

Gloria Muñoz García
Directora del Laboratorio de Control del Dopaje

LA IMPOSIBILIDAD DE ALCANZAR LA CERTEZA ABSOLUTA EN CUALQUIER MEDICION EN LOS LABORATORIOS ANALITICOS DE CONTROL OFICIAL

1. Por qué el conocimiento y los juicios científicos no pueden ser absoluta y totalmente ciertos y existe un margen para el error.

Por definición, el método científico es un método de investigación usado principalmente en la producción de conocimientos en las ciencias y que permite llegar a un conocimiento que pueda ser considerado válido. Este método se sustenta sobre dos pilares fundamentales:

- La **reproducibilidad**, es decir, la capacidad de repetir un determinado experimento en cualquier lugar y por otra persona, obteniendo un resultado similar.
- La **falsabilidad o refutabilidad**, es decir, toda proposición científica tiene que ser susceptible de ser falsada o refutada. Esto significa que las leyes o teorías obtenidas a través del método científico, puedan ser reevaluadas y que con tiempo, al contar con más evidencias y más desarrollo tecnológico que permita más conocimiento se pueda hallar que dichas leyes o teorías sean inexactas.

Es por esta naturaleza de revisión permanente del conocimiento científico que plantea el segundo pilar del método científico, que la ciencia, aun no pudiendo generar certezas totales y absolutas, en sentido estricto, sí puede proporcionar respuestas válidas y de gran fiabilidad, que en la práctica, pueden tratarse como absolutas.

En sentido estricto, en ciencia no es posible llegar a una fiabilidad del 100%, ya que en el futuro puede surgir un nuevo experimento que obligue a revisar una teoría científicamente consolidada. No es posible, por razones prácticas, económicas, tecnológicas, etc...realizar infinitos experimentos para demostrar una teoría, y aun realizándolos, siempre podría realizarse un único experimento suficientemente comprobado que refutase esa teoría.

Eso no significa que en ciencia todo sea incierto y que los científicos no puedan emitir juicios o desarrollar teorías, sino que es necesario entender que siempre existirá un margen para la falsabilidad o el error, pero que también existen mecanismos para controlar y reducir este margen y generar la suficiente confianza en la comunidad científica como para emitir certezas "cuasi" absolutas.

De hecho, la fiabilidad en una teoría aumenta a medida que nuevos experimentos la corroboran, por lo que, con el tiempo, la ciencia consolidada se acerca a esa seguridad o confiabilidad



absoluta del 100%, haciendo que el posible margen de falsabilidad o error sea mucho menor o en la práctica inexistente.

Esta misma consideración podría trasladarse no solo a la elaboración de una teoría, sino al desarrollo de un método, técnica o mecanismo de análisis para dar respuesta a diferentes cuestiones en las que pueda necesitarse un conocimiento o juicio científico (para determinar cumplimiento o no de especificaciones legales, para establecer un valor monetario, para diagnosticar enfermedades o patologías, etc...). Inicialmente, los experimentos de validación pretenden consolidar el método, evaluar que sirve para el fin previsto y que es reproducible. A medida que el método, técnica o mecanismo se aplica en diversos laboratorios, a diferentes muestras con diferentes características, el método se consolida y su margen de falsabilidad se reduce.

El primer pilar del método científico, la reproducibilidad, tampoco está exento de ese margen de error. En ciencia, en relación a la capacidad de repetir un determinado experimento siempre se manejan dos conceptos:

- Reproducibilidad, que sería la capacidad de repetir ese experimento modificando algunas condiciones como el laboratorio en que se realiza, la persona que lo realiza, los lotes de reactivos o materiales que se utilizan, etc... Esto, per se, ya indica que aunque el resultado sea similar, no puede ser absoluta y exactamente igual, ya que para la realización del experimento se han modificado algunos factores y al ser las condiciones distintas, los resultados pueden ser muy similares, pero no exactamente iguales.
- Repetibilidad, que sería la capacidad de repetir ese experimento en las mismas condiciones (mismo laboratorio, misma persona, mismos lotes, etc..). No obstante, y pesar de que un laboratorio es un ambiente en que ciertas condiciones son controlables, no se pueden controlar los infinitos factores que podrían estar afectando a un experimento concreto y por lo tanto, el experimento no se puede repetir en exactamente (en sentido literal) las mismas condiciones.

Es por ello, que cuando se realiza un experimento, o podríamos decir una medición si queremos transponerlo al ámbito del laboratorio analítico, no se puede tener una certeza absoluta, en el sentido estricto, de que ese resultado obtenido sea 100% verdadero y siempre existirá un cierto margen de error.

Pero a pesar de ese margen de error, de nuevo, existen mecanismos para controlarlo. Por un lado, la validación del método permite estimar ese margen de error y los controles de calidad que deben aplicarse en los métodos permiten corroborar de forma continua y en el tiempo que ese margen de error se mantiene y no se incrementa.

2. Cómo, a pesar del margen de la falsabilidad se puede confiar en la fiabilidad de los resultados de los laboratorios analíticos.

A pesar de que los laboratorios analíticos, y en especial aquellos que se dedican al control oficial en diferentes ámbitos (agroalimentario, forense, antidopaje, etc...) son conscientes del margen de falsabilidad de sus resultados, también tienen muy en cuenta el impacto que los resultados que emiten tienen en la toma de decisiones o en las actuaciones y acciones posteriores que se puedan realizar, que generalmente suelen ser sanciones legales o administrativas y en ocasiones

dirigidas a sujetos individuales, o incluso acciones que generen grandes pérdidas económicas a determinados sectores.

Por eso, teniendo en cuenta que este tipo de decisiones se basan en resultados analíticos, estos laboratorios hacen importantes esfuerzos en garantizar la calidad de los resultados, y el grado de confianza en estos para el propósito en cuestión.

Es habitual que estos laboratorios introduzcan medidas de aseguramiento de la calidad para garantizar su capacidad y proporcionar datos de la calidad requerida. Dichas medidas incluyen: la utilización de métodos de análisis validados, el uso de procedimientos internos de control de calidad, la participación en esquemas de ensayos de aptitud, la acreditación basada generalmente en la Norma ISO/IEC 17025, y el establecimiento de la trazabilidad de los resultados de las mediciones.

Además, los laboratorios no suelen actuar de forma aislada, sino que laboratorios de diversos ámbitos se relacionan entre sí a través de convenios, sociedades, organizaciones, etc... Ya que un mecanismo muy potente para crear confiabilidad en un método o técnica consiste en que esta haya sido discutida y testeada por diversos laboratorios. De hecho, un alto grado de certeza se alcanza después de un gran esfuerzo colectivo de investigación y de transferencia del conocimiento.

En los ámbitos en los que operan los laboratorios de control oficial, que están sometidos a una gran presión para demostrar la calidad de sus resultados y para demostrar la confianza que puede ser depositada en estos, suelen existir consensos internacionales u organizaciones internacionales que llevan a cabo diversas acciones y actividades para garantizar que los métodos y técnicas aplicadas en estos laboratorios estén consolidadas, y que la aplicación de ellas permita emitir resultados cercanos a una certeza del 100%.

Así por ejemplo, en el ámbito agroalimentario los consensos internacionales se consiguen dentro de la Unión Europea. Fruto de ese consenso se ha elaborado una serie de reglamentos y documentos de referencia, como:

- El Reglamento de Ejecución de la (UE) 2021/808 de la Comisión de 22 de marzo de 2021, que recoge los criterios de funcionamiento de los métodos de análisis aplicados, así como la interpretación de los resultados obtenidos en dichos métodos ⁽¹⁾.
- El documento SANTE/12682/2019⁽²⁾, elaborado por el European Union Reference Laboratory for Pesticides Residues in Fruits and Vegetables y que establece los criterios de control de calidad y de validación de los métodos analíticos.

Además, al amparo de este consenso se ha creado una red de laboratorios de control oficial estructurado en tres niveles: laboratorios oficiales, laboratorios nacionales de referencia, y laboratorios de referencia de la Unión Europea.

Los laboratorios de referencia se encargan de la coordinación técnica de la red de laboratorios. Así, los laboratorios de referencia de la UE coordinan a los laboratorios nacionales de los Estados Miembros de la UE, y estos a su vez a los laboratorios oficiales. Se consigue, de esta forma, que los métodos de diagnóstico y análisis utilizados sean conformes al desarrollo científico-técnico, sean fiables y estén armonizados para que los resultados sean homologables.



En el ámbito toxicológico forense la Oficina de las Naciones Unidas para las Drogas y el Crimen (UNODC) armoniza las actividades de análisis de drogas y ciencias forenses en todo el mundo. Esto permite constituirse en un foro en el que participan organismos forenses nacionales, asociaciones de institutos de ciencias forenses y expertos en análisis de drogas y ciencia forense y que posibilita el desarrollo de estándares y mejores prácticas y facilita la prestación eficiente de servicios científicos y forenses en todo el mundo. Esta organización también emite diversas guías sobre los métodos más adecuados para ser utilizados para el análisis de las diversas drogas.

También en este ámbito se encuadra la European Workplace Drug Testing Society. Esta sociedad se define ⁽³⁾ como un foro independiente para todos los temas relacionados con las pruebas de drogas en los lugares de trabajo y para garantizar que en Europa estas pruebas se realicen con un estándar de calidad definido y de manera legalmente seguro. Si bien es cierto, que esta es una sociedad cuya participación es voluntaria, lo cierto es que las Guías de Referencia que elaboran de análisis de drogas en saliva y orina ^(4,5) son seguidas en muchos laboratorios de control oficial (por ejemplo, los que realizan análisis de drogas en saliva en los controles realizados por policías municipales y Guardia Civil).

En el ámbito del control del dopaje, en el cual la persona que suscribe este informe es experta, la organización encargada de armonizar la actividad técnica de la red de laboratorios antidopaje es la Agencia Mundial Antidopaje (AMA). Esta organización se creó en 1999 como principal organismo para impulsar la lucha contra el dopaje en el ámbito internacional. Es una organización oficialmente reconocida por todos los Gobiernos del mundo ya que la Convención Internacional contra el dopaje en el deporte de 2003 de la UNESCO, firmada por España en 2005 y ratificada en 2006, establece que los Estados firmantes deberán utilizar las medidas legislativas, reglamentos, políticas o disposiciones administrativas necesarias para cumplir con los criterios fijados por la Agencia Mundial Antidopaje, comprometiendo la lucha contra el dopaje por parte de todos los gobiernos firmantes, con el apoyo de una herramienta esencial para garantizar un deporte limpio, el Código Mundial Antidopaje⁽⁶⁾, documento fundamental en el que se basa el Programa Mundial Antidopaje en el deporte. El propósito de dicho Código es promover la lucha contra el dopaje mediante la armonización colectiva y universal de los esfuerzos de organizaciones nacionales antidopaje y de federaciones deportivas internacionales.

Entre las acciones de esta Agencia se encuentra la de elaborar una serie de documentación técnica, que tiene el fin armonizar y consolidar la actividad analítica de la red de laboratorios acreditados, con el fin de garantizar la mayor certeza y fiabilidad en los resultados emitidos por estos laboratorios. En ese sentido, el documento más relevante es el Estándar Internacional de Laboratorios⁽⁷⁾ (ISL), que incluye, los requisitos y el procedimiento para obtener y mantener la acreditación de la Agencia Mundial Antidopaje para ser reconocido como laboratorio oficial antidopaje.

En el artículo 06 de este Código Mundial Antidopaje⁽⁶⁾ se indica que para las muestras de control del dopaje:

- *A los efectos de acreditar directamente un Resultado Analítico Adverso conforme al artículo 2.1, las Muestras se analizarán únicamente en laboratorios acreditados o aprobados por la AMA.*



Es, decir, solo aquellos laboratorios, que han demostrado cumplir con los requisitos del Estándar de Laboratorios de la WADA pueden llevar a cabo los análisis de control del dopaje, si bien es cierto que

los hechos relativos a las infracciones de las normas antidopaje podrán acreditarse por cualesquiera medios fiables, entre ellos, Controles de laboratorio u otros Controles forenses fiables realizados fuera de los laboratorios acreditados o aprobados por la AMA (art. 6 del Código Mundial Antidopaje⁽⁶⁾).

El apartado 5.3.5 de Selección y validación de métodos de ensayo del ISL⁽⁷⁾ indica que los laboratorios deben seleccionar, validar y documentar sus métodos analítico y demostrar que son adecuados para el uso previsto. Este documento indica los parámetros que es necesario validar en función del objetivo del método.

Con relación a la Selección de métodos, el ISL⁽⁷⁾ añade en el apartado 4.4.2, se indica que *“cualquier método que sea nuevo para el ámbito antidopaje debe ser aprobado como adecuado para el uso por la AMA antes de su implementación por cualquier laboratorio. La AMA debería utilizar cualquier medio considerado apropiado, incluyendo consultas formales con grupos de expertos científicos, publicaciones en revistas científicas sujetas a revisión por pares o participación en estudios inter-laboratorio colaborativos o EQAS (intercomparativos) organizados por WADA para evaluar si el método es adecuado para el uso previsto antes de la aprobación final.”*

Estos dos preceptos lo que pretenden de nuevo es dar confiabilidad en el método, demostrar que este ha sido lo suficientemente testeado y refrendado como para que la falsabilidad sea lo menor posible (o prácticamente inexistente).

Los requerimientos específicos que deben demostrar los métodos de análisis así como los controles de calidad a establecer y los requisitos mínimos exigibles a estos, la toma de decisión sobre el resultado final de la muestra teniendo en cuenta el conjunto de resultados analíticos obtenidos y la forma de reporte de estos se establece en otros Documentos Técnicos (por ejemplo, el documento técnico de Límites de Decisión⁽⁸⁾ que aplica a sustancias con nivel umbral TD2022DL, o el documento técnico de Mínimos Niveles de Rendimiento que aplica a sustancias sin nivel umbral, TD2022MRPL⁽⁹⁾, etc). La implementación de los requisitos técnicos contenidos en estos documentos es de obligado cumplimiento por todos los laboratorios antidopaje acreditados, pudiendo los laboratorios ser sancionados en caso de incumplimiento en la implementación de estos requisitos técnicos. De hecho, el apartado 1.1.2 del Estándar de Laboratorios⁽⁷⁾ en referencia a estos documentos técnicos indica que

Un fallo por un Laboratorio [...] en implementar un Documento Técnico o una Carta Técnica en la fecha efectiva de entrada en vigor podría resultar en la imposición de una Restricción Analítica contra el Laboratorio para un Procedimiento de Análisis concreto o la suspensión de la acreditación la Agencia Mundial Antidopaje del Laboratorio.

La elaboración de estos Documentos Técnicos corresponde a la AMA, quien lo hace en colaboración del Grupo de Expertos de Laboratorios (que incluye científicos relevantes de diversos ámbitos, no solo del control del dopaje) o con un Grupo de Expertos de un área muy

determinada. Una vez elaborado ese documento, el borrador es discutido por toda la comunidad científica del ámbito antes de su aprobación definitiva.

A diferencia de otros ámbitos de control oficial, en el ámbito del control del dopaje, la actuación de la AMA no se restringe solamente a la elaboración de documentación de referencia, o a la coordinación de la actividad de los laboratorios antidopaje, sino que al requerir además de la acreditación de la ISO/IEC 17025 (artículo 4.4.2 ISL⁽⁷⁾) la acreditación de la propia AMA, se constituye como organización reguladora y evaluadora de la actividad de los laboratorios. De hecho, en su papel de evaluador, la AMA monitoriza de diversas maneras anualmente el cumplimiento de los requisitos técnicos por parte de los laboratorios, mediante visitas de auditoria, EQAS, cuestionarios, acciones correctivas, seguimiento del reporte de los resultados en la plataforma ADAMS gestionada por la propia AMA, etc.

En el ámbito de control del dopaje, el medio de prueba de que se ha producido una infracción de las normas antidopaje puede ser *cualquier medio fiable*⁽⁶⁾. En el caso de que el medio de prueba de que se trate sea el resultado de un laboratorio el artículo 3.2 del Código Mundial Antidopaje⁽⁶⁾ indica:

3.2.1 Se presume la validez científica de los métodos analíticos o límites de decisión aprobados por la AMA que hayan sido objeto de consultas con la comunidad científica pertinente o de una evaluación inter pares.

3.2.2 Se presume que los laboratorios acreditados y otros aprobados por la AMA realizan análisis de Muestras y aplican procedimientos de custodia que son conformes al Estándar Internacional para Laboratorios.

Esto pretende dar respuesta al principal quebradero de cabeza de un juez (tal como explica Jordi Nieva-Fenoll en su capítulo Repensando Dauber: la Paradoja de la Prueba Pericial⁽¹⁰⁾): la valoración de la prueba pericial.

De hecho, es necesario considerar el trabajo realizado por WADA para conferir a los resultados emitidos por los laboratorios antidopaje de la mayor confiabilidad posible y de proporcionar a los jueces mecanismos para evaluar la confiabilidad del método aplicado. Si se analiza con detenimiento todas esas actividades en su conjunto, podría decirse que se alinea con los preceptos de la doctrina Daubert⁽¹⁰⁾, los cuales pretenden proporcionar criterios orientadores a los jueces para valorar la fiabilidad o no de un dictamen técnico. Tales criterios son:

- Que la técnica utilizada haya sido probada suficientemente frente a errores. (Véase lo expuesto anteriormente en relación a lo indicado en los artículos 5.3.5 y 4.3.2 del ISL en relación a la Selección y validación de métodos de análisis).
- Que la técnica haya revisada por otros científicos y en su caso, publicada (Véase los artículos 4.3.2 y 1.1.2 del ISL).
- Si se conoce el grado de error de la técnica.
- Justificación del mantenimiento de estándares en el uso de la técnica (ténganse en cuenta todas las actividades de elaboración de documentación técnica y de seguimiento de la implementación de esos requisitos técnicos que realiza la AMA.)

- Consenso en la comunidad científica sobre la fiabilidad de la técnica (téngase en cuenta el proceso de aprobación de un nuevo método por la AMA y el proceso de revisión de la documentación técnica de referencia, entre otras actividades).

3. **Cómo entender la fiabilidad de los resultados emitidos por los laboratorios analíticos.**

Una vez argumentado que los métodos aplicados por los laboratorios analíticos de control oficial, y en especial, los laboratorios antidopaje son métodos consolidados, contrastados y capaces de producir resultados con alta fiabilidad y una vez argumentado que en esos ámbitos existen suficientes mecanismos y controles como para que la aplicabilidad de esos métodos se haga de acuerdo a estándares de calidad que permitan obtener resultados con elevada certeza, es necesario explicar el rango de error conocido o posible que se puede atribuir a cada tipo de resultado emitido por estos laboratorios.

La primera distinción que podríamos hacer sobre los métodos analíticos aplicados y sobre los resultados emitidos es en base a que se trate de métodos cualitativos, en los que solamente se requiera evaluar la existencia o no de una determinada sustancia en una muestra o que se trate de métodos cuantitativos, en los que se trata de emitir un valor numérico sobre una determinada propiedad, generalmente la concentración de una sustancia presente en la muestra.

En el caso de resultados cuantitativos en los que se requiere determinar el valor de una determinada propiedad, ya sea la concentración de una sustancia en una muestra o el valor numérico de una relación o de otro tipo de parámetro (por ejemplo el valor de $\Delta\delta^{13}\text{C}\%$, en el caso de análisis mediante cromatografía de gases espectrometría de masas de relaciones isotópicas) se entiende en todos los ámbitos, el jurídico/legal también, que es imposible obtener un valor estrictamente verdadero de dicho parámetro, y que por esa razón es necesario reportar el valor obtenido para dicho parámetro en la medición, e informar de la incertidumbre (margen de error) asociada a ese método. De hecho, los laboratorios invierten realmente muchos recursos y esfuerzos en conseguir (y mantener) que esos márgenes de error sean los menores posibles.

En muchas ocasiones, el resultado se reporta como " $x \pm L$ ", donde x sería el valor del parámetro y L la incertidumbre asociada.

En este caso, se entiende que el verdadero valor del parámetro medido en la muestra se encuentra en el intervalo definido entre $x-L$ y $x+L$.

En ocasiones, al laboratorio solo se le pide que reporte ese valor del parámetro y la incertidumbre asociada a él y es posteriormente el receptor del informe analítico o un tercero quien tomará una decisión en base a ese resultado.

Pero generalmente, es el propio laboratorio quien en base a ese resultado debe emitir un juicio o una declaración sobre la conformidad o no de una muestra (simplificando, sobre si una muestra es positiva o negativa).

En general, siempre que se requiera un método de análisis cuantitativo, la regulación aplicable ha establecido un punto de corte o un límite de decisión, a partir del cual la muestra pasa de ser positiva a negativa, o conforme a no conforme o negativa a adversa en función de la forma de expresión de resultados que requiera cada ámbito.

Además, en cada ámbito suele haber establecida una regla que debe ser tenida en cuenta para realizar esa declaración (si se tiene en cuenta solo el valor de concentración sin tener en cuenta

la incertidumbre para determinar si se sobrepasa un nivel de corte, si se tiene en cuenta que el valor de $x-L$ supere cierto nivel de corte, etc...).

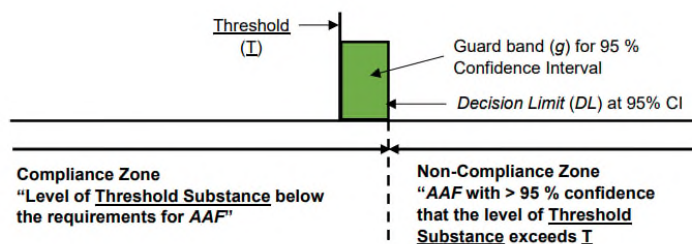
En el ámbito del control del dopaje, el documento TD2022DL⁽⁸⁾ de la AMA establece esa regla de decisión para aquellas sustancias para las que sea necesario tener en cuenta si se sobrepasa o no un determinado nivel de concentración (nivel umbral).

De acuerdo a este documento, para cada sustancia se ha establecido un nivel umbral de concentración. También en base a los diferentes ejercicios de intercomparación realizados en la red de laboratorios antidopaje acreditados por la AMA a lo largo de los años, se ha establecido el margen de error (o la incertidumbre máxima) de las metodologías aplicadas en esta red. Y adicionalmente, en base al valor umbral y teniendo en cuenta esta incertidumbre máxima, se ha establecido un Límite de Decisión.

Como criterio de validación de los métodos analíticos aplicados para esta cuantificación se requiere que la incertidumbre del método sea menor o igual a la incertidumbre máxima.

En cuanto a la toma de decisión, además de la identificación de la sustancia, se establece que se tratará de un Resultado Analítico Adverso si el valor de concentración medido en la muestra, sin tener en cuenta su incertidumbre es superior al DL. Pero si se encuentra entre el valor umbral y el DL, aunque la concentración medida supere el nivel umbral, deberá emitirse como resultado Negativo, ya que si tiene en cuenta el margen de error máximo del método ($x-L_{max}$) existe la posibilidad de que el valor verdadero de la muestra sea inferior al valor umbral.

La figura siguiente extraída del TD2022DL⁽⁸⁾ muestra visualmente esta explicación:



El establecimiento de esos puntos de corte o límites de decisión no es arbitrario y se llega a ellos tras la investigación, el estudio y la validación científica.

Un ejemplo de esto es el desarrollo del método de detección para la detección del uso de hormona de crecimiento en deporte basado en el inmunoensayo diferencial de las isoformas de la hormona del crecimiento y en el posterior establecimiento de los límites de decisión para el ámbito del control del dopaje, para lo cual se llevaron a cabo diferentes estudios, entre los que cabe destacarlos los estudios de Zida Wu et al⁽¹¹⁾ y Wallace JD et al⁽¹²⁾, que fueron los que desarrollaron este método y realizaron los primeros experimentos de validación para demostrar su eficacia. Posteriormente, los estudios realizados por Hanley J.A et al⁽¹³⁾, entre otros, posibilitaron establecer dichos límites de decisión.

Es precisamente por el conocimiento y la validación científica que hay detrás de cada metodología aplicada en el ámbito del control del dopaje, de la que este método es un ejemplo, por la que la AMA incluye en el artículo 3.2 del Código Mundial Antidopaje⁽⁶⁾, la asunción de la "validez científica de los métodos analíticos o límites de decisión aprobados por la AMA".

Esta misma asunción es recogida en el artículo 39d) de la Ley Orgánica 11/2021, de 28 de diciembre, de lucha contra el dopaje en el deporte⁽¹⁴⁾, donde: "Se presume, salvo prueba en contrario, la validez científica de los métodos analíticos y de los límites de decisión que apliquen los laboratorios de control antidopaje debidamente autorizados" (es necesario aclarar que los

laboratorios de control del dopaje debidamente autorizados son los laboratorios acreditados por la AMA y todos ellos deben seguir las regulaciones de la AMA).

En los métodos cuantitativos es muy claro entender la existencia de ese margen de error y cuantificarlo.

Sin embargo, los métodos cualitativos también presentan un margen de error, aunque este no se informe en el resultado y se considere intrínsecamente contenido en él.

El Reglamento de Ejecución de la (UE) 2021/808 de la Comisión de 22 de marzo de 2021⁽¹⁾, que se aplica fundamentalmente en el ámbito agroalimentario, pero que puede ser utilizado como guía de referencia en otros ámbitos, incluye una buena evaluación sobre este aspecto.

Así el Reglamento define:

- *«error alfa (α)»: probabilidad de que la muestra analizada sea conforme, aunque se haya obtenido un resultado de medición no conforme*
- *«error beta (β)»: probabilidad de que la muestra analizada sea en realidad no conforme, aunque se haya obtenido un resultado de medición conforme*

Y establece como requisito para que el método (cualitativo o lo que el Reglamento llama métodos para *sustancias farmacológicamente activas no autorizadas o prohibidas*) demuestre que se ajusta al propósito requerido que:

- *En los métodos de confirmación el error α (lo que se podría denominar falso positivo) será igual o inferior al 1 %.*
- *En los métodos de cribado el porcentaje de falsos conformes (error β , o falsos negativos) sea igual o inferior al 5 %.*

En la validación del método se deben evaluar los siguientes parámetros, que servirán además como límites de referencia para tomar una decisión (conforme/no conforme) sobre el resultado analítico:

- *«límite de decisión para confirmación ($CC\alpha$)»: límite en el cual, y a partir del cual se puede concluir con una probabilidad de error de α que una muestra no es conforme, y el valor $1 - \alpha$ significa certeza estadística en porcentaje de que se ha superado el límite permitido.*
- *«capacidad de detección del cribado ($CC\beta$)»: contenido mínimo de analito que puede ser detectado [...] en una muestra con una probabilidad de error de β , es decir, es la concentración más baja en la que un método es capaz de detectar [...], con una certeza estadística de $1 - \beta$, muestras que contengan residuos de sustancias prohibidas o no autorizadas.*

Esto significa que, en el caso de los métodos de cribado (primer filtro o la primera toma de decisión sobre el resultado de la muestra) cuando se determina que la muestra es *conforme* este resultado no representa una certeza absoluta o del 100%, ya que intrínsecamente el método se ha validado para demostrar que cuando se emite ese resultado conforme, ese resultado tiene una fiabilidad mínima del 95%. Es decir, existe un 5% de muestras para las que se emite ese resultado conforme que en realidad podrían contener residuos de sustancias prohibidas al nivel de concentración determinado como $CC\beta$; a medida que la concentración fuese más elevada respecto de ese nivel $CC\beta$ el error disminuiría. En el caso de los métodos de confirmación,

cuando se emite un resultado *no conforme* este resultado también lleva intrínseco un margen de error. En este caso, la fiabilidad mínima es del 99%, si bien podría existir un 1% de muestras a nivel de concentración $CC\alpha$ para las que se emite un resultado no conforme que en realidad no contienen residuos de sustancias prohibidas.

En el ámbito del control del dopaje, existen conceptos similares, aunque con ciertos matices, si bien no se encuentran tan claramente detallados en la documentación técnica aplicable. En el Documento Técnico TD2022MRPL⁽⁹⁾ Minimum required performance levels and applicable minimum reporting levels for non-threshold substances analyzed by chromatographic - mass spectrometric analytical methods se define el Límite de Detección (límite para la toma de decisión del método de cribado) como la mínima concentración del analito que puede ser detectado, con un mínimo rango de detección del 95%. Es decir, al igual que en el ámbito agroalimentario, cuando se determina que una muestra es Negativa existe la posibilidad de que un 5% de muestras que pudieran tener sustancias prohibidas a nivel de concentración del Límite de Detección fuesen no detectadas. Este rango de detección se incrementa a medida que la sustancia prohibida se encuentra a mayor concentración en la muestra.

El TD2022MRPL⁽⁹⁾ establece el mismo rango de error para el Límite de Identificación (límite para la toma de decisión del Resultado Analítico Adverso) que para el Límite de Detección: el 5%. Pero en este caso, y a diferencia del ámbito agroalimentario, ese 5% se refiere errores en la no detección de sustancias prohibidas. Es decir, de nuevo, al nivel de Límite de Identificación existiría un 5% de muestras para las cuales no sería posible identificar las sustancias prohibidas y serían reportadas como falsos negativos. Igualmente, esta tasa de error disminuye (o el rango de identificación se incrementa) a medida que la sustancia prohibida se encuentra a mayor concentración en la muestra.

En el ámbito del control del dopaje se requiere que los métodos de confirmación demuestren que la tasa de probabilidad de producir un falso positivo sea del 0%, es decir, que no sea posible. Un método de confirmación que no fuese capaz de demostrar esa tasa de error no sería aceptable para el objetivo previsto. Es decir, se requiere que en la emisión de un Resultado Analítico Adverso de acuerdo a un método concreto no exista margen de error.

Sin embargo, aunque este requerimiento sea una garantía de que el grado de fiabilidad con que se emiten los resultados en los laboratorios antidopaje es cercano al 100%, la emisión de un Resultado Analítico Adverso no puede estar exenta de cierto margen de error, aunque ese margen pueda ser mínimo.

Existe cierta polémica sobre la verdadera ratio de casos falsos negativos y falsos positivos entre deportistas testeados, ya que este no se conoce y por eso genera mucha especulación y debate. Si bien esta discusión lleva la evaluación de esta ratio no solo hacia la tasa de error de un método analítico concreto aplicado, sino a todo el procedimiento de control del dopaje (incluyendo la selección de deportistas, la planificación y el desarrollo de la toma de muestra, el análisis del laboratorio y la gestión del resultado posterior)⁽¹⁵⁾.

Si se atiende solamente a los Resultados Analíticos Adversos emitidos por los laboratorios antidopaje en las estadísticas que la AMA⁽¹⁶⁾ publica anualmente se puede comprobar que en los últimos años 2020-2021 el número de casos adversos reportados se encuentra entre 1000-

1500 en toda la red de laboratorios. Haciendo una aproximación muy burda podríamos concluir que la tasa de error de emisión de falsos positivos por los laboratorios antidopaje es muy improbable o inexistente, ya que para que esta fuera del 1% se requeriría que al menos 10-15 casos reportados hubieran sido falsos positivos, lo cual no es el caso. Con un solo caso falso positivo por año el error en porcentaje se situaría entre el 0,06-0,1%.

Sin embargo, aunque es cierto que es posible, por diversas razones que se explicaran más abajo, emitir Falsos Resultados Analíticos Adversos la emisión de este tipo de resultados es realmente excepcional y muy improbable, por lo que cabe pensar que la tasa de error es mucho menor que ese 0,06% (que sería la tasa de error en un punto ciertamente extremo).

Esto tiene cierta lógica ya que los laboratorios antidopaje están altamente presionados para trabajar con estándares de calidad de excelencia. El propio Estándar de Laboratorios⁽⁷⁾ indica en su artículo 7.2.1 que *“Un Falso Resultado Analítico Adverso no es admisible para ninguna muestra de intercomparación ciega o doble ciega ni durante los análisis rutinarios llevados a cabo por el Laboratorio.”* Además, en el supuesto de producirse y desde el momento que se descubriese que se ha reportado un Falso Resultado Analítico Adverso, de acuerdo al ISL⁽⁷⁾, el Laboratorio debería cesar la actividad analítica. En función de la causa del error y de las actividades realizadas posteriormente por el Laboratorio para corregir y remediar ese error el Laboratorio podría ser sancionado con la pérdida de su acreditación.

Teniendo en cuenta estas consideraciones, cuando un laboratorio antidopaje emite un Resultado Analítico Adverso este puede considerarse casi fiable al 100%.

No obstante, es cierto que, aunque mínima, existe una tasa de error y no se puede tener el 100% de certeza, ya que pueden existir diferentes factores que escapan al control.

Por poner un ejemplo, el proceso en el laboratorio desde la recepción de muestra hasta la emisión de un Resultado Analítico Adverso conlleva muchas etapas en las que participan muchos operadores. A pesar de en todas las etapas existen controles para detectar y evitar posibles errores, podría existir la posibilidad de, bien por errores administrativos, bien por errores de aplicación de la metodología aprobada, se pudieran emitir un falso positivo. De hecho en toda la redacción del artículo 7.0 Evaluación de los ejercicios de intercomparación de Laboratorios y desempeño de los procedimientos rutinarios de análisis del ISL⁽⁷⁾ de la AMA se considera esta posibilidad.

Además, existen otras posibles fuentes de contribución a la tasa de error sobre los resultados analíticos, generadas precisamente por ese carácter de falsabilidad que define al método científico.

Así, a modo de ejemplo, la identificación de la sustancia en la muestra, incluso utilizando una técnica tan inequívoca como la cromatografía acoplada a la espectrometría de masas podría conducir a un potencial error.

En la identificación por cromatografía y espectrometrías de masas se trata de comparar la señal de cromatografía y espectrometría de masas de la sustancia en la muestra y del estándar de la sustancia prohibida en una muestra control y comprobar que ambas tienen las mismas características.



Para realizar esa comparación, muchos ámbitos han establecido regulaciones. Así el Reglamento de Ejecución de la (UE) 2021/808 de la Comisión de 22 de marzo de 2021⁽¹⁾, el Documento Técnico de la AMA TD2023IDCR⁽¹⁷⁾ o el documento SANTE/12682/2019⁽²⁾ han establecido primeramente el número de puntos de identificación (o parámetros a medir y comparar para esa identificación) que se pueden obtener por cada parámetro/técnica aplicada y el número mínimo de puntos de identificación a considerar.

Como parámetros a evaluar para la identificación por cromatografía y espectrometría se incluyen: el tiempo de retención del pico cromatográfico, la masa/carga nominal de cada ion diagnóstico en la muestra comparados con la del estándar, el número mínimo de iones diagnóstico a tener en cuenta, la tolerancia de las relaciones de abundancia iónica entre ellos, el error de masa con el que extraer iones de diagnóstico cuando se utilizan detectores de masa exact, etc.

Al tratarse de mediciones del valor de un parámetro, y como se ha explicado anteriormente, es imposible conocer el verdadero. Por lo que cuando se aplica la comparación de determinados parámetros en las señales de la sustancia en la muestra y de un estándar de referencia conocido en la muestra control se establece un margen de tolerancia alrededor de los valores obtenidos en el estándar de referencia dentro del cual se debe encontrar el valor del parámetro de la sustancia detectada en la muestra.

Esto ya de por sí, está sometido al propio error intrínseco de las mediciones analíticas por técnicas instrumentales, el cual oscila entre 0,6 y 3,0%⁽¹⁸⁾.

Uno de los parámetros necesarios para la validación de un método es la selectividad, que es la capacidad de discriminación entre el analito de interés y sustancias similares o estructuralmente similares. La selectividad tiene especial relevancia en la identificación, ya que se trata de identificar la sustancia prohibida en la muestra por comparación con el estándar de referencia y discriminándolo de cualquier otro compuesto del planeta que pudiera presentar unas características o parámetros similares

Este parámetro se evalúa durante la etapa de validación del método conforme a las regulaciones contenidas en el TD2023IDCR⁽¹⁷⁾. Al plantear estas regulaciones estableciendo un mínimo número de identificación y un rango de tolerancia con una amplitud determinada se ha considerado una situación de equilibrio con el objetivo de no requerir unos criterios de comparación demasiado restrictivos como para aumentar la probabilidad de los falsos negativos, ni tampoco unos criterios demasiado laxos como para incrementar la posibilidad del falso positivo. Es por esto, por lo que B.J.A. Berendsen et al⁽¹⁹⁾ indican que existe una “incertidumbre” en la selectividad o en la capacidad de confirmar una sustancia por espectrometría de masas.

Esta discusión no pretende generar dudas sobre el carácter inequívoco de la cromatografía/espectrometría de masas para la identificación, ya que esta junto con la resonancia magnético nuclear son las herramientas analíticas de identificación de sustancias más potentes conocidas hasta la fecha, sino que solo se pretende ilustrar que incluso en la aplicación de esta técnica (y como ocurre con cualquier técnica de laboratorio) existe cierta tasa de error. De nuevo, existen mecanismos para disminuir ese margen de error, como es el

incremento en la exigencia de puntos mínimos de identificación o en la reducción de los márgenes de tolerancia para la comparación, que es lo que explica B.J.A. Berendsen et al⁽¹⁹⁾ en su trabajo.

Pero la “incertidumbre” en la identificación, no solo podría venir dada solo por las propias mediciones de la técnica en sí, sino también por la interpretación que pueda hacerse de los resultados analíticos obtenidos.

En el control del dopaje se trata de determinar el uso de sustancias o métodos prohibidos en el deporte en muestras fisiológicas, generalmente orina.

Una de las dificultades de los análisis de control del dopaje es que las sustancias prohibidas que son consumidas por los deportistas son modificadas en el organismo y lo que se excreta en la orina son los productos modificados (metabolitos o productos de degradación), sin que en ocasiones queden restos de la sustancia prohibida que se administró.

Aunque el Código Mundial Antidopaje permite demostrar el uso de sustancias prohibidas a través de los marcadores o metabolitos, una de las posibles fuentes de error es que esos marcadores o productos de degradación pueden proceder no solo de una sustancia prohibida, sino de otra sustancia, que incluso a veces puede ser una sustancia cuyo uso este permitido. Algunas de estas posibilidades ya son conocidas por la red de laboratorios antidopaje y están documentadas en las Cartas Técnicas de la AMA.

Como, por ejemplo, el ácido 4-clorofenoxiacético (4-CPA) que es el principal producto de degradación del meclofenoxato, pero que también puede proceder de herbicidas y de diferentes cosméticos que contengan clorfenesina⁽²⁰⁾.

Otro ejemplo, se puede encontrar en la para-hidroxi-anfetamina. Este es un estimulante que por sí mismo es sustancia prohibida, pero también es un metabolito de la sustancia prohibida anfetamina. Además, puede proceder de la sustancia no prohibida mebeverina⁽²¹⁾.

Generalmente, en estos casos ya documentados previo a emitir un Resultado Analítico Adverso es necesario realizar análisis adicionales para determinar si el origen de la presencia de ese marcador en la muestra se debe al uso o administración de una sustancia prohibida o a otra causa.

Sin embargo, y aquí es donde entra el juego la “incertidumbre”, no se conocen todas las potenciales sustancias que podrían producir metabolitos o marcadores comunes a los que se determinan para demostrar el uso de una sustancia prohibida y eso obliga a que previo a emitir un resultado, los laboratorios hagan una tarea de investigación para la correcta interpretación de los resultados, e incluso puedan solicitar una segunda opinión a otro laboratorio acreditado, especialmente para aquellos casos en los que el laboratorio se enfrenta por primera vez a la detección de una sustancia que nunca antes detectó.

Además, y esto ayuda a para minimizar este riesgo, la AMA indica en el ISL⁽⁷⁾ que es requisito de acreditación la transferencia de conocimiento entre laboratorios. Por lo que cuando un laboratorio encuentra y demuestra un hallazgo de este tipo debe compartirlo con el resto de laboratorios. La propia AMA sirve de organismo regulador y controlador que permite la

transferencia e implantación de este nuevo conocimiento que se pudiera generar entre todos los laboratorios de la red.

Estas medidas por supuesto minimizan la posibilidad de error de reportar un falso positivo, pero no pueden eliminarlo al 100%.

Además de las posibles fuentes de error en la identificación de sustancias prohibidas, existen otras fuentes de error cuyo origen sea el propio segundo principio del método científico, el de la falsabilidad. Es decir, que aparezca un nuevo experimento que refute o haga reconsiderar las teorías anteriores.

Un ejemplo de posibles fuentes de error de este tipo lo tendríamos en el caso del clenbuterol. El clenbuterol ha estado incluido en la Lista de Sustancias Prohibidas ⁽²²⁾ de la AMA desde la primera lista emitida en el año 2003. Su mera presencia en la muestra de orina era suficiente demostración de que se trataba de un Resultado Analítico Adverso, ya que se suponía que su presencia en la muestra se debía a una administración externa con fines dopantes. En los últimos años se ha descubierto que en determinados países (China, Guatemala y México) se utiliza de forma no controlada en el engorde del ganado, por lo que la ingesta de carne de ganado contaminada con clenbuterol es el origen de muchos de los hallazgos de clenbuterol en la población (incluidos deportistas) de esas regiones ⁽²³⁾. Es por ello, que desde que se conoce este factor de confusión, en la toma de decisión sobre si se ha cometido o no una infracción de una norma antidopaje, cuando un laboratorio detecta clenbuterol en una muestra de orina de un deportista por debajo de cierto nivel el resultado debe ser reportado como Resultado Anómalo y no como Resultado Analítico Adverso. Eso significa, que se requieren realizar investigaciones adicionales con el fin de determinar si el origen de ese hallazgo se debe al consumo de carne contaminada o al consumo intencionado de esta sustancia.

Otro posible ejemplo se encuentra en el reciente descubrimiento de la variante c.577del en el gen humano de la EPO en algunos deportistas chinos ⁽²⁴⁾. En una población testeada de 5656 muestras se ha comprobado que la frecuencia de aparición de esta variante es del 0.39%. Este gen tiene la particularidad de producir una EPO endógena mutada que presenta una cadena de 27 aminoácidos mayor que la EPO endógena natural. El peso molecular de esta EPO endógena mutada es aproximadamente igual al de la EPO recombinante, por lo que su comportamiento en el método de detección mediante SAR-PAGE para la detección de ERAS, que se basa en una discriminación por peso molecular, es similar al de la EPO recombinante, dificultando la discriminación entre una y otra.

Este nuevo hallazgo, no invalida el método anterior (especialmente porque de inicio la tasa de frecuencia de aparición de la variante es baja), pero lógicamente sí hace necesaria la realización de investigaciones adicionales. De hecho, y en este sentido, el documento técnico TD2022EPO⁽²⁵⁾ indica en su anexo B toda una serie de investigaciones adicionales a realizar para discernir si el deportista es portador de la variante c.577del EPO o no.

4. La fiabilidad del Pasaporte Biológico del Deportista.

En cuanto a la fiabilidad de los resultados emitidos por los laboratorios antidopaje, en este caso aquellos que tienen una Unidad de Gestión del Pasaporte del Deportista (Athlete Passport



Management Unit, APMU) de acuerdo al TD2021APMU ⁽²⁶⁾ de la AMA, mención aparte requieren los resultados relativos al Pasaporte Biológico del Deportista ⁽²⁷⁾.

El concepto de pasaporte biológico del deportista (ABP) se desarrolló a principios de los años 2000, aunque la metodología oficialmente no se introdujo hasta el año 2009 y gracias numerosos estudios e investigaciones que convergieron, constituyéndose desde entonces como una metodología suficientemente sensible y específica como proveer de la suficiente evidencia científica para probar el dopaje a través de marcadores indirectos.

Con el módulo hematológico del ABP, la sangre no fue únicamente utilizada para la detección directa de parámetros, sino también como un marcador robusto de la estimulación eritropoyética en un seguimiento individual longitudinal.

Esta aproximación de detección indirecta se considera adecuada porque estos marcadores biológicos indirectos pueden revelar variaciones anormales en la sangre potencialmente inducidas por dopaje. Desde los años 90 se sabía que algunos marcadores de la sangre como la hemoglobina o el hematocrito respondían a todas las formas de dopaje, o que el uso de EPO recombinante afectaba a los reticulocitos.

Posteriormente, se desarrolló una aproximación multiparamétrica combinando todos los datos en un único perfil sanguíneo del deportista utilizando técnicas específicas de clasificación de patrones.

El Abnormal Blood Profile Score (ABPS) (P Sottas et al ⁽²⁸⁾, F. Schütz ⁽²⁹⁾) es un marcador multiparamétrico universal que se basa en 12 diferentes marcadores sanguíneos indirectos que podrían verse modificados por la administración de EPO recombinante o por el uso de transfusiones sanguíneas. Aunque en la práctica solo 7 de esos marcadores (porcentaje de reticulocitos, concentración de hemoglobina, nivel de hematocrito, conteo de glóbulos rojos, volumen medio corpuscular, hemoglobina media corpuscular, concentración corpuscular media de hemoglobina) son utilizados y combinados en una única puntuación (score). De acuerdo a la publicación original del ABPS (P Sottas et al ⁽²⁸⁾), esta puntuación combinada es más sensible para el mismo nivel de especificidad para detectar el dopaje que cualquiera de esos parámetros sanguíneos evaluados de forma separada. En particular, el ABPS por si solo permite la detección de diferentes tipos de dopaje sanguíneo (P Sottas et al ⁽²⁸⁾).

Actualmente el módulo hematológico está compuesto además de por el ABPS por otros dos marcadores primarios que son monitorizados con remarcablemente estrictas guías de operación: la concentración de hemoglobina ([Hb]) y el índice de estimulación de eritropoyesis "OFF-Score" (OFFs), calculado como $OFFs = [Hb] - 60 \times \sqrt{Ret\%}$, incluyendo la medida de los reticulocitos como porcentaje (Ret%) y la concentración de hemoglobina en $g L^{-1}$. ⁽³⁰⁾ (Anexo C ISRM AMA ⁽³¹⁾).

Estos dos marcadores primarios son los marcadores más específicos para detectar el dopaje, mientras que los marcadores secundarios (el ABPS y el porcentaje de reticulocitos) proporcionan evidencias que permiten apoyar el dopaje bien por sí mismos, bien por combinación con otros marcadores.

El ABP se implementa con un modelo probabilístico adaptativo basado en la aproximación Bayesiana para determinar la probabilidad de que las variaciones hematológicas de un deportista tengan un origen prohibido (Sottas et al. ⁽²⁷⁾).

Por ello, el efecto del dopaje no se examina vía la modificación directa de los parámetros biológicos, sino indirectamente vía la variación en la media y desviación estándar de esos biomarcadores a través del seguimiento longitudinal individual de dichos parámetros.



Todos los biomarcadores sanguíneos presentan una variabilidad intraindividual tal que posibilitan el seguimiento longitudinal de dichos parámetros. Este seguimiento se construye de la siguiente manera: Cuando se toma una primera muestra, los valores umbral superior e inferior se determinan con puntos de referencia basados en medias poblacionales. Estos límites individuales son posterior y subsecuentemente adaptados basándose en cada uno de los valores del deportista a medida que otras muestras se van recogiendo.

Además, la estructura del modelo Bayesiano es lo suficientemente flexible como para tener en cuenta ciertos efectos de factores heterogéneos, los cuales tienen afectan a la definición de los límites individuales. Estos factores pueden ser dependientes del tiempo o independientes del tiempo. Entre los factores heterogéneos se incluyen el género, el origen étnico, la edad, el tipo de deporte, la exposición a la altitud.

Este modelo cuantitativo Bayesiano permite detectar lo que se denomina un Resultado de Pasaporte Atípico (ATPF). Este resultado se produce cuando un marcador cae fuera del rango intraindividual del deportista, asumiendo una condición física normal (es decir, que respecto del resto de valores de ese mismo marcador ese valor concreto sea considera un outlier). El nivel de especificidad para considerar un outlier es del 99%. Esto quiere decir que, existe un margen de fiabilidad del 99% de que ese valor no se deba a un origen fisiológico natural y que por lo tanto su origen se deba al uso de una sustancia o de un método prohibido en deporte (Anexo C de ISRM de la AMA ⁽³¹⁾). Cuando se produce una secuencia de desviaciones o de valores outliers, esta especificidad aumenta y se sitúa en el 99,9% (es decir, que la oportunidad de que esos valores outliers se deban a una variación fisiológica normal se reduce al 1:1000) ⁽³¹⁾.

A pesar de esto, y con el objeto de demostrar con mayor de fiabilidad que se ha cometido dopaje, el Estándar Internacional de Gestión de Resultados de la WADA (ISRM) describe en su anexo C los pasos necesarios para revisar un Resultado de Pasaporte Atípico hasta concluir si se confirma o no que se trata de una Resultado de Pasaporte Adverso (Adverse Passport Finding). Cuando la APMU detecta un Resultado de Pasaporte Atípico, debe enviar este a un Experto del panel de Expertos.

Para estos Expertos se exigen ciertos requisitos de cualificación muy específicos (TD2023APMU ⁽²⁶⁾): tener cualificaciones en uno o más de los campos de hematología clínica o de laboratorio, medicina del deporte y fisiología del ejercicio y su aplicación al dopaje sanguíneo. Además, conocimientos complementarios tal que se representen todos los campos relevantes.

Este Experto evalúa la información contenida en el Pasaporte Biológico del Deportista, así como otra información relevante para la evaluación que pudiera estar disponible (por ejemplo, el calendario de competiciones, u otra información adicional que se recoge cuando se toma la muestra al deportista, como entrenamientos en altitud, entrenamientos simulando condiciones de altitud, si ha sido sometido a transfusiones sanguíneas con fines médicos o ha sufrido pérdidas importantes de sangre debido a accidente o donación, incluso condiciones especiales de temperatura durante la toma de muestra si las hubo, etc).

En base a toda esa información el Experto aplica un “likelihood ratio” ⁽³²⁾ (LR, lo que en español se traduce como cociente de probabilidades o razón de verosimilitud). Este es un método estadístico objetivo ampliamente utilizado en los ámbitos clínico (para el diagnóstico) y forense para interpretar evidencias y obtener conclusiones (por ejemplo, se emplea en la toma de decisiones sobre la prueba de ADN (Prieto L ⁽³³⁾), ya que es un método que permite asegurar la evidencia en relación a dos hipótesis simples estadísticas (en este caso la probabilidad de dopaje



o la probabilidad de condición médica). El LR no permite una interpretación arbitraria, ya que es una forma objetiva de cuantificar cuanto más ciertos datos apoyan una hipótesis sobre la otra.

Prieto ⁽³³⁾ lo explica muy claramente en un ejemplo de interpretación de una evidencia de ADN aplicando el LR:

El perito forense puede evaluar los resultados de su analítica desde dos puntos de vista contrarios (desde la perspectiva de la acusación y desde la perspectiva de la defensa)[...]. Para calcular un LR es necesario enunciar al menos dos hipótesis sobre los hechos que deben ser excluyentes (si una es cierta, la otra debe ser falsa), por ejemplo:

Ha (hipótesis de la acusación) = el perfil genético hallado en la escena del crimen pertenece al acusado

Hd (hipótesis de la defensa) = el perfil genético hallado en la escena del crimen NO pertenece al acusado.

El LR [...] mide cuántas veces es más probable haber obtenido esos resultados genéticos si suponemos que el acusado es el donante en comparación al supuesto de que otro individuo sea el donante de ese perfil genético hallado en la escena del delito. Y se formula de la siguiente forma:

$$LR = \frac{P(E|Ha) \text{ probabilidad de obtener los resultados de la prueba genética suponiendo que el perfil pertenece al acusado}}{P(E|Hd) \text{ probabilidad de obtener los resultados de la prueba genética suponiendo que el perfil NO es del acusado}}$$

siendo E = evidencia (el resultado genético en la muestra hallada en la escena y en la muestra biológica del acusado), P = probabilidad y el símbolo “|” = suponiendo.

Supongamos que el perfil genético hallado en la escena coincide perfectamente con el perfil hallado en la muestra indubitada del acusado. Evidentemente, bajo el supuesto de que el acusado dejó su perfil (Hp), encontraremos ese perfil en la escena con probabilidad 1 (con probabilidad del 100% en forma de porcentaje, es decir, siempre) pues no puede aparecer en la evidencia un perfil genético distinto al del acusado si él es el donante. Por tanto, el numerador del LR será 1 en este caso: P(E|Hp) = 1.

Pero bajo el supuesto de que el perfil de la escena NO pertenece al acusado, la probabilidad de la evidencia cambia. Si el perfil no es del acusado tiene que pertenecer a alguien de iguales características al acusado (con el mismo perfil genético) y por tanto esta probabilidad se traduce en la frecuencia con que ese perfil genético aparece en la población (por ejemplo en 5 de cada 10 millones, es decir 5/10.000.000 = 0,0000005). Por tanto, el denominador del LR será en este ejemplo: P(E|Hd) = 0,0000005. Ahora sólo tenemos que calcular el LR total: LR = 1/0,0000005 = 2.000.000.

Pero ¿qué significa realmente este resultado? Significa que es 2 millones de veces más probable hallar el perfil genético encontrado en la escena si suponemos que lo dejó el acusado (Hp) que si suponemos que lo dejó otra persona (Hd). Es decir, la prueba muestra un resultado a favor de la hipótesis de la acusación (2 millones de veces más a favor de la acusación respecto a la defensa).

El Experto evalúa el hallazgo del Resultado de Pasaporte Atípico en base a la hipótesis de que se deba a dopaje frente a la hipótesis de que se deba a una condición médica. En función de la

cuanto mayor sea la probabilidad de una hipótesis frente a la otra, el Experto categoriza el hallazgo del Resultado de Pasaporte Atípico en (Anexo C de ISRM de la AMA ⁽³¹⁾):

“Normal”, “Sospechoso”, “Probable dopaje”, “Probable condición médica”.

En caso de que la ratio permita determinar el hallazgo se debe con una alta probabilidad al uso de una sustancia o método prohibido, frente a una muy baja probabilidad de que se deba a una condición médica, la conclusión es de “probable dopaje”.

Al tratarse este de un cualitativo, en la expresión del resultado no se indica ni la tasa de error, ni la relación de probabilidad con que se emite ese resultado, ya que a la propia conclusión de ese resultado (“probable dopaje”) se ha llegado una vez se ha visto que la ratio de probabilidad es tal que “cuasi” es cierto que el hallazgo se debe a dopaje. De lo contrario, el resultado a emitir hubiera sido otro (sospechoso, normal o probable condición médica).

Con el fin de incrementar el grado de garantía y confianza sobre ese resultado, el ISRM de WADA requiere aun subsiguientes evaluaciones antes de considerar el hallazgo como un Pasaporte Biológico Adverso.

Así, la misma información que evaluó el primer Experto es remitida a otros dos Expertos adicionales, que de forma independiente y separada procederán a su revisión de acuerdo a la misma sistemática de toma de decisión que aplicó el primer Experto, es decir, a través de la evaluación del likelihood ratio.

Para continuar con el proceso necesario para la emisión de un Resultado de Pasaporte Adverso, se requiere que los tres (3) Expertos de forma separada e independientemente emitan la conclusión “Probable dopaje”, en base a la misma herramienta de toma de decisión definida anteriormente.

Si los tres Expertos emiten esa conclusión, esto se le notificará al deportista y se le solicitará que aporte cualquier información o explicación pertinente que pueda justificar ese hallazgo de Resultado de Pasaporte Atípico que motivó toda la posterior investigación.

Esa información es remitida al Panel de Expertos, los cuales de nuevo en base a esa nueva información volverán a emitir una conclusión. Solamente si los tres Expertos concluyen que existe una alta probabilidad de que ese Resultado de Pasaporte Atípico, que se había obtenido con una especificidad del 99% en caso de un hallazgo puntual y de un 99.9% en el caso de una secuencia de hallazgos, se deba a dopaje frente a una baja probabilidad de que se deba a una condición médica, se podrá emitir un Resultado de Pasaporte Adverso.

Es por ello, que todo este mecanismo de interpretación y de toma de decisión sobre el resultado del Pasaporte Biológico del Deportista permite emitir conclusiones con una alta certeza, similares a los de otros resultados del ámbito del control del dopaje o de otros ámbitos (como el forense).

Otro de los aspectos del Pasaporte Biológico del Deportista que en ocasiones es cuestionado es que no permite determinar la identidad de la sustancia o del método prohibido que se ha utilizado para cometer dopaje.



Es cierto que los métodos de detección directa de sustancias o de uso de un método tienen ciertas limitaciones, porque no siempre es posible llevar a cabo esa determinación. Un claro ejemplo se encuentra en las dificultades técnicas para la discriminación entre el origen endógeno o exógeno (por administración) de ciertas sustancias que están presentes de forma natural en el organismo (por ejemplo, la hormona del crecimiento, los esteroides como la testosterona) o en las dificultades técnicas para detectar las transfusiones autólogas.

Por ello, la aplicación de métodos de detección indirecta permite a los laboratorios antidopaje extender su metodología para dar respuesta a las necesidades que se les exigen, es decir proveer de evidencias científicas sobre si se ha producido dopaje o no.

El Pasaporte Biológico del Deportista, que como se ha explicado más arriba se basa en el efecto que el uso de ciertas sustancias o ciertos métodos produce sobre determinados marcadores sanguíneos, no es el único método de detección indirecta que se aplica en el control del dopaje.

Así por ejemplo, el Ensayo de Biomarcadores de la Hormona de la Crecimiento⁽³⁴⁾ es otro método de detección indirecta autorizado y reconocido por la AMA que no aísla e identifica la hormona del crecimiento, tal como hacen otros métodos clásicos de control del dopaje, incluido el Inmunoensayo Diferencial de las Isoformas de la hormona del crecimiento⁽³⁵⁾, sino que este método se basa en medir los efectos que la administración de hormona del crecimiento produce sobre dos biomarcadores: el Factor de Crecimiento-I similar a la insulina (IGF-I) y el pro péptido N-terminal de colágeno tipo III (P-III-NP). Además, la toma de decisión no se basa en la variación directa en esos marcadores, sino en una puntuación (o score) que se obtiene mediante la combinación de las concentraciones de dichos marcadores y que tienen en cuenta además el género y la edad.

El análisis de confirmación mediante Cromatografías de Gases/Combustión/ Espectrometría de masas de relaciones isotópicas (GC/C/IRMS)⁽³⁶⁾ utilizado en la confirmación de Resultados Atípicos de Pasaportes Esteroides se basa también en una medición indirecta y no permite identificar la sustancia administrada.

En este caso, la medición indirecta se basa en determinar el incremento en el contenido en carbono-13 (¹³C) que presentan ciertos esteroides endógenos marcadores. El contenido en ¹³C en el organismo es constante, mientras que ciertos esteroides, que también se producen de forma endógena, cuando son administrados de forma exógena, tienen un contenido en ¹³C menor al del organismo. Esto es debido al menor contenido en ¹³C de los compuestos de origen que se utilizan para sintetizar estos compuestos (diosgenin y estigmasterol, que proceden de plantas). Esos esteroides administrados de forma exógena siguen la vía de metabolización de la testosterona y de otros marcadores esteroideos, haciendo que el contenido en ¹³C de estos varíe. En el análisis de confirmación mediante GC/C/IRMS se trata de determinar ese contenido en ¹³C de los marcadores del perfil esteroideo (que son solo cinco) mediante la medición del $\delta^{13}\text{C}\text{‰}$ y la comparación con el valor obtenido para un compuesto endógeno de referencia⁽³⁷⁾.

En este caso, el Resultado Analítico Adverso se emite cuando ese contenido (medido como $\Delta\delta^{13}\text{C}$) supera cierto nivel. La confirmación no permite la identificación de la sustancia prohibida concreta empleada y tan solo se puede concluir que el resultado obtenido es consistente con la

administración exógena de esteroides anabolizantes androgénicos endógenos (lo cual engloba toda una categoría de sustancias).

Finalmente, y como conclusión, toda la argumentación anterior tiene el objetivo de no generar dudas sobre la fiabilidad de los resultados emitidos por los laboratorios analíticos, y en especial los del ámbito del control del dopaje, sino hacer entender que es imposible que los laboratorios emitan resultados totalmente ciertos (en sentido estricto) y que la indubitabilidad de estos siempre va a estar acompañada intrínsecamente de cierto margen de error. El objetivo de los científicos, en especial cuando en base a esos resultados es necesario llevar a cabo actividades con alto impacto en las personas individuales, en la sociedad, en la economía, etc, es aplicar todos los conocimientos científicos y tecnología necesaria para garantizar que se puedan emitir juicios, resultados y conclusiones lo más cercanas a la verdad posible, es decir con una tasa mínima de error o inexistente.

5. Bibliografía

- ⁽¹⁾ Reglamento de Ejecución de la (UE) 2021/808 de la Comisión de 22 de marzo de 2021, que recoge los criterios de funcionamiento de los métodos de análisis aplicados, así como la interpretación de los resultados obtenidos en dichos métodos.
- ⁽²⁾ Document N^o SANTE/12682/2019. European Union Reference Laboratory for Pesticides Residues in Fruits and Vegetables. Analytical Quality Control and Method Validation Procedures for Pesticide Residues Analysis in Food and Feed. Valid since 1 January 2020.
- ⁽³⁾ European Workplace Drug Testing Society Bylaws. Valid from 2017.
- ⁽⁴⁾ European Guidelines for Workplace Drug Testing in Urine. 2022-10 Version 3.0 FINAL
- ⁽⁵⁾ European Guidelines for Workplace Drug Testing in Oral Fluid. 2022-11-01 Version 3.0 FINAL
- ⁽⁶⁾ Código Mundial Antidopaje. World Antidoping Agency. Valid since 01 de enero de 2021.
- ⁽⁷⁾ International Standard for Laboratories. World Antidoping Agency. Valid since 01 enero 2021.
- ⁽⁸⁾ TD2022DL v 1.0. World Antidoping Agency. Valid since 01 January 2022.
- ⁽⁹⁾ TD2022MRPL v 1.0. World Antidoping Agency. Valid since 01 January 2022.
- ⁽¹⁰⁾ Repensando Dauber: la Paradoja de la Prueba Pericial. Jordi Nieva-Fenoll. Civil Procedure Review, v.9, n.1: 11-26, jan.-apr., 2018 ISSN 2191-1339 – www.civilprocedurereview.com
- ⁽¹¹⁾ Wu Z, Bidlingmaier M, Dall R, Strasburger CJ. Detection of doping with human growth hormone. Lancet 353: 895 (1999).
- ⁽¹²⁾ Wallace JD et al. Changes in non-22-kilodalton (kDa) isoforms of growth hormone (GH) after administration of 22-kDa recombinant human GH in trained adult males. J Clin Endocrinol Metab 86:1731– 1737 (2001).
- ⁽¹³⁾ Hanley JA, Saarela O, Stephens DA, Thalabard JC. hGH isoforms differential immunoassays applied to blood samples from athletes: decision limits for anti-doping testing. Growth Horm IGF Res 24: 205, 2014.

- (14) Ley Orgánica 11/2021, de 28 de diciembre, de lucha contra el dopaje en el deporte.
- (15) Doping in elite sports assessed by randomized-response surveys. Ulrich R, Pope HG Jr., Cléret L, Petrozci A, Nepusz T, Schaffer J, Kanayama G, Comstok RD, Rabin O, Simon P.
- (16) 2021 Anti-doping testing figures. World Antidoping Agency. January 2023.
- (17) TD2023IDCR v 1.0. World Antidoping Agency. Valid since 01 enero de 2023.
- (18) Quimiometría. Capítulo 1. Guillermo Ramis Ramos, M^a Cecilia García Álvarez-Coque. Editorial Síntesis. (1999).
- (19) The Assessment of Selectivity in Different Quadrupole-Orbitrap Mass Spectrometry Acquisition Modes. Bjorn J. A. Berendsen BJA, Wegh RS, Meijer T, Nielen MWF. J Am Soc Mass Spectrom. 2015 Feb;26(2):337-46. doi: 10.1007/s13361-014-1021-x.
- (20) WADA Technical Letter – TL01. Meclofenoxate. v 4.0. 24 November 2021.
- (21) WADA Technical Letter – TL02 Meveberine metabolism. v 3.0. 01 April 2021.
- (22) The List of Prohibitions 2023. World Antidoping Agency. Valid since 01 January 2023.
- (23) WADA Technical Letter – TL23 Minimum Reporting Level for certain substances known to be potential meat contaminants. v 1.0. 01 June 2021.
- (24) Discovery of c.577del in EPO: Investigations into endogenous EPO double-band detected in blood with SAR-PAGE. Zhou X, He S, Zezhou L, Jiayu W, Zhou W, Liu X, Zhao M, Zhang L. Drug Test Anal. 2022 Apr;14(4):622-633. doi: 10.1002/dta.3200. Epub 2021 Nov 22.
- (25) TD2022EPO v 1.0. World Antidoping Agency. Valid since 01 January 2022.
- (26) TD2023APMU v 1.0. World Antidoping Agency. Valid since 01 January 2023.
- (27) The Athlete's Biological Passport and Indirect Markers of Blood Doping. Sottas P, Robinson N, Saugy M. Handb Exp Pharmacol. 2010;(195):305-26. doi: 10.1007/978-3-540-79088-4_14.
- (28) Statistical classification of abnormal blood profiles in athletes. Sottas P, Robinson N, Giraud S, Taroni F, Kamber M, Mangin P, Saugy M. Int J Biostat. 2006. 2(1):1011-1011. doi:10.2202/1557-4679.1011.
- (29) ABPS: An R Package for calculating the Abnormal Blood Profile Score. Schütz F, Zollinger A. Front. Physiol., 21 November 2018. Sec. Integrative Physiology. Volume 9 – 2018. doi.org/10.3389/fphys.2018.01638
- (30) Factors Confounding the Athlete Biological Passport: A Systematic Narrative Review. Krumm B, Faiss R. Sports Med - Open (2021) 7:65. doi.org/10.1186/s40798-021-00356-0.
- (31) International Standard for Results Management. World Antidoping Agency. Valid since 01 enero 2023.
- (32) How to use likelihood ratios to interpret evidence from randomized trials. Perneger TV. J Clin Epidemiol. 2021 Aug;136:235-242. doi: 10.1016/j.jclinepi.2021.04.010.



- ⁽³³⁾ El valor de la prueba de ADN. Prieto L, Carracedo A. Capítulo 18. Genética forense, del laboratorio a los tribunales. (2018) ISBN 978-84-9052-213-4, págs. 445-470.
- ⁽³⁴⁾ WADA-Laboratory Guidelines. Human Growth Hormone (hGH) Biomarkers Test. V 3.0 January 2021.
- ⁽³⁵⁾ TD2021GH. v 1.0. World Antidoping Agency. Valid since 01 April 2021.
- ⁽³⁶⁾ TD2022IRMS. v 1.0. World Antidoping Agency. Valid since 01 January 2022.
- ⁽³⁷⁾ Performance characteristics of Carbon Isotope Ratio method for detecting doping with testosterone based on urine diols: controls and athletes with elevated testosterone/epitestosterone ratios. Aguilera R, Chapman TE, Starcevic B, Hatton CK, Catlin DH. Clin.Chem. (2001) 47:2, 292-300.