. Porque siendo la peligrosidad el presupuesto para la aplicación de las normas que prevén medidas de seguridad, la posibilidad de determinarla con un razonable grado de certeza es el requisito esencial para imponer con legitimidad esta clase de consecuencias jurídicas del delito.⁸ Como trataré de

⁶ Cfr., por todos, SANZ MORÁN, *Las medidas de corrección y de seguridad en el Derecho penal*, 2003, pp. 93 y ss., donde alude no sólo a éste sino también a otros problemas que plantea el concepto de peligrosidad, como por ejemplo que la emisión del pronóstico de peligrosidad supone siempre una estigmatización del ciudadano sobre el que recae, que puede además favorecer que efectivamente el sujeto termine adaptando consciente o inconscientemente su comportamiento a la etiqueta que ha recibido (*self-fulfilling prophecy*); también se ha aducido que la libertad de decisión humana y el azar serían factores que invalidarían siempre cualquier posibilidad de pronóstico fiable.

⁷ Cfr. SANZ MORÁN, *Las medidas de corrección y de seguridad en el Derecho Penal*, 2003, pp. 102 a 106, y, con detalle, ALBRECHT, «Sinn und Unsinn der Prognose von Gewaltkriminalität» en HEITMEYER/SOEFFNER (eds.) *Gewalt. Entwicklungen, Strukturen, Analyseproblemen*, 2004.

⁸ Como afirma STRENG, «die Tragfähigkeit jeder Maßregelrechtfertigung steht und fällt [...] mit der Sicherheit der Gefährlichkeitsfeststellung» (STRENG, «Das Legitimations-Dilemma sichernden Freiheitsentzugs – Überlegungen zur neueren Rechtsentwicklung» en DÖLLING (ed.), *Ius humanum: Grundlagen des Rechts und Strafrecht; Festschrift für Ernst-Joachim Lampe zum 70. Geburtstag*, 2003, p. 621). En sentido parecido, KAISER, *Befinden sich die Kriminalrechtlichen Maßregeln in der Krise?*, 1990, pp. 16 y 17; SCHÖCH, *LK*, 12ª ed., 2008, Vor §61, nm. 142 a 144, considerándolo el «problema central» de las medidas de seguridad; ZIFFER, *Medidas de seguridad. Pronósticos de peligrosidad en Derecho*

demostrar a lo largo de las páginas que siguen, la enorme incertidumbre de los pronósticos de peligrosidad obliga en mi opinión no sólo a cuestionar la viabilidad de medidas de seguridad adicionales a la pena para los delincuentes imputables peligrosos, sino a replantearse varios aspectos del régimen jurídico de las medidas imponibles en lugar de la pena.⁹

Antes de empezar, una última advertencia: este trabajo se ocupa del concepto de peligrosidad y de las posibilidades y límites de realizar un pronóstico sobre la futura comisión de delitos desde la óptica de las medidas de seguridad y de sus presupuestos de aplicación, entre los cuales la constatación de la peligrosidad del sujeto es seguramente el esencial. Desde luego la realización de pronósticos sobre el futuro comportamiento de un delincuente no es necesaria sólo en este ámbito, sino que es un presupuesto para la adopción de otras muchas decisiones en la aplicación de consecuencias jurídico-penales, como la concesión o no de la suspensión de la ejecución de la pena,¹⁰ la concesión de la libertad condicional¹¹, la clasificación penitenciaria,¹² etc. Por ello las consideraciones que se efectuarán en las líneas que siguen son en buena medida aplicables a estos otros ámbitos también. Sin embargo, no todas lo son, ni lo son de manera automática: si bien los problemas empíricos de la valoración del riesgo de reincidencia que a continuación se explicarán son comunes a cualquier decisión jurídica que tenga que adoptarse sobre la base de los mismos, la manera de solucionar el problema que este estado de cosas deja planteado no tiene por qué serlo, tal y como intentaré dejar al menos apuntado en los últimos epígrafes de este trabajo. Por esta razón quiero advertir que el análisis que aquí efectuaré está hecho pensando principalmente, y salvo que expresamente se indique otra cosa, en las medidas de seguridad.

penal, 2008, p. 151: «la determinación de la conducta futura del autor constituye el punto crítico del derecho de las medidas».

⁹ El objetivo de este trabajo es por tanto modesto, limitado a subrayar los grandes márgenes de incertidumbre y error que rodean el pronóstico de peligrosidad, y no entra en las críticas que cabe efectuar en general al modelo de gestión de la criminalidad basado en la idea de incapacitación selectiva. Dicho de otro modo: aun en el caso de que las predicciones de peligrosidad pudieran ser hechas con grados de acierto mucho mayores de los que evidencian en la actualidad, utilizar la peligrosidad y en general las predicciones (tanto individuales como grupales) como criterio sobre el que diseñar la política criminal plantea otros problemas importantes, sobre los que ha llamado la atención con argumentos muy interesantes HARCOURT (*Against Prediction: profiling, policing and punishing in an actuarial age*, 2007), pero en los que, insisto, no entraré aquí, salvo por algunas referencias muy puntuales y limitadas a las medidas de seguridad en los epígrafes 8 y 9.

¹⁰ Art. 80 CP: «1. Los jueces o tribunales podrán dejar en suspenso la ejecución de las penas privativas de libertad no superiores a dos años mediante resolución motivada. En dicha resolución se atenderá fundamentalmente a la peligrosidad criminal del sujeto, así como a la existencia de otros procedimientos penales contra éste.»

¹¹ Es requisito para la conexión de la libertad condicional, entre otras cosas, que los penados «hayan observado buena conducta y exista respecto de los sentenciados un pronóstico individualizado y favorable de reinserción social» (art. 90.1.c) CP).

¹² Art. 10 LOGP: «No obstante lo dispuesto en el número Uno del artículo anterior, existirán establecimientos de cumplimiento de régimen cerrado o departamentos especiales para los penados calificados de peligrosidad extrema o para casos de inadaptación a los regímenes ordinario y abierto [...]. Según el art. 62 LOGP, el tratamiento penitenciario «Guardará relación directa con un diagnóstico de personalidad criminal y con un juicio pronóstico inicial [...]. Cfr., además, el art. 67 de la misma norma: «Concluido el tratamiento o próxima la libertad del interno, se emitirá un informe pronóstico final, en el que se manifestarán los resultados conseguidos por el tratamiento y un juicio de probabilidad sobre el comportamiento futuro del sujeto en libertad, que, en su caso, se tendrá en cuenta en el expediente para la concesión de la libertad condicional.»

2. Peligrosidad – valoración del riesgo (risk assessment) – falsos positivos

La peligrosidad criminal es la probabilidad de comisión de delitos futuros por parte de una determinada persona. Considero que esta sencilla definición es compartida por la generalidad de la doctrina,¹³ con independencia de que después haya importantes discrepancias en torno a si el concepto es útil o no, peligroso o necesario, científicamente fundamentado o una mera cobertura ideologizada para la criminalización del disidente o desviado.¹⁴ Pero dejando de momento a un lado la discusión sobre estas cuestiones, sobre algunas de las cuales habrá ocasión de volver más adelante, es necesario efectuar en este punto una precisión, que inicialmente puede resultar meramente terminológica pero que en realidad creo que refleja dos modos distintos de concebir la peligrosidad criminal.

Hace ya años que en psicología y psiquiatría el término «peligrosidad» está siendo sustituido por la expresión «valoración (o estimación) del riesgo», o más en concreto, valoración del riesgo de violencia (en inglés, «*violence risk assessment*»)¹⁵ El cambio en la terminología implica según los especialistas de estas disciplinas un cambio en la manera de concebir aquello que se predice.¹⁶ El concepto de peligrosidad estaría ligado, según afirman, a un entendimiento de ésta como una propiedad o atributo inherente al individuo, la cualidad subjetiva de ser violento; designaría por tanto una tendencia personal a la comisión de delitos vinculada inicialmente (en las formulaciones del positivismo italiano) a estados mentales patológicos – vinculación que en cierta forma ha

¹³ Cfr. ya JORGE BARREIRO, *Las medidas de seguridad en el Derecho español*, 1976, pp. 248, 257. En la actualidad, por todos, URRUELA MORA, *Las medidas de seguridad y reinserción social en la actualidad. Especial consideración de las consecuencias jurídico-penales aplicables a sujetos afectados de anomalía o alteración psíquica*, 2009, p. 67. Sobre las diferencias entre la peligrosidad criminal y la peligrosidad social, en las que no es necesario entrar aquí por estar suficientemente claras en la doctrina cfr. por ej. URRUELA MORA, *Las medidas de seguridad y reinserción social en la actualidad*, pp. 66 y s., o SANZ MORÁN, *Las medidas de corrección y de seguridad en el Derecho penal*, 2003, p. 89. Por desgracia, la diferencia entre peligrosidad criminal (la única relevante en el Código penal vigente) y peligrosidad social no es sin embargo trazada de manera tan nítida por nuestros tribunales, cfr. por ej., STS 345/2007, de 24 de abril, FD segundo, que utiliza indistintamente ambos términos como si fueran sinónimos: «Peligrosidad criminal: esto es, que una persona se considere potencialmente idónea para cometer acciones «antisociales», o dañosas [...] la medida de seguridad no se impone -sin más-, como un remedio terapéutico para el enfermo mental, inimputable penalmente, sino en función de la peligrosidad social del sujeto, y del pronóstico de reincidir en su comisión criminal. En definitiva, se trata de la adopción de una medida de seguridad socialmente defensiva, que se adopta dada la peligrosidad criminal del reo». En términos parecidos, STS núm. 65/2011 de 2 febrero, o STS núm. 216/2012 de 1 febrero.

¹⁴ En este último sentido, muy contundente, LÓPEZ-REY, «Algunas reflexiones sobre el peligro de la noción de peligrosidad», *Cuadernos de Política Criminal*, (19), 1983, pp. 19 y ss.

¹⁵ ANDRÉS PUEYO afirma que no sólo es un cambio terminológico en la literatura especializada, sino que diversos países del ámbito anglosajón y otros han incorporado el concepto de riesgo a sus legislaciones en sustitución de las referencias a la peligrosidad (ANDRÉS PUEYO, «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*, 2013, p. 489). Por otro lado, si se hace la prueba de introducir los conceptos «*violence risk assessment*» y «*dangerousness*» en los campos de título o materia en la base de datos SSCI de Web of Knowledge lo cierto es que son bastante más numerosas las referencias que aparecen en relación con la primera expresión que con la segunda.

¹⁶ Cfr. sobre lo que sigue ANDRÉS PUEYO/REDONDO ILLESCAS, «Predicción de la violencia: entre la peligrosidad y la valoración del riesgo de violencia», *Papeles del Psicólogo*, (28-3), 2007, pp. 158 a 165, y STEADMAN, «From Dangerousness to Risk Assessment of Community Violence: taking stock at the turn of the century», *Journal of the American Academy of Psychiatry and the Law* (28:265-71), 2000, pp. 267 y ss. También ANDRÉS PUEYO en DEMETRIO CRESPO (dir.), *Neurociencias y Derecho Penal*, 2013, pp. 483 y ss.

pervivido en la tendencia que sigue estando presente de considerar más peligrosos a quienes padecen un trastorno mental que a quienes no lo sufren –, pero que con el tiempo se fue alejando de estos determinantes biológicos, aunque se mantuvo como un atributo individual, y como tal estable en el tiempo, y además dicotómico, en el sentido de que o se es peligroso o no se es.

La idea de valoración del riesgo de violencia, por el contrario, parte de la base de que lo que se predice es la probabilidad de ocurrencia del acto violento, no la cualidad de ser violento inherente al individuo.¹⁷ Mientras que con la realización del pronóstico de peligrosidad tradicionalmente se ha clasificado a los individuos en personas peligrosas y no peligrosas, cuando se habla de valoración del riesgo lo que se hace es emitir un juicio acerca de la probabilidad de ocurrencia de un evento futuro, ocurrencia que estará ligada desde luego a factores subjetivos individuales (como la personalidad, el carácter, la educación recibida, los hábitos adquiridos, etc.) pero que también dependería de las circunstancias ambientales y situacionales en que se desarrolle el comportamiento futuro.¹⁸ Estos factores contextuales, externos al sujeto y determinados por el momento y lugar concretos de la acción, tendrían más peso en la predicción del comportamiento futuro si ésta se realiza desde la perspectiva de la valoración del riesgo de violencia que si se hace desde una concepción centrada en la peligrosidad como atributo o propensión individual.¹⁹

Por otro lado, se critica también al concepto de peligrosidad el ser muy impreciso y vago, ya que ni especifica suficientemente los factores cuya presencia convierte a un individuo en peligroso, ni concreta la clase de acto delictivo y/o violento que se espera que cometa el sujeto, sino que se refiere en general a la comisión de «delitos» en el futuro. Ambas cosas condicionarían muy negativamente la capacidad predictiva del constructo «peligrosidad», pues de un lado la clave para una evaluación objetiva, fiable y válida de algo sería una definición lo más clara y precisa posible de ese algo, y por otro lado la predicción de un suceso con un mínimo grado de fiabilidad requeriría asimismo definirlo de la manera más exacta posible.²⁰ Frente a este entendimiento, el enfoque centrado en la valoración del riesgo, que sería el propio de los procedimientos de predicción estructurados (ya sean actuariales o de juicio clínico estructurado, sobre estos conceptos cfr. *infra* en el epígrafe 3.), se basa en las correlaciones existentes entre factores de riesgo claramente establecidos, y el criterio a predecir, también definido con precisión – determinado tipo de comportamiento, violento y/o delictivo –, relaciones que estarían demostradas empíricamente o, al menos, serían estadísticamente

¹⁷ ANDRÉS PUEYO/REDONDO ILLESCAS, *Papeles del Psicólogo*, 2007, p. 159.

¹⁸ Como afirma ANDRÉS PUEYO: «las causas de la conducta violenta no son exclusivamente atribuibles a factores del individuo. Así, por ejemplo, un sujeto fumador no es peligroso en sí mismo, pero sí lo es si está trabajando en una gasolinera y especialmente si fuma cuando está rellenando el depósito de un automóvil.» (ANDRÉS PUEYO en DEMETRIO CRESPO (dir.), *Neurociencias y Derecho Penal*, 2013, p. 498).

¹⁹ Cfr. de nuevo ANDRÉS PUEYO/REDONDO ILLESCAS, *Papeles del Psicólogo*, 2007, pp. 159, 163, 164, y STEADMAN, *J Am Acad Psychiatry Law*, 2000, quien señala que las estimaciones de probabilidad propias del enfoque basado en la valoración del riesgo tienen dos características principales: «(1) probabilities may change over time; and (2) the probability is not a trait of the person but is rather a product of an assessment by a clinician or clinical team that reflects their best judgments about how that person's presenting characteristics, history and future interactions with his or her environment together produce a likelihood of future harm» (p. 268).

²⁰ Especialmente crítico en este sentido ANDRÉS PUEYO en DEMETRIO CRESPO (dir.), *Neurociencias y Derecho Penal*, 2013, pp. 495 y ss. y *passim*.

significativas. Como cada tipo de conducta tiene sus propios predictores específicos, se afirma que la capacidad predictiva cuando se procede de esta manera aumenta considerablemente.²¹

Por último, la perspectiva de la valoración del riesgo tendría frente al concepto tradicional de peligrosidad dos ventajas ulteriores: en primer lugar, que la valoración del riesgo permite tomar en consideración no sólo los factores de riesgo (aquéllos que con su presencia favorecen la comisión de delitos) sino también lo que se conoce como factores protectores (aquéllos cuya presencia disminuye la probabilidad de reincidencia), y ello, combinado con la insistencia que se pone desde esta perspectiva en la mutabilidad de las circunstancias en que puede hallarse una persona (así como en la posibilidad de modificar también determinados rasgos individuales), favorecería actuaciones tendentes no sólo a diagnosticar el riesgo, sino también a incidir sobre los factores que lo condicionan para disminuirlo. En otras palabras, el enfoque de la valoración del riesgo permitiría no sólo el diagnóstico del peligro sino también la gestión del mismo, lo que significa contribuir a diseñar estrategias que ayuden al sujeto a evitar la comisión de delitos en el futuro.²² En segundo lugar, a diferencia de la peligrosidad que sería una variable discreta (sí/no), el riesgo de violencia o de comportamiento delictivo futuro es un constructo continuo que ofrece grados de probabilidad, y que pone esta probabilidad en relación con la ocurrencia de concretos comportamientos en el futuro. Esto constituiría una ventaja, en la medida en que «hay sujetos que únicamente son peligrosos para determinado tipo de víctimas y no para otras» y que «el juicio clínico dicotómico (peligrosidad sí o no) provoca una pérdida de la información relativa y temporal que tanto afecta a las predicciones, al no existir graduaciones intermedias relacionadas con el nivel de probabilidad de aparición futura de la conducta criminal».²³

En mi opinión, algunos aspectos de esta crítica destacan efectivamente problemas inherentes al concepto de peligrosidad que los juristas haríamos bien en tener presentes, mientras que otras de las supuestas ventajas que se señalan como propias del enfoque basado en la valoración del riesgo de violencia quizá no sean tales. Empezando por estas últimas, desde luego que el concepto de peligrosidad es tremendamente impreciso, pero por desgracia seguramente ocurre algo parecido con el de valoración del riesgo, porque habría que concretar: riesgo, ¿de qué? Si lo que valoramos es el riesgo de violencia en general, desde el campo de la psicología y psiquiatría se advierte de que también éste es un concepto vago y amplio donde los haya, con manifestaciones muy diferentes (pues hay distintas clases de violencia: sexual, intrafamiliar, verbal o física, habitual u ocasional, etc.) y de significado difícil de precisar.²⁴ Si atendemos a la variable que predicen los modernos instrumentos elaborados desde el enfoque de la valoración del riesgo, los hay de muy diversa configuración: si bien algunos son específicos para la predicción de algunas clases de violencia (violencia sexual, en el caso del SVR-20, o violencia contra la pareja en el caso del SORAG), otros

²¹ ANDRÉS PUEYO/REDONDO ILLESCAS, *Papeles del Psicólogo*, 2007, pp. 162, 164.

²² ANDRÉS PUEYO/REDONDO ILLESCAS, *Papeles del Psicólogo*, 2007, p. 165; ANDRÉS PUEYO en DEMETRIO CRESPO (dir.), *Neurociencias y Derecho Penal*, 2013, p. 501.

²³ ANDRÉS PUEYO en DEMETRIO CRESPO (dir.), *Neurociencias y Derecho Penal*, 2013, p. 500.

²⁴ Cfr. las consideraciones y advertencias que efectúan al respecto los propios autores ANDRÉS PUEYO/REDONDO ILLESCAS, *Papeles del Psicólogo*, 2007, pp. 159 a 162.

predicen el comportamiento violento en general (como el PCL-R), o incluso el comportamiento delictivo como tal (el LSI-R, por ejemplo), lo que es aún más amplio.²⁵ Y hay que tener en cuenta que cuando el Código penal español exige realizar un pronóstico de peligrosidad lo define generalmente como pronóstico de comisión de delitos futuros, exige por tanto una predicción más imprecisa que la predicción de la violencia, porque muchos tipos delictivos no la requieren: piénsese no sólo en los delitos de hurto o contra la salud pública, por ejemplo, para los que el Proyecto de reforma presentado al Congreso prevé introducir la libertad vigilada, sino también en otros, como algunos en el ámbito de la protección penal de la libertad sexual, que se caracterizan por no incorporar violencia, ni, a veces, siquiera intimidación como elementos típicos. Luego, en resumen, ni todos los instrumentos modernos de valoración del riesgo están específicamente diseñados para pronosticar un tipo concreto de comportamiento, ni está claro que sea esta predicción más específica la que requiere el legislador. Por supuesto esto no impide reconocer, sin duda alguna, que en los instrumentos estructurados de valoración del riesgo los factores a valorar y la forma en que han de ser tenidos en cuenta están mucho más especificados y objetivizados que en las tradicionales apreciaciones «artesanales» de la peligrosidad, pero lo que creo que no puede negarse es que, en definitiva, también desde este enfoque sigue resultando bastante imprecisa la delimitación de aquello que se predice. En todo caso, sobre este punto abundaré un poco más en el epígrafe dedicado a las limitaciones de los instrumentos de predicción (*infra*, epígrafe 5.), por lo que remito a ese lugar para una consideración más detallada.

Por otro lado, y siguiendo con aquéllos aspectos en los que entiendo que la concepción de la peligrosidad no se diferencia demasiado de la de la valoración del riesgo, creo que las estimaciones de probabilidad que se obtienen con este último método no difieren en último término demasiado de la conclusión (dicotómica) tradicional sobre si la peligrosidad está presente o no. Si bien la probabilidad de reincidencia (o de reincidencia violenta, *cfr. supra*) podrá asumir un valor X entre el 0 y el 100%, pasando por todos los puntos intermedios, lo cierto es que todos estos instrumentos contienen puntos de corte, a partir de los cuales se considera la puntuación obtenida por un sujeto como indicativa de un riesgo de violencia alto o bajo (o, si se quiere, dividido en más niveles: extremo, alto, moderado, bajo o muy bajo, etc.).²⁶ Y pienso que no hay excesiva diferencia entre decir que el sujeto es de riesgo bajo (o bajo/moderado – moderado – moderado/alto – alto – muy alto...) y afirmar que es muy poco peligroso (o moderadamente peligroso, bastante peligroso, muy peligroso, extremadamente peligroso, etc.).²⁷ Con todo, sí es cierto que el hecho de que desde el enfoque de la

²⁵ Cfr. la tabla que con información sobre todos estos instrumentos incluyen ANDRÉS PUEYO/REDONDO ILLESCAS, *Papeles del Psicólogo*, 2007, p. 167.

²⁶ Cfr. sobre la necesidad de establecer estos puntos de corte SCHUMANN, «Prognosen in der strafgerichtlichen Praxis und deren empirische Grundlagen» en FRISCH/VOGT, *Prognoseentscheidungen in der strafrechtlichen Praxis*, 1994, pp. 39 y s., donde tras explicar que existen modelos matemáticos para calcular dónde deben establecerse los puntos de corte de manera que se minimicen las tasas de falsos positivos y de falsos negativos, añade que sin embargo desde el punto de vista jurídico estos puntos de corte a veces se modifican precisamente porque no se quieren asumir las tasas de errores estadísticamente óptimas; sobre esto volveremos en el epígrafe 5. Cfr. Igualmente POLLÄHNE, *Kriminalprognostik. Untersuchungen im Spannungsfeld zwischen Sicherheitsrecht und Rechtssicherheit*, 2011, pp. 238 y s.

²⁷ Señala también que todos los métodos de pronóstico alcanzan el mismo tipo de conclusiones ALBRECHT, «Kriminalprognosen – Entwicklungen und Stand der Forschung» en FREUND et al. (eds.), *Grundlagen und Dogmatik des gesamten Strafrechtssystems. Festschrift für Wolfgang Frisch zum 70. Geburtstag*, 2013, p. 1067. También POLLÄHNE destaca

valoración del riesgo se utilicen instrumentos operacionalizados que arrojan valores numéricos de probabilidades entre el 0 y el 1 (o entre el 0% y el 100%) ayuda a poner de manifiesto la existencia de un número importante de casos donde la estimación del riesgo conduce a un resultado poco claro, cuando los índices de probabilidades están cercanos al 50% (de posibilidades de que la persona cometa un nuevo delito). Seguramente con este enfoque queda más patente que con el de la peligrosidad que hay casos, numerosos además, en los que la forma quizá más exacta de resumir el resultado de la predicción es decir que no podemos predecir si seguramente delinquirá o no.²⁸

Sin embargo, hay un punto en el que considero que sí es importante la matización que introduce el enfoque centrado en la predicción del riesgo, y es cuando subraya que lo que se predice no es una disposición o tendencia inherente al sujeto, sino la probabilidad de que se lleven a cabo acciones (violentas, o criminales), y que éstas resultan de una interacción entre las propiedades del individuo y las circunstancias del entorno en el que se encuentra, por lo que la predicción de la conducta futura debe tomar en consideración también los factores ambientales, y su variabilidad. Que esta visión subjetiva de la peligrosidad es la que está presente – de manera explícita o no – en la doctrina penal (y seguramente en la jurisprudencia)²⁹ creo que se pone de manifiesto por ejemplo en la definición de un autor como ROMEO CASABONA, según la cual «La peligrosidad es [...] una cualidad, una aptitud personal, unas características que definen la personalidad del sujeto».³⁰ Siguiendo a este autor, cuya concepción en la actualidad se sigue defendiendo³¹, en el juicio de peligrosidad hay que diferenciar dos momentos: el diagnóstico de la peligrosidad y la prognosis criminal. El primero consistirá en «la comprobación de la cualidad sintomática de peligroso», mientras que el segundo es el «intento de determinar con el mayor rigor la probabilidad de que el sujeto cometa nuevos delitos en el futuro».³² Vemos aquí cómo en puridad bastaría con el segundo momento para dar una definición de peligrosidad como la que hemos ofrecido al inicio de este epígrafe, y sin embargo se antepone una comprobación anterior sobre la cualidad inherente al individuo, de ser peligroso.

que aunque sólo en los métodos estadísticos puede hablarse de puntos de corte en sentido estricto, el problema normativo de establecer la frontera entre lo aún no (suficientemente) peligroso y lo ya peligroso (al menos de manera relevante) es común a todos los métodos de predicción (*Kriminalprognostik*, 2011, p. 237).

²⁸ Ha destacado especialmente este extremo FRISCH, por ejemplo en «Strafrechtliche Prognoseentscheidungen aus rechtswissenschaftlicher Sicht» en FRISCH/VOGT, *Prognoseentscheidungen in der strafrechtlichen Praxis*, 1994, p. 70.

²⁹ Y, desde luego, históricamente en nuestras leyes de peligrosidad; cfr. al respecto por ej. TERRADILLOS BASOCO, *Peligrosidad social y Estado de Derecho*, 1981, pp. 123 y ss.

³⁰ ROMEO CASABONA, *Peligrosidad y Derecho Penal preventivo*, 1986, pp. 24 y s., donde el autor enseguida pone de manifiesto que lo que interesa no es «la cualidad de la persona en sí misma, sino el riesgo de delito que pone de manifiesto, riesgo que se pretende eliminar» (ob. cit., p. 25), pero ello no obsta a que según este autor el riesgo de comisión de delitos futuros descansa en dicha cualidad personal del individuo, razón por la cual descarta otras concepciones de la peligrosidad más objetivas como la de Petrocelli (*ibidem*).

³¹ Así, expresamente, URRUELA MORA, *Las medidas de seguridad y reinserción social*, 2009, pp. 70 y ss. También GRACIA MARTÍN, «Las medidas de seguridad y reinserción social» en EL MISMO (coord.), *Tratado de las consecuencias jurídicas del delito*, 2006, pp. 457 y s., donde además de adherirse a la concepción de Romeo aclara que «la peligrosidad es una cualidad eminentemente *personal*, y no puede confundirse dicha cualidad con las causas externas que pueden determinar que una persona llegue a ser peligrosa, como por ejemplo determinadas condiciones objetivas que puedan influir en el comportamiento del sujeto» (cursiva en el original).

³² Las citas son de URRUELA MORA, *Las medidas de seguridad y reinserción social*, 2009, pp. 70 y 72; la descripción original de ROMEO CASABONA está en su obra *Peligrosidad y Derecho Penal preventivo*, 1986, pp. 32 a 39.

En la medida en que en todos los métodos de predicción de la peligrosidad los elementos inherentes al sujeto (entendiendo por tales cosas como los antecedentes delictivos, la edad a la que cometió el primer delito, posibles patologías mentales que padezca, rasgos de personalidad, etc.) son fundamentales para efectuar el pronóstico, podría pensarse que carece de trascendencia este matiz en el que aquí estamos insistiendo, de si la peligrosidad se concibe como una cualidad personal o como una probabilidad de ocurrencia de delitos futuros. Sin embargo, creo que hay al menos un punto importante donde este matiz sí puede tener consecuencias que merece la pena destacar, y es el siguiente: si en la valoración de la peligrosidad se pone el acento en las características subjetivas de la personalidad del individuo (estables y relativamente inmutables), no es ilógico pensar que mientras tales características concurren sería correcto considerar al individuo como peligroso, y seguir considerándolo tal aunque dicha peligrosidad no llegara a manifestarse en la comisión de ulteriores delitos (durante un tiempo determinado, o incluso nunca). Por el contrario, si lo decisivo es, según la perspectiva de la valoración del riesgo, la probabilidad de que se produzca la conducta delictiva o violenta, el hecho de que trascurra el tiempo para el que se había hecho la predicción y el delito no se cometa ha de considerarse como evidencia de que el pronóstico fue equivocado.

Este extremo es importante porque está muy relacionado con el significado que se reconozca a un fenómeno muy conocido en la predicción de la delincuencia (y sobre cuya relevancia volveremos en las páginas que siguen), cual es el de los falsos positivos. En toda predicción de comportamientos futuros pueden producirse dos clases de errores al comparar lo predicho con lo efectivamente acaecido. Por un lado están los falsos positivos, que son los sujetos respecto de los cuales se ha predicho la ocurrencia de un fenómeno X (en nuestro caso, que delinquirá en el futuro), y sin embargo dicho evento finalmente no se produce. Los falsos negativos, por el contrario, son aquellos sujetos respecto de los que se predijo que el fenómeno X no ocurriría (en nuestro caso: que no delinquirían), y sin embargo en realidad sí que tiene lugar. Como veremos en las líneas que siguen, la existencia o no de ambas clases de errores en los pronósticos de reincidencia, y la frecuencia con que se produzcan, puede poner seriamente en cuestión la fiabilidad y validez del pronóstico de peligrosidad.

Pues bien, algunos autores han criticado la relevancia que generalmente se reconoce a la existencia de esta clase de errores en los pronósticos de peligrosidad. En concreto, se ha argumentado que el hecho de que durante un periodo determinado de tiempo un sujeto con pronóstico desfavorable de peligrosidad no haya reincidido no significa que sea un falso positivo, porque puede muy bien seguir siendo peligroso, sólo que su peligrosidad no se ha manifestado todavía en un hecho concreto. Se trataría de un caso similar al del conductor que conduce bebido, que sigue siendo peligroso aunque no haya causado hasta ahora ningún accidente.³³ De este modo, el hecho de que tras la realización de una predicción se compruebe que cierta cantidad de personas a las que se había

³³ Así, URBANIOK, «Validität von Risikokalkulationen bei Straftätern – Kritik an einer methodischen Grundannahme und zukünftige Perspektiven-» en EGG (ed.), „Gefährliche Straftäter“. Eine Problemgruppe der Kriminalpolitik?, 2005, pp. 87 y ss., y p. 94.

considerado peligrosas no han llegado a delinquir no tendría por qué poner en cuestión ni la validez del pronóstico realizado, ni la del método empleado para hacerlo.³⁴

Sin embargo, a mi juicio aceptar este argumento comporta negar el propio concepto de falso positivo: en efecto, según este punto de vista todos los pronósticos de peligrosidad tendrían que considerarse siempre acertados, porque sería imposible demostrar lo contrario ya que la ausencia de conducta delictiva no sería una evidencia suficiente. Llevado al extremo, ni siquiera probaría la ausencia de peligrosidad el hecho de que el sujeto llegara a morir tras muchos años sin haber vuelto a cometer ningún delito, pues quién sabe si habría delinquido finalmente de haber vivido unos cuantos años más. Adviértase que esta concepción del falso positivo, que está ligada a un entendimiento de la peligrosidad como cualidad subjetiva³⁵ – que puede manifestarse, o no hacerlo, en la comisión de delitos – convierte en realidad al juicio de peligrosidad en algo prácticamente imposible de ser refutado, y considero que con ello se corre el peligro de abandonar el terreno de lo científico, pues las tesis planteadas de tal modo que no son susceptibles de refutación dejan de pertenecer a la ciencia.³⁶

Frente a ello hay que recordar que lo decisivo para evaluar si un pronóstico ha sido acertado o equivocado es que el evento esperado acontezca o no en un plazo determinado (aquél para el que se hizo la predicción), y el evento esperado en los pronósticos de peligrosidad criminal es la comisión de nuevos delitos. Si estamos hablando de lo que científicamente es posible predecir, todos los métodos de predicción contrastados en la psicología y criminología parten de la base de que: (1) las predicciones siempre se hacen para un periodo de tiempo determinado; no es posible predecir el riesgo de comisión de delitos para un futuro indeterminado, y además la probabilidad de acierto del pronóstico es tanto mayor cuanto más corto sea el periodo respecto del que se formula;³⁷ (2) es respecto del tiempo de validez de la respectiva predicción en relación con el cual hay que comprobar el acierto o el error; si una predicción estimó una alta probabilidad de que el sujeto cometiera delitos sexuales en un periodo de dos años y esos dos años pasan sin reincidir, el sujeto sigue siendo un falso positivo en relación con esa predicción,³⁸ incluso aunque finalmente cometa un delito ocho años después (pues sobre esta eventualidad nada se había dicho en la predicción inicial, limitada a un periodo más corto); y (3) la existencia de falsos positivos y falsos negativos no significa que el método de predicción empleado sea malo, o que se haya utilizado de manera equivocada en el caso concreto; por el contrario, la existencia de falsos positivos y falsos negativos es inherente a todo método de pronóstico de la peligrosidad, que por definición ofrece sólo probabilidades (y no seguridades) de que los delitos se cometan. La existencia de falsos positivos es, desde el punto de

³⁴ Cfr., con más detalles, URBANIOK en EGG (ed.), „Gefährliche Straftäter“. Eine Problemgruppe der Kriminalpolitik?, 2005, p. 93, y pp. 94 y s.

³⁵ Especialmente claro en URBANIOK en EGG (ed.), „Gefährliche Straftäter“. Eine Problemgruppe der Kriminalpolitik?, 2005, pp. 90 y s.

³⁶ VIVES ANTÓN, «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)*, 1974, pp. 396 y s., p. 406.

³⁷ NEDOPIL, «Prognosebegutachtungen bei zeitlich begrenzten Freiheitsstrafen – Eine sinnvolle Lösung für problematische Fragestellungen?», *NSiZ*, 2002, p. 348; SCHUMANN en FRISCH/VOGT (eds.), *Prognoseentscheidungen in der strafrechtlichen Praxis*, 1994, p. 33.

³⁸ Así, acertadamente en mi opinión, KINZIG, *Die Legalbewährung gefährlicher Rückfalltäter. Zugleich ein Beitrag zur Entwicklung des Rechts der Sicherungsverwahrung*, 2ª ed., 2010, pp. 155 y s.

vista estadístico, simplemente una realidad a tener en cuenta. Desde el punto de vista del Derecho penal lo que importa es precisamente tenerla muy presente para, a partir de ella y del conocimiento lo más exacto posible de su dimensión, aplicar las consideraciones normativas que parezcan adecuadas y extraer de todo ello consecuencias para el tratamiento jurídico que deba darse al pronóstico de peligrosidad.

Todo ello no impide reconocer una parte de razón al argumento que estamos discutiendo. Imaginemos que se efectúa una valoración de riesgo para un sujeto concreto y se le asigna, pongamos por caso, un 70% de probabilidades de reincidir. Esto significa que, si evaluáramos a 100 sujetos parecidos al nuestro con ese mismo método, y la evaluación fuera acertada, alrededor de 70 de ellos terminarían reincidiendo y 30 no. Como la probabilidad de reincidencia del 70% afecta por igual a los 100 miembros del grupo, en el momento de emitir el pronóstico es imposible decir si el sujeto al que estamos evaluando – perteneciente a dicho grupo – estará entre los 70 que reincidirán o entre los otros 30 que no lo harán, pero con independencia de que este sujeto delinca finalmente o no, el pronóstico que respecto de él se hizo (70% de probabilidades de reincidir) era en aquel momento igualmente acertado.³⁹ El problema sin embargo es que, siendo todo esto cierto, en Derecho penal la adscripción de un sujeto a un grupo con determinado nivel de riesgo lleva consigo la decisión sobre si adoptar, respecto de ese concreto individuo, consecuencias muy gravosas (por ejemplo la imposición de una medida de seguridad privativa de libertad).⁴⁰ De manera que lo decisivo no es si la predicción fue correcta para el grupo de 100 personas, sino si fue correcta para este individuo en particular. Por eso resulta de importancia decisiva saber con qué frecuencia al imponer la medida estamos limitando la libertad de una persona que no iba a delinquir, y con qué frecuencia podemos estar dejando en libertad a un sujeto que sí va a hacerlo. Y de ahí que el problema de los falsos positivos resulte de ineludible consideración.

En definitiva, y para concluir este apartado, considero que en Derecho penal podemos seguir utilizando el concepto de peligrosidad, ante todo porque de momento es el que el legislador sigue empleando y porque me parece que no impide la comunicación con los especialistas de otras disciplinas como la psicología, psiquiatría o criminología, en las que se va imponiendo la perspectiva de la valoración del riesgo, ya que el parecido entre ambas formulaciones es notable. Ello no obstante, creo que sí es importante tener presentes las críticas que desde estas disciplinas se han formulado al concepto de peligrosidad tal y como ha sido entendido tradicionalmente en la dogmática penal y en el ámbito forense, entre otras cosas porque quizá algunas de ellas puedan aplicarse también a estos nuevos enfoques de la valoración del riesgo, como a continuación trataré de mostrar.

³⁹ URBANIOK en EGG (ed.), „Gefährliche Straftäter“. Eine Problemgruppe der Kriminalpolitik?, 2005, p. 93.

⁴⁰ Tal y como destaca ALBRECHT en «Psychiatrie, Gefährlichkeit und Prognose» en YUNDINA et al. (eds.), *Forensische Psychiatrie als interdisziplinäre Wissenschaft. Festschrift zum Geburtstag von Norbert Nedopil*, 2013, p. 5.

3. Cómo se predice la peligrosidad

A lo largo del tiempo han existido diversas maneras de predecir la peligrosidad. A grandes rasgos pueden clasificarse en tres grandes grupos: por un lado, los métodos actuariales (también denominados estadísticos), cuya utilización ha experimentado un enorme crecimiento desde los años 90 sobre todo en Estados Unidos y en el ámbito anglosajón; en el extremo opuesto se encontraría el método clínico no estructurado, y en cierta forma a mitad de camino podría situarse lo que se conoce como juicio clínico estructurado.⁴¹ El método clínico puro o no estructurado consiste en la predicción de peligrosidad que realiza un experto basándose en su conocimiento y experiencia personales, sin sujetarse a ningún protocolo o regla explícita sino con entera libertad para valorar los factores que estime oportunos y asignarles la relevancia que le parezca adecuada. Ello no significa, sin embargo, que se trate necesariamente de una apreciación arbitraria o un juicio no fundamentado. En realidad dentro del método clínico cabría distinguir dos clases: la primera sería una forma de proceder meramente intuitiva, cuando el juicio se formula sobre la base de la impresión subjetiva que al evaluador le ha causado el sujeto a valorar, y la predicción se basa en su experiencia y conocimiento personales pero sin efectuar un análisis detallado y exhaustivo de factores o elementos concretos. La segunda sería la valoración clínica no estructurada en sentido estricto, o lo que en el ámbito alemán se conoce como método clínico idiográfico, o empírico-individual.⁴² Aquí sí se efectúa un estudio detallado del sujeto, que puede incluir la administración de tests u otros instrumentos objetivos de evaluación, así como un repaso detallado de sus antecedentes vitales y delictivos, un estudio de su personalidad, etc., pero sin seguir un protocolo o regla explícita, y los datos obtenidos son procesados sin atender a ningún estándar objetivo, sino sobre la base de la experiencia clínica del evaluador.

Los métodos actuariales, por el contrario, estructuran todo el proceso de estimación de la peligrosidad.⁴³ Estos instrumentos descansan sobre una serie de premisas: la primera es que la

⁴¹ Cfr. utilizando también estas tres categorías ANDRÉS PUEYO/REDONDO ILLESCAS, *Papeles del Psicólogo*, 2007, pp. 165 y ss., así como SINGH/GRANN /FAZEL, «A comparative study of violence risk assessment tools: a systematic review and metaregression analysis of 68 studies involving 25.980 participants», *Clinical Psychology Review*, (31), 2011, p. 501. Una clasificación algo distinta ofrece EBEC RODRÍGUEZ, que identifica cinco etapas en la evolución de los métodos de pronóstico de la peligrosidad (cfr. «Valoración de la peligrosidad criminal (riesgo-violencia) en psicología forense. Aproximación conceptual e histórica», *Psicopatología Clínica, Legal y Forense*, (3-2), 2003, pp. 52 y ss.) También dividen los métodos de predicción en cinco grupos, aunque atendiendo no a la evolución histórica sino al grado de estructuración de los componentes, SKEEM/MONAHAN, «Current directions in violence risk assessment», *Current Directions in Psychological Science*, (20-1), 2011, p. 39. En el ámbito alemán, y por las razones que a continuación se explican en el texto, suelen clasificarse los métodos de predicción en intuitivo, clínico y actuarial (o estadístico), cfr. por ej. JOST, *Gefährliche Gewalttäter? Grundlagen und Praxis der Kriminalprognose*, 2012, pp. 79 y ss.); ALBRECHT en HEITMEYER/SOEFNER (eds.) *Gewalt. Entwicklungen, Strukturen, Analyseproblemen*, 2004, pp. 477 y ss.

⁴² DAHLE/LEHMANN, «Grundlagen und Methoden der Kriminalprognose» en EGG (ed.), *Psychologisch-psychiatrische Begutachtung in der Strafrecht*, 2012, pp. 157 y ss.; RASCH, «Verhaltenswissenschaftliche Kriminalprognosen» en FRISCH/VOGT, *Prognoseentscheidungen in der strafrechtlichen Praxis*, 1994, p. 24.

⁴³ Cfr. una descripción panorámica de varios de estos instrumentos de predicción, así como también de los métodos de juicio clínico estructurado a los que a continuación me referiré en ARMAZA ARMAZA, *El tratamiento penal del delincuente imputable peligroso*, 2013, pp. 108 a 118. ANDRÉS PUEYO/REDONDO ILLESCAS ofrecen también en su trabajo en *Papeles del Psicólogo*, 2007, p. 167, un listado de protocolos y guías de valoración del riesgo, así como de la

probabilidad de que suceda un evento futuro depende de la combinación de factores que determinaron su aparición en el pasado. Pero no interesa conocer las relaciones de causa-efecto, o cómo un factor incide sobre la ocurrencia o no de un determinado resultado, sino únicamente la probabilidad estadísticamente asociada a que, dada la presencia de un factor o suma de factores, ocurra determinado suceso.⁴⁴ En consecuencia, los instrumentos se construyen seleccionando muestras de sujetos, observando qué circunstancias concurren en ellos y observando su comportamiento. Se reúne información detallada sobre los factores que han estado asociados, en dicho grupo, a la ocurrencia del evento que queremos predecir (en este caso, el delito), y con ello se construyen listas de ítems cuya presencia se considera que está asociada, de manera estadísticamente significativa, a la comisión de delitos. Cuando se quiere valorar la peligrosidad de un individuo concreto hay que repasar la lista cerrada de factores de riesgo, y asignar a cada uno de ellos un valor numérico, en función de que esté presente o no en el individuo examinado. Estas cifras después son combinadas mediante un algoritmo estadístico que da como resultado una determinada puntuación para el sujeto, y esa puntuación es comparada después con la escala de valoración que el propio instrumento de medida contiene. En función de en qué nivel de la escala esté la puntuación del sujeto, se le asigna a éste una determinada probabilidad de reincidencia. Estos métodos dan por supuesta por tanto la validez del razonamiento analógico, en el siguiente sentido: la premisa mayor diría que en la muestra utilizada para construir el test, el 52% de la gente que obtuvo una puntuación entre 5 y 10 delinquirió en el período de seguimiento (por ejemplo, de 3 años); la premisa menor sería que el sujeto S tiene la puntuación 7, y la conclusión sería que el sujeto S tendrá una probabilidad similar al 52% de delinquir en un periodo de tres años.⁴⁵

Como puede observarse, la predicción sobre el riesgo de violencia que ofrecen los métodos o instrumentos actuariales no es en realidad una predicción individual, en el sentido de que esté hecha para el individuo concreto teniendo en cuenta sus circunstancias particulares, sino que consiste en aplicar al individuo que estamos valorando los valores calculados respecto de cohortes de personas que presentan una serie de factores similares a los de este individuo. Como acertadamente señalan ANDRÉS PUEYO/REDONDO ILLESCAS, «el peligro de equivocarse [al aplicar al individuo la regla descubierta para el grupo] está en relación directa con la variabilidad interindividual de los grupos».⁴⁶ El hecho de que las predicciones que ofrecen los métodos actuariales no sean en realidad predicciones sobre el individuo concreto sino estimaciones de riesgo para grupos de personas es una de las críticas más importantes que se han hecho contra estos métodos de estimación de la peligrosidad, especialmente en el ámbito alemán, donde no sólo la doctrina seguramente mayoritaria

disponibilidad - en su caso - de una versión adaptada para su uso en España, aunque sin diferenciar en la enumeración los métodos puramente actuariales de los de juicio clínico estructurado.

⁴⁴ ANDRÉS PUEYO/REDONDO ILLESCAS, *Papeles del Psicólogo*, 2007, p. 168.

⁴⁵ Ejemplo tomado de HART/COOKE, «Another look at the (im-)precision of individual risk estimates made using actuarial risk assessment instruments», *Behavioral Sciences and the Law*, (31), 2013, p. 82. Sobre los problemas que plantea no sólo esta premisa, sino también otras asunciones de las que parten los métodos estadísticos de predicción de la peligrosidad, cfr. las interesantes consideraciones críticas de ALBRECHT en HEITMEYER/SOEFNER (eds.), *Gewalt. Entwicklungen, Strukturen, Analyseproblemen*, 2004, pp. 480 y ss., críticas que por otra parte no son en absoluto nuevas, sino que muchas de ellas hace décadas que se vienen formulando; cfr. VIVES ANTÓN en VVAA, *Peligrosidad social y medidas de seguridad*, 1974, pp. 410 y ss.

⁴⁶ ANDRÉS PUEYO/REDONDO ILLESCAS, *Papeles del Psicólogo*, 2007, p. 168.

sino también la jurisprudencia del BGH consideran que una predicción de este tipo no puede ser nunca un dato suficiente para decidir sobre la peligrosidad del sujeto, sino que tendrá que ser complementada por un estudio completo e individualizado, en el que se tengan en cuenta todas las características relevantes para la valoración de la personalidad del autor concreto y de sus hechos.⁴⁷

Frente a los métodos actuariales, los métodos que se conocen como de juicio clínico estructurado (también conocidos en ocasiones como instrumentos de la tercera generación)⁴⁸ dejan más libertad al evaluador para estimar el concreto grado de peligrosidad del individuo, pero tratan de mejorar la fiabilidad y la transparencia del método clínico puro. Contienen, como los instrumentos actuariales, listas de ítems (factores de riesgo) que son los que se han de valorar, y también protocolizan la forma de recoger la información (por ejemplo asignar «no» si el ítem no está presente, «?» si posiblemente lo está, y «sí» si lo está sin duda alguna). Pero frente a los instrumentos actuariales clásicos, los de juicio clínico estructurado suelen recoger información no sólo sobre datos estáticos e inmodificables (número de delitos previamente cometidos, edad de comisión del primer delito, sexo, edad actual, etc.) sino también sobre ítems más complejos y además dinámicos (por ej., competencias sociales, actitudes antisociales, si tiene empleo o no, etc.), y, sobre todo, dejan en manos del experto la decisión final sobre la valoración del riesgo, en el sentido de que por ejemplo, cabrá concluir que a pesar de que en el sujeto no concurren muchos factores de riesgo, el peligro de violencia futura es no obstante alto porque uno de los factores presentes tiene un significado extraordinariamente importante.⁴⁹

De todos modos la clasificación de cada uno de los instrumentos de estimación del riesgo diseñados en los últimos años en cada una de estas tres categorías no siempre es fácil, pues por ejemplo métodos que en principio son para realizar juicios clínicos estructurados pueden ser utilizados de forma actuarial, asignando a los ítems puntuaciones numéricas (0/1/2) y usando la puntuación global obtenida por el sujeto para estimar directamente su peligrosidad,⁵⁰ o bien la probabilidad numérica de riesgo de violencia obtenida tras la aplicación de un método actuarial puede complementarse tomando en consideración factores adicionales a los enumerados en el test antes de tomar la decisión definitiva sobre el grado de peligrosidad del sujeto.⁵¹

⁴⁷ EGG por ejemplo afirma que las exigencias de justicia individualizada que derivan de la idea de Estado de Derecho impiden una aplicación meramente esquemática de catálogos de criterios, y exigen por el contrario una valoración individual diferenciada, que tenga en cuenta las específicas particularidades del caso concreto (EGG, «Forensisch-psychologische Begutachtung in der Strafjustiz – Entwicklung und Perspektiven in Deutschland» en EL MISMO (ed.), *Psychologisch-psychiatrische Begutachtung in der Strafjustiz*, 2012, p. 19. En sentido similar DAHLE/LEHMANN en EGG (ed.), *Psychologisch-psychiatrische Begutachtung*, 2012, p. 166, y, en relación con la jurisprudencia reciente del BGH, ALBRECHT en YUNDINA et al. (eds.), *FS-Nedopil*, 2013, pp. 10 y s.

⁴⁸ ANDREWS/BONTA, *The Psychology of criminal behavior*, 4ª ed., 2006, pp. 288 y ss.; DAHLE/LEHMANN en EGG (ed.), *Psychologisch-psychiatrische Begutachtung*, 2012, pp. 162 y ss.

⁴⁹ SKEEM/MONAHAN, *Current Directions in Psychological Science*, 2011, p. 39.

⁵⁰ Puede verse un ejemplo de esto último, en relación con el SVR-20, en PÉREZ RAMÍREZ et al., «Predicción de riesgo de reincidencia en agresores sexuales», *Psicothema*, (20-2), 2008, pp. 207 y ss.

⁵¹ Cfr. las consideraciones de HART y COOKE sobre los enfoques que denominan «ortodoxo» y «latitudinario» en relación con el uso de los instrumentos actuariales (HART/COOKE, *Behav. Sci. Law*, 2013, pp. 82 y s).

Durante años se ha discutido si alguno de estos métodos ofrece mejores predicciones que otros, en el sentido de más fiables y con menores tasas de errores, y también si dentro de los instrumentos actuariales algunos son mejores que otros, pero no parece que de momento se haya alcanzado un resultado concluyente.⁵² Si bien hay bastante acuerdo en considerar que el método intuitivo (si es que a una forma intuitiva de funcionar se le puede llamar método) está en la actualidad totalmente desacreditado, la cuestión es más problemática en relación con los demás sistemas. Al juicio clínico no estructurado, que seguramente es el más extendido en estos momentos en España⁵³ y también tiene una gran tradición en Alemania, se le ha criticado que su problema principal es la falta de fiabilidad y transparencia, ya que en la medida en que la decisión final sobre la probabilidad de comisión de delitos no es fruto del seguimiento de reglas estandarizadas, un segundo experto podría llegar a conclusiones distintas sobre la peligrosidad de un sujeto concreto sin que hubiera criterios objetivos para decidir cuál de los dos tiene más razón.⁵⁴ Por otro lado, los métodos actuariales (y hasta cierto punto los de juicio clínico estructurado) adolecen del problema de la falta de individualización al que ya hemos aludido, así como de otros inconvenientes a los que más abajo haremos referencia. Pero de momento, y antes de profundizar en estas cuestiones, interesa exponer algunos datos empíricos sobre el acierto de las predicciones de peligrosidad.

4. Algunos datos empíricos

De las dos clases de errores que pueden derivar de una predicción de peligrosidad, los falsos negativos son los más fáciles de detectar: cuando el sujeto al que se consideró no peligroso, y respecto del cual por tanto no se adoptó ninguna medida, vuelve a delinquir, no sólo da lugar a un delito del que nuevamente se ocupará la Administración de Justicia, sino que con frecuencia es objeto de una atención desmesurada por parte de los medios de comunicación, con consecuencias

⁵² Cfr. SKEEM/MONAHAN, *Current directions in Psychological Science*, 2011, pp. 39 y s. Hans-Jörg ALBRECHT señala que sí se ha demostrado empíricamente la superioridad de los métodos actuariales sobre las valoraciones clínicas, y que aquello sobre lo que no existe evidencia es sobre si algunos de los instrumentos actuariales predicen el comportamiento violento mejor que otros (*FS-Nedopil*, 2013, pp. 7 y 8). FAZEL/SINGH/DOLL/GRANN, tras un meta-análisis sobre 73 muestras de sujetos, concluyen que los instrumentos actuariales no predicen mejor que los métodos de juicio clínico estructurado, en contra de lo que habían afirmado algunos estudios anteriores («Use of risk assessment instruments to predict violence and antisocial behaviour in 73 samples involving 24 827 people: systematic review and meta-analysis», *British Medical Journal*, (345), 2012, p. 5).

⁵³ PUJOL ROBINAT/PUIG BAUSIL, «Concepto de peligrosidad criminal. Evolución histórica del concepto», *Cuadernos de Política Criminal*, (94), 2008, p. 279, quienes afirman también que en los últimos años ha aumentado el uso del juicio clínico estructurado.

⁵⁴ ANDRÉS PUEYO/REDONDO ILLESCAS, *Papeles del Psicólogo*, 2007, p. 167: «presenta una notable dificultad para encontrar justificaciones empíricas y sistemáticas ya que existen bajos niveles de acuerdo inter-jueces, poca precisión y una débil justificación teórica». Cfr. asimismo SINGH /GRANN/FAZEL, quienes afirman que estudios recientes sugieren que esta forma de predecir la peligrosidad tiene poca validez predictiva (*Clinical Psychology Review*, 2011, p. 501), y también SKEEM/MONAHAN, *Current directions in Psychological Science*, 2011, p. 39, afirmando que es el método que menor soporte empírico tiene. Otros autores, sin embargo, aunque reconocen que «debido a la complejidad del enfoque idiográfico resulta prácticamente imposible una estandarización, lo que en último término limita la deseable transparencia de la elaboración del juicio», afirman que ello no obstante se ha demostrado que en determinadas circunstancias el método clínico ofrece estimaciones más fiables del riesgo de reincidencia que los métodos estadísticos (DAHLE/LEHMANN en EGG (ed.), *Psychologisch-Psychiatrische Begutachtung*, 2012, p. 172).

desastrosas en cuanto a las condiciones de posibilidad de llevar a cabo una política criminal racional.⁵⁵

Pero tratar de comprobar empíricamente cuántas de las personas a las que sí se calificó como peligrosas en realidad no lo eran, presenta la dificultad de que si la consecuencia de dicha calificación es la imposición o el mantenimiento de una privación de libertad (ya sea en forma de medida de seguridad, ya en forma de no concesión de libertad condicional o permisos, o bien no revisión de la cadena perpetua, etc.), la propia situación de privación de libertad impide al sujeto demostrar que si hubiera estado en libertad no habría cometido delitos.⁵⁶ Por ello son relativamente escasas las ocasiones en que ha podido hacerse un seguimiento riguroso de personas diagnosticadas como peligrosas en libertad, aunque cuando se han dado las condiciones para un «experimento natural» de estas características los resultados han sido bastante reveladores.

4.1 En otros países

En EEUU fue el caso Baxstrom el que desencadenó la preocupación por la posible sobrevaloración de la peligrosidad criminal. En el año 1966 una sentencia del Tribunal Supremo Federal de los EEUU (Baxstrom vs Herold, 383 U.S. 107) ordenó la liberación, o el traslado a hospitales psiquiátricos civiles, de 967 personas que permanecían recluidas en establecimientos para enfermos mentales criminales en el Estado de Nueva York. Se trataba de criminales convictos, a los que no se había dejado en libertad al término de su condena porque tras un examen psiquiátrico se los había considerado, además de mentalmente enfermos, peligrosos para la comunidad. La sentencia entendió que no se habían respetado las garantías del debido proceso y como consecuencia de ello los pacientes fueron o bien directamente liberados o bien trasladados a hospitales psiquiátricos ordinarios. Tras un seguimiento hecho a estos pacientes durante cuatro años se puso de manifiesto que sólo una minoría había evidenciado comportamiento violento o agresivo: sólo 24 personas tuvieron que volver a ingresar en centros de alta seguridad para criminales mentalmente enfermos, lo que da como resultado una tasa de falsos positivos de más del 97%. Por otro lado, a 199 pacientes se les hizo un seguimiento más detallado para comprobar si, con independencia del reingreso o no

⁵⁵ Sobre las interacciones entre medios de comunicación y agentes políticos en la evolución de la política criminal cfr. VARONA GÓMEZ, «Medios de comunicación y punitivismo», *InDret*, (1), 2011. Como afirma POLLÄHNE, especialmente cuando los delitos que se cometen son graves atentados con violencia o contra la libertad sexual, el que se subraye que las tasas reales de reincidencia son muy bajas aparece ante los ojos de las víctimas individuales (y reales) de estos delitos como un puro cinismo, y cada caso es para la opinión pública un caso de más cuya no evitación constituye un fracaso (*Kriminalprognostik*, 2011, pp. 234 y s). Por supuesto, tampoco podemos olvidar que la cifra negra de la criminalidad hace que los falsos negativos que llegamos a conocer sean menores que los falsos negativos reales, pero en la medida en que el mismo fenómeno se produce respecto de la delincuencia primaria, no me parece un argumento decisivo en relación con el significado de la cifra negra para la evaluación de la calidad de los pronósticos de reincidencia tras la comisión de un primer delito, aparte de que la propia predicción también genera efectos de signo contrario, cuando el pronóstico de peligrosidad contribuye a que llegue a convertirse en reincidente el que de otro modo quizá no lo habría llegado a ser - *self-fulfilling-prophecy* - (cfr. POLLÄHNE, *Kriminalprognostik*, 2011, p. 234).

⁵⁶ Es más: como señala Günther ALBRECHT, si tras décadas de prisión o de internamiento en un hospital psiquiátrico el falso positivo es finalmente liberado y no comete más delitos en su vejez, esto se interpretará (ironías de la vida) incluso como un éxito del internamiento o del tratamiento; ALBRECHT en HEITMEYER/SOEFFNER (eds.) *Gewalt. Entwicklungen, Strukturen, Analyseproblemen*, 2004, p. 487.

en hospitales de alta seguridad, manifestaron o no comportamientos anormales. Sólo 30 los evidenciaron en un periodo de seguimiento de cuatro años mientras estuvieron ingresados en hospitales civiles, y, de 98 pacientes (de estos 199) que fueron puestos en libertad, 20 volvieron a ser detenidos, pero sólo dos lo fueron por la comisión de actos violentos (2%).⁵⁷ Importa resaltar que estas bajísimas tasas de reincidencia violenta en el caso Baxstrom se produjeron en un grupo de casi mil delincuentes convictos, a los que las autoridades habían considerado previamente tan peligrosos como para mantenerlos indefinidamente recluidos tras el cumplimiento de sus condenas.

Pocos años después otra sentencia judicial parecida volvió a crear las condiciones para un «experimento natural» similar al que supuso Baxstrom: en el «caso Dixon» 586 convictos considerados enfermos mentales y peligrosos fueron trasladados a hospitales civiles, y el seguimiento de estos sujetos evidenció que sólo el 14,5% volvió a cometer delitos violentos o actos violentos que comportaran el reingreso.⁵⁸ Como conclusión del estudio de seguimiento realizado sobre estos pacientes, los autores declararon que «si quisiéramos utilizar esta información para predecir estadísticamente la peligrosidad nuestra mejor estrategia sería asumir que ninguno de los pacientes era peligroso. Con ello, nos equivocaríamos en 14 de los 98 casos. Cualquier otro método incrementaría nuestro error».⁵⁹

Como consecuencia de los estudios realizados sobre los casos Baxstrom y Dixon, en los Estados Unidos se extendió a principio de los años setenta en el ámbito psicológico y psiquiátrico un notable escepticismo sobre la viabilidad de predecir con un mínimo grado de certeza la futura conducta criminal.⁶⁰ Sin embargo, la administración de justicia seguía reclamando la realización de pronósticos de peligrosidad y no parecía haber otros grupos de profesionales que estuvieran en mejor situación para emitir tales pronósticos, de modo que sobre todo a partir de la década de los noventa del siglo pasado la investigación cobró un nuevo impulso, tratando de diseñar instrumentos que corrigieran los fallos de las predicciones anteriores y consiguieran mayores tasas de acierto, lo que condujo al

⁵⁷ Todos los datos están extraídos de KINZIG, *Die Sicherungsverwahrung auf dem Prüfstand. Ergebnisse einer theoretischen und empirischen Bestandsaufnahme des Zustandes einer Maßregel*, 1996, p. 87.

⁵⁸ Datos extraídos de AUERHAHN, *Selective incapacitation and public policy, Evaluating California's imprisonment crisis*, 2003, p. 82, quien a su vez cita el trabajo original de THORNBERRY/JACOBY sobre el caso Dixon (*The Criminally Insane: A Community Follow-up of Mentally Ill Offenders*, 1979), trabajo este último que no he podido consultar directamente. KINZIG, en su obra citada en la nota anterior, afirma que fueron 438 los pacientes liberados en el caso Dixon, pero coincide con AUERHAHN en señalar que el porcentaje de pacientes que volvieron a mostrar un comportamiento violento fue del 14% (cfr. KINZIG, *Die Sicherungsverwahrung auf dem Prüfstand*, 1996, pp. 87 y s.)

⁵⁹ Cita extraída de AUERHAHN, *Selective incapacitation and public policy*, 2003, pp. 82 y 83; la cita es de THORNBERRY/JACOBY, *The criminally insane: a community follow-up of mentally ill offenders*, 1979.

⁶⁰ Véase la afirmación que hacía STEADMAN en 1974 tras finalizar el estudio sobre el caso Baxstrom: «Sencillamente no existe ninguna evidencia empírica de que un psiquiatra o cualquier otro profesional médico esté en mejor situación, sobre la base de un juicio clínico, de identificar a aquellas personas que llevarán a cabo un comportamiento criminal, de lo que lo está cualquier otra persona que se base directamente en el azar.» (STEADMAN/COCOZZA, *Carreers of the criminally insane*, 1974, cita reproducida en RASCH en FRISCH/VOGT (eds.) *Prognoseentscheidungen in der strafrechtlichen Praxis*, 1994, p. 17, por donde se cita). En 1981 MONAHAN concluyó que las predicciones de peligrosidad hechas por psiquiatras y psicólogos sobre delincuentes con trastornos mentales sólo eran acertadas en un caso de cada tres (STEADMAN, *J Am Acad Psychiatry Law*, 2000 p. 268).

boom de los métodos actuariales.⁶¹ Pero antes de analizar los pros y los contras de estos sistemas de predicción es necesario hacer referencia a otros estudios empíricos, más recientes, sobre la criminalidad realmente evidenciada en libertad por grupos de sujetos considerados peligrosos.

Recientemente se han publicado en Alemania algunos trabajos que arrojan también resultados preocupantes sobre la cifra de falsos positivos a los que se puede haber estado privando de libertad en los últimos años. En uno de ellos se hizo un seguimiento de 33 pacientes, anteriormente ingresados según la regulación vigente en la RDA, a los que el Tribunal Constitucional alemán ordenó dejar en libertad (o, en una pequeña proporción, mantener internados pero en dependencias psiquiátricas ordinarias) por estimar que era inconstitucional prolongar sus internamientos tras la reunificación alemana sobre la base de la regulación de las medidas de seguridad contenida en el StGB. El seguimiento realizado a estos sujetos reveló que en un margen de tiempo de 5,72 años sólo 8 de 32 personas liberadas (el 25%) reincidieron, y de estos 8 sólo 5 cometieron delitos violentos que condujeron de nuevo al encarcelamiento o internamiento. De los 15 sujetos respecto de los cuales existían pronósticos negativos en el momento de la puesta en libertad sólo reincidieron 4 (el 27%). En este caso, si se hubiera mantenido el internamiento sobre la base del pronóstico de peligrosidad, por cada sujeto verdaderamente peligroso que hubiera permanecido internado se habría encerrado además a otros tres que en realidad no eran peligrosos.⁶² O, dicho de otra manera, más del 70% de los informes que afirmaban la presencia de peligrosidad eran falsos positivos.

Otra investigación reciente en Alemania ha evidenciado datos que parecen aún más preocupantes, esta vez en relación con la custodia de seguridad impuesta a posteriori.⁶³ En dos estudios consecutivos, ALEX ha hecho un seguimiento de delincuentes para los que, tras el cumplimiento de la condena, la Fiscalía había solicitado la imposición de la custodia de seguridad a posteriori, existiendo además en la mayor parte de los casos un informe pericial que apoyaba la solicitud tras haber constatado una elevada peligrosidad, pero respecto de los cuales los tribunales por diferentes razones no habían acordado la medida. Pues bien, en un primer grupo de 77 sujetos puestos en libertad hasta el 31.12.2006 y tras consultar hasta enero de 2009 el registro federal de penados (*Bundeszentralregister*), sólo 32 habían sido condenados de nuevo, mientras que de más de la mitad

⁶¹ STEADMAN, *J Am Acad Psychiatry Law*, 2000, p. 268; NEDOPIL, *Prognosen in der forensischen Psychiatrie – Ein Handbuch für die Praxis*, 3ª ed., 2006, pp. 43 y s.

⁶² El estudio es el de RUSCHE, *In Freiheit gefährlich? Eine Untersuchung zu Häufigkeit und Gründen falscher Kriminalprognosen bei psychisch kranken Gewaltverbrechern*, 2004, pp. 123 y s, de donde se han extraído todos los datos citados en el texto. El propio autor apunta que los resultados no son mejores que los que arrojaron los casos Baxstrom y Dixon (ob. cit., p. 125).

⁶³ ALEX, *Nachträgliche Sicherungsverwahrung – ein rechtsstaatliches und kriminalpolitisches Debakel*, 2010. Aquí citamos el trabajo por la segunda edición, aparecida en 2013. Denominamos custodia de seguridad «posterior» o «a posteriori» a lo que en alemán se conoce como *nachträgliche Sicherungsverwahrung*. Esta modalidad de la custodia de seguridad, introducida en Alemania en 2004, consiste en que la medida no se impone en la sentencia condenatoria, sino que el tribunal puede acordarla después, siempre que – además de cumplirse una serie de requisitos legales en cuanto a gravedad del delito cometido etc. – después de la condena se hayan puesto de manifiesto nuevas circunstancias de las que quepa deducir que el sujeto constituye un gran peligro para la sociedad.

(n=42) no se tenía noticia de que hubieran cometido ningún nuevo delito.⁶⁴ Pero aparte de las cifras globales sobre reincidentes y no reincidentes resulta muy interesante atender a la clase de delito que se volvió a cometer: de los 32 reincidentes sólo 14 (el 18,18% del total de 77) habían sido condenados a una pena privativa de libertad sin suspensión condicional, mientras que el resto había cometido delitos mucho más leves castigados únicamente con multa o con penas privativas de libertad cuya ejecución había quedado en suspenso. Y a su vez, de estas 14 condenas sólo 5 se debían a la comisión de un delito de los que por su gravedad justifican la imposición de la custodia de seguridad.⁶⁵

Para la segunda edición del trabajo que estamos comentando el autor intentó obtener los datos oficiales de reincidencia de los mismos sujetos hasta 2011, para comprobar si el transcurso de un lapso de tiempo más amplio había hecho aumentar el número de delitos cometidos por la cohorte de sujetos estudiados hasta 2009. Sin embargo, esta vez las autoridades alegaron diversos problemas supuestamente relacionados con la protección de datos de carácter personal, y no fue posible obtener dicha información.⁶⁶ Sí se logró efectuar un nuevo estudio, sobre 52 sujetos excarcelados entre los años 2007 y 2009, respecto de los cuales los resultados fueron los siguientes: del 59% (31) no constaban datos de nuevas condenas hasta agosto de 2012. De los 21 que sí reincidieron, sólo trece (25%) cometieron un delito condenado con pena privativa de libertad sin suspensión condicional (el resto de delitos, por tanto, fueron más leves y resultaron castigados sólo con pena de multa o con privación de libertad que fue dejada en suspenso), y a su vez de entre estos últimos sólo 8 son delitos que alcanzan la gravedad suficiente como para permitir la imposición de una custodia de seguridad según la (entonces vigente) regulación del StGB.⁶⁷

Si se pone en relación la reincidencia realmente ocurrida con la peligrosidad predicha, sumando los resultados de los dos estudios llevados a cabo por ALEX los números totales de falsos y verdaderos positivos y falsos y verdaderos negativos serían los siguientes:⁶⁸

⁶⁴ ALEX, *Nachträgliche Sicherungsverwahrung*, 2ª ed., 2013, p. 117. Los otros tres sujetos que faltarían para completar la suma de 77 habían fallecido entre su puesta en libertad y el momento de comprobación de los datos en el *Bundeszentralregister*. El periodo medio transcurrido entre excarcelación y comprobación de datos sobre nuevas condenas fue de 33 meses (ibídem, p. 111)

⁶⁵ Todos los datos extraídos de las pp. 117 y 118 de la ob. cit. en la nota anterior.

⁶⁶ En opinión del autor (*Nachträgliche Sicherungsverwahrung*, 2ª ed., 2013, pp. 110 a 112), las dificultades alegadas eran meras excusas que no se correspondían con verdaderos problemas legales de protección de la intimidad, y lo que revelan es una reacción a los resultados publicados en la primera edición del trabajo en 2010. ALEX afirma que dar a conocer datos reales sobre tasas de reincidencia bajas en un grupo de sujetos considerado altamente peligroso simplemente es un mensaje que no encaja bien en el panorama político actual, y que constituye una razón suficiente para que las autoridades intenten escamotear datos de este tipo en la medida de lo posible (ob. cit., p. 111). No obstante, el autor pudo reunir de todos modos algunos datos sobre nuevos delitos de aquella primera cohorte de sujetos a partir de otras fuentes (jurisprudencia, medios de comunicación), y estos datos fueron incluidos en la segunda edición del trabajo, la que estoy citando en este artículo, de manera que los datos que he reproducido en el texto incluyen, cuando pudo disponerse de ella, esa nueva información (en concreto, cinco nuevos delitos respecto de los datos ofrecidos en la edición de 2010), cfr. *ibídem*.

⁶⁷ ALEX, *Nachträgliche Sicherungsverwahrung*, 2ª ed., 2013, pp. 144 y s.

⁶⁸ Fuente: ALEX, *Nachträgliche Sicherungsverwahrung*, 2ª ed., 2013, p. 148. En esta tabla se recogen sólo los datos de los 73 casos (pertenecientes tanto al primer como al segundo estudio) en los que los dictámenes de los peritos fueron coincidentes al valorar la peligrosidad.

		Reincidencia observada		
		grave	inexistente o leve	suma
Peligrosidad pronosticada	alta	17 verdaderos positivos	39 falsos positivos	56
	baja	1 falsos negativos	16 verdaderos negativos	17
Suma		18	55	

Lo que se observa en esta tabla es que los pronósticos de peligrosidad tienen un grado elevado de acierto cuando concluyen que el sujeto es poco peligroso, mientras que yerran mucho más cuando afirman la elevada peligrosidad del sujeto; en concreto, se evidencia una sobreestimación notable de la peligrosidad, ya que de 56 sujetos considerados por los peritos como muy peligrosos⁶⁹ sólo lo fueron en realidad 17. El número de falsos positivos es dos veces mayor que el de verdaderos positivos, o, dicho de otra forma, de los 56 sujetos respecto de los que se predijo una alta peligrosidad, el 70% eran falsos positivos.

También KINZIG ha ofrecido datos sobre el seguimiento de personas puestas en libertad en Alemania tras haber cumplido custodia de seguridad a pesar de que existían respecto de ellos pronósticos de peligrosidad negativos. En un trabajo de 2010 este autor aporta datos sobre 22 personas a las que se puso en libertad, ya fuera porque hasta 1998 la custodia de seguridad impuesta por primera vez tenía un límite máximo de 10 años de duración (fuera cual fuera el pronóstico de peligrosidad), o bien (para los puestos en libertad a partir de 1998), porque a pesar de no existir para estos sujetos un pronóstico favorable no se cumplían otras exigencias legales del parágrafo 67d apartado 3 primer inciso⁷⁰ del Código penal alemán. Pues bien, de estas 22 personas puestas en libertad reincidieron en total sólo 8, y a su vez de estos 8 sólo 2 cometieron delitos graves con violencia o peligro para las personas (robo violento e incendio), es decir, el tipo de delincuencia para cuya evitación está prevista la custodia de seguridad.⁷¹ Dicho de otra manera, para impedir la comisión de un delito grave con violencia o peligro para las personas, por cada sujeto realmente peligroso se habría mantenido encerrados a otros 10 que no lo eran.

⁶⁹ Y eso que como «reincidencia grave» se han considerado en esta tabla todos los casos en que como consecuencia del nuevo delito (o delitos) cometidos se impuso una pena privativa de libertad sin suspensión condicional, y no únicamente los casos de delitos graves contra la vida, integridad física o libertad sexual de la víctima.

⁷⁰ Según el cual, tras la reforma introducida por la Ley para la Lucha contra la delincuencia sexual y otros delitos peligrosos de 1998 (versión ya derogada en la actualidad), la custodia de seguridad debía declararse extinguida tras diez años de cumplimiento cuando no existiera el peligro de que el interno, debido a su tendencia, vaya a cometer delitos graves que produzcan en las víctimas graves daños corporales o anímicos («wenn nicht die Gefahr besteht, dass der Untergebrachte infolge seines Hanges erhebliche Straftaten begehen wird, durch welche die Opfer seelisch oder körperlich schwer geschädigt werden»).

⁷¹ KINZIG, *Die Legalbewährung gefährlicher Rückfalltäter*, 2010, pp. 196 y ss., y p. 306.

4.2 En España

Habida cuenta de que no existen en España medidas de seguridad privativas de libertad para imputables peligrosos no tiene sentido buscar datos similares a los que hemos expuesto para Alemania, pero sí existe algún estudio en el que se ha contrastado la reincidencia real de sujetos excarcelados (ya sea por quedar en libertad tras extinguir su condena, o bien por otros motivos), respecto de los cuales se había efectuado un pronóstico previo de reincidencia.

Hace unos años se aplicó en Cataluña un instrumento de evaluación del riesgo de violencia de jico clínico estructurado (el *Sexual Violence Risk-20*, conocido como SVR-20) a un grupo de 163 agresores sexuales que habían quedado en libertad tras cumplir condena.⁷² Cada uno de ellos fue evaluado con este instrumento y se le asignó una probabilidad de reincidencia, y después se comprobó cuántos de ellos habían reincidido en un periodo de seguimiento promedio de cuatro años. Los autores del estudio pretendían con ello ir explorando las posibilidades de aplicar estos instrumentos, tan consolidados ya en el mundo anglosajón, a población española, y contrastar si es posible obtener con su uso predicciones correctas de peligrosidad.⁷³ Los resultados del estudio mostraron que de 24 sujetos que realmente reincidieron el test había identificado como futuros reincidentes a 17 (70,8% de clasificaciones correctas de sujetos reincidentes), y que de 139 sujetos que no reincidieron el test había identificado correctamente a 111 (79,9% de clasificaciones correctas de sujetos no-reincidentes). Los autores afirmaron que el total de clasificaciones correctas fue de 78,5% y concluyeron que «el pronóstico de la violencia sexual consigue valores de aciertos destacables», y que el SVR-20 «puede constituir una buena ayuda técnica para predecir el riesgo de reincidencia sexual».⁷⁴ Sin embargo, estos datos resultan matizados si en lugar de medir cuántos sujetos de los que verdaderamente reincidieron había identificado correctamente el instrumento, atendemos a cuántos sujetos de los que el instrumento había predicho como futuros reincidentes, después lo fueron realmente. Se trata de dos cosas muy diferentes: lo primero se conoce como sensibilidad y especificidad del instrumento de predicción, y lo segundo como valor predictivo. La diferencia se hace evidente en la siguiente tabla:⁷⁵

⁷² Cfr. REDONDO ILLESCAS/PÉREZ/MARTÍNEZ, «El riesgo de reincidencia en agresores sexuales: investigación básica y valoración mediante el SVR-20», *Papeles del Psicólogo*, (28-3), 2007, pp. 187 a 195, donde se ofrece una visión general de los resultados, y especialmente PÉREZ RAMÍREZ et al., *Psicothema*, 2008, pp. 205 a 210, donde se ofrecen todos los datos detallados del estudio.

⁷³ Los autores afirman que se trata del «primer estudio empírico en España que investiga la eficacia predictiva de un instrumento de estimación de riesgo de violencia sexual, en concreto del *Sexual Violence Risk-20*» (PÉREZ RAMÍREZ et al., *Psicothema*, 2008, p. 210).

⁷⁴ REDONDO ILLESCAS/PÉREZ/MARTÍNEZ, *Papeles del Psicólogo*, 2007, p. 193, y PÉREZ RAMÍREZ et al., *Psicothema*, 2008, p. 210.

⁷⁵ Reproduzco a continuación la tabla de resultados que ofrecen PÉREZ RAMÍREZ et al., *Psicothema*, 2008, en la p. 209, a la que he añadido la última columna, pues el cálculo del valor predictivo no estaba incluido en la tabla original.

		Reincidencia sexual observada		Porcentaje de predicciones correctas (valor predictivo)
		SÍ	NO	
Reincidencia sexual pronosticada	SÍ	17 Verdaderos positivos	28 Falsos positivos	37,7% (valor predictivo positivo)
	NO	7 Falsos negativos	111 Verdaderos negativos	94,07% (valor predictivo negativo)
Porcentaje de detecciones correctas		70,8% (sensibilidad)	79,9% (especificidad)	

Como se advierte fácilmente, si a lo que se atiende es al valor predictivo del test, existe un número de falsos positivos que casi duplica al de los verdaderos positivos, pues según los datos aportados por los autores, por cada sujeto verdaderamente peligroso calificado como tal (n=17) fueron también considerados peligrosos, falsamente, casi otros dos (n=28). Es decir, de los 45 sujetos predichos como peligrosos, en el 62% de los casos dicha predicción no fue corroborada por el comportamiento real posterior de los sujetos, fueron falsos positivos.

El 20 de septiembre de 2012 se presentó en el Colegio Oficial de Psicólogos de Madrid un «Informe sobre la eficacia de las predicciones de peligrosidad de los informes psicológicos forenses emitidos en los Juzgados de Vigilancia Penitenciaria de Madrid»⁷⁶. En dicho informe se estudian 150 expedientes en los que el Psicólogo forense de los Juzgados de Vigilancia Penitenciaria de Madrid había emitido informe entre los años 2003 y 2006 (que suponen la totalidad de los informes realizados sobre delitos graves en dichos años), y se compara la predicción de peligrosidad efectuada con la reincidencia real evidenciada después por dichos sujetos. Para obtener estos últimos datos se consultó el Sistema de Información Penitenciaria, hasta 2012. El periodo de seguimiento fue por tanto de entre cinco y seis años. Del trabajo no queda del todo claro con ocasión de qué decisión del JVP se solicitaba el informe: si era para permisos de salida, libertad condicional o tercer grado, pero en todo caso los datos que se ofrecen sobre los aciertos y errores en las predicciones efectuadas son los siguientes⁷⁷:

⁷⁶ El informe, elaborado por la Psicóloga Forense María del Rocío GÓMEZ HERMOSO, puede consultarse como documento pdf en la página web del Colegio Oficial de Psicólogos de las Islas Baleares, <http://www.copib.es/pdf/Vocalies/Juridica/2012-11-18%20predicciones.pdf> (consultada el día 21 de octubre de 2013)

⁷⁷ Datos extraídos del informe citado en la nota anterior, p. 7.

	Reincide	No reincide	Total
Peligrosidad sí (informe desfavorable)	15 Verdaderos positivos	77 Falsos positivos	92
Peligrosidad no (informe favorable)	3 Falsos negativos	55 Verdaderos negativos	58
Total	18	132	150

El número de falsos positivos es cinco veces superior al de verdaderos positivos, o, dicho de otra manera, en el 80% de los informes en los que se consideró que el sujeto era peligroso la realidad desmintió la predicción.

En definitiva, creo que a partir de los datos aportados hasta ahora puede corroborarse la sospecha, tradicional en la doctrina penal, de que los pronósticos de peligrosidad esconden tasas importantes de falsos positivos. Si bien la predicción de peligrosidad suele ser bastante acertada en la detección de los no peligrosos, la magnitud de las tasas de falsos positivos evidencia que la peligrosidad es por lo común fuertemente sobreestimada. La medida de esta sobreestimación parece variar bastante de unos estudios a otros, pero siempre los falsos positivos son superiores a los verdaderos positivos detectados mediante la predicción de peligrosidad, llegándose en algunos casos a porcentajes del 70 o incluso del 80%.

A la vista de este estado de cosas, dos preguntas surgen inmediatamente. La primera, a qué se debe esta sobreestimación sistemática de la peligrosidad, es decir, si se debe a una mala utilización de los instrumentos de predicción, o si es un problema inherente al funcionamiento (correcto) de los mismos. Y, la segunda, si es posible mejorar estas tasas de error en alguna medida, ya sea utilizando mejor los instrumentos de los que disponemos o elaborando otros mejores.

5. Limitaciones inherentes a los métodos estadísticos de predicción de la peligrosidad^{*/**}

En la actualidad la posibilidad de predicción de la conducta delictiva y/o violenta se afronta con más optimismo que hace cuarenta años, si bien se pone mucho énfasis en subrayar las limitaciones inherentes a cualquier predicción de comportamiento futuro. En general se defiende que los instrumentos de predicción modernos ofrecen estimaciones fiables sobre el peligro de reincidencia, con un grado de probabilidad superior al mero azar.⁷⁸ Pero mientras algunos autores se muestran relativamente confiados en que las predicciones de riesgo de violencia puedan equipararse a las que son comunes en otras áreas de conocimiento como la meteorología o la economía,⁷⁹ otros afirman que a pesar de toda la investigación y los estudios publicados en las últimas décadas, no estamos hoy mucho mejor en cuanto a la posibilidad de predecir el comportamiento violento o criminal de lo que estábamos hace 30 años.⁸⁰

Personalmente coincidimos con quienes señalan que el progreso real que ha supuesto la investigación sobre la valoración del riesgo de violencia en los últimos 30 o 40 años, en especial con

* El epígrafe 5 ha sido redactado en coautoría con el Prof. Dr. Francisco Montes Suay, Catedrático de Estadística e Investigación Operativa de la Universidad de Valencia.

** En este epígrafe sólo aludimos a los métodos estadísticos (ya sean estrictamente actuariales, ya de juicio clínico estructurado) porque los pronósticos de peligrosidad emitidos según el método clínico, al no seguir protocolos estandarizados y estar centrados en el caso concreto (interpretando los factores de riesgo y factores protectores en relación con las particularidades específicas del supuesto individual), no ofrecen tasas de sensibilidad, especificidad o valores predictivos conocidos que puedan ser objeto de análisis o contraste. Como mucho, cabría comparar el porcentaje de acierto de los pronósticos hechos a lo largo de su vida por cada experto, pero hasta donde sabemos no existen sobre esto datos disponibles. Por otro lado, queremos subrayar que para llevar a cabo la exposición que sigue no hemos pretendido hacer ni mucho menos una revisión exhaustiva de toda la bibliografía extrajurídica existente (abundantísima, especialmente en el ámbito anglosajón) sobre el estado actual de la discusión en torno a los instrumentos actuariales de predicción; sólo hemos pretendido subrayar algunos de los problemas más comúnmente atribuidos a estos instrumentos (así como algún otro menos conocido, pero no por ello menos preocupante), basándonos en trabajos que cumplieran fundamentalmente los siguientes criterios: que fueran de reciente publicación, y que aportaran datos sobre el mayor número de instrumentos o de muestras posibles, para poder ser considerados representativos del estado actual de la discusión.

⁷⁸ ALBRECHT en YUNDINA et al. (eds.), *FS-Nedopil*, 2013, p. 7, y también en FREUND et al. (eds.), *FS-Frisch*, 2013, p. 1071; DIITMANN, «Lieber würfeln oder eine Münze werfen? Über Risiken, wissenschaftliche Erkenntnis und Prognosen» en YUNDINA et al. (eds.), *Forensische Psychiatrie als interdisziplinäre Wissenschaft. Festschrift zum Geburtstag von Norbert Nedopil*, 2013, p. 56. Mucho más optimista que en su estudio sobre el caso Baxstrom se muestra en la actualidad STEADMAN: «even with the various limitations of current knowledge, there is room for an optimism that would have been misplaced in 1970» (*J Am Acad Psychiatry Law*, 2000, p. 270). También subraya que las predicciones de peligrosidad actuales ofrecen resultados mucho mejores que el azar SLOBOGIN, «Dangerousness and expertise redux», *Emory Law Journal*, (56-2), 2006, p. 293.

⁷⁹ ANDRÉS PUEYO/REDONDO ILLESCAS, *Papeles del Psicólogo*, 2007, p. 169: «es posible afirmar que las predicciones de riesgo de violencia, realizadas por procedimientos rigurosos como los descritos, son perfectamente comparables en su calidad y posibilidades a las de predicciones hechas en otras áreas profesionales reconocidas y útiles como la predicción meteorológica, las predicciones en ingeniería civil, las predicciones económicas o sociológicas.»

⁸⁰ FAZEL/SINGH/DOLL/GRANN, *BMJ* 2012, p. 5: «even after 30 years of development, the view that violence, sexual, or criminal risk can be predicted in most cases is not evidence based. This message is important for the general public, media, and some administrations who may have unrealistic expectations of risk prediction for clinicians». También bastante escéptico ALBRECHT en HEITMEYER/SOEFFNER (eds.) *Gewalt. Entwicklungen, Strukturen, Analyseproblemen*, 2004, pp. 510 a 512. Cfr. asimismo STRENG, «Die Zukunft der Sicherungsverwahrung nach der Entscheidung des Bundesverfassungsgerichts», *Juristenzeitung*, (17), 2011, p. 829.

el desarrollo de los métodos actuariales, no radica tanto en una mejora notable de la capacidad predictiva de estos nuevos instrumentos respecto a la que es capaz de alcanzar el juicio clínico, sino en que ofrecen una fiabilidad, inteligibilidad y transparencia mucho mayores, y por tanto permiten ser más conscientes de las limitaciones y los márgenes de error con que se efectúan las predicciones.⁸¹ Merece la pena por tanto hacer un somero repaso de estas limitaciones, para saber exactamente qué tipo de información (y de qué calidad) aporta el informe forense sobre el que se apoya la decisión del juez acerca de la peligrosidad del sujeto. Pues seguramente hay una diferencia notable entre las posibilidades que realmente existen de efectuar pronósticos fiables sobre la conducta delictiva futura de las personas y las expectativas que sobre ello parece tener el legislador cuando supedita la adopción de determinadas decisiones a la comprobación de la peligrosidad del sujeto.⁸²

5.1 Sensibilidad - especificidad - valor predictivo

Ante todo es importante para comprender bien lo que a continuación se va a explicar tener presente la diferencia entre la sensibilidad y la especificidad de un instrumento de predicción, y el valor predictivo del mismo. La sensibilidad es la capacidad de un instrumento de predicción para detectar a las personas que tienen la característica que estamos buscando, en nuestro caso, para detectar a los delincuentes que sí reincidirán. La especificidad es un valor complementario al anterior: es la capacidad del instrumento para detectar correctamente a los que no reincidirán. Ambas categorías son complementarias en el sentido de que, cuanto mayor es una de ellas, generalmente menor es la otra: cuanto más amplios sean los criterios para clasificar a un delincuente como peligroso, mayor será la sensibilidad (i.e., menos peligrosos se «escaparán» del diagnóstico), pero menor será la especificidad, porque aumenta la probabilidad de incluir como peligrosas a personas que en realidad no lo son. Y al contrario, si son muy estrictos los criterios para clasificar a alguien como peligroso tendremos menos fallos de este segundo tipo (pocos no-peligrosos serán erróneamente considerados peligrosos), pero habrá personas que sí iban a delinquir en el futuro que se nos habrán quedado fuera del diagnóstico.

Por otro lado está el valor predictivo. Mientras que la sensibilidad y la especificidad miden cuántos sujetos de los que reinciden o no lo hacen ha sido capaz de identificar correctamente el instrumento, el valor predictivo mide cuántos sujetos de los que el instrumento había predicho como futuros reincidentes, después lo han sido realmente o no. La sensibilidad dice: sabiendo cuántos han delinquido, vamos a ver cuántos de esos habíamos sido capaces de identificar con nuestro test. El valor predictivo, por el contrario, dice: sabiendo cuántos hemos identificado como peligrosos, vamos a ver cuántos de ellos realmente han delinquido después. Y aquí es importante señalar que una alta sensibilidad puede coexistir con un bajo valor predictivo del instrumento, si éste tiende a

⁸¹ ALBRECHT en YUNDINA et al. (eds.), *FS-Nedopil*, 2013, p. 11.

⁸² Véase la advertencia de FAZEL/SINGH/DOLL/GRANN, *BMJ*, 2012, p. 5 (reproducida en la nota 83), así como en parecido sentido ZIFFER, *Medidas de seguridad*, 2008, p. 161: «La confianza ingenua que parece tener el legislador en la capacidad de los peritos, jueces y funcionarios de ejecución penitenciaria de formular pronósticos confiables sobre la peligrosidad y las posibilidades de reincidencia resulta sorprendente y produce un fuerte contraste con el escepticismo existente en esta materia».

sobreestimar la peligrosidad: pocos de los peligrosos se habrán escapado de la detección, pero ello habrá sido al precio de incluir además a muchos que en realidad no lo fueron después, y de ahí que el valor predictivo resulte bajo. La tabla que hemos reproducido *supra* con los resultados del estudio llevado a cabo en Cataluña es una muestra de ello, y un ejemplo aún más extremo lo tenemos en el estudio de los informes emitidos por los psicólogos de los Juzgados de Vigilancia Penitenciaria de Madrid cuyos resultados hemos reproducido en el epígrafe anterior,⁸³ como se comprueba fácilmente si a los datos absolutos añadimos el cálculo de los valores predictivos:

	Reincide	No reincide	Total	Porcentaje de predicciones correctas (valor predictivo)
Peligrosidad sí (informe desfavorable)	15 Verdaderos positivos	77 Falsos positivos	92	16,3% (valor predictivo positivo)
Peligrosidad no (informe favorable)	3 Falsos negativos	55 Verdaderos negativos	58	94,8% (valor predictivo negativo)
Total	18	132	150	
Porcentaje de detecciones correctas	83,3% (sensibilidad)	41,6% (especificidad)		

Si bien con este estudio se ha logrado detectar al 83,3% de los futuros reincidentes (sensibilidad), ello ha sido al precio de que por cada futuro reincidente correctamente identificado se ha calificado erróneamente también como tales a otros 5 que no lo eran. Dicho de otro modo: en el 80% de los informes en los que se predijo que el sujeto sería peligroso la realidad desmintió la predicción. Y un repaso a los datos que arrojan los diversos estudios sobre casos reales que hemos recopilado en el epígrafe 4 indica que el valor predictivo de los métodos utilizados para pronosticar la peligrosidad ha sido en todos ellos también muy bajo, dadas las elevadas tasas de falsos positivos.

5.2 Tasa de prevalencia del fenómeno (*base rate*) y errores que estamos dispuestos a asumir

Las razones a las que obedece lo que parece ser una sistemática sobreestimación de la peligrosidad son de dos tipos. Las primeras son de naturaleza estadística, y derivan simplemente de las características inherentes al fenómeno que queremos predecir, y en concreto de la baja prevalencia

⁸³ GÓMEZ HERMOSO, *Informe sobre la eficacia de las predicciones de peligrosidad*, 2012, p. 7.

del mismo: la conocida como fórmula de Bayes⁸⁴ indica que, si la sensibilidad y la especificidad permanecen constantes, cuanto más baja es la prevalencia de un fenómeno en una población, tanto mayor será la tasa de falsos positivos que se obtenga al predecirlo. El siguiente ejemplo ilustra en qué proporción: supongamos que disponemos de un instrumento de predicción de la peligrosidad que tiene un porcentaje de acierto del 70%, y queremos evaluar con él a un grupo de 100 delincuentes para el que la prevalencia (el porcentaje de reincidencia) es del 50%. Como la prevalencia es del 50%, sabemos que de estos 100, 50 reincidirán. Pero como el método que estamos utilizando tiene una tasa de fallos del 30%, para 15 de estos 50 sujetos emitiremos erróneamente un pronóstico favorable (falsos negativos), y respecto de los 50 sujetos no reincidentes el método clasificará erróneamente como reincidentes a otros 15 (falsos positivos). Ahora imaginemos que la tasa de prevalencia del fenómeno (la tasa de reincidencia en la clase de delincuentes que estamos evaluando) fuera sólo del 20%. Suponiendo que evaluamos nuevamente a un grupo de 100, de estos 100, 20 reincidirán. Como nuestro instrumento de predicción sigue teniendo un porcentaje de acierto del 70%, de estos 20 reincidentes habremos predicho erróneamente como inofensivos al 30%, es decir, a 6 (falsos negativos). Y de los 80 que no van a reincidir, habremos clasificado erróneamente como peligrosos también al 30%, que son 24 (falsos positivos). Como puede verse, a igual número de sujetos evaluados, y a igual capacidad predictiva del instrumento, si bien el número de predicciones erróneas es el mismo (30), la distribución de estos errores en 'falsas alarmas' (falsos positivos) y 'omisiones' (falsos negativos) no lo es, y el dato de la baja prevalencia del fenómeno hace que la tasa de falsos positivos se multiplique.⁸⁵

En relación con la predicción de la violencia este dato es trascendental⁸⁶ porque ocurre que, por un lado, los datos empíricos que se conocen sobre criminalidad violenta indican que es un fenómeno

⁸⁴ $VP = S \times BR / [S \times BR + (1-E) \times (1-BR)]$, donde VP = valor predictivo (la probabilidad de que una persona clasificada como reincidente realmente lo sea); BR = *base rate* (prevalencia del fenómeno, proporción de individuos reincidentes); S = sensibilidad; E = especificidad

⁸⁵ He utilizado en este ejemplo porcentajes de acierto en los instrumentos de predicción y de prevalencia del fenómeno que, como veremos en los próximos epígrafes, se corresponden en buena medida con la realidad. Pero el número de falsos positivos aumenta mucho más en relación con fenómenos, como los delitos de homicidio, en los que la tasa de prevalencia es más baja: siguiendo el ejemplo que pone ALBRECHT, si el número de homicidios cometidos en una determinada población fuera del 0,1%, y dispusiéramos de un instrumento de predicción de este comportamiento que alcanzara un 95% de acierto, y lo aplicáramos a 100.000 personas, el resultado sería de 95 verdaderos positivos y 4.995 falsos positivos (ALBRECHT en HEITMEYER/SOEFNER (eds.) *Gewalt. Entwicklungen, Strukturen, Analyseproblemen*, 2004, pp. 485 y s).

⁸⁶ La relevancia de este fenómeno ha sido destacada en numerosos estudios sobre la fiabilidad de las predicciones de peligrosidad, cfr. por ej., ALBRECHT en HEITMEYER/SOEFNER (eds.) *Gewalt. Entwicklungen, Strukturen, Analyseproblemen*, 2004, pp. 485 y s.; NEDOPIL, *NStZ*, 2002, pp. 346 y ss.; EL MISMO, *Prognosen in der forensischen Psychiatrie*, 2006, pp. 50 y ss.; KÖNIG, «Der Nutzen standardisierter Risikoprognoseinstrumente für Einzelfallentscheidungen in der forensischen Praxis», *Recht & Psychiatrie*, (28), 2010, pp. 69 y ss.; YANG/WONG/COID, «The efficacy of violence prediction: a meta-analytic comparison of nine risk assessment tools», *Psychological Bulletin*, (36-5), 2010, p. 741; BUCHANAN/LEESE, «Detention of people with dangerous severe personality disorders: a systematic review», *The Lancet*, (8-358), 2001, pp. 1955 y ss., quienes exponen que, suponiendo que la tasa de prevalencia para los actos violentos cometidos por personas afectadas por trastornos severos de la personalidad sea de 9,5%, y teniendo en cuenta los porcentajes de sensibilidad y especificidad y el valor predictivo de los instrumentos actualmente existentes para valorar el riesgo de violencia, sería necesario poner bajo custodia a 6 sujetos para prevenir la comisión de un solo acto violento. Cfr. asimismo SZMUCKLER, «Risk assessment: 'numbers' and 'values'», *The Psychiatrist (formerly The Psychiatric Bulletin)* 2003, 27, p. 205, donde ofrece una tabla en la que se evidencia cómo varía el valor predictivo de diferentes instrumentos de predicción de la violencia en función de la tasa de prevalencia del fenómeno a predecir.

poco frecuente, y que la reincidencia en la comisión de delitos violentos también lo es,⁸⁷ pero por otro lado es igualmente cierto que no existen datos suficientes sobre la prevalencia de cada clase delito en las distintas poblaciones,⁸⁸ y los estudios que existen arrojan en ocasiones resultados muy dispares.⁸⁹ Esto hace que a la hora de aplicar un determinado instrumento de valoración del riesgo a un sujeto nos puede faltar un dato muy importante para poder saber *ex ante* si tenderá a sobreestimar o subestimar su peligrosidad, aunque en todo caso los datos existentes tienden a indicar que en la medida en que las tasas de prevalencia son bajas, la peligrosidad será sobreestimada.⁹⁰

La segunda de las razones que influye en la sobreestimación de la peligrosidad no es de naturaleza estadística sino político-criminal, pues está relacionada con la cantidad de falsos positivos y de falsos negativos que estamos dispuestos a asumir. Para entender esto es necesario partir de lo siguiente: los instrumentos (estructurados) de valoración del riesgo asocian probabilidades numéricas de reincidencia a las puntuaciones obtenidas por un sujeto en una escala, probabilidades que van desde el 0 hasta el 1. Pero en algún punto de dicha escala es necesario establecer la frontera entre lo que vamos a considerar riesgo alto y riesgo bajo, es decir, hay que determinar el punto de corte (o los puntos, si en lugar de sólo dos categorías queremos tres, o más: riesgo bajo – moderado – alto, etc.). Pues bien, estadísticamente es posible calcular cuántos falsos positivos y cuántos falsos negativos aparecerán si se establecen esos puntos de corte entre alta y baja peligrosidad en unos valores

⁸⁷ Como afirma ANDRÉS PUEYO, «la delincuencia violenta es mucho menos frecuente que la general (en torno al 10-12% frente al 50% de la delincuencia contra la propiedad o el 30% contra la salud pública), no está en aumento (está en descenso o estabilizada en la mayoría de los países) y la reincidencia no es muy elevada (aprox. 18% la violenta vs. 40% la no violenta)» (en DEMETRIO CRESPO (dir.), *Neurociencias y Derecho Penal*, 2013, p. 488).

⁸⁸ Por ej., NEDOPIL afirma que no se conocen las tasas de reincidencia de delincuentes respecto de los que se ha de tomar una decisión sobre su puesta en libertad cuando tienen ya una edad relativamente elevada (alrededor de 55 años), siendo así que se trata de un grupo de personas en los que el pronóstico de peligrosidad es muy relevante, en la medida en que en buena parte serán internos que solicitan la revisión de una pena de cadena perpetua; NEDOPIL, *NSfZ*, 2002, p. 348.

⁸⁹ En cuanto a concretos grupos de delitos, cfr. por ejemplo los resultados de diversos estudios sobre reincidencia de homicidas que sintetizan GROß/NEDOPIL, «Basisraten für kriminelle Rückfälle – Ergebnisse einer Literaturübersicht» en NEDOPIL, *Prognosen in der forensichen Psychiatrie*, 2006, p. 79: mientras que la comisión de nuevos delitos de homicidio oscila entre 0 y 6% (si bien este 6% no está referido a nuevas condenas por homicidio, sino a detenciones), la tasa de reincidencia cometiendo cualquier otro delito oscila en los diversos estudios entre 3 y 35,7% (si tenemos en cuenta tasa de nuevas condenas, y no de nuevas detenciones). En cuanto a la reincidencia de los delincuentes sexuales, se afirma que globalmente oscila en torno al 20%, (por ej. PÉREZ RAMÍREZ et al., *Psicothema* 2008, p. 208); HERRERO da cuenta de recientes meta-análisis en el ámbito internacional según los cuales la reincidencia sexual en quienes han sido condenados por un delito sexual oscila entre el 15 y el 20%, si bien conviene diferenciar entre comisión de nuevos delitos contra la libertad sexual y reincidencia en general (así, según un estudio de 2005, tras un periodo de seguimiento de 5 años la tasa de reincidencia sexual es del 13.7%, la de cualquier tipo de violencia (incluyendo sexual) del 14.3% y la de reincidencia en general (cualquier tipo de delito) el 36.2%), cfr. HERRERO, «¿Por qué no reincide la mayoría de los agresores sexuales?», *Anuario de Psicología Jurídica*, 2013, 23, p. 72. Problemas comunes a los estudios sobre tasas de prevalencia de los delitos en general, así como prevalencia de reincidencia, son, entre otros, las cifras negras de delitos no detectados que en algunos ámbitos se consideran importantes, los periodos de seguimiento no suficientemente prolongados, o la falta de equivalencia entre los criterios utilizados para definir lo que se considera reincidencia entre unos estudios y otros (por ej., si ha de considerarse tal una nueva detención, o si es necesaria una nueva condena; si se considera reincidencia la evidencia de realización de cualquier nuevo delito, o sólo de delitos del mismo tipo que los que generaron la condena inicial, etc.).

⁹⁰ ALBRECHT llega a afirmar que la prevalencia de los delitos violentos más graves, es decir los de homicidio, es tan baja (en Alemania, alrededor de 700 homicidios en el año 2010 en una población de unos 82 millones de personas) que cualquier predicción, a no ser que se limite a periodos de tiempo muy cortos y bajo circunstancias del autor y del contexto muy específicas, nunca podrá ser mejor que el azar (*FS-Nedopil*, 2013, p. 8).

numéricos u otros, y también es posible calcular cuál de estos valores es el que reduce en mayor medida ambas clases de errores.⁹¹ Por ejemplo, en el estudio que citábamos anteriormente de PÉREZ RAMÍREZ y otros, en el que se utiliza el SVR-20 para valorar el riesgo de reincidencia en una muestra de agresores sexuales en Cataluña, se propone la puntuación de .20 en el SVR-20 como aquella que ofrece el mejor equilibrio posible entre sensibilidad y especificidad, y por ello genera un menor número tanto de falsos positivos como de falsos negativos. Sin embargo, la utilización de este valor como punto de corte para clasificar a los sujetos en futuros reincidentes y futuros no reincidentes sigue generando un 6% de falsos negativos (de los 118 sujetos a los que se pronosticó un riesgo bajo de reincidencia, 7 sí volvieron a delinquir). En consecuencia, si político-criminalmente esta cifra pareciera inasumible, sería posible elegir otro valor (inferior a .20) como punto de corte, con lo que, al aumentar el número de sujetos pronosticados como reincidentes, disminuiría el número de falsos negativos. Ahora bien, esta decisión aumentaría al mismo tiempo el número de falsos positivos, que como ya vimos alcanzaba en la predicción inicial un porcentaje del 62,3% respecto del total de los predichos como reincidentes. Y a la inversa, si fuera esta última cifra la que nos pareciera excesiva, podríamos aumentar la puntuación exigida en el SVR-20 para entender que existe a partir de ella el peligro de reincidencia, pero ello automáticamente elevaría el número de sujetos reincidentes que escaparían de la detección (falsos negativos).

Estas decisiones evidentemente no dependen ya de razones matemáticas, sino que obedecen a criterios normativos y político-criminales sobre qué tipo de errores se prefiere minimizar.⁹² Y todo apunta a que cuando se trata de adoptar medidas sobre sujetos eventualmente peligrosos predomina el deseo de minimizar las 'omisiones', aunque ello sea a costa de asumir un importante número de 'falsas alarmas'. Creo que resulta ilustrativo de esta tendencia la valoración que contiene el *Informe sobre predicciones de peligrosidad de los informes psicológicos forenses emitidos en los Juzgados de Vigilancia Penitenciaria de Madrid*, al que nos hemos referido *supra*, y en el que si bien se detectaron correctamente el 83% de las reincidencias, había una altísima tasa de falsos positivos (de los 92 informes en que se efectuó un pronóstico negativo, 77 sujetos no reincidieron, es decir, el porcentaje de falsos positivos fue del 83%). Pues bien, en el informe se afirma que debe extremarse la rigurosidad a la hora de emitir informes favorables:

«Teniendo en cuenta las graves consecuencias sociales y personales de estos delitos, no se deben realizar propuestas favorables si no se cuenta con todas las condiciones evaluadas a favor, ya que en estos casos debe primar la seguridad. El riesgo que se asuma debe ser mínimo. Es necesario controlar al máximo el riesgo para

⁹¹ Se trata fundamentalmente de la que se conoce como curva ROC (*receiver operating characteristics curve*), que es el resultado de la combinación de la tasa de verdaderos positivos (sensibilidad) contra la tasa de falsos positivos (1 menos especificidad) para cada punto de corte del instrumento de predicción.

⁹² Subrayan que esta segunda decisión ya no es una cuestión clínica NEDOPIL/STADTLAND, «Das Problem der falsch Positiven: Haben wir unsere prognostische Kompetenz seit 1966 verbessert?» en LÖSEL/BENDER/JEHLE (eds.), *Kriminologie und wissenschaftsbasierte Kriminalpolitik. Entwicklungs- und Evaluationsforschung*, 2007, pp. 545 y ss. Los autores exponen los diferentes números de falsos positivos que aparecen en una muestra de 262 sujetos evaluados con el PCL-R y el HCR-20, en función de que el punto de corte para la afirmación de la peligrosidad se establezca en unas u otras puntuaciones, y concluyen señalando que el experto sólo puede exponer el conocimiento empírico que existe y advertir de las limitaciones del mismo, y que las decisiones que haya que tomar sobre dicha base ya son siempre de naturaleza normativa (ibidem, p. 549).

evitar nuevas víctimas. Obviamente, al establecer criterios rigurosos puede darse el caso de que penados que ya pudiesen obtener beneficios penitenciarios no los obtengan en un primer momento. Sin embargo, como hemos dicho, consideramos que debe prevalecer en estos casos el valor de seguridad».⁹³

Hasta aquí hemos visto dos razones (una matemática, la otra no) a las que puede obedecer un elevado número de falsos positivos en las predicciones de peligrosidad. Pero llegados a este punto interesa preguntarse cuáles están siendo, en la realidad de su aplicación práctica, los valores predictivos de los instrumentos de valoración del riesgo.⁹⁴

5.3 Valores predictivos de los instrumentos de valoración del riesgo

Los instrumentos estructurados de valoración del riesgo se construyen, como hemos dicho antes, utilizando una muestra de sujetos cuyas características se observan y cuyo comportamiento se registra cuidadosamente, para determinar después, siempre en relación con ese grupo de sujetos, qué características han resultado asociadas en mayor medida al comportamiento delictivo, y en consecuencia qué factores pueden considerarse mejores predictores del mismo. Que la capacidad predictiva del instrumento sea mejor o peor depende de cómo se comporte cuando se aplica a sucesivas – nuevas – muestras de sujetos: si aplicado a nuevos grupos muy parecidos al inicial y siguiendo exactamente el mismo protocolo las predicciones resultan corroboradas de nuevo por el comportamiento real, el instrumento demuestra su buena capacidad predictiva; si, por el contrario, en estas nuevas muestras el comportamiento real no coincide con lo predicho, hay que concluir que el instrumento no resulta fiable como medio de predicción. Estos estudios sucesivos (denominados estudios de validación) son por tanto decisivos para conocer la bondad de cada instrumento de valoración del riesgo de violencia y/o reincidencia, pero para que pueda considerarse que un resultado parecido obtenido en un nuevo estudio es una corroboración de la capacidad predictiva del instrumento en cuestión es necesario que el nuevo estudio se haya desarrollado siguiendo estrictamente los protocolos del estudio inicial (o los estándares que los autores del mismo hayan establecido en el manual de uso que acompaña a cada test). Para entenderlo: supongamos que un instrumento se ha construido sobre una muestra de jóvenes de entre 14 y 18 años, varones, internados en centros cerrados de cumplimiento de medidas penales, y que ofrece buenos valores predictivos de reincidencia violenta para un periodo de seguimiento de 5 años. Si posteriormente esa prueba se administra a un grupo de mujeres, de mediana edad, entre la población general, para predecir la comisión de cualesquiera delitos en un plazo de 10 años, el hecho de que en este segundo estudio los resultados obtenidos sean muy diferentes de los del primero no pone en absoluto en cuestión la capacidad predictiva del instrumento de valoración.

⁹³ GÓMEZ HERMOSO, *Informe sobre la eficacia de las predicciones de peligrosidad*, 2012, p. 3.

⁹⁴ Sobre el controvertido tema de si los instrumentos estructurados predicen mejor que el juicio clínico, y/o si los instrumentos estrictamente actuariales predicen mejor o peor que los métodos de juicio clínico estructurado, ya hemos aportado algunos datos *supra* al final del epígrafe 3. y al inicio del epígrafe 5., por lo que nos remitimos a lo que allí se dijo. En todo caso, como ahora nos interesa la capacidad predictiva general de los distintos métodos, es una cuestión que no es necesario abordar con mayor detalle a los efectos de este trabajo.

Pues bien, aunque los autores de los distintos instrumentos de valoración del riesgo afirman que existen numerosos estudios que han replicado sus resultados, por lo que estaría acreditada su capacidad predictiva, algunos trabajos recientes sugieren que la validez de estos instrumentos no puede considerarse enteramente establecida, porque los estudios de validación no respetan en muchos casos las características esenciales de los estudios originales: por ejemplo, se aplican a muestras de población que no son similares a aquélla para la que el instrumento se diseñó (en edad, sexo, raza, número de antecedentes delictivos, etc.), o no utilizan el mismo periodo de seguimiento que el estudio inicial, o no se basan en las mismas fuentes para obtener la información relevante (por ej., el estudio inicial consideró como reincidencia los nuevos delitos relatados por el propio sujeto, mientras que el estudio de seguimiento sólo considera reincidencia la aparición de nuevas inscripciones en el registro oficial de antecedentes), etc., o simplemente, al no contener información sobre estos extremos, los resultados que se obtienen no pueden ser considerados definitivamente ni como corroboración ni como refutación de la capacidad predictiva de los instrumentos que administran.⁹⁵

El hecho es, no obstante, que la mayor parte de los estudios posteriores sobre la validez predictiva de los instrumentos de valoración suelen encontrar asociaciones significativas entre las puntuaciones consideradas de riesgo y la variable a predecir (reincidencia, violencia), por lo que suelen ser interpretados como pruebas de la robustez de los instrumentos aplicados. Sin embargo, cuando las diferencias entre el estudio inicial y los posteriores afectan a aspectos esenciales, como por ejemplo la clase de resultado que se comprueba (delincuencia violenta – delincuencia no violenta), el hecho de que sea posible utilizar el mismo instrumento para predecir con similar grado de acierto comportamientos diferentes hace surgir la duda de qué es lo que en realidad se está midiendo con estos tests: como apuntan ROSSEGER y otros, «si los instrumentos actuariales de valoración del riesgo correlacionan con cualquier clase de comportamiento bajo diferentes condiciones, en diferentes contextos, usando distintos periodos de seguimiento y bases de datos, esto no puede servir automáticamente como una corroboración de la (presunta) solidez del instrumento, porque no está claro qué es lo que el instrumento está midiendo. Si el rasgo latente que está siendo valorado por el instrumento de predicción es el comportamiento inadaptado y no la tendencia a cometer delitos sexuales o violentos, resulta cuestionable que este instrumento deba usarse ante los tribunales para valorar el riesgo de comisión repetida de delitos sexuales o violentos. Mientras que algunas desviaciones respecto del estudio original podrían ser indicios de la solidez del modelo, una desviación mayor pone en peligro su validez, especialmente si la desviación afecta a la variable dependiente».⁹⁶

⁹⁵ Así, ROSSEGER et al., «Current obstacles in replicating risk assessment findings: a systematic review of commonly used actuarial instruments», *Behavioural Sciences and the Law*, (31), 2013, pp. 154 a 164 (los instrumentos cuyos estudios de replicación fueron tenidos en cuenta en este trabajo fueron el VRAG, el SORAG y el Static-99). En la p. 162 los autores resumen los resultados obtenidos afirmando que de las 108 muestras que investigaron, ninguna podía considerarse una replicación perfecta del estudio original, y que alrededor de la mitad de las muestras sólo aplicaban correctamente dos tercios o menos de los criterios relevantes del instrumento de predicción. Añaden que en ocasiones no está claro si simplemente los autores de los estudios posteriores olvidaron consignar parámetros que sí habían respetado, o si directamente no se ajustaron en esos puntos a las indicaciones del estudio original.

⁹⁶ ROSSEGER et al., *Behav. Sci. Law*, 2013, p. 163.

Relacionada con este problema hay también otra cuestión: la de si algunos de los instrumentos de valoración del riesgo existentes predicen mejor que otros o no. Sobre esto ya hemos indicado que se han realizado numerosos estudios⁹⁷ y aunque no hay resultados definitivos, parece que no se han detectado diferencias muy significativas entre ellos.⁹⁸ Suponiendo que efectivamente esto sea así, es decir, que los distintos instrumentos tengan similar capacidad predictiva, cabe entonces preguntarse a qué se debe esta «intercambiabilidad» entre unos y otros. SKEEM y MONAHAN apuntan que los distintos instrumentos de valoración existentes miden – aunque sea en formas diferentes – factores comunes de riesgo, que estarían todos ellos relacionados con cuatro dimensiones: el historial delictivo, un estilo de vida irresponsable, psicopatía y actitudes criminales, y problemas relacionados con el abuso de sustancias. Es más, relatan los resultados obtenidos por un grupo de autores que realizaron el siguiente experimento: cogieron todos los ítems de cuatro conocidos instrumentos de valoración de la peligrosidad, los metieron en una lata de café, y después de sacudirla extrajeron al azar los ítems para crear cuatro nuevos instrumentos de predicción. Y encontraron que los instrumentos creados con la lata («coffee can instruments») predecían la delincuencia violenta y no violenta igual de bien que los instrumentos originales.⁹⁹ SKEEM y MONAHAN concluyen que, con independencia de diferencias superficiales, todos los instrumentos bien validados de predicción de la delincuencia estarían reflejando «un patrón arraigado de interacciones interpersonales disfuncionales y agresivas y de estilos de vida antisociales e inestables que son comunes a muchos autores de actos de violencia».¹⁰⁰ Pero esta conclusión no haría sino reforzar la advertencia que hacían ROSSEGGER et. al. sobre que, aunque aparentemente los distintos instrumentos estén diseñados para predecir tipos específicos de conducta violenta o de conducta delictiva, en realidad todos ellos pueden estar en realidad midiendo una tendencia más general e inespecífica al comportamiento antisocial.¹⁰¹

Pues bien, una vez hechas estas advertencias sobre la – en apariencia – similar capacidad predictiva de los distintos instrumentos de valoración del riesgo de violencia, sigue pendiente la pregunta de cuál es dicha capacidad predictiva. En un reciente meta-análisis en el que se comparaban los valores obtenidos por diferentes estudios de validación de los instrumentos de valoración del riesgo de violencia más usados internacionalmente, los resultados fueron los siguientes:¹⁰²

⁹⁷ Cfr. por ej., las referencias que aportan SKEEM/MONAHAN, *Current Directions in Psychological Science*, 2011, p. 40., así como las referencias que se indicaron ya *supra*, al inicio del epígrafe 5.

⁹⁸ Así lo afirman SKEEM/MONAHAN, *Current Directions in Psychological Science*, 2011, p. 40; también YANG/WONG/COID, *Psychological Bulletin*, 2010, p. 759. SINGH y otros, sin embargo, opinan que sí cabe detectar diferencias sustanciales entre unos instrumentos y otros, en el sentido de que los instrumentos diseñados para valorar el riesgo de violencia en poblaciones específicas evidencian mejor capacidad predictiva que los diseñados para poblaciones más generales (SINGH/GRANN/FAZEL, *Clinical Psychology Review*, 2011, p. 509).

⁹⁹ SKEEM/MONAHAN, *Current Directions in Psychological Science*, 2011, p. 40. El estudio al que se refieren es el de KRONER /MILLS/MORGAN (2005): «A coffee can, factor analysis, and prediction of antisocial behavior: the structure of criminal risk», *International Journal of Law & Psychiatry*, (28), 2005, pp. 360 a 374.

¹⁰⁰ YANG/WONG/COID, *Psychological Bulletin*, 2010, p. 759.

¹⁰¹ Lo cual pondría en cuestión una de las supuestas ventajas que – como vimos páginas atrás – los instrumentos de valoración del riesgo reclaman tener frente a la óptica tradicional de predicción de la peligrosidad, en el sentido de que como los primeros se ocupan de la predicción de comportamientos muy claros y bien definidos, su capacidad predictiva sería superior.

¹⁰² Datos extraídos de FAZEL/SINGH/DOLL/GRANN, *BMJ* 2012, tabla contenida en la p. 10.

	Delincuencia violenta ¹⁰³	Delincuencia sexual ¹⁰⁴	Delincuencia en general ¹⁰⁵
Sensibilidad ¹⁰⁶	0,92 (0,88 - 0,94)	0,88 (0,83 - 0,92)	0,41 (0,28 - 0,56)
Especificidad	0,36 (0,28 - 0,44)	0,34 (0,20 - 0,51)	0,80 (0,67 - 0,8)
Valor predictivo positivo ¹⁰⁷	0,41 (0,27 - 0,60)	0,23 (0,09 - 0,41)	0,52 (0,32 - 0,59)
Valor predictivo negativo	0,91 (0,81 - 0,95)	0,93 (0,82 - 0,98)	0,76 (0,61 - 0,84)
Número de los que hace falta detener ¹⁰⁸	2 (2-4)	5 (2-11)	2 (2-3)
Número de los que pueden ser liberados ¹⁰⁹	10 (4-18)	14 (5-48)	3 (2-6)

A la hora de valorar estos resultados conviene recordar de nuevo la diferencia entre sensibilidad y valor predictivo a la que nos referíamos supra (epígrafe 5.1.). Los instrumentos evaluados en este estudio ofrecen tasas altas de sensibilidad: al menos los diseñados para predecir el riesgo de delincuencia violenta y delincuencia sexual consiguen detectar correctamente a alrededor del 90% de los delincuentes que efectivamente llevarán a cabo estas conductas. Pero estos valores elevados de sensibilidad contrastan con otros muchos más bajos en cuanto al valor predictivo, que es precisamente el criterio que más interesa si lo que queremos es saber con qué grado de acierto podemos predecir la reincidencia. Como dijimos supra, el valor predictivo positivo indica el porcentaje de personas que, habiendo sido clasificadas como peligrosas por el instrumento de

¹⁰³ Los instrumentos de valoración del riesgo de conducta violenta que empleaban los estudios revisados para este meta-análisis fueron HCR-20; SARA, SAVRY y VRAG.

¹⁰⁴ Los instrumentos de valoración del riesgo de delincuencia sexual que empleaban los estudios revisados para este meta-análisis fueron SORAG, Static-99 y SVR-20.

¹⁰⁵ Los instrumentos de valoración del riesgo de delincuencia en general que empleaban los estudios revisados para este meta-análisis fueron LSI-R y PCL-R. Información resumida sobre todos estos instrumentos de valoración del riesgo puede encontrarse en ARMAZA ARMAZA, *El tratamiento del delincuente imputable peligroso*, 2013, pp. 115 y 109 para el HCR-20 y el VRAG, respectivamente; pp. pp. 110, 112 y 116, para el SORAG, Static-99 y SVR-20, respectivamente, y pp. 114 para el PCL-R. También puede consultarse la tabla que ofrecen ANDRÉS PUEYO/REDONDO ILLESCAS, *Papeles del Psicólogo*, 2007, p. 167, y el resumen contenido en YANG/WONG/COID, *Psychological Bulletin*, 2010, pp. 766 y s. Para información más detallada sobre el HCR-20 véase también ARBACH/ANDRÉS PUEYO, «Valoración del riesgo de violencia en enfermos mentales con el HCR-20», *Papeles del Psicólogo*, (28-3), 2007, pp. 180 y ss.; y en relación con el SVR-20, cfr. REDONDO ILLESCAS/PÉREZ/MARTÍNEZ, *Papeles del Psicólogo*, 2007, pp. 191 y ss.

¹⁰⁶ Los valores de sensibilidad y especificidad se ofrecen con un intervalo de confianza del 95% (cuyos límites mínimo y máximo son los indicados entre paréntesis).

¹⁰⁷ Los valores predictivos positivos y negativos son medianas y las cifras indicadas entre paréntesis corresponden al intervalo intercuartílico, cuyos extremos superior e inferior son el primer y tercer cuartil, respectivamente.

¹⁰⁸ Número de personas consideradas peligrosas que hace falta mantener detenidas para evitar un delito.

¹⁰⁹ Número de personas consideradas poco peligrosas que podrían dejarse en libertad antes de que se produjera un solo delito.

valoración del riesgo, efectivamente delinquen. Un valor predictivo positivo de 0.5 no es mejor que el mero azar como criterio predictivo: significa que de todos los sujetos que predijimos que sí volverían a delinquir, la mitad no lo hicieron. Pues bien, adviértase que en los estudios revisados en el trabajo que estamos citando dicho valor es en casi todos ellos inferior a 0.5, llegando incluso, en los instrumentos de valoración del riesgo de reincidencia sexual, a sólo el 23% - esto es, utilizando dichos instrumentos, del total de sujetos que predijimos que sí cometerían delitos sexuales, el 77% no lo hicieron -. Únicamente en los instrumentos de valoración del riesgo que miden la probabilidad de comisión de delitos en general (en este estudio, la PCL-R y el LSI-R) el valor es superior al 0.5, pero únicamente por dos décimas, lo que prácticamente lo equipara al azar.

Frente a estos valores predictivos positivos tan bajos, llama la atención que los valores predictivos negativos sean mucho más elevados: del orden del 90% en los dos primeros grupos de instrumentos de predicción, y cercanos al 80% en el tercero. Esto significa que a la hora de predecir quienes *no* volverán a cometer delitos, los instrumentos aciertan mucho más.

¿Qué conclusiones cabe extraer de estos datos? Los autores de este estudio afirman que debido a sus altas tasas de sensibilidad estos instrumentos de valoración del riesgo pueden ser una ayuda útil a la hora de tomar decisiones sobre tratamiento y gestión, pues logran detectar moderadamente bien a los individuos peligrosos.¹¹⁰ Por ejemplo, a la hora de decidir a qué internos ofrecer en prisión la posibilidad de seguir un tratamiento reeducativo para delincuentes sexuales, el test identificaría a una población en riesgo alto de reincidencia de la que sólo escaparían un 10 o un 20% de los delincuentes peligrosos. Esta alta sensibilidad se obtiene desde luego a costa de una especificidad mucho menor (es decir: en el grupo al que ofreceríamos el tratamiento habría también muchos sujetos que seguramente no lo necesitan porque no van a volver a reincidir). Sin embargo, en la medida en que ofrecer un tratamiento no vulnera ni restringe los derechos de los internos, e incluso a los no peligrosos les podría resultar en alguna medida beneficioso, la tasa de falsos positivos no es un inconveniente grave para este tipo de decisiones. Sin embargo, el problema se plantea de forma muy diferente cuando de lo que se trata es de resolver otras cuestiones. Como afirman los autores del estudio citado, la utilidad de estos instrumentos si se quisieran usar como criterio decisivo a la hora de determinar la pena a imponer, o de tomar decisiones negativas sobre la libertad del sujeto, está muy limitada por sus bajos valores predictivos positivos:¹¹¹ si nos guiáramos por las puntuaciones obtenidas en los tests analizados, por cada delincuente sexual al que le negáramos la libertad por miedo a su reincidencia mantendríamos indebidamente en prisión a otros 5 (como media). La situación es sin embargo la opuesta cuando de lo que se trata es de determinar a quién sí puede dejarse en libertad: los instrumentos resultan mucho más fiables a la hora de predecir la ausencia de riesgo de reincidencia, pues quien hubiera obtenido la puntuación indicativa de bajo riesgo en un

¹¹⁰ FAZEL/SINGH/DOLL/GRANN, *BMJ* 2012, p. 4.

¹¹¹ En el mismo sentido se pronuncian YANG/WONG/COID, *Psychological Bulletin*, 2010, p. 761, quienes concluyen su estudio afirmando que, debido a sus moderados niveles de eficacia predictiva, los instrumentos de valoración del riesgo no deberían utilizarse como el único ni como el principal medio para tomar decisiones clínicas o judiciales que dependen de un alto nivel de exactitud predictiva, como la «preventive detention».

test sería poco peligroso realmente en 9 de cada 10 casos (algo menos en los instrumentos que miden el riesgo de reincidencia general).

Es cierto que otras investigaciones han encontrado niveles algo mejores de capacidad predictiva que los de la tabla anterior.¹¹² Por ejemplo, en otro meta-análisis realizado por SINGH y otros sobre la eficacia predictiva de diversos instrumentos, los valores predictivos positivos y negativos obtenidos fueron los siguientes:¹¹³

Instrumento	Valor predictivo positivo: mediana (IQR)	Valor predictivo negativo: mediana (IQR)
SAVRY	0,76 (0,42 - 0,85)	0,76 (0,49 - 0,91)
HCR-20	0,71 (0,55 - 0,85)	0,67 (0,51 - 0,70)
VRAG	0,66 (0,37 - 0,79)	0,74 (0,62 - 0,84)
LSI-R	0,57 (0,44 - 0,70)	0,53 (0,44 - 0,65)
SARA	0,53 (0,31 - 0,62)	0,79 (0,67 - 0,92)
PCL-R	0,52 (0,45 - 0,75)	0,68 (0,39 - 0,82)
SORAG	0,38 (0,33 - 0,86)	0,64 (0,59 - 0,90)
Static-99	0,33 (0,18 - 0,56)	0,82 (0,71 - 0,94)
SVR-20	0,33 (0,16 - 0,60)	0,65 (0,56 - 0,87)

¹¹² En su trabajo de 2006 Slobogin cita varios artículos de los años 80 y 90 en los que los porcentajes de falsos positivos serían muchos más reducidos que los que acabamos de exponer: entre un 25 y un 47% (SLOBOGIN, *Emory Law Journal*, 2006, p. 291 y nota 73). Sin embargo, en la medida en que el trabajo de FAZEL/SINGH/DOLL/GRANN en *BMJ* 2012 es más reciente, y tiene un alcance más amplio (revisa estudios realizados sobre 73 muestras que comprendieron en total 24.827 sujetos) he considerado más apropiado destacar sus resultados en el texto. Por otro lado, la representatividad de algunos de los trabajos que cita Slobogin parece bastante limitada: por ej. uno de los estudios que cita, el trabajo de APPERSON/MULVEY /LIDZ, «Short-term clinical prediction of assaultive behaviour: artifacts of research methods», *The American Journal of Psychiatry*, (150-9), 1993, pp. 1374 y ss., en el que para uno de los grupos estudiados se obtuvo una tasa de falsos positivos de sólo el 25%, se estudió a 136 pacientes psiquiátricos durante sólo 4 meses, y se midió la conducta violenta que evidenciaron dentro de la institución. Es sabido que las predicciones de peligrosidad tienen una probabilidad mucho mayor de acierto cuando se hacen para periodos cortos de tiempo, como era el caso de este estudio, que cuando se refieren a periodos más largos (uno o varios años; en el estudio de FAZEL/SINGH/DOLL/GRANN, *BMJ* 2012, cuyos resultados he expuesto en el texto el periodo de seguimiento medio fue de 49,6 meses), y también es sabido que resulta más fácil hacer predicciones cuando se refieren a un contexto concreto (comportamiento dentro de la institución) que cuando se refieren a contextos más variables e indeterminados, como lo es la vida en libertad.

¹¹³ SINGH/GRANN/FAZEL, *Clinical Psychology Review*, 2011; los valores indicados a continuación están extraídos de la tabla 4 en la p. 507, y corresponden a una clasificación de los sujetos en «peligrosos» y «no peligrosos» en la que se ha considerado peligrosos sólo a aquellos que según el instrumento correspondiente obtenían una puntuación indicativa de riesgo alto, y no peligrosos tanto a los que obtenían puntuaciones indicativas de riesgo moderado como a los que las obtenían indicativas de riesgo bajo.

Como puede observarse, también aquí son pocos los instrumentos que evidencian valores predictivos positivos claramente superiores al 50% (sólo tres de los nueve analizados), y además llama la atención que los instrumentos que mejores valores ofrecen en este segundo estudio (SAVRY, HCR-20 y VRAG) son de los que daban un valor predictivo positivo de sólo 0.41 en el trabajo de FAZEL et al. cuyos valores hemos reproducido en la tabla anterior, y que se publicó sólo un año después que éste que ahora estamos analizando de SINGH et. al. Además, en ambos trabajos hay que tomar en consideración las cifras que constan entre paréntesis, y que indican el intervalo intercuartílico. Es decir: en realidad los distintos estudios de predicción llevados a cabo aplicando cada uno de los instrumentos de valoración del riesgo han obtenido distintos valores predictivos positivos (y negativos, pero ahora nos interesan los primeros, que son los que han acertado en la predicción de la existencia de peligrosidad), la mitad de los cuales oscilaban entre las dos cifras indicadas en el paréntesis, y siendo el número que precede al paréntesis la mediana de dichos datos.¹¹⁴ En consecuencia, incluso en los casos en que la mediana del valor predictivo positivo - en este segundo estudio - está situada en el entorno del 70% (insistimos: sólo tres de los nueve instrumentos analizados), no hay que perder de vista que con ese mismo instrumento de predicción se han obtenido en algunos de los estudios de validación revisados valores inferiores o superiores, la mayor parte de los cuales se sitúan entre el mínimo y el máximo indicados en el paréntesis. Y, salvo en uno de los instrumentos, el mínimo queda siempre por debajo del 50%. Luego ni siquiera en estos casos puede asegurarse que las predicciones hechas utilizando estos instrumentos de valoración de la peligrosidad sean siempre mejores que el azar.

5.4 Los márgenes de error y la traslación de los resultados grupales a la predicción del comportamiento individual

Relacionado en parte con lo que acabamos de explicar, queda además un último problema que abordar. Como hemos indicado en repetidas ocasiones en las páginas anteriores, cualquiera de estos métodos estadísticos de los que estamos hablando (ya sean los estrictamente actuariales, ya los de juicio clínico estructurado cuando se utilizan para obtener estimaciones numéricas de probabilidad) ofrecen únicamente probabilidades de reincidencia o de comportamiento violento para grupos de

¹¹⁴ Otros trabajos también han destacado que las tasas reales de reincidencia que muestran los sujetos a los que se ha clasificado como de alto riesgo varían muchísimo de unas muestras a otras, y que ello ocurre incluso cuando a los sujetos se les ha aplicado el mismo instrumento de valoración del riesgo; así SINGH/FAZEL/GUEORGUEVA/BUCHANAN, «Rates of violence in patients classified as high risk by structured risk assessment instruments», *The British Journal of Psychiatry*, (204-3), 2014, pp. 180 a 197, y HELMUS/HANSON/THORNTON/BABCHISHIN/HARRIS, «Absolute recidivism rates predicted by Static-99R and Static-2002R sex offender risk assessment tools vary across samples: a meta-analysis», *Criminal Justice and Behaviour*, (39), 2012, pp. 1148 a 1171. Ambos trabajos recomiendan por ello ser muy cautos a la hora de asignar a un individuo una probabilidad numérica de reincidencia como resultado de su clasificación en una u otra categoría de riesgo de estos instrumentos. El primero de ellos subraya que esta variabilidad demuestra la relevancia que tienen las diferentes tasas de prevalencia (*base rates*) en la predicción del comportamiento delictivo, algo a los que ya nos hemos referido *supra* en este trabajo (epígrafe 5.2), y concluye diciendo que «it does not seem possible to use structured risk assessment instruments to assign reliably a predetermined numerical probability to the potential for an individual to act violently. [...] These findings support recommendations for caution, given the present state of knowledge, in the use of such probabilistic risk estimates to influence decisions related to individual liberty and public safety. The results of individual risk assessments should be reported with explicit acknowledgement of the possible sources of error associated with their use.»

individuos, y no predicciones concretas para un sujeto determinado. Como indicábamos en el epígrafe 3. de este trabajo, estos instrumentos presuponen la validez de la aplicación analógica de los resultados del grupo a la conducta del individuo, de manera que la puntuación obtenida por el sujeto concreto al que administramos el test se corresponde con un rango de valores que en el grupo de referencia han estado asociados a una probabilidad de reincidencia de X%, y ésta es la que se ofrece como guía para estimar la probabilidad de reincidencia del individuo al que estamos evaluando.

Todos estos valores numéricos, sin embargo, están sujetos a determinados márgenes de error. Cuando un instrumento de valoración del riesgo afirma, por ejemplo, que en relación con los sujetos sobre los cuales se realizó el estudio, aquéllos cuyas puntuaciones entran en la categoría de «alto riesgo» tuvieron una tasa de reincidencia del 33%, esa cifra de 33 es un parámetro que define la probabilidad media de reincidencia asociada a ese grupo, pero con un margen de error que puede ir por ejemplo desde 16% hasta 50% para un intervalo de confianza del 95%.¹¹⁵ Dicho de otro modo: el valor real de la tasa de reincidencia para ese grupo de individuos adoptará cualquier valor entre 16% y 50% el 95% de las veces, o sea, 19 veces de cada 20. Lo cual revela ya una apreciable carga de incertidumbre en estos pronósticos grupales, porque desde el punto de vista del hipotético juez que tuviera que decidir por ejemplo sobre la imposición de una medida de seguridad adicional a la pena, no sería en absoluto irrelevante que la probabilidad esperable de reincidencia del grupo en el que se encuadra el sujeto sea del 16 o del 50%, y sobre todo que la probabilidad real pueda adoptar, el 95 por cien de las veces, cualquier valor entre estos dos extremos.

Pero, aún más importante que esto, cuando utilizamos ese instrumento para evaluar a un nuevo individuo, y esa persona obtiene una puntuación que entra dentro del rango considerado de alto riesgo, ¿podemos partir de que la probabilidad de reincidencia para este individuo concreto tiene los mismos márgenes de error que los valores medios para el grupo?

Normalmente los estudios sobre la capacidad predictiva de los diferentes instrumentos de valoración del riesgo no incluyen ninguna referencia a esta cuestión, lo que supone que asumen implícitamente que el margen de error para la predicción individual es el mismo que el margen de error grupal del modelo. Pero trabajos recientes afirman que esto no es cierto, y sostienen que la estimación de la probabilidad de reincidencia para cada nuevo individuo tiene un margen de error que ya no se denomina intervalo de confianza, sino intervalo de predicción, y que es muy distinto del intervalo de confianza grupal tanto conceptualmente como en cuanto a su magnitud: el intervalo de predicción es siempre mayor que el intervalo de confianza.¹¹⁶ ¿Cuánto mayor? Los estudios en los que se han estimado estos márgenes de error arrojan resultados sorprendentes. Por ejemplo, en un trabajo en el que se estimaron estos intervalos para medir la probabilidad de reincidencia a partir de las

¹¹⁵ El ejemplo está sacado del trabajo de HART/COOKE, *Behav. Sci. Law*, 2013, p. 92. En el mismo ejemplo de Hart y Cooke, la tasa de reincidencia del grupo considerado de bajo riesgo fue de 10%, con un margen de error de entre 2% a 18% para un intervalo de confianza de 95%.

¹¹⁶ COOKE/MICHIE, «Violence risk assessment. Challenging the illusion of certainty» en MCSHERRY/KEIZER (eds.), *Dangerous people. Policy, prediction and practice*, 2011, p. 151 (y passim).

puntuaciones obtenidas por distintos individuos en la escala de psicopatía de Hare (PCL-R), la probabilidad media de reincidencia para un individuo que hubiera obtenido una puntuación de 12,5 en dicha escala era del 14%, pero el margen de error iba de 0 a 98 para un intervalo de confianza del 95%. Para una puntuación de 25, el margen de error iba desde 0 a 99, y para una puntuación de 30, de 0 a 99,5 (en todos los casos para un intervalo de confianza del 95%).¹¹⁷ Como puede observarse, si se tienen en cuenta estos márgenes de error el valor que cabe dar a la cifra que hayamos obtenido como probabilidad media de reincidencia no puede ser muy alto, si hay que asumir que en el 95% de los casos puede adoptar en realidad cualquier valor situado entre 1% y 99%. Y la separación de los individuos cuyo comportamiento se pronostica en categorías de alto riesgo o de bajo riesgo en función de las probabilidades medias que les haya asignado el test pierde toda capacidad discriminadora, en la medida en que estos márgenes de error tan amplios se solapan por completo.¹¹⁸

Estos trabajos ponen radicalmente en cuestión la capacidad de efectuar, en términos numéricos de probabilidades, predicciones mínimamente acertadas sobre individuos concretos, y lo que es más, advierten de que así como los márgenes de error de los parámetros aplicables a los grupos pueden reducirse (aumentando los tamaños de las muestras, y alcanzando mejores definiciones de los factores de riesgo), no cabe esperar lo mismo en relación con la tremenda incertidumbre que rodea las estimaciones de probabilidad para los individuos, porque lo que ésta refleja – como, por otra parte, era de esperar – es la extraordinaria e irreductible variabilidad interindividual.¹¹⁹

Estas conclusiones han sido contestadas, desde luego, por otros autores.¹²⁰ Algunas de estas críticas alegan el paralelismo con las empresas aseguradoras: se afirma que es necesario partir de que a los individuos les son aplicables las probabilidades calculadas para los grupos, de la misma manera que se hace cuando se determina el precio de un seguro de vida, pues de otra manera, y si esto no fuera posible, no habría forma de determinar el precio de los seguros.¹²¹ El hecho de que las aseguradoras no quiebren, la gente contrate los seguros y las empresas funcionen, demuestra que la aplicación al individuo de las probabilidades calculadas para el grupo al que pertenece tiene que estar hecha sobre bases correctas. Pero a esta crítica se ha replicado a su vez que la analogía con el funcionamiento de las aseguradoras es falaz: estas empresas hacen negocio si calculan adecuadamente las probabilidades para los grupos, es decir, la proporción de gente que va a morir en un determinado periodo de tiempo dentro de una población con unas características dadas. Pero les es absolutamente

¹¹⁷ COOKE/MICHIE, «Limitations of diagnostic precision and predictive utility in the individual case: a challenge for forensic practice», *Law and Human Behaviour*, (34), 2010, p. 268.

¹¹⁸ Cfr. también otros resultados, similares a los descritos en el texto, en el trabajo de HART/COOKE, *Behav. Sci. Law*, 2013, cit., p. 93.

¹¹⁹ COOKE/MICHIE en MCSHERRY/KEIZER (eds.), *Dangerous people*, 2011, p. 157.

¹²⁰ Por ej. HARRIS/RICE/QUINSEY, «Shall evidence-based risk assessment be abandoned?», *The British Journal of Psychiatry* (192), 2008, p. 154 (en esta misma página está publicada también la réplica de Hart, Michie y Cooke a esa crítica de Harris et al.), o SKEEM/MONAHAN, *Current Directions in Psychological Science*, 2011, p. 40. En su artículo de 2013 (HART/COOKE, *Behav. Sci. Law* 2013), los propios autores enumeran las críticas que sus trabajos han recibido, con referencias.

¹²¹ SKEEM/MONAHAN, *Current Directions in Psychological Science*, 2011, p. 40. Cfr. asimismo QUINSEY et al., *Violent offenders. Appraising and managing risk*, 2ª ed., 2006, p. 207, aduciendo también el ejemplo de los seguros y el de los pronósticos en medicina.

indiferente que sus predicciones resulten ciertas o equivocadas para cada uno de los individuos pertenecientes a esa población, y reconocen la imposibilidad de predecir, respecto de cada uno de sus asegurados, si efectivamente fallecerá en un periodo de 5 años o no; lo que le importa a la compañía aseguradora es que, de las 1.000 personas que tiene aseguradas, fallezca en ese periodo la proporción que tiene calculada, sean cuales sean los individuos concretos que mueran¹²². Mientras que al juez lo que le importa no es que de los 1.000 sujetos que obtendrán en un año el tercer grado reincida la misma proporción que el año pasado, sino si reincidirá el sujeto concreto, con nombre y apellidos, que tiene frente a él y al que tiene que decidir si se lo concede o no.

En realidad incluso desde un punto de vista no especializado en estadística sino meramente lógico cabe considerar cuestionable la atribución sin más a individuos concretos de las probabilidades aplicables a grupos de personas: como señalan COOKE y MICHIE,¹²³ cuando trasladamos sin más una propiedad de un grupo a un individuo incurrimos en la falacia de la división, que consiste en inferir indebidamente que algo es verdadero acerca de una o varias de las partes de un todo, porque es verdadero acerca del compuesto del que forma parte. Utilizando el mismo ejemplo que ponen estos autores: de la misma manera que es falaz concluir que, como en general las personas más inteligentes ganan más dinero que las menos inteligentes, Juan con un CI de 120 seguro que gana más que Pepe con uno de 100, también es falaz concluir que, como en general las personas con puntuaciones más altas en la PCL-R reinciden con mayor probabilidad que las que tienen puntuaciones bajas en dicha escala, Juan, que tiene una puntuación de 30, va a cometer más delitos, más graves, o más pronto, que Pepe, que tiene una de 15.

Algunos autores señalan que negarse a extrapolar al individuo las probabilidades de reincidencia conocidas respecto del grupo al que se asimila es renunciar a utilizar el conocimiento científico que poseemos sobre el comportamiento humano.¹²⁴ Sin embargo, en nuestra opinión hay que distinguir diferentes contextos de actuación y distintos tipos de decisiones que se pueden tomar en función de los pronósticos de peligrosidad basados en conocimientos científicos: como trataremos de mostrar en el epígrafe 8., mientras que puede haber ámbitos en el Derecho penal en los que sea adecuado utilizar el conocimiento que nos proporcionan los juicios de pronóstico, aunque estos tengan todas las limitaciones que hemos estado señalando hasta ahora, hay otras decisiones, y en particular la imposición de medidas de seguridad inocuizadoras adicionales a la pena impuesta por el injusto culpable, cuya adopción no puede basarse en estos juicios.

¹²² COOKE/MICHIE en MCSHERRY/KEIZER (eds.), *Dangerous people*, 2011, p. 154, y COOKE/MICHIE, *Law Hum Behav*, 2010, p. 271.

¹²³ COOKE/MICHIE en MCSHERRY/KEIZER (eds.), *Dangerous people*, 2011, p. 154.

¹²⁴ QUINSEY et al., *Violent offenders*, 2006, p. 207: «Cuando un médico dice, «tiene usted un 80% de probabilidades de sobrevivir», lo que significa es que «De todas las personas como usted que tienen esa enfermedad, el 80% sobrevive». La gente quiere conocer estas estadísticas para poder tomar decisiones informadas sobre qué hacer. En nuestra opinión, es perfectamente apropiado, ético y legal usar los datos grupales de esta manera. En principio, el uso de un instrumento actuarial para valorar el riesgo de reincidencia violenta está basado exactamente en el mismo razonamiento. Intentar tratar a un delincuente como si fuera único en sentido estricto (es decir, no miembro de ningún grupo en absoluto) equivaldría a ignorar toda la investigación relevante.»

5.5 El problema del «riesgo moderado»

Tradicionalmente se ha considerado que los instrumentos actuariales de valoración del riesgo de reincidencia o de violencia plantean otro problema importante: el de qué hacer respecto de los individuos cuya probabilidad de riesgo oscila en torno al 50%. Estas cifras no le son al operador jurídico de mucha ayuda porque en realidad están indicando que es imposible ofrecer respecto de ese individuo una predicción sustancialmente mejor que el mero azar. El problema es importante porque además el número de casos a los que afecta no es pequeño: generalmente cuando se evalúa el riesgo de reincidencia (violenta o no) que presenta un grupo de personas sólo unos pocos alcanzan los valores más extremos en la escala, mientras que la mayoría queda en una zona intermedia en la que al ser el riesgo moderado o intermedio la predicción no arroja resultados concluyentes.

Sin embargo, no me voy a extender en relación con este punto (ni en los criterios normativos que se han ofrecido para orientar las decisiones en estos casos)¹²⁵, porque habida cuenta de lo que hemos expuesto en las líneas que anteceden sobre la arbitrariedad inherente al establecimiento de los puntos de corte, los márgenes de error a los que están sometidos tanto los pronósticos grupales como sobre todo los pronósticos individuales de reincidencia, la baja capacidad predictiva en general de los distintos instrumentos de predicción, y los problemas que plantea el desconocimiento de las tasas de prevalencia de los fenómenos que se quieren predecir, no consideramos que la decisión del operador jurídico se tome en condiciones muy distintas en estos casos de riesgo «moderado» que en los pretendidamente más claros casos de «riesgo extremo».¹²⁶

6. La cuestión en los tribunales

El hecho de que los pronósticos de peligrosidad estén rodeados de este tremendo grado de incertidumbre no ha preocupado hasta ahora mucho a los tribunales, que no consideran que ello sea un obstáculo para la imposición de medidas de seguridad. Hasta donde alcanzo, no se han producido en España pronunciamientos expresos sobre esta cuestión. En nuestro país – donde, recuérdese, desde 1995 sólo es posible imponer medidas de seguridad a sujetos que padecen trastornos mentales y han sido declarados por ello inimputables o semiimputables – en general los tribunales suelen derivar la existencia de peligrosidad a partir únicamente de la gravedad del hecho cometido y de la acreditación de la enfermedad mental del sujeto. En ocasiones las fundamentaciones de las sentencias no permiten averiguar si se ha llegado siquiera a pedir opinión al perito sobre la peligrosidad del acusado, y lo que parece deducirse de ellas es que el tribunal considera necesaria la imposición de medida de seguridad siempre que el perito haya considerado necesario el sometimiento del acusado a un tratamiento para su enfermedad.¹²⁷ Por otro lado, el dato

¹²⁵ Sobre este último extremo véase especialmente FRISCH, *Prognoseentscheidungen im Strafrecht. Zur normativen Relevanz empirischen Wissens und zur Entscheidung bei Nichtwissen*, 1983, y EL MISMO en FRISCH/VOGT (eds.), *Prognoseentscheidungen in der strafrechtlichen Praxis*, 1994, pp. 55 y ss.

¹²⁶ En sentido similar, POLLÄHNE, *Kriminalprognostik*, 2011, p. 241.

¹²⁷ Es muestra de esta forma de proceder, por ej., la SAP Alicante (Sección 2ª) núm. 710/2010 de 19 octubre, que absuelve al acusado de un delito de asesinato en grado de tentativa por haber actuado bajo la influencia de un brote

de que la jurisprudencia reciente del Tribunal Supremo siga sin distinguir con claridad la peligrosidad criminal de la peligrosidad social permite deducir, creo, que ni siquiera termina de estar claro cuál es la clase de peligro que con la medida de seguridad se trata de conjurar.¹²⁸ En consecuencia, si ya la fundamentación de la existencia de peligrosidad es por lo general tan parca, y el objeto del juicio de peligro tan indeterminado, no es de extrañar que ni siquiera se aluda a la cuestión de la fiabilidad de los métodos utilizados para predecir la peligrosidad.

En otros países, sin embargo, sí existen algunas resoluciones interesantes a este respecto. En los Estados Unidos de América la doctrina sobre la admisibilidad como prueba de los informes periciales emitidos por psiquiatras sobre la peligrosidad del sujeto fue establecida en el caso *Barefoot vs Estelle*, de 1983 (463 U.S. 880), y no ha sido modificada hasta la fecha.¹²⁹ Se trataba de un caso en el que el reo había sido declarado culpable de asesinar a un policía, un crimen castigado con la pena capital, y después de dictado este veredicto el jurado tenía que decidir si procedía imponerle o no la pena de muerte. La legislación aplicable exigía que para resolver esta cuestión el jurado se pronunciara sobre si «existe una probabilidad de que el acusado cometa crímenes violentos que constituirían una amenaza continuada para la sociedad».¹³⁰ La Fiscalía presentó a dos peritos psiquiatras que, sin haber examinado personalmente al acusado, afirmaron que existía dicha probabilidad, llegando a decir uno de ellos que la comisión de esos futuros actos de violencia por

psicótico, y tras afirmar que «Informan los Médicos Forenses que el procesado sigue estando aquejado de su trastorno disociativo, trastorno que debe ser debidamente abordado con la correspondiente psicoterapia» añaden sin solución de continuidad que «En tanto subsista el problema psíquico que presenta el acusado concurre la situación de peligrosidad, subsistiendo el riesgo de que puedan repetirse conductas como la aquí enjuiciada, entendiéndose la Sala necesario, para proteger su persona y la de terceros, la medida de seguridad de internamiento en hospital psiquiátrico penitenciario mientras su estado clínico no cambie.» (FD octavo).

¹²⁸ Resulta ilustrativa la STS núm. 65/2011 de 2 febrero, una de las que con más detalle abordan la cuestión de la peligrosidad como fundamento de las medidas de seguridad, en relación con la cual afirma que: «En cuanto a los fines y función de la medida a adoptar, ha de ponderarse, de una parte, la protección del propio acusado, quien mediante el correspondiente tratamiento médico-terapéutico puede controlar los impulsos de su enfermedad mental y acabar haciendo una vida normalizada, objetivo de rehabilitación social que acabará repercutiendo también en beneficio de la comunidad. Y se protege también con la medida a la sociedad, salvaguardándola de los riesgos que genera una persona que ya tiene acreditada una peligrosidad objetivada en el hecho enjuiciado, evitando la reiteración de tales actos en el futuro». Respecto del caso concreto, reproduce algunas de las afirmaciones contenidas en los informes médicos, uno de los cuales concluye que «dada la ausencia de conciencia de la enfermedad y que puede representar un peligro para terceros, se considera necesario su ingreso en una unidad de psiquiatría para control y tratamiento, si procede», y en otro se afirma que «en relación con los hechos que se le atribuyen las bases psicobiológicas de la imputabilidad del informado presentan una afectación completa». En consecuencia, la Sala concluye que «Estos datos permiten inferir un *peligro social* que ha de ser combatido mediante la medida de seguridad que ha decidido la Sala sentenciadora de instancia», que era la de internamiento en centro psiquiátrico, insistiendo a continuación en que en función de la evolución del sujeto se podrá modificar en un futuro la medida según lo que dispone el art. 97 CP (todas las citas son del FD quinto, todas las cursivas añadidas). Llama la atención en esta sentencia, de un lado, que prácticamente el único criterio al que se alude para fundamentar la peligrosidad sea el hecho de que ya se haya cometido un delito. Y, por otro lado, que siendo una de las que con más detalle tratan la cuestión de la peligrosidad, ni siquiera esté claro qué tipo de peligro (por no hablar ya del grado del mismo) trataría de combatir la medida de seguridad, pues se sigue aludiendo a un peligro «social». Unas líneas más arriba la misma sentencia había especificado que el juicio de peligrosidad «tiene por finalidad prever la posibilidad de que la persona concernida cometa nuevos hechos dañinos para la sociedad, según se recoge en el art. 951.2º del C. Penal» (FD quinto, cursiva añadida).

¹²⁹ SLOBOGIN, *Emory Law Journal*, 2006, p. 279.

¹³⁰ Article 37.071 Texas Code on Criminal Procedure: «there is a probability that the defendant would commit criminal acts of violence that would constitute a continuing threat to society».

parte del acusado podía predecirse con un 100% de seguridad. Sus afirmaciones convencieron al jurado, y se decretó la imposición de la pena de muerte. En su defensa, y aparte de otras cuestiones, el reo había alegado dos argumentos para impugnar el testimonio de los psiquiatras. El primero, que la utilización de psiquiatras para efectuar predicciones sobre la conducta futura de una persona debía considerarse inconstitucional: no estarían capacitados, ni como colectivo ni individualmente, para emitir esta clase de afirmaciones, en la medida en que éstas evidenciarían unas tasas de error tan altas que conducirían en muchos casos a sentencias equivocadas, lo que constituiría una violación de la octava y la decimocuarta enmiendas de la Constitución de los EEUU. En segundo lugar, alegó que sería también inconstitucional admitir como prueba afirmaciones de los peritos sobre la peligrosidad hechas en respuesta a preguntas hipotéticas planteadas por el jurado y sin haber examinado al sujeto en cuestión. El Tribunal Supremo, sin embargo, rechazó por mayoría ambas alegaciones y confirmó la sentencia, ejecutándose la pena capital poco después.

En este procedimiento se había presentado como *amicus brief*¹³¹ un informe de la Asociación Americana de Psiquiatría (APA) según el cual era un hecho reconocido dentro de la profesión que las predicciones a largo plazo de futura peligrosidad no son fiables, y que en el mejor de los casos dos de cada tres predicciones a largo plazo sobre la peligrosidad de un individuo están equivocadas¹³². La mayoría del tribunal, sin embargo, consideró que la petición del acusado, de que no debería ser admisible como prueba sobre la peligrosidad de una persona ningún juicio clínico emitido por psiquiatras, era algo así como pedirle al tribunal que «desinventara la rueda»¹³³. Aparte de aludir a resoluciones anteriores en las que se habrían considerado admisibles estos informes periciales como prueba, el Tribunal Supremo arguyó fundamentalmente tres razones para sostener su negativa: (1) que poner en cuestión la admisibilidad de la prueba pericial psiquiátrica sobre la peligrosidad debido a su falta de fiabilidad conllevaría cuestionar también el admitirla en otros contextos en los que se realizan continuamente predicciones sobre la futura conducta del acusado, como por ejemplo el internamiento civil; (2) que la opinión expresada por la APA en su escrito no era compartida por la totalidad de los profesionales en esta disciplina, en la medida en que muchos psiquiatras no tenían inconveniente en emitir su juicio sobre la peligrosidad de los acusados; (3) que lo que el acusado y la APA manifestaban no era que las predicciones de peligrosidad fueran siempre incorrectas, sino sólo que lo eran en la mayor parte de los casos, y que por lo tanto el asunto no era un problema de admisibilidad de la prueba sino del valor probatorio que debería reconocérsele, y que esto debía dejarse al criterio del jurado, el cual, tras un proceso contradictorio en el que el acusado tenía la oportunidad de cuestionar estos informes periciales, sería capaz de formarse una opinión acertada

¹³¹ *Amicus brief* o *amicus curiae brief*: literalmente, «amigo del tribunal», designa en la terminología jurídico penal anglosajona el informe que una persona individual, una organización o un empleado público puede entregar al tribunal, sin ser parte en el juicio, para exponer una opinión sobre un asunto que considera de relevancia. Suelen presentarse en casos de amplia relevancia pública, como los relativos a derechos civiles, y su contenido no es en absoluto vinculante para el tribunal.

¹³² Recuérdese que en aquellas fechas (la sentencia *Barefoot vs Estelle* es de 1983) hacía pocos años que se habían publicado los trabajos hechos sobre los casos *Baxstrom* y *Dixon*, a los que me he referido *supra*.

¹³³ *Barefoot vs Estelle*, 463 U.S. 880, 896: «The suggestion that no psychiatrist's testimony may be presented with respect to a defendant's future dangerousness is somewhat like asking us to disinvent the wheel».

sobre esta cuestión tan discutida, «separando el trigo de la paja».¹³⁴ Por otro lado, el Tribunal no consideró tampoco contrario a la Constitución el hecho de que las opiniones de los psiquiatras se hubieran emitido sin examinar personalmente al acusado, en la medida en que en general se admite como prueba la opinión de expertos sobre cuestiones que no han examinado personalmente, y que el hecho de que en este caso estuviera en juego la vida del acusado no era razón para alterar las reglas generales sobre la prueba.

La sentencia contiene un voto particular disidente redactado por el Juez Blackmun y suscrito por otros dos magistrados más. En esencia, estos magistrados opinaban que el grado de fiabilidad de los juicios clínicos emitidos por psiquiatras sobre la peligrosidad de una persona era demasiado bajo (siendo así que se los puede considerar erróneos en dos de cada tres casos), especialmente para supuestos en los que estaba en juego la aplicación de una pena de muerte. Si la ley en Tejas exige que se pruebe más allá de toda duda razonable que el sujeto continuará cometiendo delitos y constituirá una amenaza continuada para la sociedad, dijeron, es imposible considerar que unos informes periciales que yerran más a menudo de lo que aciertan puedan tener ningún valor probatorio. Por otro lado, este voto particular ponía particular énfasis en rebatir el argumento utilizado por la mayoría, en el sentido de que el jurado estaría en condiciones de formarse una opinión adecuada sobre la tan discutida fiabilidad de los informes gracias al cuestionamiento de los mismos que podía llevar a cabo la defensa en un procedimiento contradictorio. El Juez Blackmun arguyó que el aura de respetabilidad y prestigio que rodea a los médicos hace que no sólo los jueces profesionales se vean inclinados a aceptar sus opiniones sin someterlas a un análisis crítico suficiente, sino que esto ocurre en mucha mayor medida con ciudadanos particulares, que son los que integran el jurado: confrontados con lo que se les presenta como la opinión de un reconocido experto, médicamente fundamentada, sobre un tema del que ellos no saben nada, y siendo así que incluso en el propio campo de la psiquiatría la cuestión es muy discutida, es imposible que los jurados estén en condiciones, como aducía la opinión mayoritaria, de «separar el trigo de la paja».¹³⁵ El voto particular destacaba también que el hecho de que la opinión de los peritos presentados por la fiscalía fuera sometida a contradicción por la defensa no podía ser un remedio suficiente para este problema, en la medida en que un perito de la defensa nunca podría afirmar que el acusado no constituye peligro alguno, ya que esta aseveración es igualmente insostenible en términos científicos. Así, los jurados sólo podrían confrontar la declaración de unos peritos médicos de que el acusado es muy peligroso,

¹³⁴ *Barefoot vs Estelle*, 463 U.S. 880, 899 y nota al pie 7: «Petitioner's entire argument, as well as that of Justice Blackmun's dissent, is founded on the premise that a jury will not be able to separate the wheat from the chaff. We do not share in this low evaluation of the adversary process».

¹³⁵ Quizá merece la pena señalar que en el caso *Barefoot* se daba también la circunstancia de que uno de los peritos presentados por la fiscalía, y en concreto el que llegó a afirmar que estaba seguro, «con un 100% de probabilidades», de que el sujeto constituiría una amenaza constante para la sociedad, fue el Dr. James Grigson, conocido con el sobrenombre de «Dr. Death» por la facilidad con que testificaba en contra del reo en casos de pena capital, lo que llevó a decenas de jurados a dar por probada la peligrosidad del acusado, y que finalmente fue expulsado de la Asociación Americana de Psiquiatría precisamente debido a su comportamiento en estos juicios, en los que realizaba afirmaciones con grados de seguridad que no estarían respaldados por los estándares científicos y omitiendo el deber de estudiar personalmente al sujeto sobre el que emitía su juicio; cfr. AMNISTÍA INTERNACIONAL: *Documento - EE. UU. (Texas). Pena de muerte/preocupación jurídica*, Índice AI: AMR 51/028/2003/s, EXTRA 15/03 (citado según <http://www.amnesty.org/es/library/asset/AMR51/028/2003/es/f903b222-d72a-11dd-b0cc-1f0860013475/amr510282003es.html>, consultado por última vez el 10 de diciembre de 2013).

con las de otros expertos que tendrían que limitarse a decir que esta clase de afirmaciones son poco fiables y están sometidas a grandes márgenes de error, lo cual seguramente haría poco para disminuir el efecto decisivo que las primeras habrían tenido ya sobre ellos.¹³⁶ Por otro lado, no habría que perder de vista que en casos castigados con pena capital se pide al jurado su veredicto sobre la peligrosidad del acusado cuando éste ya ha sido declarado culpable de la comisión de un delito por lo general horrible, lo que seguramente contribuye a incrementar la predisposición de los jurados a aceptar aquellas opiniones de expertos que les confirmen sus previos prejuicios sobre la maldad del sujeto. En definitiva, la pretendida pericia médica sobre la peligrosidad no sería tal, en opinión de los jueces disidentes, porque el conocimiento de los psiquiatras sobre la peligrosidad de un individuo no sería mejor que el que pudiera tener un profano, de modo que presentar sus opiniones revestidas del privilegio de la «sacralidad» de lo científico sería engañoso para los jurados.

Por supuesto, el contexto al que hace referencia esta sentencia no es en absoluto comparable al español, donde ni existe la pena de muerte ni son jurados los que tienen que evaluar la peligrosidad del sujeto,¹³⁷ por lo que podría pensarse que tiene escasa relevancia en relación con el problema que puede suponer la valoración de la peligrosidad en nuestro ordenamiento jurídico. Pero ello no obstante la referencia a esta sentencia me parece pertinente por dos razones. De un lado, porque ilustra la resistencia de los tribunales a reconocer ningún tipo de relevancia jurídica al grado de incertidumbre que rodea a las predicciones de peligrosidad, incluso en supuestos tan graves como aquéllos en los que está en juego la vida del reo, y siendo así que la alternativa en casos como estos – es decir, si se negara la existencia de peligrosidad – no sería en modo alguno la puesta en libertad del reo, sino su reclusión de por vida en un centro penitenciario.¹³⁸ La doctrina contenida en *Barefoot vs Estelle* sigue manteniéndose, además, a pesar de que existen estudios que demuestran que los reos condenados a muerte por los jurados debido a su peligrosidad no evidencian en realidad una conducta más violenta que otros internos ingresados en cárceles de alta seguridad.¹³⁹

¹³⁶ *Barefoot vs Estelle*, 463 U.S. 880, 934 y nota 13: «Given a choice between an expert who says that he can predict with certainty that the defendant, whether confined in prison or free in society, will kill again, and an expert who says merely that no such prediction can be made, members of the jury charged by law with making the prediction surely will be tempted to opt for the expert who claims he can help them in performing their duty, and who predicts dire consequences if the defendant is not put to death». En definitiva, el voto particular concluía que «In a capital case, the specious testimony of a psychiatrist, colored in the eyes of an impressionable jury by the inevitable untouchability of a medical specialist's words, equates with death itself» (463 U.S. 880, 916).

¹³⁷ Incluso en los delitos que en nuestro país están sometidos a la Ley del Jurado, éste se pronuncia sólo sobre los hechos de los que depende la calificación jurídica, no sobre los presupuestos de aplicación de las medidas de seguridad; cfr. arts. 3, 4, 52 y 68 LOTJ.

¹³⁸ Se ha puesto de relieve además que, al contrario de lo que era común en la época en que se dictó la sentencia de *Barefoot* (1983), cuando el tiempo real de permanencia en prisión de los condenados a cadena perpetua estaba alrededor de 12 años, en la actualidad y tras el endurecimiento que ha experimentado el sistema penal en los EEUU en general la alternativa a la pena de muerte es una cadena perpetua sin posibilidad de suspensión condicional (*life without parole*, LWOP), o, en los pocos Estados en que es una cadena perpetua con posibilidad de revisión (*life sentence with parole*, LWP), dicha posibilidad no está permitida por regla general hasta que han pasado al menos 40 años de cumplimiento «a pulso» (cfr. EDENS et al., «Predictions of Future Dangerousness in Capital Murder Trials: Is It Time to «Disinvent the Wheel?»», *Law and Human Behaviour*, (29-1), 2005, p. 58)

¹³⁹ EDENS et al., *Law and Human Behaviour*, 2005, passim.

Y hay otra cuestión relacionada con esta sentencia que me parece interesante resaltar: los jueces que suscribieron el voto particular opinaban que los juicios *clínicos* emitidos por psiquiatras sobre la peligrosidad del reo (al menos en los casos de pena capital) no debían ser admitidos como prueba no sólo por su escasísima fiabilidad, sino sobre todo por su efecto tendencioso, en la medida en que favorecerían la adopción de decisiones por los jurados basadas en prejuicios. Sin embargo, expresaron una opinión más favorable en relación con otro tipo de pruebas, como el testimonio de personas no profesionales (*lay testimony*), o la referencia a correlaciones estadísticamente demostradas entre factores de riesgo presentes en el reo y futuro comportamiento violento, para lo cual no se requieren conocimientos especiales en psiquiatría.¹⁴⁰ SLOBOGIN ha expuesto recientemente una tesis similar a ésta. Este autor parte de la base (discutible en mi opinión, por todas las razones que he apuntado en el epígrafe 5., pero no es eso lo que ahora me interesa reiterar) de que la posibilidad de predicción acertada de la violencia es en la actualidad satisfactoria o, al menos, mejor que el mero azar, lo que le parece suficiente para que pueda ser admisible cuando sea necesario estimar la peligrosidad en un juicio. Pero, por razones similares a las que expusiera en su día el juez Blackmun en el voto particular de *Barefoot vs Estelle*, entiende que el juicio clínico emitido por un médico psiquiatra (o por un psicólogo) conlleva un peligro muy notable de condicionar indebidamente la decisión de los jurados. En opinión de SLOBOGIN, este peligro es tan grande que excede los beneficios que de dicha prueba pudieran extraerse en orden a proporcionar información fiable sobre la peligrosidad del sujeto, y considera por ello que no debe permitirse a la acusación aportarla como prueba de la peligrosidad del reo. Sí deberían ser admisibles, por el contrario, los métodos actuariales o de juicio clínico estructurado, pues estos instrumentos aportan una estimación de la peligrosidad que se ha obtenido de manera objetiva, aplicando un protocolo estandarizado cuyos márgenes de error se conocen, y cuya credibilidad no aumenta ni disminuye por la autoridad que como médico (o como experto en general) pueda tener quien la expone. En opinión de SLOBOGIN, sólo si es el propio acusado quien considera que le conviene presentar juicios clínicos para acreditar una baja peligrosidad, entonces estaría abierta también la puerta a que la acusación utilizara ese mismo método para tratar de probar lo contrario.¹⁴¹

Estas consideraciones relacionadas con la prueba merecen ser completadas con algunas otras, ya no derivadas estrictamente de la naturaleza incierta de los pronósticos de peligrosidad, sino en general de la situación en que se encuentra la defensa del acusado cuando trata de rebatir pruebas presentadas en su contra sobre su peligrosidad. La prueba pericial psiquiátrica tiene una trascendencia diferente cuando lo que se trata de probar es la peligrosidad, que en la otra situación en la que suele hacerse uso de ella, esto es, cuando la aporta la defensa para reclamar la aplicación de una eximente por enajenación mental. En este último caso el acusado se encuentra desde el principio en una posición de desventaja, porque para solicitar dicha eximente tiene que haber admitido,

¹⁴⁰ 463 U.S. 880, 922 y nota 4: «Statistical prediction is clearly more reliable than clinical prediction, [...] and prediction based on statistics alone may be done by anyone».

¹⁴¹ SLOBOGIN, *Emory Law Journal*, 2006, *passim*, y en especial pp. 296 y ss., 311 y ss. Lo cierto es que, sea como fuere, en los últimos lustros parece que en los casos de pena capital cuando la decisión de imponer la pena de muerte depende de la constatación de la peligrosidad del reo (cosa que ocurre en muchos estados de los EEUU) está aumentando la tendencia a apoyar dichos juicios en probabilidades calculadas con ayuda de instrumentos actuariales de predicción, antes que en juicios clínicos (HARCOURT, *Against prediction*, 2007, pp. 100 a 103).

aunque sea sólo de manera implícita o condicional, que ha cometido el hecho, y además el punto de partida es la presunción general de que las personas están mentalmente sanas a menos que consigan demostrar lo contrario. Sea cual sea la naturaleza de la prueba que el acusado pretenda presentar sobre su situación psíquica, la acusación no tendrá por lo general muchos problemas para presentar pruebas en contrario al menos igual de convincentes a priori. Pero en las predicciones de peligrosidad la situación es distinta, y cuando es la acusación la que introduce periciales psiquiátricas para demostrar la existencia de peligro de comisión de futuros delitos la defensa lo tiene mucho más difícil para contrarrestarlas. En materia de culpabilidad, además, el acusado está protegido en todo momento por la presunción de inocencia y los estándares de prueba que han de alcanzarse para desvirtuarla son elevados; en materia de peligrosidad, sin embargo, no existe ninguna presunción parecida, ni legalmente reconocida ni que opere en la práctica; el estándar que deben alcanzar las pruebas para demostrar esta circunstancia no está determinado (y la práctica demuestra que no llega ni de lejos al necesario para demostrar la culpabilidad), y, en definitiva, el acusado se ve en la tesitura de tener que demostrar que no hará algo en el futuro – cosa mucho más complicada que demostrar que no hizo algo en el pasado –, y además en una situación en la que, en la medida en que ya se le ha considerado culpable de algún delito, acaba de demostrar que al menos una vez sí ha sido ya capaz de hacerlo.¹⁴²

Sin que sea posible abordar ahora la cantidad de problemas de orden procesal que este panorama deja planteados, únicamente he querido evidenciar que la incertidumbre de los pronósticos es sólo uno de los problemas que presenta la determinación de la peligrosidad, al cual hay que añadir, al menos, todos estos otros, y sin embargo sigue siendo una cuestión a la que a los tribunales hasta ahora (ni en nuestro país ni en Estados Unidos) han dado mucha importancia.

La doctrina denuncia que algo parecido podría decirse de Alemania. En las sentencias recientes en las que el Tribunal Constitucional Alemán¹⁴³ ha tenido ocasión de pronunciarse sobre la relevancia jurídica que puede tener la incertidumbre de los pronósticos de peligrosidad (sentencias que se han dictado fundamentalmente en relación con la custodia de seguridad o *Sicherungsverwahrung*) el tribunal no ha entrado con detalle en el problema. El tribunal parte de la base de que el internamiento prolongado que supone la custodia de seguridad no vulnera la dignidad de la persona – garantizada en el art. 1 de la Constitución Federal Alemana – siempre que resulte imprescindible debido al peligro continuado que supone el sujeto. La comprobación de dicho peligro deviene, pues, esencial para la legitimidad de la medida, pues la grave limitación del derecho fundamental a la libertad que supone la custodia de seguridad sólo puede asumirse mientras sirva para proteger intereses esenciales de terceras personas, y siempre que se mantenga dentro de las exigencias del principio de proporcionalidad.¹⁴⁴ Pero, en opinión del Tribunal, «las incertidumbres existentes en

¹⁴² SLOBOGIN, *Emory Law Journal*, 2006, pp. 312 y s.

¹⁴³ Sobre el tratamiento de la cuestión por el Tribunal Supremo Federal (*Bundesgerichtshof*, BGH), cfr., ALBRECHT en FREUND et al. (eds.), *FS-Frisch*, 2013, pp. 1074 y s.

¹⁴⁴ Cfr. sentencia del TC alemán de 5 de febrero de 2004 (BVerfG, 2 BvR 2029/01), números marginales 74 y 96. Esta sentencia consideró constitucional que la custodia de seguridad no tuviera límite máximo de duración cuando fuera impuesta por primera vez a un sujeto, así como aplicar la custodia de seguridad indeterminada a sujetos a los que se

relación con el pronóstico, que es la base para el internamiento, no hacen desaparecer ni la adecuación ni la necesidad de limitar la libertad [con la imposición de la custodia de seguridad]»,¹⁴⁵ y «las decisiones que se toman sobre la base de pronósticos (*Prognoseentscheidungen*) siempre implican el riesgo de que el pronóstico sea equivocado, pero siguen siendo imprescindibles en Derecho. El pronóstico es y seguirá siendo indispensable como base para cualquier protección frente a peligros, aun cuando en el caso concreto pueda ser deficiente».¹⁴⁶ Seguramente es la sentencia de 5 de febrero de 2004 la que se ha ocupado más detenidamente del problema. El Tribunal adujo que había recabado las opiniones de expertos en la materia (Profs. Marneros y Nedopil), según los cuales el conocimiento acerca de los factores de riesgo (de comisión de nuevos delitos) habría mejorado notablemente en los últimos años, lo que permitiría efectuar pronósticos relativamente buenos y fiables sobre una parte de los delincuentes, hasta el punto de que habría algunos delincuentes en los que confluían tal cantidad de factores de riesgo que el peligro podría pronosticarse «con seguridad».¹⁴⁷ El Tribunal concluyó que el pronóstico constituye una base apropiada para la decisión precisamente en los casos, poco frecuentes, de peligro muy elevado, a los que se refiere el § 67d apartado 3 del CP alemán (el cual, en la regulación cuya constitucionalidad se cuestionaba – y que ya no es la vigente en la actualidad – permitía prolongar indefinidamente la duración de la custodia de seguridad).¹⁴⁸

La doctrina penal alemana se ha mostrado crítica con estas afirmaciones, considerando que las pocas consideraciones sobre la fiabilidad y la incertidumbre de los pronósticos de peligrosidad hechas en estas sentencias fueron decepcionantes,¹⁴⁹ y que en realidad escamotearon el problema principal, pues dieron por supuesta en lugar de entrar a analizar la cuestión central de la que depende la legitimidad de las medidas de seguridad, esto es, si es posible o no efectuar pronósticos con el grado suficiente de certeza, dados los numerosos problemas que en relación con su fiabilidad ha puesto de manifiesto la doctrina, tanto médica como jurídica.¹⁵⁰ La verdad es que no parece ciertamente que la conclusión del tribunal se extraiga sin más ni siquiera de los trabajos de los expertos en cuya opinión basó la sentencia sus optimistas afirmaciones sobre la viabilidad de los pronósticos de peligrosidad. En la sentencia de 5 de febrero de 2004, por ejemplo, se cita expresamente un artículo de NEDOPIL¹⁵¹

había impuesto esta medida bajo la legislación anterior, en la que existía para estos casos un límite máximo de 10 años.

¹⁴⁵ Sentencia del TC alemán de 4 de mayo de 2011 (BVerfG, 2 BvR 2365/09), número marginal 99. Esta sentencia declaró la inconstitucionalidad de la regulación de la custodia de seguridad tras el pronunciamiento del TEDH de 17 de diciembre de 2009 (Case M. v. Germany), que consideró contraria a los artículos 5 y 7 del CEDH la imposición a posteriori de la custodia de seguridad a un ciudadano alemán.

¹⁴⁶ Sentencia del TC alemán de 5 de febrero de 2004 (BVerfG, 2 BvR 2029/01), número marginal 101.

¹⁴⁷ «[E]in bestimmter und bestimmbarer Anteil der Probanden versammle eine derartige Häufung von Risikofaktoren auf sich, dass eine Gefahr sicher prognostiziert werden könne» (BVerfG, 2 BvR 2029/01, de 5.2.2004, número marginal 102).

¹⁴⁸ BVerfG, 2 BvR 2029/01, de 5.2.2004, número marginal 102.

¹⁴⁹ «Enttäuschend», en palabras de KINZIG, *Die Legalbewährung gefährlicher Rückfalltäter*, 2010, p. 43

¹⁵⁰ KINZIG, *Die Legalbewährung gefährlicher Rückfalltäter*, 2010, p. 43, así como STRENG, *Juristenzeitung*, (17), 2011, p. 828, o KREUZER/BARTSCH, «Anmerkung I» a BVerfG, Urt. v. 04.05.2011, *Strafverteidiger*, (8), 2011, pp. 479 y s. POLLÄHNE considera la actitud del TC Alemán un resignado encogimiento de hombros, que evidenciaría un preocupante pragmatismo ante una situación en la que el Tribunal no vislumbra alternativas viables al pronóstico de peligrosidad (§ 61 n.m. 54 en *Nomos Kommentar*, 4ª ed., 2013).

¹⁵¹ NEDOPIL, *NStZ*, 2002.

en el que efectivamente se afirma, como sostiene el Tribunal, que el conocimiento sobre los factores de riesgo de reincidencia ha mejorado notablemente en los últimos años, pero el autor añade que también son cada vez más claras las fronteras de lo que es posible predecir, e insiste en las limitaciones que plantean las tasas de prevalencia bajas o desconocidas, en la imposibilidad de realizar predicciones a largo plazo, en la insuficiencia de los conocimientos empíricos existentes, y en las dificultades inherentes a la predicción de fenómenos tan complejos como el comportamiento humano.¹⁵² No deja de llamar la atención, por otro lado, que el Tribunal Constitucional cite este trabajo en concreto en una sentencia que concluye afirmando la constitucionalidad de la custodia de seguridad impuesta a posteriori, cuando precisamente en este artículo NEDOPIL se muestra muy crítico con dicha posibilidad desde el punto de vista de los pronósticos psiquiátricos sobre la peligrosidad del sujeto: en efecto, el autor advierte expresamente de que si el reo no ha tenido la posibilidad de disfrutar de permisos o de salidas durante la ejecución de la pena, la base para efectuar un pronóstico favorable tras el cumplimiento de la misma será mínima, y que en consecuencia los tribunales podrán descargar sobre el informe negativo de los psiquiatras la responsabilidad de decretar a continuación el internamiento en custodia de seguridad, cuando puede haber sido la propia administración de justicia – al denegar las salidas o la libertad condicional – la que haya creado las condiciones que hagan imposible el pronóstico favorable, con lo que el informe pericial sobre la peligrosidad podría convertirse en último término más bien en un instrumento para apaciguar conciencias, antes que ser un medio para comprobar empíricamente las probabilidades de reincidencia.¹⁵³

Por otra parte, la doctrina también ha señalado que si bien podrían compartirse las afirmaciones del Tribunal Constitucional en el sentido de que la custodia de seguridad es efectivamente un medio idóneo e incluso necesario para hacer frente a los peligros que emanan del autor,¹⁵⁴ quedó sin analizar en estas sentencias la tercera de las exigencias del principio de prohibición de exceso, esto es, la proporcionalidad en sentido estricto. Y como afirman KREUZER y BARTSCH, resulta como mínimo cuestionable desde este último canon de enjuiciamiento el que la protección de la sociedad exija privar de libertad a través de la custodia de seguridad a un número no pequeño de sujetos que en realidad no son peligrosos.¹⁵⁵

¹⁵² NEDOPIL, *NStZ*, 2002, pp. 346 a 348.

¹⁵³ NEDOPIL, *NStZ*, 2002, p. 349.

¹⁵⁴ El juicio de necesidad lo entendió el TC alemán (tal y como viene haciéndolo también el TC español) en el sentido de que no exista otro medio igualmente eficaz de evitar el daño, lo que evidentemente hace que resulte muy difícil negar la necesidad de la custodia de seguridad: «A pesar de la posible existencia de otras alternativas para la protección [de los bienes para los que el autor supone un peligro] no hay objeciones desde la perspectiva constitucional a la supresión del límite máximo de la custodia de seguridad. La privación de libertad a través de la custodia de seguridad indeterminada ofrece en el caso concreto una protección evidentemente mayor que cualquier otra medida de tratamiento imaginable con una vigilancia más relajada» (BVerfG, 2 BvR 2029/01, de 5.2.2004, número marginal 103).

¹⁵⁵ KREUZER/BARTSCH, *Strafverteidiger*, (8), 2011, p. 480.

7. Conclusiones

De todo lo dicho hasta ahora en los epígrafes 4, 5 y 6 pueden extraerse resumidamente las siguientes conclusiones:

1. Los diversos métodos de predicción de la peligrosidad, incluso en sus formulaciones más novedosas y sin perjuicio del notable avance que se ha conseguido en los últimos años en relación con el mejor conocimiento de factores de riesgo y factores protectores, siguen evidenciando una capacidad predictiva muy baja cuando de lo que se trata es de predecir la probabilidad de comisión de futuros delitos por parte de un sujeto en concreto. A pesar de que algunos instrumentos evidencien niveles de sensibilidad y especificidad altos para determinados grupos de personas, y de que los valores predictivos negativos sean por lo general también elevados, el valor predictivo positivo, que es el parámetro que más interesa como base para el juicio de pronóstico requerido para la imposición de las medidas de seguridad (porque es el que nos dice con qué probabilidad, cuando consideramos peligroso a un sujeto, su comportamiento posterior confirmará la predicción) sigue siendo muchas veces inferior al mero azar, es decir, inferior a 0.5, lo cual significa que nos equivocamos más veces de las que acertamos en nuestras predicciones de peligrosidad, llegando en ocasiones las tasas de falsos positivos al 70 o al 80%. Además, estos valores oscilan entre márgenes de error bastante amplios, y si se pretende trasladar estos valores a la predicción del comportamiento de un individuo concreto seguramente tales márgenes de error son todavía mayores.

2. Las limitaciones estructuralmente inherentes a la posibilidad de predecir el comportamiento futuro de los individuos (limitaciones que son de tipo psicológico, estadístico y político-criminal), unidas al hecho de que en los últimos cuarenta años en realidad se haya avanzado bastante poco por lo que respecta a la capacidad predictiva – positiva – de los diversos instrumentos, hacen aparecer como poco probable que en un futuro previsible este estado de cosas vaya a mejorar de manera sustancial.¹⁵⁶

3. Los datos empíricos disponibles evidencian una notable y sistemática sobreestimación de la peligrosidad por parte de los peritos y de los tribunales.

4. Los tribunales hasta ahora han ignorado en gran medida este estado de cosas, o al menos no han considerado que de él deban derivarse consecuencias para la regulación o la interpretación de las disposiciones que permiten la imposición de medidas restrictivas de derechos para los sujetos a los que se considera peligrosos.

¹⁵⁶ Así lo afirman diversos autores, por ej. YANG/WONG/COID, *Psychological Bulletin*, 2010, p. 759: «After almost five decades of developing risk prediction tools, the evidence increasingly suggests that the ceiling of predictive efficacy may have been reached with the available technology. Other approaches such as tree modeling (Steadman et al., 2000) and Neural Networks (Price et al., 2000) require further exploration, but it is unlikely that a very high level of predictive accuracy is achievable because of theoretical constraints.» En parecido sentido, ALBRECHT en FREUND et al. (eds.), *FS-Frisch*, 2013, pp. 1072 y 1076, afirmando en este último lugar que el estado actual de la investigación apunta a que en un futuro cercano no van a poder resolverse de manera satisfactoria ni el problema de los falsos positivos ni el de los falsos negativos.

8. Consecuencias para la dogmática de las medidas de seguridad y para algunas penas

A mi juicio, de todo lo expuesto hasta ahora derivan consecuencias que afectan tanto a la legitimidad como a los límites de las medidas de seguridad. Debe tenerse en cuenta que a continuación me voy a referir especialmente a medidas imponibles a delincuentes imputables considerados peligrosos para que se ejecuten tras el cumplimiento de la pena, es decir, medidas impuestas más allá de lo que permitiría el marco del injusto culpable. En consecuencia, y como habrá ocasión también de ir advirtiendo, las conclusiones podrían ser distintas en el caso de otras consecuencias jurídicas diferentes, aunque estén basadas también sobre juicios de pronóstico de peligrosidad (en especial, medidas de seguridad a imponer a semiimputables según el sistema vicarial, o decisiones que afecten a la flexibilización del régimen de ejecución de una pena impuesta a un individuo culpable). Por último, extraeré también alguna conclusión respecto de la constitucionalidad de la pena de prisión permanente revisable.

Casi resulta ocioso advertir también, por otro lado, que debido a elementales razones de espacio las consideraciones que a continuación se efectuarán necesariamente han de limitarse a exponer algunas ideas generales básicas, todas ellas necesitadas de mayor concreción y desarrollo ya que afectan a problemas complejos, desarrollo en el que sin embargo ya no es posible profundizar en los limitados márgenes de este trabajo.¹⁵⁷

8.1 Fundamento de las medidas de seguridad

Por lo que respecta al fundamento de las medidas de seguridad, éste radica según opinión probablemente mayoritaria en la actualidad tanto en España como en Alemania en el principio de ponderación de intereses y en la salvaguarda del interés preponderante, idea que fue formulada inicialmente por NOWAKOWSKI: por un lado estaría el interés del individuo en no ser limitado en sus derechos fundamentales más allá de lo que permite la pena basada en la culpabilidad, y por otro lado estaría el interés de la sociedad en protegerse frente a la comisión de delitos futuros por parte de ese sujeto.¹⁵⁸ Adoptar esta perspectiva para la fundamentación de las medidas de seguridad implica

¹⁵⁷ Por ejemplo, no puede abordarse la valoración que, a la luz de todo lo expuesto en este trabajo, merece la interesante aportación de FRISCH a este debate (desarrollada sobre todo en su monografía *Prognoseentscheidungen im Strafrecht*, 1983, y expuesta también en otros trabajos como en FRISCH/VOGT (eds.), *Prognoseentscheidungen in der strafrechtlichen Praxis*, 1994, pp. 55 y ss.).

¹⁵⁸ Cfr., entre otros, URRUELA MORA, *Las medidas de seguridad y reinserción social*, 2009, pp. 11 y ss.; SANZ MORÁN, *Las medidas de corrección y de seguridad*, 2003, pp. 81 y ss.; KAISER, *Befinden sich die Kriminalrechtlichen Maßregeln in der Krise?*, 1990, pp. 12, y pp. 48 y s.; VAN GEMMEREN, *MK*, 2ª ed., 2012, § 61 nm. 2; SCHÖCH, *LK*, 12ª ed., 2008, § 61, n.m. 38 y s. La propuesta original de Nowakowski se encuentra en NOWAKOWSKI, «Zur Rechtsstaatlichkeit der vorbeugenden Maßnahmen» en WELZEL et al. (eds.), *Festschrift für Hellmuth von Weber zum 70. Geburtstag*, 1963, pp. 103 y ss. Hay autores que, sin cuestionar completamente la idea de salvaguarda del interés preponderante como justificación de las medidas de seguridad, ponen el acento también en otros aspectos, como por ejemplo la existencia de un deber del Estado de proteger a sus ciudadanos frente a peligros, cfr. en relación con este último aspecto FRISCH, «Die Maßregeln der Besserung und Sicherung im strafrechtlichen Rechtsfolgensystem. Straftheoretische Einordnung, inhaltliche Ausgestaltung und rechtsstaatliche Anforderungen», *ZStW*, (102), 1990, pp. 367 y ss. Sobre las distintas propuestas en relación con el fundamento de las medidas de seguridad cfr. también, ampliamente, ZIFFER, *Medidas de seguridad*, 2008, pp. 71 y ss.

que tiene que ser posible, de alguna forma racional, comparar el daño – actual y perfectamente cuantificable – que sufren los bienes jurídicos del sujeto con la imposición de la medida de seguridad, con el daño – eventual e indeterminado – que sufriría la sociedad con los delitos que el sujeto cometerá, probablemente, en el futuro. Pero, ¿realmente es posible hacer esa comparación cuando el segundo término de la misma es imposible de determinar de forma mínimamente fiable?¹⁵⁹

Es decir, para resolver sobre la procedencia de la imposición de una medida de seguridad desde el punto de vista de la salvaguarda del interés preponderante necesitaríamos saber qué delitos, cuántos, y en qué espacio de tiempo, va a cometer el sujeto, y con qué grado de probabilidad; esto es lo que nos podría ofrecer criterios para decidir hasta qué punto debe prevalecer el derecho fundamental del sujeto a la libertad, y a partir de qué momento, y en qué condiciones, podría verse limitado en atención a los intereses de protección de la sociedad.¹⁶⁰ Pero esto es precisamente lo que no estamos en condiciones de averiguar, habida cuenta de las limitaciones de los pronósticos de seguridad que han quedado expuestas en las páginas anteriores.

A pesar de que en ocasiones se ha trazado expresamente un paralelismo entre el fundamento de las medidas de seguridad y el de las causas de justificación, cuya razón de ser obedece según opinión mayoritaria también a esta idea de la salvaguarda del interés preponderante, hay que advertir que en relación con las medidas de seguridad la comparación entre los intereses en conflicto no puede hacerse siguiendo las reglas a las que estamos acostumbrados en materia de justificación. Ante todo porque la situación que legitima la agresión a bienes jurídicos de otras personas en cualquiera de las causas de justificación es una situación de conflicto *inminente* entre los bienes, que permite efectuar una comparación exacta entre la magnitud del daño que se provoca con la conducta justificada y el daño que gracias a ella se evita, mientras que esa inminencia está

¹⁵⁹ Advierte la dificultad que esto plantea STRENG, *Juristenzeitung*, (17), 2011, p. 829: «Pero aquí no es posible efectuar una simple ponderación entre ambos intereses [el de la sociedad y el del autor] según el esquema del «interés preponderante», porque no puede perderse de vista que no son posibles pronósticos de peligrosidad realmente fiables», y añade que esta cuestión casi nunca se aborda de forma sistemática. También ZIFFER reconoce que legitimar las medidas de seguridad sobre la base de la idea de prevalencia del interés preponderante convierte en decisiva la cuestión de la incertidumbre de los pronósticos de peligrosidad, aunque termina considerando soluble el problema acudiendo a la normativización del presupuesto de aplicación de las medidas de seguridad según la idea de la distribución de riesgos que ha defendido FRISCH (ZIFFER, «Begriff der Strafe und Sicherungsverwahrung» en FREUND/MURMANN/BLOY/PERRON (eds.), *Grundlagen und Dogmatik des gesamten Strafrechtssystems. Festschrift für Wolfgang Frisch zum 70. Geburtstag*, 2013, pp. 1084 y ss., donde se refiere a la opinión de FRISCH en FRISCH/VOGT (eds.), *Prognoseentscheidungen in der strafrechtlichen Praxis*, 1994, pp. 73 y ss.)

¹⁶⁰ SÁNCHEZ LÁZARO ha hecho un intento de concretar en términos numéricos la difusa ponderación entre bienes que supone la decisión de imponer una medida de seguridad de internamiento psiquiátrico, utilizando como ejemplo un caso real extraído de una sentencia de nuestro Tribunal Supremo (SÁNCHEZ LÁZARO, «Deconstruyendo las medidas de seguridad», *InDret*, (2), 2010, pp. 13 y ss. y también 17 y ss.). Sin embargo, a pesar del innegable mérito que merece el intento, y de que comparto algunos de los criterios que utiliza para asignar valores numéricos a los intereses en cuestión (en especial la relativización del peso que merece la amenaza a los bienes jurídicos que se verían afectados por la conducta del sujeto, dado que se trata de meras probabilidades y no de lesiones reales, cfr. *ibidem*, p. 12) considero que dicha conversión en cifras no puede dejar de ser en gran medida arbitraria, dada la ausencia de información empírica sobre, entre otras cosas, las probabilidades de reincidencia (no consta que en el caso en cuestión se hubiera emitido un pronóstico de peligrosidad que concretara este parámetro) y también sobre los delitos que volvería a cometer el sujeto (SÁNCHEZ LÁZARO asume que serían delitos contra la vida e integridad física porque esos fueron los comportamientos que dieron lugar a los hechos por los que se valora la imposición de la medida, pero ya sabemos que un homicida no necesariamente volverá a cometer, en caso de reincidir, ese mismo delito, pues las tasas de reincidencia en delitos contra la vida son muy bajas).

ausente cuando se toma la decisión sobre la imposición de medidas de seguridad, en la que se pondera un peligro de comisión de delitos futuros que en el mejor de los casos se calcula a meses, cuando no a años, vista. Por otro lado, me parece que la ponderación de intereses sólo puede efectuarse de manera mínimamente clara y comprensible cuando se ponderan magnitudes que tienen una naturaleza similar o, al menos, conocida: por ejemplo, la ponderación en las causas de justificación tiene lugar entre bienes jurídicos concretos y determinables, que muchas veces constituyen además derechos fundamentales o al menos derechos subjetivos (vida - libertad sexual; integridad corporal - propiedad, etc.), e incluso para quienes admiten la posibilidad de actuar en estado de necesidad para defender bienes colectivos, sigue tratándose de conflictos entre bienes determinados y conocidos (el orden público, por ejemplo), que van a ser objeto de una clase concreta de ataque y de manera inmediata. Nada de esto ocurre, sin embargo, en las medidas de seguridad: frente a la privación efectiva de la libertad (si la medida es de internamiento) o la restricción de la misma si se trata de medidas de otra naturaleza, encontramos como segundo término para la ponderación un difuso e inconcreto interés de la sociedad en protegerse frente a la comisión de futuros y desconocidos delitos.

El propio NOWAKOWSKI era consciente de que la incertidumbre de los pronósticos de peligrosidad podía ser una objeción importante a la viabilidad de su propuesta, pero consideró que era una crítica superable porque en su opinión la biología criminal y la estadística criminal evidenciaban que respecto de ciertos tipos de casos al menos sí era posible efectuar pronósticos «relativamente seguros».¹⁶¹ En mi opinión este optimismo es, sin embargo, infundado, porque no se corresponde con los datos reales que aporta a día de hoy la criminología, tal y como he intentado demostrar en los epígrafes 4 y 5 de este trabajo.

Recientemente GRACIA MARTÍN ha sostenido una opinión parecida a la que estoy defendiendo aquí, si bien centrada en la custodia de seguridad como medida de internamiento a ejecutar una vez cumplida la pena. Este autor afirma que tal medida no resulta legítima según el principio de ponderación de intereses debido en buena parte a la incertidumbre del pronóstico de peligrosidad: en efecto, GRACIA parte de que la prognosis de peligrosidad para sujetos imputables para largos periodos de tiempo es «insostenible en el plano científico de la experiencia», y concluye que «El hecho de que el mal que se trata de evitar [los delitos que puede cometer en el futuro el sujeto puesto en libertad] sea uno de producción incierta respecto del cual no puede formularse más que un pronóstico de peligro más o menos fundado, mientras que el mal que se produce con la aplicación de la medida de custodia de seguridad es un menoscabo prolongado e indefinido del derecho fundamental a la libertad real, efectivo y actual, además de irreparable, es demostrativo por sí solo de la considerable mayor gravedad del mal causado con respecto al que se trata de evitar».¹⁶²

¹⁶¹ NOWAKOWSKI, *FS-von Weber*, 1963, pp. 113 a 116. En la p. 114 afirma que aunque puede ser efectivamente muy difícil predecir la conducta futura en general de un sujeto cualquiera, sin embargo sí son «relativamente seguros muchos pronósticos que derivan de específicas trayectorias vitales y estructuras de personalidad de sujetos criminales. La experiencia judicial diaria lo demuestra. No se le exige demasiado al juez si se le pide que en estos casos efectúe un juicio de probabilidades».

¹⁶² GRACIA MARTÍN, «Sobre la legitimidad de medidas de seguridad contra delinquentes imputables peligrosos en el Estado de Derecho» en GARCÍA VALDÉS et al., *Estudios Penales en Homenaje a Enrique Gimbernat*, t. 1, 2008, pp. 997 (la primera cita) y 998 y s. (la segunda). GRACIA concluye que por ello la custodia de seguridad vulneraría el principio de proporcionalidad (ibídem), pero en el contexto de la ponderación de intereses; por ello he preferido citar su opinión en este apartado del texto y no en el siguiente, donde me referiré a la proporcionalidad de las medidas de seguridad en sentido estricto (clase y duración de la medida en función de la gravedad del hecho cometido y de la peligrosidad), y donde GRACIA no comparte la opinión que yo defenderé.

Es más, lo que el estado actual de la investigación sobre instrumentos de predicción del riesgo sí nos indica es que en este juicio de ponderación hay que introducir un tercer elemento, porque no sólo están implicados la libertad del sujeto peligroso y el interés de la sociedad en protegerse frente a él, sino también la libertad de otros sujetos, no peligrosos, que se verán afectados por la aplicación de medidas de seguridad como consecuencia de la inevitable presencia de falsos positivos. Retomemos el ejemplo que utilicé en el epígrafe 5.2, y supongamos que disponemos de un instrumento de predicción de reincidencia en la comisión de delitos sexuales cuya sensibilidad es del 70%, y que sabemos que la tasa de reincidencia sexual en este grupo de autores es de alrededor del 20%. Supongamos también que respecto del individuo concreto cuya peligrosidad estamos evaluando este instrumento nos hubiera dado una puntuación indicativa de un alto riesgo de reincidencia. Pues bien, la probabilidad de que se trate de un falso positivo es del 63%.¹⁶³ En este caso, y sea cual fuere el valor que quisiéramos darle al interés de la sociedad en protegerse frente a la comisión de nuevos delitos sexuales graves, en el otro platillo de la balanza sabemos que 1,7 personas más (de media) resultarían privadas de libertad indebidamente por cada sujeto peligroso que mantuviéramos encerrado siguiendo este método de predicción de la peligrosidad. Los números son aún más desfavorables si la predicción ha de hacerse en relación con el peligro de comisión de delitos como los de homicidio o asesinato, en los cuales al ser la tasa de prevalencia aún menor la proporción de falsos positivos, incluso hecha con instrumentos de predicción con valores altos de sensibilidad, es todavía mayor. De ahí que tampoco consigan eludir esta objeción los autores que abogan por introducir medidas de seguridad adicionales a la pena sólo de manera muy excepcional, para casos extraordinariamente graves en los que exista una probabilidad de reiteración de delitos muy graves por un autor extraordinariamente peligroso¹⁶⁴, porque precisamente cuanto más grave es el delito más baja es la *base rate*, y por tanto mayores tasas de falsos positivos hay que asumir en la predicción de la peligrosidad.¹⁶⁵

Personalmente, veo muy difícil que como resultado de un juicio de ponderación hecho en estos términos pudiera prevalecer el interés de la sociedad a su propia protección. ¿Puede en un Estado de Derecho que proclama la libertad como valor superior del ordenamiento jurídico admitirse que sea de mayor valor el interés de la sociedad en protegerse frente a la eventual comisión de indeterminados delitos graves (no sabemos cuántos serían, ni cuáles en concreto, ni cuándo se cometerían, ni si en realidad llegarían finalmente a cometerse), que el derecho a la libertad no sólo del autor peligroso sino también el de otros dos sujetos no peligrosos? La cuestión puede ser opinable,¹⁶⁶ pero a mi juicio la respuesta ha de ser negativa, al menos si se trata de medidas de seguridad privativas de libertad para delincuentes imputables tras el cumplimiento de la pena.¹⁶⁷

¹⁶³ Los datos de sensibilidad y tasa de prevalencia podrían perfectamente corresponderse con los reales, según las cifras ofrecidas *supra* en los epígrafes 5.2 y 5.3. La cifra del 63% de falsos positivos deriva de que, tal y como se calculó *supra* cuando se expuso este mismo ejemplo, el instrumento clasificaría en total como peligrosos a 38 sujetos, de los cuales sólo 14 reincidirán en realidad, quedando 24 sujetos como falsos positivos.

¹⁶⁴ Por ej. ZUGALDÍA ESPINAR, *RDPC*, 2009, p. 205

¹⁶⁵ Y en todo caso recuérdese que las tasas de falsos positivos *empíricamente comprobadas* por los diversos estudios citados en el epígrafe 4 alcanzan en ocasiones el 90%.

¹⁶⁶ SCHÖCH por ejemplo, aunque también subraya que las altas tasas de falsos positivos deben tenerse en cuenta al valorar la proporcionalidad de las medidas de seguridad, afirma sin embargo que en relación con los delitos más

8.2 Proporcionalidad de las medidas de seguridad

La incertidumbre del pronóstico de peligrosidad no sólo debilita en mi opinión el fundamento de las medidas de seguridad, sino que genera una notable incertidumbre a la hora de determinar su proporcionalidad. Porque si la medida, en su clase y duración, ha de ser proporcional a la peligrosidad del autor (clase y número de los delitos que se espera que cometa), pero este parámetro no se puede determinar, ¿en relación con qué efectuamos entonces el juicio de proporcionalidad?¹⁶⁸ En mi opinión la insolubilidad de este problema proporciona un argumento muy importante a favor de mantener los límites estrictos de duración de las medidas de seguridad en relación con la gravedad y duración de la pena prevista para el delito que se haya cometido, tal y como están en la redacción original (y en el momento en que se escriben estas líneas, aún vigente) del Código penal de 1995. Como es de todos conocido, estos límites han sido muy criticados por algunos autores,¹⁶⁹ a los que asiste la razón cuando arguyen que la lógica propia de las medidas de seguridad exigiría que la clase y duración de las mismas dependiera únicamente de la subsistencia del presupuesto que justifica su imposición, esto es, la peligrosidad del sujeto. Ahora bien, siendo esto cierto, no lo es menos que dicha peligrosidad no es posible determinarla de una manera mínimamente fiable tal y como he tratado de demostrar en las páginas anteriores (salvo quizá para periodos de tiempo muy cortos y circunstancias muy determinadas, como comentaré enseguida). Y, siendo esto así, y a pesar de que desde un punto de vista estrictamente lógico suponga introducir un límite extraño a la esencia de las medidas de seguridad, referir la naturaleza y duración de las mismas a la pena correspondiente al hecho cometido es el único criterio que proporciona un límite máximo cierto y

graves (se refiere en concreto a asesinos sádicos) debe bastar para la imposición de la custodia de seguridad incluso una probabilidad de sólo el 10% de que el sujeto vuelva realmente a reincidir (y asumir por tanto una tasa de falsos positivos del 90%), mientras que en relación con delitos menos graves como robo con violencia, lesiones peligrosas o abusos sexuales la solución no podría ser la misma (SCHÖCH, «Das Urteil des Bundesverfassungsgerichts zur Sicherungsverwahrung», *GA*, 2012, pp. 30 y s).

¹⁶⁷ Además de todo lo expuesto hasta aquí, en mi opinión hay otro peligro que subyace en la tesis de la ponderación de intereses como justificación para la imposición de las medidas de seguridad, y es la proximidad que existe entre esta idea y otra que últimamente también se maneja como criterio en este contexto, cual es la de distribución de riesgos entre individuo y sociedad. Aunque se habla de «distribución» de riesgos, en realidad y tal y como ha señalado Díez Ripollés a mi juicio con razón, en este enfoque subyace la idea de que «la sociedad rechaza hacerse cargo de los costes derivados de los riesgos de reincidencia delictiva», y «es un discurso que se sirve de una terminología tecnocrática para ocultar la insolidaridad social que lo inspira. El punto de referencia revelador de su naturaleza se encuentra en que la sociedad no admite, o restringe notablemente, sus responsabilidades en la génesis y abordaje de la delincuencia.» (DÍEZ RIPOLLÉS, *RECPC*, 2005, pp. 17 y s.) A mi juicio en la idea de ponderación de intereses subyace este mismo peligro, en la medida en que el conflicto entre bienes del que se habla es una situación tan indeterminada, futura e imprecisa, que siempre podrá aparecer como justificado el sacrificio de los bienes del individuo si se consigue describir el peligro que éste representa de manera suficientemente dramática. De nuevo he de limitarme a dejar apuntada esta idea, que por falta de espacio no puedo analizar más ampliamente.

¹⁶⁸ De otra opinión, sin embargo, CEREZO MIR, «Reflexiones críticas sobre algunas manifestaciones de la moderna tendencia a incrementar el rigor en la exigencia de responsabilidad criminal», *Revista Penal*, (22), 2008, p. 18, que afirma que «tampoco el margen de error en el pronóstico de la conducta futura del delincuente afecta, por sí mismo, a la proporcionalidad de la medida».

¹⁶⁹ Cfr., por todos, SANZ MORÁN, *Las medidas de corrección y de seguridad*, 2003, pp. 180 y ss. (con una relación pormenorizada de autores a favor y en contra del criterio de proporcionalidad recogido en el Código penal en la nota 52 de la p. 181). Defendiendo la posición asumida por el legislador penal en 1995, y con referencias doctrinales más recientes, ALONSO RIMO, «Medidas de seguridad y proporcionalidad con el hecho cometido (a propósito de la peligrosa expansión del Derecho penal de la peligrosidad)», *EPC*, v. XXIX, 2009, *passim*.

conocido, que al menos está en relación con un elemento importante como es el hecho cometido, y por lo tanto garantiza un mínimo irrenunciable de seguridad jurídica. Si consideramos necesario, como exigencia ineludible del principio de legalidad, la previa comisión del delito para aplicar la medida, a mi juicio son también elementales exigencias de seguridad jurídica las que, habida cuenta de la imposibilidad de determinar la peligrosidad con un mínimo de fiabilidad, exigen establecer límites determinados de duración a las medidas.¹⁷⁰

En este sentido me parecen interesantes las consideraciones que efectúa STRENG en relación con el mandato de proporcionalidad contenido en el párrafo 62 del StGB, según el cual «no podrá acordarse una medida de corrección y seguridad cuando no resulte proporcionada al significado de los delitos cometidos por el autor y los esperables de él, así como al grado del peligro que de él emana».¹⁷¹ Aunque desde luego no llega a extraer la consecuencia de que la duración de la pena correspondiente al hecho cometido tenga que servir de límite a la duración de la medida de seguridad, STRENG argumenta que de los dos parámetros en relación con los cuales el § 62 StGB obliga a efectuar el juicio de proporcionalidad (hecho cometido – hechos esperables y peligro), la orientación preventivo-especial hacia el futuro que es propia de las medidas de seguridad conduciría a poner el acento en los delitos que se espera que el sujeto vuelva a cometer. Sin embargo, añade que la incertidumbre que rodea al pronóstico de peligrosidad introduce una notable inseguridad en la comprobación de la relación entre razón de la intervención (hecho, peligro) y gravedad de la misma, y que por ello razones de seguridad jurídica obligan a tomar en consideración ante todo el hecho cometido, que es el único criterio seguro acerca de la peligrosidad del sujeto.¹⁷²

Con estas afirmaciones no pretendo desconocer los problemas que también plantea establecer la proporcionalidad de la medida en relación con el hecho cometido, tanto en abstracto como en la

¹⁷⁰ Como ha señalado entre nosotros con acierto ALONSO RIMO, y aunque hay acuerdo en considerarlo una ineludible exigencia de seguridad jurídica, también es en realidad contrario a la «lógica» de las medidas de seguridad tener que esperar a que el delito se haya cometido para comenzar la intervención sobre el sujeto: si está claro el peligro que supone, y lo que intentamos es precisamente prevenirlo, ¿por qué esperar a que se produzca el daño? (ALONSO RIMO, *EPC*, 2009, pp. 117 y ss.) Por otro lado, HASSEMER ha advertido recientemente de que ante la desordenada expansión de aspiraciones securitarias que está invadiendo el Derecho penal quizá no tarde en empezar a cuestionarse incluso el requisito de la postdelictualidad de las medidas, ya que supone desde luego un obstáculo a la prevención más eficaz (HASSEMER, «Glanz und Elend der Maßregeln im Strafrecht» en JOECKS (ed.), *Recht – Wirtschaft – Strafe. Festschrift für Erich Samson zum 70. Geburtstag*, 2010, p. 42).

¹⁷¹ § 62 StGB: «Eine Maßregel der Besserung und Sicherung darf nicht angeordnet werden, wenn sie zur Bedeutung der vom Täter begangenen und zu erwartenden Taten sowie zu dem Grad der von ihm ausgehenden Gefahr außer Verhältnis steht».

¹⁷² STRENG, *Strafrechtliche Sanktionen. Die Strafzumessung und ihre Grundlagen*, 3ª ed., 2012, p. 168. Es más, y siguiendo con el ejemplo de Alemania, según STRENG hay estudios empíricos que han demostrado que el internamiento en centros psiquiátricos como medida de seguridad para inimputables o semiimputables que han cometido delitos de naturaleza sexual (medida que en este país no tiene un límite máximo de duración) conduce a privaciones de libertad bastante más prolongadas que las que sufren autores de delitos similares a los que se impone una pena. Este desequilibrio no parece obedecer a una mayor peligrosidad de los internos, si ésta se mide en relación con el daño causado a la víctima y con los antecedentes penales, lo que, en opinión del citado autor, permite constatar que la legitimación de las medidas de seguridad sobre la idea de la protección de los intereses preponderantes (y, añadido yo, la ausencia de límites estrictos de duración) conduce a un déficit de garantías en relación con la capacidad limitadora que el principio de culpabilidad tiene respecto de la pena (STRENG, *Strafrechtliche Sanktionen*, 2012, p. 165. Lo que yo he traducido por 'déficit de garantías' aparece en el original como «ein Begrenzungsdefizit im Vergleich mit dem Schuldprinzip»).

concreta regulación por la que optó el legislador penal en 1995, y que la doctrina se ha ocupado de denunciar.¹⁷³ Lo que quiero resaltar es que la opción por la peligrosidad como parámetro al que ligar la duración de la medida, cuando ello suponga desbordar los límites que marca el injusto (en su caso, culpable) realizado, resulta a mi juicio inviable por la imposibilidad de formular a este respecto juicios mínimamente fiables y empíricamente sostenibles. Y de nuevo hay que insistir en que en esta materia, si el legislador no establece límites infranqueables, la tendencia de los operadores jurídicos a sobreestimar el peligro y el temor a que la persona puesta en libertad vuelva a cometer delitos conduce inevitablemente a sacrificar los intereses del individuo en aras de la seguridad de la sociedad, como a lo largo de las páginas anteriores he intentado poner de manifiesto.

8.3 Pronósticos en diferentes contextos. Relación entre culpabilidad y peligrosidad

La incertidumbre inherente a los pronósticos no tiene por qué significar que estos resulten completamente inadmisibles en cualquier ámbito del Derecho penal. Tienen razón la doctrina y la jurisprudencia cuando señalan que la legislación vigente los exige en muchos momentos, y que cuando es preciso tomar decisiones en relación con el comportamiento que pensamos que una persona desarrollará en el futuro, puede ser conveniente o necesario contar con un pronóstico pericial que aporte información empírica (aunque sea con todas las limitaciones que se quieran, y que nos hemos ocupado de destacar en este trabajo). Pero la cuestión es que no todas las decisiones son iguales, ni limitan los derechos del afectado con la misma intensidad, ni resultan legitimables en la misma medida.

1) Por un lado, considero que plantea pocos problemas de legitimidad la realización de pronósticos cuando su objetivo es identificar a aquéllos sujetos a los que se va a ofrecer un tratamiento o una posibilidad (en lugar de aplicar una sanción o eliminar un beneficio). Por ejemplo, si los recursos de que disponen las Instituciones Penitenciarias para ofrecer tratamientos a determinados tipos de delincuentes (agresores sexuales, agresores contra la pareja) es limitado, puede ser muy útil emplear algún instrumento que nos permita identificar, dentro de todos los internos condenados por esta clase de delitos, aquéllos que evidencian un mayor riesgo de reincidir, y a los que por lo tanto es más conveniente tratar. Un método que evidencie una sensibilidad elevada sería seguramente muy útil incluso a pesar de que su valor predictivo fuera bajo, pues el hecho de que dentro de la población a la que ofreceremos el tratamiento haya «falsos positivos» no representa un problema importante en la medida en que no les privamos de ningún derecho sino que se les ofrece una oportunidad, que quizá sirva de ayuda para mejorar su comportamiento incluso a quienes no son peligrosos. Lo que interesa sobre todo en decisiones de este tipo es que la tasa de falsos negativos sea reducida, de manera que dejemos fuera del tratamiento al menor número posible de delincuentes peligrosos, y puede asumirse sin mayores problemas una tasa elevada de falsos positivos.

¹⁷³ Por todos, SANZ MORÁN, *Las medidas de corrección y de seguridad*, 2003, pp. 184 y ss., o GRACIA MARTÍN en EL MISMO, *Tratado de las consecuencias jurídicas del delito*, 2006, pp. 459 y ss.

2) También son necesarios juicios de pronóstico para la adopción de medidas cautelares, y en estos casos parece preferible disponer de predicciones hechas con instrumentos contrastados que ayuden a fundamentar racionalmente la decisión y a minimizar los riesgos para la víctima. Por ejemplo ante la denuncia de una mujer de que ha sido objeto de violencia por parte de su pareja las autoridades han de tomar la decisión de si aplican las medidas cautelares que la legislación pone a su disposición, y para ello han de valorar el riesgo que tiene esa mujer de ser objeto de nuevas agresiones. En estos casos existen instrumentos de valoración del riesgo que pueden ayudar a tomar la decisión de si procede imponer una medida cautelar drástica como la prisión provisional, la medida de alejamiento, o no hace falta adoptar estas precauciones¹⁷⁴. Pero téngase en cuenta que se trata de predicciones hechas para un periodo de tiempo corto, que no necesariamente implican la privación de libertad del afectado, que van a ser revisadas en breves plazos de tiempo, y que se refieren a un contexto concreto y conocido: al peligro de reiteración de violencia de género en el marco de una concreta relación de pareja, y no a la probabilidad de comisión de delitos en general y en cualquier contexto. Aquí hay una persona determinada, que podría ser objeto de una agresión grave contra su vida, integridad o libertad sexual en las próximas horas, o días, y la situación se parece mucho más al presupuesto de aplicación de las causas de justificación. En consecuencia, la adopción de medidas de prevención para evitar dicho ataque es a mi juicio más fácil de legitimar con la idea de la prevalencia del interés preponderante. Además, en la medida en que se está actuando en el marco de un procedimiento para juzgar los actos de violencia que ya se hayan producido, incluso aunque se adoptara como medida cautelar la más drástica de todas, la prisión provisional, si el sujeto llegara a ser condenado el tiempo que hubiera estado privado de libertad de manera cautelar se abonaría para descontarlo de la duración de la pena privativa de libertad impuesta como condena. Creo que todas estas razones hacen que resulte en principio asumible, desde el punto de vista del sacrificio que supone para los derechos de acusado, la adopción de medidas perjudiciales para él sobre la base de un juicio de pronóstico.

3) En tercer lugar, resulta en mi opinión más fácil legitimar la adopción de decisiones negativas para el sujeto sobre la base incierta de los pronósticos de peligrosidad cuando las consecuencias que va a sufrir el individuo se mantienen dentro del marco penal correspondiente al injusto culpable, o, como mínimo, al injusto realizado. Por ejemplo, la decisión sobre si se concede o no la suspensión de la pena, uno de cuyos criterios rectores según lo dispuesto en el art. 80.1 CP es la valoración de la peligrosidad del sujeto, resulta mucho menos problemática que la decisión sobre si se impone a un individuo una custodia de seguridad posterior al cumplimiento de la pena porque, aunque ambas decisiones pueden suponer una limitación de su derecho fundamental a la libertad de similar intensidad, en un caso se trata de una injerencia que el sujeto ha provocado responsablemente, en cierto sentido «se ha merecido» al cometer el hecho, y si el juez opta por no conceder la suspensión de la pena la ejecución de la misma es algo que de todos modos entendemos abarcado al menos por la legitimidad que confiere el principio de culpabilidad. En sentido similar, la legitimidad de la imposición de medidas de seguridad a sujetos semiimputables, cuando la ejecución de las mismas se

¹⁷⁴ Cfr. ECHEBURÚA/FERNÁNDEZ-MONTALVO/DE CORRAL, *Predicción del riesgo de homicidio y de violencia grave en la relación de pareja. Instrumentos de evaluación del riesgo y adopción de medidas de protección*, 2009, especialmente los tres primeros capítulos.

guía por el sistema vicarial, no plantea especiales problemas de legitimación (a pesar de que el pronóstico de peligrosidad que se efectúa está sometido a las mismas limitaciones e incertidumbres que cualquier otro) porque al abonarse el tiempo de cumplimiento de la medida para la duración de la pena en ningún caso va a sobrepasarse con la aplicación de ambas consecuencias jurídicas el marco del injusto culpable.¹⁷⁵

4) Es cuando tratamos de justificar la imposición de consecuencias jurídicas limitadoras de derechos fundamentales más allá de lo permitido por el principio de culpabilidad, y fuera de las precauciones que puede ser necesario adoptar en el marco de la investigación de hechos concretos ya acaecidos, cuando la incertidumbre del pronóstico de peligrosidad resulta especialmente difícil de asumir. Aunque culpabilidad y peligrosidad sean dos conceptos diferentes, y cada uno de ellos sirva de fundamento a una clase diferente de consecuencias jurídicas del delito, a mi juicio es innegable que la culpabilidad ofrece un fundamento mucho más sólido para la imposición de la pena de lo que la peligrosidad (y la idea de prevalencia del interés preponderante) está en condiciones de hacer en relación con la medida de seguridad, debido a la enorme incertidumbre de los pronósticos de peligrosidad. Por ello, me parece que un Derecho penal respetuoso de las garantías que para el ciudadano representan los principios penales (y no me refiero sólo al de culpabilidad, sino también y especialmente al de seguridad jurídica) no tiene más remedio que acomodar el régimen de las medidas de seguridad en algún grado a la medida del injusto culpable, o, al menos, a la del injusto realizado, por mucho que ello signifique «alterar» en cierta forma la lógica de esta institución.

Con esto estoy afirmando la existencia de conexiones o paralelismos entre penas y medidas de seguridad, y entre culpabilidad y peligrosidad, que quizá cabe relacionar con otras que se han puesto de manifiesto, por contradictorio que ello pudiera parecer a primera vista, en las medidas imponibles a sujetos imputables, como lo demuestran algunas consideraciones realizadas sobre el fundamento de la custodia de seguridad en Alemania. Por una parte, no cabe olvidar que el TEDH ha afirmado que desde un punto de vista material la custodia de seguridad ha de ser considerada una pena, al menos por lo que respecta a la aplicación a la misma de la garantía de irretroactividad de las leyes penales¹⁷⁶. Pero también la doctrina alemana ha apuntado algunas similitudes en cuanto al fundamento de la custodia de seguridad y de la pena. En efecto, al tratarse de una medida de seguridad privativa de libertad, de duración indeterminada, y que se impone a quien ya ha saldado sus cuentas con la sociedad tras el cumplimiento de la pena, resulta especialmente difícil (y a la vez necesario) encontrar razones que justifiquen lo que a priori podría ser considerado simple y llanamente una instrumentalización del individuo, al que se limitan drásticamente sus derechos fundamentales en aras del aseguramiento de la sociedad. Y en este sentido se ha acudido en ocasiones a la idea de que quien ha demostrado a través de la repetida comisión – culpable – de delitos que no está dispuesto o no está en condiciones de adaptar su comportamiento a las normas establecidas para proteger la libertad de todos, ha de tolerar que se restrinja su propia libertad en

¹⁷⁵ Apunta también consideraciones en un sentido parecido STRENG, *Strafrechtliche Sanktionen*, 2012, p. 383 y 165.

¹⁷⁶ STEDH de 17 de diciembre de 2009, caso M. v. Germany (Application no. 19359/04), a la que siguieron después algunas otras sentencias resolviendo en el mismo sentido casos similares.

orden a la protección de la sociedad¹⁷⁷. Pero aludir a la repetida comisión de ilícitos culpables por parte de un sujeto responsable como fundamento para la medida de seguridad recuerda a la idea de una culpabilidad intensificada, en un sentido similar a una culpabilidad por la conducción de vida (*Lebensführungsschuld*), como han señalado algunos autores,¹⁷⁸ y eso supone un acercamiento notable entre el fundamento de las medidas, al menos las meramente asegurativas, y el de las penas, así como paralelismos con algunas agravaciones de la pena suficientemente conocidas que obedecen a este tipo de consideraciones (como la reincidencia). Por otro lado, desde el punto de vista del pronóstico de peligrosidad, la repetida comisión de delitos en el pasado es sin duda uno de los factores con mayor poder predictivo de reincidencia futura, como han puesto de manifiesto numerosísimos estudios.

¿No podría ser que lo que no queremos admitir como castigo de la culpabilidad nos cueste mucho menos de asumir si lo disfrazamos de peligrosidad? Profundizar en estas cuestiones es materia que merece un estudio detallado, que por evidentes razones de espacio no es posible efectuar aquí, pero abordarlo es uno de los retos a mi entender más importantes que tiene planteados en la actualidad la dogmática penal.

8.4 Prisión permanente revisable

Por último, quisiera añadir una reflexión en relación no con una medida de seguridad, sino con una pena, pero que por el régimen jurídico de que la dota el Proyecto de reforma del Código Penal de octubre de 2013¹⁷⁹ presenta una dependencia muy especial de los resultados de los pronósticos de peligrosidad: me refiero a la pena de prisión permanente revisable tal y como viene proyectada en dicho texto, y la tesis que sostengo es que las dificultades del pronóstico de peligrosidad, y sobre todo del pronóstico de reincidencia en relación con delitos graves cuya prevalencia es muy baja (por ejemplo los delitos contra la vida), ponen seriamente en cuestión, a mi juicio, la constitucionalidad de esta pena de prisión permanente revisable. Y ello porque el propio Proyecto hace depender la compatibilidad de esta pena con los arts. 15 y 25.2 CE de la posibilidad de revisión periódica,¹⁸⁰

¹⁷⁷ Por ej. STRENG, *Strafrechtliche Sanktionen*, 2012, p. 167. Sobre la idea, como fundamento en general para las medidas de seguridad, de que la repetida comisión de delitos evidenciaría una pérdida de la «libertad interior», que justificaría que a la persona se la prive de su «libertad exterior», cfr. por ej. FRISCH, *ZStW*, (102), 1990, pp. 365 y ss.

¹⁷⁸ Especialmente KÖHLER, «Die Aufhebung der Sicherungsmaßregeln durch die Strafgerechtigkeit» en PAWLIK/ZACZYK (eds.), *Festschrift für Günther Jakobs zum 70. Geburtstag am 26. Juli 2007*, 2007, pp. 286 y ss.; también lo reconoce STRENG, aunque rechazando la posición de Köhler porque le parece demasiado exagerada («zu weitgehend»), en la medida en que disolvería la diferencia entre pena y medida de seguridad (por ej. en *Juristenzeitung*, (17), 2011, pp. 829 y s. y nota 33). Entre nosotros SILVA se ha hecho eco de este problema, destacando la contradicción inherente a la custodia de seguridad (en tanto que trata al sujeto a la vez como ser racional y no racional), y sugiriendo que sería conveniente un análisis detallado para ver en qué medida una concepción de la culpabilidad de este tipo puede operar como fundamento de la custodia de seguridad o medidas similares, y en qué medida ello afectaría a la naturaleza – de pena o de medida – de estas privaciones de derechos (SILVA SÁNCHEZ, «¿Es la custodia de seguridad una pena?», *InDret*, (2), 2010 (editorial)).

¹⁷⁹ BOCG. Congreso de los Diputados Núm. A-66-1 de 04/10/2013.

¹⁸⁰ Véase el apartado II de la Exposición de Motivos del Proyecto, donde se dice que «lo que determina la inhumanidad de una pena es la falta de un horizonte de libertad que, en la regulación de la prisión permanente revisable, garantiza la existencia de un procedimiento judicial continuado de revisión. [...] La pena de prisión permanente revisable no constituye, por ello, una suerte de «pena definitiva» en la que el Estado se desentiende del

revisión que se acordará – cumplidos otros requisitos, especialmente el transcurso de determinados periodos mínimos de cumplimiento efectivo – en función del pronóstico de peligrosidad que presente el sujeto. Según el texto del Proyecto, la revisión periódica que habrá que hacer cada dos años sólo conducirá a la libertad (condicionada al cumplimiento de una serie de reglas) si hay un pronóstico favorable: «pronóstico favorable de reinserción social», se dice en el art. 92.1.c), pero después el art. 92.3, al regular cuándo procederá revocar la suspensión de la ejecución de la pena de prisión permanente revisable, dispone que el Juez de Vigilancia Penitenciaria «revocará la suspensión de la ejecución del resto de la pena y la libertad condicional concedida cuando se ponga de manifiesto un cambio de las circunstancias que hubieran dado lugar a la suspensión que no permita mantener ya *el pronóstico de falta de peligrosidad* en que se fundaba la decisión adoptada» (cursiva añadida). Pues bien, con independencia de otros problemas que también se han puesto ya de manifiesto¹⁸¹, el sometimiento de la pena de prisión permanente a una revisión regulada en estos términos introduce como ha señalado ya la doctrina una enorme incertidumbre en relación con la duración efectiva de la pena, que resulta muy problemática desde la perspectiva del principio de legalidad y la exigencia de seguridad jurídica que éste implica, aplicable a toda la actividad del Estado pero especialmente al ejercicio del *ius puniendi*.¹⁸² Tras lo expuesto en los epígrafes anteriores creo que ha quedado claro que el grado que alcanza esta incertidumbre es notabilísimo, y que especialmente en estos delitos con tasas de prevalencia tan bajas el número de falsos positivos es especialmente elevado. Por ello, el juicio que se ofrezca sobre la peligrosidad tiene un elevadísimo riesgo de incurrir en la arbitrariedad y el decisionismo, con las gravísimas consecuencias que tiene la decisión en estos casos, pues supondría prolongar una privación de libertad ya considerablemente larga.

Por otro lado, repárese en que para la concesión de la suspensión no se exige que no haya un pronóstico negativo, sino que se requiere un pronóstico de «falta de peligrosidad».¹⁸³ La interpretación que haya que dar a esta expresión, si es que el Proyecto llega a aprobarse en estos términos, deberá ser especialmente cuidadosa, porque de la misma manera en que no es posible

penado. Al contrario, se trata de una institución que compatibiliza la existencia de una respuesta penal ajustada a la gravedad de la culpabilidad, con la finalidad de reeducación a la que debe ser orientada la ejecución de las penas de prisión.»

¹⁸¹ Por ejemplo, lo difícil que resulta considerar compatible con el mandato constitucional de reinserción social contenido en el art. 25.2 CE una pena cuya primera revisión no puede efectuarse, sea cual sea la evolución del penado, hasta que han transcurrido 25, 30, 28 o incluso 35 años (según los casos), aparte de las restricciones que existen también para el acceso a permisos de salida o a la concesión del tercer grado; cfr. JUANATEY DORADO, «Una «moderna barbarie»: la prisión permanente revisable», *RGDP*, (20), 2013, pp. 8 y s.

¹⁸² Cfr., por ej., JUANATEY DORADO, *RGDP*, 2013; LASCURAÍN SÁNCHEZ, «Ni aunque sea revisable», artículo de opinión publicado en el diario *El País*, 3.10.2013; CUERDA RIEZU, *La cadena perpetua y las penas muy largas de prisión: por qué son inconstitucionales en España*, 2011, pp. 83 y ss., con ulteriores argumentos. También lo advertía el CONSEJO GENERAL DEL PODER JUDICIAL en su *Informe* de fecha 16.01.2013 al Anteproyecto que precedió al proyecto de Código penal (pp. 40 – 42).

¹⁸³ Que además, según la propia redacción del art. 92.3 en el Proyecto, tuvo que concurrir también previamente para acceder a la suspensión de la ejecución de la prisión permanente revisable. Se confirma así lo que auguraba ROBLES PLANAS en un conocido artículo cuando hace unos años apuntaba que «[m]ás bien la tendencia parece ser la de que sólo la *garantía de no peligrosidad* impide la intervención coactiva. La suficiente seguridad sólo queda garantizada si la puesta en libertad del autor sólo acontece cuando no exista ningún riesgo (más) de reincidencia» («Sexual predators. Estrategias y límites del Derecho penal de la peligrosidad», *InDret*, (4), 2007, p. 15).

asegurar con un 100% de seguridad que un reo reincidirá, tampoco es posible asegurar con absoluta certeza que no lo hará, de modo que si los tribunales pretendieran supeditar la revisión de la prisión permanente a la existencia de un pronunciamiento pericial que se exprese con este grado de contundencia, la prisión permanente no se revisaría en la práctica nunca. En efecto, si la reforma se aprobara en estos términos los tribunales deberían contentarse, como base para conceder la revisión y consiguiente suspensión de la ejecución, con informes que afirmen que la probabilidad de reincidencia es baja, o que el peligro es débil, porque difícilmente encontrarán peritos dispuestos a afirmar que es inexistente. Pero si bien es cierto, como se ha señalado en el epígrafe 5. de este trabajo, que los métodos de predicción de la peligrosidad son muchísimo más fiables cuando predicen la falta de peligrosidad que cuando afirman la existencia de la misma, no podemos perder de vista lo que pesan en decisiones como ésta (puesta en libertad, aunque sea condicional y tras periodos largos de cumplimiento, de personas condenadas por asesinatos muy graves o asesinatos terroristas) el temor a equivocarse con un falso negativo, y en general el temor a reacciones exageradas de la opinión pública y de los medios de comunicación – temor que además, por lo que vemos y oímos continuamente en la prensa, no es en absoluto infundado –. De modo que existen buenas razones para temer que, precisamente en relación con personas declaradas culpables de la comisión de delitos tan graves, el riesgo que querrán correr las instituciones penitenciarias y los jueces al decidir sobre la libertad será nulo, las estimaciones de probabilidades de reincidencia bajas en los informes periciales serán utilizadas como coartada para no conceder la revisión, y la pena de prisión será, en consecuencia, permanente de verdad.¹⁸⁴

9. Consideraciones finales

Aunque existen varios excelentes trabajos sobre la materia, la doctrina se ha ocupado hasta ahora relativamente poco de los problemas inherentes a las medidas de seguridad, o como mínimo se ha ocupado menos de esta parcela que del Derecho penal «de las penas». Sin embargo, esto debería cambiar si queremos estar a la altura de lo que se nos avecina: la sociedad reclama cada vez más que el Derecho penal actúe como medio eficaz de prevención de peligros, especialmente a través de medidas de seguridad, el legislador responde en consecuencia, y los penalistas deberíamos reaccionar ante esta demanda empezando por poner en cuestión si es cierto que el Derecho penal debe responder, y en su caso de qué manera y hasta qué punto, a lo que al parecer se exige de él. Los problemas y contradicciones que plantea en la actualidad el Derecho penal de la peligrosidad y las medidas de seguridad son muchos y desde luego ocuparse de ellos no puede ser objeto de este trabajo,¹⁸⁵ por lo que me limitaré a añadir sólo una última consideración final.

¹⁸⁴ Otro problema en modo alguno menor que plantea también la redacción proyectada del art. 92.3, en el que no podemos entrar aquí, es que la constatación de que ya no puede mantenerse el pronóstico de falta de peligrosidad es suficiente, sin que se requiera expresamente la comisión de un nuevo delito, para revocar la suspensión ya concedida. La inseguridad jurídica que introduce esta posibilidad de revocación de la suspensión a partir de la comprobación de un presupuesto descrito con de tal grado de vaguedad me parece verdaderamente alarmante.

¹⁸⁵ Por enumerar sólo algunos de ellos, recuérdense las dificultades que plantean tanto los medios admisibles de prueba como el estándar de exigencia de la misma en relación con la peligrosidad, que hemos apuntado en el epígrafe 6., así como la denuncia, no por repetida menos preocupante, del efecto de retroalimentación que los pronósticos de peligrosidad, especialmente los negativos, pueden tener sobre las personas a las que afectan. Por otro lado las

A pesar de los excesos en que se incurrió en los últimos años en Alemania en relación con la custodia de seguridad,¹⁸⁶ la doctrina en dicho país continúa considerando mayoritariamente necesarias las medidas de seguridad adicionales a la pena para delincuentes imputables peligrosos, precisamente para poder mantener, se dice, un Derecho penal en el que sí rija de forma estricta el principio de culpabilidad. Son frecuentes las alusiones a que el Derecho penal de la culpabilidad es insuficiente, demasiado estrecho («eng»), para dar respuesta a las ineludibles necesidades de prevención a las que hay que hacer frente hoy en día para proporcionar una protección eficaz a la sociedad, y para ello es necesario disponer de unas medidas de seguridad que, sin estar sujetas a dichas limitaciones, den una respuesta adecuada a esa demanda. Si no se dispusiera de estas medidas de seguridad desligadas de las garantías de la culpabilidad, se dice, la peligrosidad se intentaría combatir con una agravación generalizada de las penas, lo que conllevaría una intensificación excesiva de la respuesta punitiva más allá de lo tolerado por el principio de culpabilidad, con penas desproporcionadas a la gravedad del injusto culpable, y el resultado en conjunto sería peor.¹⁸⁷ Sin embargo, este argumento nos obliga a plantearnos, en mi opinión, al menos tres cuestiones.

La primera, si de verdad es cierto que la existencia en un ordenamiento penal de medidas de seguridad contundentes destinadas a neutralizar la peligrosidad de ciertos grupos de delincuentes es capaz de contener una deriva punitivista (en el sentido estricto de agravamiento de las *penas*) de dicho ordenamiento. A pesar de que existen casos, como señaladamente el español, que a primera vista podría considerarse que corroboran esta afirmación (en el sentido de que en nuestro país podría haber sido la ausencia de medidas de seguridad para los delincuentes imputables peligrosos una de las causas que ha impulsado el drástico agravamiento de las penas que venimos padeciendo desde que se aprobó el Código en 1995), creo que el problema es, por desgracia, mucho más complicado.¹⁸⁸ Por apuntar sólo dos ideas: en primer lugar, precisamente lo ocurrido en nuestro país creo que pone de manifiesto que no ha sido únicamente el temor a la delincuencia violenta de ciertos grupos de sujetos peligrosos lo que explica las reformas penales de los últimos 15 años, ya que éstas no se han limitado ni mucho menos a la delincuencia violenta, ni a la delincuencia grave, ni a la habitual. Y en segundo lugar, el caso de otros países como por ejemplo Estados Unidos pone de manifiesto que es perfectamente compatible la coexistencia en un ordenamiento jurídico de penas extraordinariamente

medidas de seguridad no privativas de libertad, como la libertad vigilada, plantean problemas parcialmente distintos que es necesario analizar. Y, por mencionar un último aspecto de carácter más general, quisiera destacar, siguiendo a Albrecht, el hecho de que cuarenta años después del surgimiento de los enfoques del *labeling approach* los estudios sobre carreras criminales sigan buscando los factores determinantes del comportamiento delictivo casi exclusivamente en las características individuales del sujeto y en las circunstancias de su contexto social inmediato, y no en las reacciones sociales formales e informales frente al delito (cosa que este autor considera «tan escandalosa como fatal», ALBRECHT en HEITMEYER/SOEFFNER (eds.) *Gewalt. Entwicklungen, Strukturen, Analyseproblemen*, 2004, p. 510.

¹⁸⁶ Puede verse un resumen de toda la evolución legal de esta medida, incluyendo la sentencia del TC alemán de 2011 y las reacciones a la misma, en ALEX, *Nachträgliche Sicherungsverwahrung*, 2013, pp. 4 a 74.

¹⁸⁷ Esta argumentación se evidencia con toda claridad por ejemplo en los trabajos de HASSEMER, *FS-Samson*, 2010, passim y especialmente p. 33; o RADTKE, «Schuldgrundsatz und Sicherungsverwahrung. Freiheitsgarantien im zweiseitigen Sanktionensystem», *GA*, 2011, pp. 636 y ss.

¹⁸⁸ Se muestra también muy escéptico en cuanto a la virtualidad que puede tener la idea de peligrosidad para contener el incremento de las penas DÍEZ RIPOLLÉS, *RECPC*, 2005, p. 31.

severas y de medidas de seguridad adicionales a la pena, y de muy diverso tipo, para sujetos considerados peligrosos.¹⁸⁹

La segunda cuestión a que me refería es si estas «exigencias» que aparentemente la sociedad plantea al Derecho penal obedecen a un aumento real de la criminalidad violenta grave que no esté siendo contenido adecuadamente a través de las penas proporcionadas al injusto culpable, o es más bien un mero discurso securitario¹⁹⁰ que se utiliza para perseguir otros fines, distintos de una protección eficaz frente a delitos graves. Es muy cierto que la seguridad de los ciudadanos en las sociedades postmodernas está afectada por numerosos riesgos, pero seguramente los más peligrosos no provienen de unas tasas de delincuencia violenta grave contra bienes jurídicos individuales que en la mayoría de los países – desde luego en el nuestro– no han experimentado en los últimos años un ascenso notable.¹⁹¹ Los penalistas conocemos las estadísticas de criminalidad, de la misma manera que reconocemos que la «inseguridad» de los ciudadanos frente a los delincuentes violentos es un sentimiento subjetivo y colectivo seguramente generalizado y en aumento, pero no debido a un aumento real de este tipo de delincuencia, y por ello es muy discutible que requiera un endurecimiento de las medidas penales para proporcionar «aseguramiento cognitivo» (por emplear la conocida expresión de JAKOBS), ya que el problema no es cognitivo, sino simbólico.¹⁹² Si somos conscientes de esto, ¿por qué se mantiene ese discurso, de que el Derecho penal tiene que proporcionar satisfacción a las necesidades de seguridad más allá de lo permitido por las limitaciones del principio de culpabilidad? En mi modesta opinión la intensificación de la respuesta penal, ya sea con agravamiento de penas o bien con medidas de seguridad para un grupo muy concreto de delincuentes «peligrosos», es en buena medida un maquillaje que trata de enmascarar la incapacidad de los gobiernos para proporcionar verdadera seguridad a sus ciudadanos en otros ámbitos (especialmente en lo que se refiere al deterioro continuado de las condiciones económicas, laborales y sanitarias, o a la generalización de la corrupción y saqueo de los recursos públicos, por mencionar algunos de los que me parecen más acusados en España). Si bien la existencia de estos riesgos no es en modo alguno una novedad, la crisis económica en la que vivimos desde hace años los ha evidenciado con especial claridad. El mantenimiento por parte de los penalistas del citado discurso proporciona una cobertura a unas medidas penales cuya legitimidad, en términos político-

¹⁸⁹ Sobre la política criminal en los EEUU cfr. la panorámica que ofrecía SILVA SÁNCHEZ en REDONDO (coord.), *Delincuencia sexual y sociedad*, 2002, pp. 147 y ss. Puede verse un resumen actualizado de la situación de las cárceles en EEUU en FERNÁNDEZ-PACHECO ESTRADA, «Sobre los peligros del punitivismo. El fenómeno de la encarcelación masiva en Estados Unidos», *InDret*, (3), 2013. Sobre el grado que además alcanza en ese país el control postpenitenciario de los delincuentes considerados peligrosos, cfr. ALONSO RIMO, «La publicidad de los antecedentes penales como estrategia de prevención del delito», *RGDP*, (17), 2012.

¹⁹⁰ Sobre el concepto de securitización cfr. BÖHM, «Securitización», *Revista Penal*, (32), 2013, *passim*.

¹⁹¹ Cfr. por ej. los datos estadísticos que aporta DÍEZ RIPOLLÉS, «Algunos rasgos de la delincuencia en España a comienzos del siglo XXI», *REIC*, (4), 2006, o la desviación entre evolución real de las tasas de delitos violentos, por un lado, y tratamiento mediático y evolución del miedo al delito en la opinión pública, por otro, que ha puesto de manifiesto VARONA GÓMEZ, *InDret*, (1), 2011. Incluso el propio legislador ha reconocido en ocasiones que la razón del endurecimiento punitivo no obedecía a aumentos reales de la criminalidad, sino al impacto social de determinados casos concretos; paradigmático es el conocido caso de la Exposición de Motivos de la LO 8/2006, de reforma de la Ley Penal del Menor.

¹⁹² Lo desataca por ej. SÁNCHEZ LÁZARO, «Alarma social y Derecho penal» en ROMEO CASABONA/SÁNCHEZ LÁZARO (eds.), *La adaptación del Derecho penal al desarrollo social y tecnológico*, 2010, pp. 53 y ss., proponiendo diversas soluciones, en cuyo análisis no es posible entrar aquí.

criminales y iusfundamentales, es más que dudosa.¹⁹³ El que articular un discurso distinto, y sobre todo comunicarlo eficazmente a la opinión pública, sea por muchos motivos difícil, no significa que no sea cada vez más necesario.

Por último, la tercera pregunta que me parece necesario plantear es si es compatible defender las garantías del principio de culpabilidad, como imprescindibles para mantener un Derecho penal respetuoso de la libertad del ciudadano, y a la vez reclamar la existencia en paralelo de un Derecho penal de la peligrosidad donde esos límites no existan, precisamente con el objeto de que, para los no peligrosos, sí los podamos mantener.¹⁹⁴ La similitud que evidencia este razonamiento con la tesis del Derecho penal del enemigo, tan denostado – con razón – por la doctrina mayoritaria en relación con otros ámbitos del Derecho penal, me parece innegable, y preocupante que la doctrina no lo advierta con la misma claridad respecto de las medidas de seguridad, en especial las destinadas a imputables peligrosos.¹⁹⁵

A mi juicio, la poca preocupación que mostramos los juristas (incluyendo a los tribunales) por las bases empíricas de la predicción de la peligrosidad, en tanto que fundamento de ese Derecho penal de la peligrosidad, puede tener algo que ver con esto. Como afirma SCHUMANN:

«El desinterés de los juristas por los fundamentos de la realización de los pronósticos podría deberse a que en realidad no se pretende obtener verdaderas predicciones. Podría pensarse que la palabra «pronóstico» no se utiliza en un sentido técnico. ¿Quizá en realidad de lo que se trata es de una clasificación entre personas, a las que procede tratar con suavidad en el castigo porque en realidad «son completamente normales», y otras, con las que debe emplearse la dureza porque «son bastante peligrosos»? Pero como estas clasificaciones de los seres humanos resultarían problemáticas, al menos si aparecieran así en la sentencia [...] hay que disfrazar esta tipología como si fuera un pronóstico. Porque formulaciones como «esta persona no va a representar un peligro en el futuro» o por el contrario «evidencia un peligro importante», son aceptables. Por un lado el perito podría equivocarse, por otro lado la persona podría cambiar. Un juicio de valor sobre las personas puede, si se disfraza de pronóstico, resultar asumible.»¹⁹⁶

¹⁹³ En relación con esto me parecen muy interesantes las conclusiones que ha extraído recientemente VARONA GÓMEZ tras analizar la información que ofrece la quinta edición de la Encuesta Social Europea sobre actitudes punitivas de los ciudadanos («Percepción y elección del castigo en España: resultados a partir de la Encuesta Social Europea (5ª ed.)», CPC, 2013, pp. 145 y ss.), y en concreto cuando afirma que los resultados de dicha encuesta proporcionan apoyo a las teorías simbólico-expresivas de las actitudes punitivas, según las cuales estas últimas derivan de temores de los ciudadanos no relacionados sólo con cuestiones directamente ligadas a la justicia penal, sino con otras inseguridades o temores contemporáneos, y concluye que hasta cierto punto «el político que responda a las (aparentes) demandas de mayor rigor penal con reformas punitivas está actuando como el médico que le da a un paciente un placebo. Pero con la relevante diferencia de que en nuestro caso el placebo no es inocuo, sino que consiste (la sanción penal) en causar daño a terceros» (ob. cit., pp. 184 y s.)

¹⁹⁴ Como afirma GARCÍA RIVAS, «parece evidente que se está utilizando esa ‘segunda vía’ para conculcar lo que la ‘primera vía’ dice proteger.» («La libertad vigilada y el Derecho penal de la peligrosidad», *RGDP*, (16), 2011, p. 12).

¹⁹⁵ JAKOBS por su parte no ha tenido inconveniente en reconocerlo, y en señalar también que la doctrina mayoritaria sin embargo no lo denuncia con la misma contundencia que respecto de otros ámbitos; cfr. JAKOBS, «Coacción y personalidad. Reflexiones sobre una teoría de las medidas de seguridad complementarias a la pena», *InDret*, (1), 2009, p. 13.

¹⁹⁶ SCHUMANN en FRISCH/VOGT (eds.), *Prognoseentscheidungen in der strafrechtlichen Praxis*, 1994, p. 41.

10. Tabla de jurisprudencia

<i>Tribunal, Sala y Fecha</i>	<i>Referencia</i>	<i>Magistrado Ponente</i>
Sentencia TEDH, Gran Sala, 21.10.2013	Application no. 42750/09	
Sentencia TEDH, Secc. 5ª, 17.12.2009	Application no. 19359/04	
STS, 2ª, 24.04.2007	nº 345/2007	Julián Sánchez Melgar
STS, 2ª, 02.02.2011	nº 65/2011	Julián Sánchez Melgar
STS, 2ª, 01.02.2012	nº 216/2012	Juan Saavedra Ruiz
SAP Alicante (Secc. 2ª) 19.10.2010	nº 710/2010	Francisco Javier Guirau Zapata
Sentencia Tribunal Supremo EEUU, 06.07.1983	463 U.S. 880	J. White
STC Alemán, 05.02.2004	BVerfG 2BvR 2029/01	
STC Alemán, 04.05.2011	BVerfG 2 BvR 2365/09	

11. Bibliografía

ACALE SÁNCHEZ (2010), *Medición de la respuesta punitiva. Tratamiento penal del delincuente imputable peligroso*, Thomson-Aranzadi, Cizur Menor.

Günter ALBRECHT (2004), «Sinn und Unsinn der Prognose von Gewaltkriminalität» en HEITMEYER/SOEFFNER (eds.), *Gewalt. Entwicklungen, Strukturen, Analyseproblemen*, Suhrkamp, Frankfurt, pp. 475 a 524.

Hans-Jörg ALBRECHT (2013), «Kriminalprognosen – Entwicklungen und Stand der Forschung» en FREUND/MURMANN/BLOY/PERRON (eds.), *Grundlagen und Dogmatik des gesamten Strafrechtssystems. Festschrift für Wolfgang Frisch zum 70. Geburtstag*, Duncker & Humblot, Berlin, pp. 1063 a 1076.

Hans-Jörg ALBRECHT (2013), «Psychiatrie, Gefährlichkeit und Prognose» en YUNDINA y otros (eds.), *Forensische Psychiatrie als interdisziplinäre Wissenschaft. Festschrift zum Geburtstag von Norbert Nedopil*, Medizinisch Wissenschaftliche Verlagsgesellschaft, Berlin, pp. 1 a 14.

ALONSO RIMO (2009), «Medidas de seguridad y proporcionalidad con el hecho cometido (a propósito de la peligrosa expansión del derecho penal de la peligrosidad)», *Estudios Penales y Criminológicos*, vol. XXIX, pp. 107 a 140.

ALONSO RIMO (2012), «La publicidad de los antecedentes penales como estrategia de prevención del delito», *Revista General de Derecho Penal*, (17), pp. 1 a 36.

ALEX (2013), *Nachträgliche Sicherungsverwahrung – ein rechtsstaatliches und kriminalpolitisches Debakel*, 2ª ed., Felix Verlag, Holzkirchen (1ª ed. de 2010).

AMNISTÍA INTERNACIONAL: Documento - EE. UU. (Texas). Pena de muerte/preocupación jurídica, Índice AI: AMR 51/028/2003/s, EXTRA 15/03 (citado según <http://www.amnesty.org/es/library/asset/AMR51/028/2003/es/f903b222-d72a-11dd-b0cc-1f0860013475/amr510282003es.html>, consultado por última vez el 10 de diciembre de 2013).

ANDRÉS PUEYO (2013), «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*, Edisofer, Madrid, pp. 483 a 503.

ANDRÉS PUEYO/REDONDO ILLESCAS (2007), «Predicción de la violencia: entre la peligrosidad y la valoración del riesgo de violencia», *Papeles del Psicólogo*, (28-3), pp. 157 a 173.

ANDREWS/BONTA (2006), *The Psychology of criminal behavior*, 4ª ed., LexisNexis.

APPERSON/MULVEY/LIDZ (1993), «Short-term clinical prediction of assaultive behaviour: artifacts of research methods», *The American Journal of Psychiatry*, (150-9), pp. 1374 a 1379.

ARBACH/ANDRÉS PUEYO (2007), «Valoración del riesgo de violencia en enfermos mentales con el HCR-20», *Papeles del Psicólogo*, (28-3), pp. 174 a 186.

ARMAZA ARMAZA (2013), *El tratamiento penal del delincuente imputable peligroso*, Comares, Granada.

AUERHAHN (2003), *Selective incapacitation and public policy. Evaluating California's imprisonment crisis*, State University of New York Press.

BÖHM (2013), «Securitización», *Revista Penal*, (32), pp. 72 a 90.

BUCHANAN/LEESE (2001), «Detention of people with dangerous severe personality disorders: a systematic review», *The Lancet*, (8-358), pp. 1955 a 1959.

CEREZO MIR (1998), «Consideraciones político-criminales sobre el nuevo Código Penal de 1995» en VVAA, *Estudios jurídicos en memoria del profesor Dr. D. José Ramón Casabó Ruiz*, v. 1, Valencia, pp. 373 a 395.

CEREZO MIR (2008), «Reflexiones críticas sobre algunas manifestaciones de la moderna tendencia a incrementar el rigor en la exigencia de responsabilidad criminal», *Revista Penal*, (22), pp. 16 a 21.

CONSEJO GENERAL DEL PODER JUDICIAL (2013), *Informe al Anteproyecto de Ley Orgánica por la que se modifica la ley orgánica 10/1995, de 23 de noviembre, del código penal*, 16.01.2013 (accesible online en la página web del Consejo General del Poder Judicial: http://www.poderjudicial.es/cgpj/es/Poder_Judicial/Consejo_General_del_Poder_Judicial/Actividad_del_CGPJ/Informes, consultada por última vez el 14 de enero de 2014).

COOKE/MICHIE (2010) «Limitations of diagnostic precision and predictive utility in the individual case: a challenge for forensic practice», *Law and Human Behaviour*, (34), pp. 259 a 274.

COOKE/MICHIE (2011), «Violence risk assessment. Challenging the illusion of certainty», en MCSHERRY/KEIZER (eds.), *Dangerous people. Policy, prediction and practice*, Routledge, New York-London, pp. 147 a 161.

CUERDA RIEZU (2011), *La cadena perpetua y las penas muy largas de prisión: por qué son inconstitucionales en España*, Barcelona, Atelier.

DAHLE/LEHMANN (2012), «Grundlagen und Methoden der Kriminalprognose» en EGG (ed.), *Psychologisch-psychiatrische Begutachtung in der Strafjustiz*, KrimZ, Wiesbaden, pp. 151 a 179.

DÍEZ RIPOLLÉS (2005), «De la sociedad del riesgo a la seguridad ciudadana: un debate desenfocado», *Revista Electrónica de Ciencia Penal y Criminología*, (07-01), pp. 1 a 37.

DÍEZ RIPOLLÉS (2006), «Algunos rasgos de la delincuencia en España a comienzos del siglo XXI», *Revista Española de Investigación Criminológica*, (4), pp. 1 a 19.

DITTMANN (2013), «Lieber würfeln oder eine Münze werfen? Über Risiken, wissenschaftliche Erkenntnis und Prognosen» en YUNDINA et al. (eds.), *Forensische Psychiatrie als interdisziplinäre Wissenschaft. Festschrift zum Geburtstag von Norbert Nedopil*, Medizinisch Wissenschaftliche Verlagsgesellschaft, Berlin, pp. 45 a 57.

ECHEBURÚA/FERNÁNDEZ-MONTALVO/DE CORRAL (2009), *Predicción del riesgo de homicidio y de violencia grave en la relación de pareja. Instrumentos de evaluación del riesgo y adopción de medidas de protección*, Centro Reina Sofía – Generalitat Valenciana, Serie Documentos 15.

EDENS/BUFFINGTON-VOLLUM/KEILEN/ROSKAMP/ANTHONY (2005), «Predictions of Future Dangerousness in Capital Murder Trials: Is It Time to “Disinvent the Wheel?”», *Law and Human Behaviour*, (29-1), pp. 55 a 86.

EGG (2012), «Forensisch-psychologische Begutachtung in der Strafjustiz – Entwicklung und Perspektiven in Deutschland» en EL MISMO (ed.), *Psychologisch-psychiatrische Begutachtung in der Strafjustiz*, KrimZ, Wiesbaden, pp. 15 a 32.

ESBEC RODRÍGUEZ (2003), «Valoración de la peligrosidad criminal (riesgo-violencia) en psicología forense. Aproximación conceptual e histórica», *Psicopatología Clínica, Legal y Forense*, (3-2), pp. 45 a 64.

FAZEL/SINGH/DOLL/GRANN (2012), «Use of risk assessment instruments to predict violence and antisocial behaviour in 73 samples involving 24827 people: systematic review and meta-analysis», *British Medical Journal*, (345:e4692), pp. 1 a 12.

FERNÁNDEZ-PACHECO ESTRADA (2013), «Sobre los peligros del punitivismo. El fenómeno de la encarcelación masiva en Estados Unidos», *InDret*, (3), pp. 1 a 23.

FRISCH (1983), *Prognoseentscheidungen im Strafrecht. Zur normativen Relevanz empirischen Wissens und zur Entscheidung bei Nichtwissen*, Decker's Verlag, Heidelberg-Hamburg.

FRISCH (1990), «Die Maßregeln der Besserung und Sicherung im strafrechtlichen Rechtsfolgensystem. Straftheoretische Einordnung, inhaltliche Ausgestaltung und rechtsstaatliche Anforderungen», *Zeitschrift für die Gesamte Strafrechtswissenschaft*, (102-2), pp. 343 a 393.

FRISCH (1994), «Strafrechtliche Prognoseentscheidungen aus rechtswissenschaftlicher Sicht» en FRISCH/VOGT (eds.), *Prognoseentscheidungen in der strafrechtlichen Praxis*, Nomos, Baden-Baden, pp. 55 a 136.

GARCÍA ALBERO (2009), «Ejecución de penas en el proyecto de reforma. Estudio de un problema concreto: ¿qué hacer con los reos habituales o reincidentes en los que subsiste la peligrosidad criminal tras el licenciamiento definitivo?» en ÁLVAREZ GARCÍA (dir.), *La adecuación del Derecho penal español al ordenamiento de la Unión Europea. La política criminal europea*, Valencia, Tirant lo Blanch, pp. 127 a 140.

GARCÍA RIVAS (2011), «La libertad vigilada y el Derecho penal de la peligrosidad», *RGDP*, (16), pp. 1 a 27.

GARLAND (2001), *The culture of control: crime and social order in contemporary society*, Chicago University Press.

GÓMEZ HERMOSO (2012), *Informe sobre la eficacia de las predicciones de peligrosidad de los informes psicológicos forenses emitidos en los Juzgados de Vigilancia Penitenciaria de Madrid* (publicado en la página web del Colegio oficial de Psicólogos de las Islas Baleares <http://www.copib.es/pdf/Vocalies/Juridica/2012-11-18%20predicciones.pdf>, consultada por última vez el día 21 de octubre de 2013).

GRACIA MARTÍN (2006), «Las medidas de seguridad y de reinserción social» en EL MISMO (coord.), *Tratado de las consecuencias jurídicas del delito*, Valencia, Tirant lo Blanch, pp. 429 a 490.

GRACIA MARTÍN (2008), «Sobre la legitimidad de medidas de seguridad contra delincuentes imputables peligrosos en el Estado de Derecho» en GARCÍA VALDÉS et al. (eds.), *Estudios Penales en Homenaje a Enrique Gimbernat*, t. 1, Edisofer, Madrid, pp. 975 a 1003.

GROß/NEDOPIL (2006), «Basisraten für kriminelle Rückfälle – Ergebnisse einer Literaturübersicht», en NEDOPIL (y cols.), *Prognosen in der forensichen Psychiatrie – Ein Handbuch für die Praxis*, 3ª ed., Pabst Science Publishers, pp. 65 a 98.

HARCOURT (2007), *Against Prediction: profiling, policing and punishing in an actuarial age*, The University of Chicago Press.

HARRIS/RICE/QUINSEY (2008), «Shall evidence-based risk assessment be abandoned?», *The British Journal of Psychiatry*, (192), p. 154.

HART/COOKE (2013), «Another look at the (im-)precision of individual risk estimates made using actuarial risk assessment instruments», *Behavioral Sciences and the Law*, (31), pp. 81 a 102.

HASSEMER (2010), «Glanz und Elend der Maßregeln im Strafrecht» en JOECKS (ed.), *Recht – Wirtschaft – Strafe. Festschrift für Erich Samson zum 70. Geburtstag*, Müller, Heidelberg, pp. 31 a 42.

HELMUS/HANSON/THORNTON/BABCHISHIN/HARRIS (2012), «Absolute recidivism rates predicted by Static-99R and Static-2002R sex ofender risk assessment tolos vary across samples: a meta-analysis», *Criminal Justice and Behaviour*, (39), pp. 1148 a 1171.

HERRERO (2013), «¿Por qué no reincide la mayoría de los agresores sexuales?», *Anuario de Psicología Jurídica*, (23), pp. 71 a 77.

JAKOBS (2009), «Coacción y personalidad. Reflexiones sobre una teoría de las medidas de seguridad complementarias a la pena», *InDret*, (1), pp. 1 a 16.

JORGE BARREIRO (1976), *Las medidas de seguridad en el Derecho español: un estudio de la ley de peligrosidad y rehabilitación social de 1970 y de la doctrina de la sala de apelación de peligrosidad*, Civitas, Madrid.

JORGE BARREIRO (2005), «Reflexiones sobre la compatibilidad de la regulación de las medidas de seguridad en el CP de 1995 con las exigencias del Estado de Derecho» en VVAA, *Homenaje al profesor Dr. Gonzalo Rodríguez Mourullo*, Thomson-Civitas, pp. 567 a 586.

JOST (2012), *Gefährliche Gewalttäter? Grundlagen und Praxis der Kriminalprognose*, Kohlhammer, Stuttgart.

JUANATEY DORADO (2013), «Una «moderna barbarie»: la prisión permanente revisable», *RGDP*, (20), pp. 1 a 13.

KAISER (1990), *Befinden sich die Kriminalrechtlichen Maßregeln in der Krise?*, CF Müller, Heidelberg.

KINZIG (1996), *Die Sicherungsverwahrung auf dem Prüfstand. Ergebnisse einer theoretischen und empirischen Bestandsaufnahme des Zustandes einer Maßregel*, Freiburg, edition iuscrim.

KINZIG (2010), *Die Legalbewährung gefährlicher Rückfalltäter. Zugleich ein Beitrag zur Entwicklung des Rechts der Sicherungsverwahrung*, 2ª ed., Berlin, Duncker & Humblot.

KÖHLER (2007), «Die Aufhebung der Sicherungsmaßregeln durch die Strafgerechtigkeit», en PAWLIK/ZACZYK (eds.), *Festschrift für Günther Jakobs zum 70. Geburtstag am 26.Juli 2007*, Carl Heymanns, Köln, pp. 273 a 292.

KÖNIG (2010), «Der Nutzen standardisierter Risikoprognoseinstrumente für Einzelfallentscheidungen in der forensischen Praxis», *Recht & Psychiatrie* (2010) 28, pp. 67-73.

KREUZER/BARTSCH (2011), «Anmerkung I» a BVerfG, Urt. v. 04.05.2011, *Strafverteidiger*, (8), pp. 472 a 480.

LASCURAÍN SÁNCHEZ (2013), «Ni aunque sea revisable», artículo de opinión publicado en el diario *El País*, 3.10.2013.

LÓPEZ-REY (1983), «Algunas reflexiones sobre el peligro de la noción de peligrosidad», *Cuadernos de Política Criminal*, (19), pp. 19 a 25.

NEDOPIIL (2002), «Prognosebegutachtungen bei zeitlich begrenzten Freiheitsstrafen – Eine sinnvolle Lösung für problematische Fragestellungen?», *Neue Zeitschrift für Strafrecht*, (7), pp. 344 a 349.

NEDOPIIL (2006), *Prognosen in der forensischen Psychiatrie – Ein Handbuch für die Praxis*, 3ª ed., Pabst Science Publishers.

NEDOPIIL/STADTLAND (2007), «Das Problem der falsch Positiven: Haben wir unsere prognostische Kompetenz seit 1966 verbessert?» en LÖSEL/BENDER/JEHLE (eds.), *Kriminologie und wissenschaftsbasierte Kriminalpolitik. Entwicklungs- und Evaluationsforschung*, Forum Verlag Godesberg, Mönchengladbach, pp. 541 a 550.

NOWAKOWSKI (1963), «Zur Rechtsstaatlichkeit der vorbeugenden Maßnahmen» en WELZEL/CONRAD/Armin KAUFMANN/Hilde KAUFMANN (eds.), *Festschrift für Hellmuth von Weber zum 70. Geburtstag*, Ludwig Röhrscheid Verlag, Bonn, pp. 98 a 120.

PÉREZ RAMÍREZ/REDONDO ILLESCAS/MARTÍNEZ GARCÍA/GARCÍA FORERO/ANDRÉS PUEYO (2008), «Predicción de riesgo de reincidencia en agresores sexuales», *Psicothema*, (20-2), pp. 205 a 210.

POLLÄHNE (2011), *Kriminalprognostik. Untersuchungen im Spannungsfeld zwischen Sicherheitsrecht und Rechtssicherheit*, De Gruyter, Berlin.

POLLÄHNE (2013), § 61 en KINDHÄUSER/NEUMANN/PAEFFGEN (eds.), *Nomos Kommentar Strafgesetzbuch*, 4ª ed., v. I, Nomos, Baden-Baden.

PUJOL ROBINAT/PUIG BAUSILI (2008), «Concepto de peligrosidad criminal. Evolución histórica del concepto», *Cuadernos de Política Criminal*, (94), pp. 255 a 283.

QUINSEY et al. (2006), *Violent offenders. Appraising and managing risk*, 2ª ed., American Psychological Association, Washington.

RADTKE (2011), «Schuldgrundsatz und Sicherungsverwahrung. Freiheitsgarantien im zweispurigen Sanktionensystem», *Goldtdammer's Archiv für Strafrecht*, pp. 636 a 652.

RASCH (1994), «Verhaltenswissenschaftliche Kriminalprognosen» en FRISCH/VOGT (eds.), *Prognoseentscheidungen in der strafrechtlichen Praxis*, Nomos, Baden-Baden, pp. 17 a 29.

REDONDO ILLESCAS/PÉREZ/MARTÍNEZ (2007), «El riesgo de reincidencia en agresores sexuales: investigación básica y valoración mediante el SVR-20», *Papeles del Psicólogo*, (28-3), pp. 187 a 195.

REQUEJO RODRÍGUEZ (2008), «Peligrosidad criminal y Constitución», *InDret*, (3), pp. 1 a 23.

ROBLES PLANAS (2007), «Sexual predators. Estrategias y límites del Derecho penal de la peligrosidad», *Indret*, (4), pp. 1 a 25.

ROMEO CASABONA (1986), *Peligrosidad y Derecho Penal preventivo*, Bosch, Barcelona.

ROSSEGGER/GERTH/SEEWALD/URBANIÖK/SINGH/ENDRASS (2013), «Current obstacles in replicating risk assessment findings: a systematic review of commonly used actuarial instruments», *Behavioural Sciences and the Law*, (31), pp. 154 a 164.

RUSCHE (2004), *In Freiheit gefährlich? Eine Untersuchung zu Häufigkeit und Gründen falscher Kriminalprognosen bei psychisch kranken Gewaltverbrechern*, Regensburg, Roderer.

SÁNCHEZ LÁZARO (2010), «Deconstruyendo las medidas de seguridad», *InDret*, (2), pp. 1 a 26.

SÁNCHEZ LÁZARO (2010), «Alarma social y Derecho penal», en ROMEO CASABONA/SÁNCHEZ LÁZARO (eds.), *La adaptación del Derecho penal al desarrollo social y tecnológico*, Comares, Granada.

SANZ MORÁN (2003), *Las medidas de corrección y de seguridad en el Derecho Penal*, Valladolid, Lex Nova

SCHÖCH (2008), Vor § 61 y § 61, en LAUFHÜTTE/RISSING - VAN SAAN/TIEDEMANN (eds.), *Strafgesetzbuch: Leipziger Kommentar. Großkommentar*, 12ª ed., v. III, De Gruyter, Berlin.

SCHÖCH (2012), «Das Urteil des Bundesverfassungsgerichts zur Sicherungsverwahrung», *Goldtdammer's Archiv für Strafrecht*, pp. 14 a 31.

SCHUMANN (1994), «Prognosen in der strafgerichtlichen Praxis und deren empirische Grundlagen», en FRISCH/VOGT (eds.), *Prognoseentscheidungen in der strafrechtlichen Praxis*, Nomos, Baden-Baden, pp. 31 a 41.

SILVA SÁNCHEZ (1997), «La regulación de las medidas de seguridad (artículo 6)» en EL MISMO, *El nuevo Código penal: cinco cuestiones fundamentales*, Barcelona, Bosch, pp. 15 a 49.

SILVA SÁNCHEZ (2002), «El retorno de la inocuización» en REDONDO (coord.), *Delincuencia sexual y sociedad*, Ariel, Barcelona, pp. 143 a 159.

SILVA SÁNCHEZ (2009), «El contexto del Anteproyecto de Reforma del Código Penal de 2008» en AAVV, *El Anteproyecto de modificación del Código Penal de 2008. Algunos aspectos (Cuadernos penales José María Lidón, Núm. 6)*, Bilbao, Universidad de Deusto, pp. 15 a 34.

SILVA SÁNCHEZ (2010), «¿Es la custodia de seguridad una pena?», *InDret*, (2), (editorial).

SINGH/FAZEL/GUEORGUEVA/BUCHANAN (2014), «Rates of violence in patients classified as high risk by structured risk assessment instruments», *The British Journal of Psychiatry*, (204-3), pp. 180 a 197.

SINGH/GRANN/FAZEL (2011), «A comparative study of violence risk assessment tools: a systematic review and metaregression analysis of 68 studies involving 25.980 participants», *Clinical Psychology Review*, (31), pp. 499 a 513.

SKEEM/MONAHAN (2011), «Current directions in violence risk assessment», *Current Directions in Psychological Science*, (20-1), pp. 38 a 42.

SLOBOGIN (2006), «Dangerousness and expertise redux», *Emory Law Journal*, (56-2), pp. 275 a 325.

STEADMAN (2000), «From Dangerousness to Risk Assessment of Community Violence: taking stock at the turn of the century», *Journal of the American Academy of Psychiatry and the Law*, (28), pp. 265 a 71.

STRENG (2003), «Das Legitimations-Dilemma sichernden Freiheitsentzugs - Überlegungen zur neueren Rechtsentwicklung» en DÖLLING, (ed.), *Jus humanum. Grundlagen des Rechts und Strafrecht. Festschrift für Ernst-Joachim Lampe zum 70. Geburtstag*, Duncker & Humblot, Berlin, pp. 611 a 642.

STRENG (2011), «Die Zukunft der Sicherungsverwahrung nach der Entscheidung des Bundesverfassungsgerichts», *Juristenzeitung*, (17), pp. 827 a 835.

STRENG (2012), *Strafrechtliche Sanktionen. Die Strafzumessung und ihre Grundlagen*, 3ª ed., Kohlhammer, Stuttgart.

SZMUCKLER (2003), «Risk assessment: 'numbers' and 'values'», *The Psychiatrist (formerly The Psychiatric Bulletin)*, (27), pp. 205 a 207.

TERRADILLOS BASOCO (1981), *Peligrosidad social y Estado de Derecho*, Akal, Madrid.

URBANIOK (2005), «Validität von Risikokalkulationen bei Straftätern – Kritik an einer methodischen Grundannahme und zukünftige Perspektiven –» en EGG (ed.), „Gefährliche Straftäter» *Eine Problemgruppe der Kriminalpolitik?*, KrimZ, Wiesbaden, pp. 85 a 108.

URRUELA MORA (2009), *Las medidas de seguridad y reinserción social en la actualidad. Especial consideración de las consecuencias jurídico-penales aplicables a sujetos afectados de anomalía o alteración psíquica*, Comares, Granada.

VAN GEMMEREN (2012), § 61 en JOECKS/MIEBACH (eds.) *Münchener Kommentar zum Strafgesetzbuch*, 2ª ed., vol. II, Beck, München.

VARONA GÓMEZ (2011), «Medios de comunicación y punitivismo», *InDret*, (1), pp. 1 a 34.

VARONA GÓMEZ (2013), «Percepción y elección del castigo en España: resultados a partir de la Encuesta Social Europea (5ª ed.)», *Cuadernos de Política Criminal*, (111), pp. 145 a 194.

VIVES ANTÓN (1974), «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, pp. 389 a 417.

YANG/WONG/COID (2010), «The efficacy of violence prediction: a meta-analytic comparison of nine risk assessment tools», *Psychological Bulletin*, (136-5), pp. 740 a 767.

ZIFFER (2008), *Medidas de seguridad. Pronósticos de peligrosidad en Derecho penal*, Hammurabi, Buenos Aires.

ZIFFER (2013), «Begriff der Strafe und Sicherungsverwahrung», en FREUND/MURMANN/BLOY/PERRON (eds.), *Grundlagen und Dogmatik des gesamten Strafrechtssystems. Festschrift für Wolfgang Frisch zum 70. Geburtstag*, Duncker & Humblot, Berlin, pp. 1077 a 1091.

ZUGALDÍA ESPINAR (2009), «Medidas de seguridad complementarias y acumulativas para autores peligrosos tras el cumplimiento de la pena», *Revista de Derecho Penal y Criminología*, (3-1), pp. 199 a 212.