

Análisis crítico por expertos de trabajos seleccionados

Resumen del trabajo seleccionado para el análisis crítico

Correlación entre tres ensayos para la determinación de hormona antimülleriana

Correlation between three assays system for anti-Müllerian hormone (AMH) determination

Hang Wun Raymond Li, Ernest Hung Yu Ng, Benancy Po Chau Wong, Richard A. Anderson, Pak Chung Ho, William Shu Biu Yeung

Department of Obstetrics & Gynaecology, Queen Mary Hospital, The University of Hong Kong, 102 Pokfulam Road, Hong Kong, People's Republic of China. Centre for Reproductive Health, The University of Edinburgh, Royal Infirmary of Edinburgh, 51 Little France Crescent, Edinburgh, EH16 4SA, Scotland, UK

J Assist Reprod Genet. 2012;29:1443-1446

Resumen

Propósito: el análisis de la hormona antimülleriana (HAM) es de reconocida importancia en medicina reproductiva, pero los ensayos no están estandarizados. Nosotros hemos evaluado la correlación entre el nuevo equipo de ELISA Genn II (Beckman-Coulter) y los antiguos equipos de ELISA Immunotech (IOT) y *Diagnostic System Laboratories* (DSL).

Método: un total de 56 muestras de suero de pacientes con subfertilidad y desórdenes endocrino-reproductivos fueron analizadas por duplicado usando los tres equipos de ELISA para determinación de HAM. Las muestras cubrieron un amplio rango de concentraciones de HAM (1,9 a 142,5 pmol/l).

Resultados: observamos buena correlación en-

tre el nuevo equipo (AMH Gen II) y los antiguos equipos de AMH IOT y DSL ($R^2 = 0,971$ y $0,930$ respectivamente). Las ecuaciones de regresión fueron AMH (Gen II) = $1,353 \times \text{AMH (IOT)} + 0,051$ y AMH (Gen II) = $1,223 \times \text{AMH (DSL)} - 1,270$ respectivamente.

Conclusiones: las concentraciones de HAM obtenidas utilizando el equipo Gen II fueron ligeramente más altas que las encontradas al utilizar los equipos IOT y DSL. La estandarización de los ensayos es un requerimiento mundial de carácter urgente pero este análisis facilita la interpretación de los valores obtenidos históricamente y en futuros estudios usando cualquiera de los tres ensayos disponibles. No se recomienda adaptar valores de corte por conversión directa según trabajos publicados previamente.

Comentarios

Dra. Patricia Otero

Bioquímica especialista en Endocrinología - Laboratorio de Endocrinología, Hospital Durand

Comentarios

Este trabajo propone como objetivo la comparación de las 3 metodologías conocidas para dosar hormona antimülleriana. Dos de ellas ya están discontinuadas (DSL e IOT) y la tercera (Gen II) es la única disponible, actualmente, en el mercado.

Esta última, en realidad, es un híbrido entre los otros 2 ensayos ya que combina los anticuerpos usados en el método DSL y los estándares de calibración de IOT.

Se utilizaron 56 sueros de mujeres de entre 18 y 40 años que presentaban subfertilidad y poliquistosis ovárica, abarcando un rango de concentración de HAM desde 1,9 pmol/l (0,26 ng/ml) hasta 142,5 pmol/l (19,8 ng/ml). Se procesaron las muestras por las 3 metodologías mencionadas.

Los test estadísticos usados para el análisis de los resultados permiten evaluar concordancia (test de Bland y Altman) y correlación (test de Passing Bablok) entre los 3 métodos en estudio.

La primera conclusión es que las 3 metodologías se correlacionan entre sí. Los coeficientes de correlación obtenidos son similares a los ya publicados en otros trabajos.

Al aplicar el test de Passing Bablok, cuando se compara Gen II vs. DSL, entre ambos métodos, se observan diferencias de tipo proporcionales, dato concordante con los publicados por Wallace y cols. (Ann Clin Biochem. 2011); y Nelson-La Marca (Reproductive Biomedicine on line. 2011).

Las mismas diferencias de tipo proporcionales se observan al correlacionar Gen II e IOT, en este caso, estos resultados no concuerdan con datos publicados en otros trabajos (Kumar y cols. *J Immunol Methods*. 2010; y, además, esto resulta llamativo ya que el nuevo ensayo Gen II está calibrado con los mismos estándares que el de Immunotech, por lo que no debería mostrar este tipo de diferencias.

Al evaluar la concordancia entre los métodos con el test estadístico de Bland y Altman, los resultados también son discordantes con trabajos publicados anteriormente.

En primer lugar, los autores encuentran que los resultados obtenidos con DSL son mayores, en un promedio del 22%, a los obtenidos con Gen II, datos completamente contradictorios a los obtenidos por otros autores, donde me incluyo con mi experiencia (datos no publicados), que encontramos resultados por DSL más bajos (hasta en el 40%) que por Gen II.

En segundo lugar, se muestran resultados con la metodología Gen II que, en promedio, son hasta el 35% más altos que con IOT, datos que los propios autores no pueden explicar dado que se usa la misma curva patrón. Estas diferencias son mayores a valores más altos de HAM (los autores publican una diferencia mayor (38%)

para valores de HAM “menores a 100 pmol/l”, estimo que debería decir “valores mayores a 100 pmol/l”).

A pesar de que se calculan factores de conversión, que surgen de las ecuaciones de la recta de regresión, y que sirven para poder comparar valores obtenidos por metodologías anteriores con el nuevo método de Beckman, se desaconseja aplicar esta corrección ya que las diferencias no son constantes, tal como lo muestra el análisis de regresión lineal de Passing Bablok, sino que aumentan a valores más altos.

Tampoco es adecuado utilizar los mismos valores de *cut-off* para las diferentes metodologías, sino que se deben estimar los valores de corte para cada metodología aplicada.

Si bien el número de muestras utilizado en el presente trabajo tiene significación estadística, quizás aumentando el n, puedan modificarse algunos de los resultados obtenidos (por ej. la correlación entre Gen II e IOT).

Desde mi punto de vista, es correcto plantear la necesidad de contar con un estándar internacional y, a partir de ello, poder estimar los valores de corte mínimos y máximos imprescindibles para la correcta interpretación de los resultados de esta herramienta de laboratorio que tanta utilidad clínica está demostrando.

Comentarios

Dras. Cecilia Zylbersztejn, Ana María Sequera

La hormona antimülleriana (HAM o AMH) es la sustancia inhibidora de los conductos müllerianos denominada también MIS (*müllerian-inhibiting substance*). Originalmente la medición de la HAM se desarrolló para determinar la presencia de tejido testicular en niños nacidos con disgenesia gonadal o anorquia. Es una glicoproteína que pertenece a la familia de factores de crecimiento transformante β , sintetizada por las células de Sertoli del testículo.

En la mujer, la HAM producida por las células de la granulosa de los folículos primarios y preantrales pequeños es el parámetro que mejor se correlaciona con el número de estos folículos medidos por ecografía, y representa el mejor marcador de la declinación gradual de la capacidad reproductiva en la mujer (1-3). La HAM inhibe la incorporación al “reclutamiento” de los folículos primordiales, disminuye la mitosis y la acción de la FSH en las CG de los folículos primarios y preantrales. Preserva la reserva folicular y asegura la monoovulación. No depende de FSH y es un marcador más precoz que FSH del envejecimiento ovárico. Su medición permite evaluar la reserva folicular y es proporcional al número de folículos antrales (NFA), contados en la fase folicular temprana por ecografía transvaginal de alta resolución.

El índice de folículos reclutados depende del tamaño del stock de folículos primordiales, estos disminuyen exponencialmente con la edad, lo que se refleja en los valores de la HAM.

El presente trabajo plantea que sobre la base de la estabilidad de la HAM durante el ciclo menstrual y las variaciones intercíclicas no significativas, su medición se emplea en el diagnóstico y en la diferenciación de distintas causas de infertilidad o de oligomenorrea secundaria. Al mismo tiempo, la HAM permite diferenciar pacientes con falla ovárica precoz y reserva ovárica disminuida (HAM bajas), pacientes con síndrome de ovario poliquístico (SOP) (HAM elevadas 2 o 3 veces). Se observó que en pacientes con hipogonadismos hipogonadotróficos e hiperprolactinemias, no se encontraron variaciones significativas de HAM respecto de su edad.

En pacientes sometidas a tratamiento de reproducción asistida, los niveles basales de HAM se correlacionan significativamente con el número de folículos obtenidos posestimulación y de ovocitos recuperados. Es un marcador de respuesta, ya sea insuficiente o excesiva (hiperestimulación ovárica en pacientes SOP). La HAM no es un marcador de predicción de embarazo, aunque sí de pronóstico de respuesta a tratamientos

de estimulación ovárica. Según otros autores, los valores más altos de HAM están asociados a un mayor número de ovocitos maduros, mayor número de embriones, mayor tasa de embarazo y mayor índice de nacidos vivos (4).

En función de este concepto, numerosos trabajos definen los valores de HAM mínimos (*cut-off*) para esperar una respuesta satisfactoria al estímulo hormonal y justificar el procedimiento. Según Muttukrishna y cols. (2005) en 108 pacientes en quienes se usó un “*cut-off*” de 0,2 ng/ml” para HAM por Immunotech (IOT), tiene una sensibilidad del 87% para obtener hasta 4 ovocitos y predecir una pobre respuesta, mientras que Tremellen y cols. (2005), con el mismo método, muestran el 75% de sensibilidad con un valor de corte de 8,1 pmol/l que equivale a 1,13 ng/ml para el mismo número de folículos estimulados. Con el método de DSL (*Diagnostic System Laboratories*), las publicaciones muestran también discrepancias en los valores de corte entre 0,35 ng/ml, 0,71 o 1,26 ng/ml (5).

Los autores plantean la comparación de tres métodos para la medición de HAM en 56 muestras de sueros provenientes de mujeres atendidas en un Centro de Fertilidad y reproducción en Edimburgo-UK. Las metodologías comparadas son tres sistemas de enzimo-inmunoensayo, uno de origen francés: IOT; otro de EE.UU.: DSL y el nuevo sistema AMH Gen II, de Beckman Coulter, EE.UU. El nuevo kit AMH Gen II reemplaza a los dos anteriores, y fue desarrollado luego de la adquisición de las empresas Immunotech y DSL por Beckman Coulter. Las características de los kits utilizados se detallan a continuación:

1. **AMH Elisa** (catálogo Nº DSL -10-14400) (DSL, Webster, TX-USA): tiene sensibilidad (S) de 0,04 pmol/l, y el inserto reporta un coeficiente de variación (CV) intra y entre-ensayos menor del 4,6% y 8,0 respectivamente.
2. **EIA AMH IOT** (catálogo Nº A16507) (IOT, Marsella-Francia): tiene S de 0,7 pmol/l, y el inserto reporta un coeficiente de variación (CV) intra y entre-ensayos menor del 12,3% y 14,2 respectivamente.
3. **AMH Gen II** (catálogo Nº A79765) (Beckman Coulter, Chaska, MN-USA): tiene S de 0,57 pmol/l, y el inserto reporta un coeficiente de variación (CV) intra y entre-ensayos menor del 5,4% y 5,6 respectivamente.

El primer método en el mercado para la medición de HAM fue el de IOT, desarrollado para el uso en pediatría. La presencia de testículos determina valores altos de HAM. Esta aplicación explica la menor sensibilidad del método en los comienzos.

La utilización de HAM en mujeres con alteraciones de su fertilidad, para evaluar su reserva

ovárica (6), determinó el incremento de su medición, lo que requirió mayor sensibilidad en los métodos. Si bien el kit de DSL presentaba una mayor sensibilidad (tal cual los resultados del presente trabajo) fueron necesarios varios años antes de que fuera empleado en trabajos de investigación por los grupos europeos. Se discutió su uso por presentar valores más bajos que los encontrados por IOT. En nuestra experiencia, comparando resultados por ambos métodos, los valores por IOT eran significativamente más elevados, siendo mayor aún su diferencia en los valores de HAM por encima de 2 ng/ml, determinados por DSL. Los valores por debajo de 1 ng/ml medidos por DSL no pudieron ser discriminados por IOT.

Los autores muestran una buena correlación entre Gen II y DSL ($R^2: 0,930, p<0,0001$) y entre Gen II y IOT ($R^2: 0,971, p<0,0001$). Las ecuaciones de regresión obtenidas (Passing y Bablock) fueron:

$$\text{AMH (Gen II)} = 1,223 \times \text{AMH (DSL)} - 1,270$$

$$\text{AMH (Gen II)} = 1,353 \times \text{AMH (IOT)} + 0,051$$

$$\text{AMH (IOT)} = 0,910 \times \text{AMH (DSL)} - 1,103$$

Esto expresa que las diferencias entre IOT y DSL son menores que las encontradas entre Gen II y DSL y entre Gen II e IOT. En este trabajo se describe el diseño del AMH Gen II, que utiliza el anticuerpo del método de DSL y los calibradores de IOT. La curva estándar de IOT se construía a partir de un estándar de alta concentración que se diluía en forma seriada, al momento del procesamiento de las muestras. Estas preparaciones resultaban inestables, de acuerdo con indicaciones del fabricante, y no podían ser guardadas. En cambio, los calibradores del equipo DSL venían preparados, listos para su uso, manteniendo su estabilidad hasta su vencimiento. Es de notar que en el AMH Gen II, la curva estándar ya viene preparada y se comercializa en forma independiente del resto de los reactivos.

El sistema Gen II, a pesar de utilizar los calibradores del IOT, presenta valores el 35% más altos que este método y el 22% más altos que DSL, lo que no pueden explicar los autores.

Los análisis por Bland y Altman confirmaron que los valores obtenidos por AMH Gen II son sistemáticamente más altos que los encontrados por IOT y DSL. La media de las diferencias entre los valores por Gen II y DSL es de 8,0 pmol/l (equivalentes a 1,11 ng/ml), entre Gen II e IOT de 15,7 pmol/l (2,19 ng/ml), mostrando la tendencia mayor para Gen II. Mientras que entre IOT y DSL el promedio de las diferencias es de -7,7 pmol/l (1,07 ng/ml), con valores mayores para IOT, como se había descripto.

En el trabajo solamente se evalúan las discrepancias en valores de HAM <100 pmol/l (equivalentes

a 13,9 ng/ml). Los valores altos presentan mayor discrepancia entre métodos, lo que puede contribuir a magnificar estas diferencias. La importancia de la medición de HAM es caracterizar a las pacientes pobres respondedoras, con valores bajos de HAM, con menor discrepancia entre métodos, que lamentablemente no ha sido evaluado por estos autores.

Llama la atención que siendo muestras de un centro de fertilidad, los valores sean tan altos. Si bien el rango de edad estudiado es muy amplio (18,6-40,6 años), posiblemente la proporción de pacientes jóvenes con SOP esté significativamente aumentada con respecto a las añosas, con reserva ovárica disminuida (población mayoritaria que consulta por problemas de fertilidad), por lo cual son escasos los resultados bajos de HAM en este estudio.

Los mismos autores plantean las limitaciones de este trabajo reconociendo que las variaciones encontradas en la zona de valores altos de HAM no tienen impacto clínico. Por otra parte, la escasa cantidad de muestras a concentraciones bajas de HAM (zona del *cut-off* para definir pobres respondedoras) deja sin evaluar la zona más crítica, que caracteriza a las pacientes que consultan por infertilidad.

Con el propósito de evitar las diferencias entre métodos, los autores proponen encontrar una estandarización universal. No recomiendan adoptar los *cut-off* de las publicaciones previas, por conversión directa y opinan que es más apropiado redefinir los *cut-off* con un estándar internacional unificado. Lamentablemente, consideramos que esa es casi una utopía, por cuanto ningún otro analito del laboratorio lo ha podido cumplir hasta el momento. Los autores proponen utilizar los del Gen II, que está reemplazando los dos métodos previos, pero como se ha visto, el mismo estándar tiene comportamientos distintos en cada sistema (IOT/Gen II).

Consideramos que a fin de lograr la estandarización universal de los resultados de los ensayos, se debería aspirar a lograr una estandarización única, utilizando un estándar patrón, con trazabilidad asegurada y con sensibilidad en la zona de relevancia clínica. De lograrse, los resultados permitirían redefinir la determinación de HAM con un *cut-off* cuyo valor es esencial tanto para estudios de investigación como para la práctica clínica a nivel general.

Conclusiones del análisis: el análisis estadístico realizado muestra que una buena correlación no

es suficiente para demostrar la validación de un sistema frente a otro. El análisis por Bland y Altman aporta, aunque sea en forma total, cuál sería el promedio de las diferencias de un método frente al otro.

Lo observado en este trabajo pone en evidencia la importancia de que el seguimiento de las pacientes se realice con la misma metodología, más aún cuando se evalúa la eficacia de los tratamientos para estimular su “rejuvenecimiento” ovárico, con DHEA u otras sustancias, o simplemente para establecer el seguimiento de su reserva ovárica.

La aparición de nuevos kits siempre es un desafío para el laboratorio. Es necesario tener en cuenta no sólo las mejores características analíticas de un sistema, sino conocer su *performance* para discriminar las zonas de relevancia clínica. Actualmente ya hay otros kits comerciales para la medición de la HAM, que están en el mercado internacional y que deberán ser evaluados.

Referencias

1. Hehenkamp WJ, Looman CW, Themmen AP, de Jong FH, Te Velde ER, Broekmans FJ. Anti-Müllerian hormone levels in the spontaneous menstrual cycle do not show substantial fluctuation. *J Clin Endocrinol Metab.* 2006;91:4057-4063.
2. Visser JA, de Jong FH, Laven JS, Themmen AP. Anti-Müllerian hormone: a new marker for ovarian function. *Reproduction.* 2006 Jan;131(1):1-9.
3. Sequera AM, Lucini C, Pasqualini S, Tomatis C, Zylbersztein C. Hormona Antimülleriana: ¿Un marcador precoz de disfunción ovárica? Congreso XVI de SAEM. RAEM. 2009;46:133.
4. Li HWR, Yeung WSB, Lau EYL, Ho PC, Ng EHY. Evaluating the performance of serum antimüllerian hormone concentration in predicting the live birth rate of controlled ovarian stimulation and intrauterine insemination. *Fertil Steril.* 2010;94(6):2177-81.
5. La Marca A, Sighinolfi G, Radi D, Argento C, Baraldi E, Carducci Arsenio A, Stabile G, Volpe A. Anti-Müllerian hormone (AMH) as a predictive marker in assisted reproductive technology (ART). *Human Reproduction Update.* 2010;16(2):113-130.
6. Zylbersztein C, Sequera AM, Cortelezzi M. Aspectos bioquímicos de la reserva ovárica. En: SAEGRE (eds.). *Avances en Endocrinología Ginecológica y Reproductiva.* Buenos Aires: Editorial Ascune. 2012; pp. 559-572.