



REVISIÓN

Modelos predictivos en reproducción asistida: revisión sistemática y análisis crítico

Joana Peñarrubia^{a,*}, Juan Antonio García-Velasco^b y Jose Landeras^c

^a IVIRMA-Barcelona, Barcelona, España

^b IVIRMA-Madrid, Madrid, España

^c IVIRMA-Murcia, Murcia, España

Recibido el 13 de abril de 2019; aceptado el 21 de mayo de 2019

Disponible en Internet el 30 de julio de 2019

PALABRAS CLAVE

Modelo predictivo;
Reproducción
asistida;
Respuesta ovárica;
Gestación

KEYWORDS

Predictive model;
Assisted
reproduction;
Ovarian response;
Pregnancy

Resumen En medicina existe un interés creciente por la predicción del riesgo individual de los pacientes para desarrollar un determinado problema de salud o para predecir su respuesta a un tratamiento. Dado que la fecundación *in vitro* (FIV) puede ser estresante física y emocionalmente, y que no está exenta de riesgos para la salud, las parejas con esterilidad candidatas a una FIV deben ser bien informadas acerca de las posibilidades de éxito antes de cada ciclo de tratamiento. Presentamos una revisión sistemática y un análisis crítico de los modelos predictivos de respuesta ovárica y de gestación en FIV, demostrando que en muchas ocasiones la calidad de dichos modelos es baja. La metodología inadecuada a la hora de desarrollar un modelo predictivo dificulta su aplicación en la práctica clínica. Es fundamental desarrollar y disponer de modelos predictivos metodológicamente adecuados con la finalidad de optimizar nuestra capacidad predictiva en reproducción asistida y alcanzar una verdadera individualización y personalización de la medicina reproductiva.

© 2019 Asociación para el Estudio de la Biología de la Reproducción y Sociedad Española de Fertilidad. Publicado por Elsevier España, S.L.U. Todos los derechos reservados.

Prediction models in assisted reproduction: Systematic review and critical analysis

Abstract In medicine, there is a growing interest in predicting the individual risk of patients to develop a specific health problem or to predict their response to a treatment. Since *in vitro* fertilisation (IVF) can be physically and emotionally stressful, and as it is not free of health risks, the couples candidates for IVF should be well informed about the chances of success before each treatment cycle. A systematic review and a critical analysis of predictive models of ovarian response to stimulation and pregnancy after IVF is presented, showing, that in many cases the quality of these models is low. The inadequate methodology when developing a predictive

* Autor para correspondencia.

Correo electrónico: Joana.Penarrubia@ivirma.com (J. Peñarrubia).

model makes it difficult to apply in clinical practice. It is essential to develop, and to have methodologically appropriate predictive models, in order to optimise predictive capacity in assisted reproduction and achieve a true individualisation and personalisation in reproductive medicine.

© 2019 Asociación para el Estudio de la Biología de la Reproducción y Sociedad Española de Fertilidad. Published by Elsevier España, S.L.U. All rights reserved.

Introducción

En los últimos años estamos asistiendo a un cambio de paradigma importante en la forma de tratar a nuestras pacientes que requieren una estimulación ovárica. Hasta hace muy poco, las decisiones las tomábamos basándonos en estudios aleatorizados que comparaban diferentes estrategias y nos aportaban soluciones para las pacientes con el perfil de las incluidas en los ensayos clínicos, que no siempre son un reflejo de la realidad clínica. Pero evidentemente no individualiza la decisión, sino que es la misma para la media de pacientes que son incluidas en el estudio. En cambio, cada vez más las decisiones son tomadas tras el análisis de miles de datos recogidos en las bases de datos que hoy manejamos, y tras un análisis riguroso de estos datos «de la vida real», podemos individualizar —ahora sí— la pauta para esta paciente en concreto. Esto es algo que ya se hace en otras áreas de la medicina, más concretamente en la oncología, donde el empleo de bases de datos y algoritmos estadísticos para identificar la probabilidad de que ocurra un resultado basándonos en datos históricos es más habitual.

La predicción con modelos estadísticos se inició en 1955 para poder saber con antelación el tiempo que iba a hacer y de ahí se ha ido aplicando a muchas otras facetas de la vida, incluida la medicina. Disponemos cada vez de más datos y tenemos más interés en analizarlos; además, los ordenadores son cada vez más rápidos y capaces, con un *software* más sencillo, lo que facilita el acceso. Estos modelos nos van a permitir no confiar tanto en la intuición o la memoria, poder tener la información siempre a mano, e integrarla con el paciente. Existen ya descritos varios modelos para predecir no solo la respuesta ovárica, sino la probabilidad de éxito tras una fecundación *in vitro* (FIV) (ivfpredict.com) o de que un embrión sea euploide (Vera-Rodríguez et al., 2015).

La necesidad de disponer de capacidad predictiva en reproducción asistida viene dada por el hecho de que somos variables porque, tal como dijo uno de los padres de la medicina moderna, sir William Osler (1849-1919), «no hay 2 caras iguales, no hay 2 cuerpos iguales, no hay 2 individuos que reaccionen por igual y se comporten de igual modo en condiciones fisiológicas o patológicas». Y este diferente comportamiento individual nos conduce a la necesidad de poder aplicar una verdadera medicina personalizada cuyo objetivo va a ser separar a los pacientes en diferentes grupos, con decisiones médicas, intervenciones o productos adaptados a cada paciente en función de la predicción

de su respuesta o su riesgo de enfermedad. Y para poder desarrollar esta medicina personalizada necesitamos poder trabajar con precisión, es decir, conseguir la misma respuesta terapéutica en diferentes ocasiones al aplicar el mismo tratamiento en el mismo individuo, y con exactitud, es decir, aproximarnos al máximo al resultado óptimo de un tratamiento.

Predicción de respuesta ovárica para fecundación *in vitro*

En los inicios de la estimulación ovárica para FIV y, ante la falta evidente de un marcador de respuesta, todas las pacientes eran estimuladas con el mismo protocolo, el mismo fármaco y a las mismas dosis. Y con frecuencia nos encontrábamos con respuestas ováricas extremas que llevaban a la cancelación del ciclo por una respuesta ovárica insuficiente o a la hiperrespuesta ovárica, con el riesgo de desarrollar una entidad clínica temida, y por entonces desconocida desde el punto de vista fisiopatológico, como es el síndrome de hiperestimulación ovárica (SHO).

En 1989, Scott et al. (1989) describieron el primer marcador hormonal de respuesta ovárica, la hormona folículoestimulante (FSH) basal, que durante mucho tiempo nos ha sido más o menos útil a la hora de predecir el grado de respuesta a la estimulación en un ciclo de FIV. Pero el valor más importante de la introducción de la FSH basal en la práctica clínica fue que abrió el camino para la investigación en el concepto de reserva ovárica y para el descubrimiento de otros marcadores de respuesta ovárica, como son la hormona antimülleriana (AMH) y el recuento de folículos antrales (AFC), los cuales nos han permitido predecir fundamentalmente las respuestas ováricas extremas, la baja respuesta y la hiperrespuesta (La Marca y Sunkara, 2014). Sin embargo, a pesar de estos grandes avances en el conocimiento de la fisiología ovárica, todavía no tenemos capacidad predictiva de la respuesta ovárica óptima, es decir, la que se asociaría a una mayor probabilidad de gestación y a una reducción de los riesgos de hipo e hiperrespuesta y que, de acuerdo con el estudio clásico de Sunkara et al. (2011), sería la recuperación de 10 a 14 ovocitos.

En la actualidad, podemos decir que no controlamos todo el proceso de la FIV. A pesar de nuestro control sobre fases tan importantes como la recuperación de ovocitos, el manejo de gametos y embriones en el laboratorio y la técnica de la transferencia embrionaria, todavía existen 2 «cajas negras» en este proceso en las que necesitamos

capacidad predictiva: la respuesta a la estimulación ovárica y la implantación embrionaria.

En lo que respecta a la respuesta a la estimulación ovárica no somos precisos ya que, de acuerdo con el estudio de [Rombauts et al. \(2015\)](#), realizado en 176 pacientes que fueron estimuladas en 3 ciclos con la misma pauta de estimulación, si bien el 75% de las pacientes se mantenía en la misma categoría de respuesta (normal, alta o baja), el 25% restante cambiaba de categoría en los diferentes ciclos. Y tampoco somos exactos, es decir, no conseguimos siempre una respuesta ovárica óptima. Y todo ello sin olvidarnos de un subgrupo importante de pacientes, las denominadas respondedoras subóptimas, que supondrían el 43,3% de los ciclos, en las que se recuperan entre 4-9 ovocitos y que van a presentar un 20-30% menos de probabilidades de gestación que las normorrespondedoras ([Sunkara et al., 2011](#)). Este subgrupo de pacientes que presentan una respuesta subóptima tiene por definición unos marcadores de reserva ovárica normales, lo que nos impide su identificación antes del inicio de la estimulación. Así, en un estudio de [Peralta et al. \(2017\)](#) sobre un total de 4.019 ciclos de FIV la respuesta ovárica fue subóptima en el 44,3% de los ciclos y uno de los marcadores con mayor capacidad predictiva de respuesta ovárica, el AFC, no permitió discriminar entre las pacientes con respuesta subóptima y las normorrespondedoras.

Todo lo expuesto nos dificulta el alcanzar una verdadera individualización en la estimulación ovárica y el poder cumplir sus objetivos, que son: 1) la personalización de la estimulación ovárica, en cuanto a protocolo, tipo de gonadotrofina y dosis inicial; 2) la reducción de las respuestas ováricas extremas (tanto la baja como la hiperrespuesta) y la respuesta subóptima; 3) la eliminación de las complicaciones iatrogénicas ligadas a la estimulación, como el SHO; 4) la optimización de las tasas de recién nacido vivo (RNV), y 5) el cumplimiento de las expectativas de la paciente.

Los esfuerzos para desarrollar un modelo predictivo de respuesta ovárica reflejan la necesidad de estos modelos en la práctica clínica. Esta necesidad se explica por la incapacidad de los marcadores de reserva ovárica existentes para predecir el grado de respuesta ovárica con un 100% de certeza. El desarrollo de modelos predictivos de respuesta a una estimulación con gonadotrofinas puede permitir aproximarnos a una medicina personalizada y a la individualización en los protocolos de estimulación. Partiendo del hecho de que ningún modelo predictivo es infalible, nuestro objetivo ha sido revisar los modelos predictivos de respuesta ovárica descritos en la literatura con la finalidad de realizar un análisis crítico de los mismos y aportar objetividad y criterio cuando se analiza la literatura existente sobre el tema.

Predicción de gestación tras fecundación in vitro

La idea de establecer un modelo predictivo para poder tener una estimación lo más real posible en cuanto a la tasa de éxito esperable tras un tratamiento de reproducción asistida (TRA) surge de la necesidad de establecer un correcto asesoramiento a la pareja subfétil antes de iniciar un ciclo de FIV. La carga emocional y física, los potenciales

riesgos clínicos y los elevados costes económicos que implican estos tratamientos llevaron a intentar desarrollar estas herramientas de estimación pronóstica y poder ayudar a los pacientes en la toma de decisiones. Por otra parte, contrariamente a la percepción general, la FIV no garantiza la consecución del embarazo ni siquiera acumulando varios ciclos de tratamiento, por ello sería de un gran valor poder asesorar objetiva y eficazmente a los pacientes acerca de sus expectativas reales de éxito en un programa de FIV antes de ser incluidos en él. También debemos tener en cuenta que en la mayoría de los casos se trata realmente de situaciones de subfertilidad y no de esterilidad, por lo que la gestación podría también ocurrir de manera espontánea y sin necesidad de realizar un TRA.

Las predicciones realizadas desde la experiencia clínica o la impresión clínica del médico resultarán probablemente muy imprecisas y empíricas; sin embargo, son las más ampliamente utilizadas por médicos y sistemas públicos de salud, centrándose principalmente en el factor edad femenino.

En este contexto, los modelos predictivos podrían tener un papel importante al intentar calcular las expectativas reales y objetivas de éxito de los pacientes, *a priori* y de forma individualizada. Aunque planteado así un modelo predictivo puede parecer la panacea en el asesoramiento médico de los pacientes, también tienen sus importantes limitaciones ya que cada modelo predictivo se ha desarrollado en un centro determinado con una población de referencia particular, y en unas circunstancias sociales y demográficas determinadas, pudiendo predecir de manera muy diferente de como lo haría en otro centro y con otras condiciones distintas. Esto sería debido a las diferencias en los resultados clínicos de cada centro de reproducción asistida en particular, a las diferencias demográficas o étnicas de la población de referencia, al desarrollo de la propia tecnología de las TRA o a la aparición de nuevos marcadores con demostrado valor predictivo a través de los continuos estudios y avances científicos que se producen en este campo. Las soluciones que se plantean para solventar estas limitaciones pasan por crear un nuevo modelo predictivo para cada centro y población de referencia o por actualizar el modelo original ajustándolo a las nuevas características poblacionales.

Desarrollo de un modelo predictivo

Todo comienza con la pregunta clínica que nos hagamos, la recopilación de los datos y tener un *software* que nos ayude. El desarrollo de un modelo predictivo en cualquier ámbito se divide en 3 fases ([fig. 1](#)):

1. Derivación del modelo: incluye la identificación de las variables predictivas, la estimación de los coeficientes de regresión de cada variable y el desarrollo de un modelo, es decir, un algoritmo matemático, mediante un análisis de regresión logística.
2. Validación del modelo: valorar la capacidad predictiva del modelo. Esta fase incluye la validación interna en el grupo de pacientes en el que se ha desarrollado el modelo, es decir, si el modelo es reproducible, y la

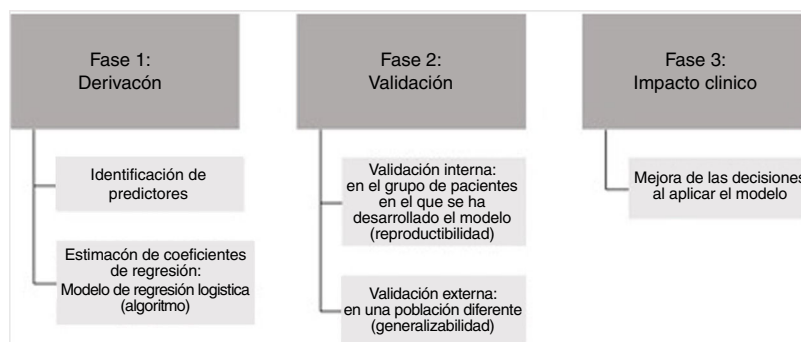


Figura 1 Desarrollo de un modelo predictivo. Modificado a partir de [Van Loendersloot, 2014](#).

validación externa en una población diferente, es decir, si el modelo es generalizable.

3. **Análisis del impacto clínico:** establecer si la aplicación del modelo en la práctica mejora las decisiones clínicas frente a lo que hubiera hecho el médico.

Fase 1. Derivación del modelo

Identificación de variables predictivas

Los predictores deben ser variables que se han escogido por su potencial capacidad predictiva y pueden incluir características demográficas, datos clínicos o físicos, resultados de laboratorio o respuesta a tratamientos previos. Por otra parte, el número de predictores no debe ser excesivo, ya que se corre el riesgo de sobrestimar la capacidad predictiva del modelo. En cuanto al tamaño de la muestra, este debe ser de al menos de 10 individuos por cada variable predictiva.

Las variables más estudiadas en la predicción de la respuesta ovárica han sido la edad de la mujer, el índice de masa corporal (IMC), los niveles de FSH, hormona luteinizante, inhibina B y AMH, datos ecográficos como el AFC y el volumen ovárico y diversos test dinámicos, como el test de clomifeno o el de test de análogos agonistas de la hormona liberadora de gonadotropinas (aGnRH).

Estimación del coeficiente de regresión

Una vez identificados los potenciales predictores, se determina el efecto cuantitativo o peso de cada predictor mediante el cálculo del correspondiente coeficiente de regresión y se desarrolla el modelo mediante un análisis de regresión logística que nos proporcionará una ecuación o algoritmo matemático con la variable o variables con más peso.

Fase 2. Validación del modelo

La segunda fase en el desarrollo de un modelo predictivo es la validación del mismo. La validación de un modelo con el fin de establecer su capacidad predictiva se realiza calculando su *capacidad de discriminación* para distinguir a pacientes con y sin el evento estudiado, o determinando su *grado de calibración*, es decir, la concordancia entre la probabilidad observada y la estimada por el modelo. Por

ejemplo, si calculamos una probabilidad del 30% de gestación tras un ciclo de FIV, la frecuencia observada de gestación debe ser aproximadamente de 30 por cada 100 mujeres.

La *capacidad de discriminación* se expresa mediante el área bajo la curva ROC (AUC_{ROC}); así, un modelo con una AUC_{ROC} de 0,5 no tiene poder de discriminación mientras que una AUC_{ROC} de 1,0 reflejaría una discriminación perfecta.

El *grado de calibración* de un modelo se calcula mediante la prueba de bonanza de ajuste de Hosmer-Lemeshow, de tal modo que si en esta prueba el valor de p es inferior a 0,05, no existe riesgo de un error de calibración. También se puede determinar gráficamente mediante un «plot» de calibración; en caso de una calibración perfecta, todos los puntos de la gráfica se encuentran en la diagonal, lo que indica que las probabilidades se corresponden con las proporciones observadas.

La validación de un modelo predictivo se subdivide en la fase de validación interna y en la de validación externa. Con la validación interna se evalúa la capacidad del modelo para predecir la respuesta a un tratamiento en el grupo de pacientes en el que se desarrolló el modelo (reproducibilidad). Sin embargo, antes de utilizar un modelo predictivo en la toma de decisiones clínicas no es suficiente haber demostrado un rendimiento bueno o razonable después de la validación interna. Muchos modelos predictivos ofrecen unos resultados demasiado optimistas y un aspecto fundamental en la predicción es que el modelo derivado a partir de un grupo de pacientes pueda ser aplicado a otro. Por lo tanto, es necesario que el modelo trabaje satisfactoriamente en una población «similar pero diferente» a la población en la que se desarrolló; es decir, es necesario realizar una validación externa (generalizabilidad). Esta validación externa puede ser temporal, en la que el modelo es validado en un nuevo grupo de pacientes del mismo centro, pero en un diferente periodo, o geográfica, en la que el modelo es validado en nuevos pacientes de un centro diferente.

Es importante tener en cuenta que en medicina reproductiva es frecuente la aparición de nuevos biomarcadores y que, en consecuencia, los modelos predictivos que se desarrollen en este ámbito deben ser regularmente actualizados con la finalidad de adaptarlos a las nuevas herramientas disponibles.

Fase 3. Análisis del impacto clínico

La última fase en el desarrollo de un modelo predictivo, y probablemente la más importante, es el análisis del impacto clínico, es decir, si la aplicación del modelo mejora las decisiones clínicas en términos de calidad o de coste-efectividad. Idealmente, este análisis debe realizarse mediante la comparación de los resultados entre pacientes asignados aleatoriamente para recibir el tratamiento guiado por el modelo y pacientes tratados sin aplicación del modelo predictivo. Una alternativa menos válida es realizar un *documento de manejo de decisiones terapéuticas*, es decir, una encuesta a los profesionales, antes (guiados por su experiencia previa o su instinto) y después de ser «expuestos» al modelo predictivo.

Metodología

Modelos predictivos de respuesta ovárica

Se realizó una revisión de los estudios publicados desde el año 2000 hasta el 2017 y que hacían referencia a la predicción de la respuesta ovárica en ciclos de FIV. Se efectuó una búsqueda sistemática en PubMed usando como palabras clave: predicción de respuesta ovárica, modelos predictivos de respuesta ovárica, marcadores de reserva ovárica, marcadores de respuesta ovárica, individualización de la estimulación ovárica, FSH, inhibina B, AMH y AFC. Las palabras clave se identificaron en el título o en el resumen de las publicaciones.

La búsqueda generó 140 citaciones. De estas, 74 fueron rechazadas con base en el título, el resumen o la lectura del artículo. Finalmente, se incluyeron 66 estudios que hacían referencia a alguna de las fases del desarrollo de un modelo predictivo expuestas anteriormente (Akanke et al., 2004; Al-Azemi et al., 2011; Anckaert et al., 2012; Andersen et al., 2011; Arce et al., 2013; Ashrafi et al., 2017; Bancsi et al., 2002; Bancsi et al., 2004a; Bancsi et al., 2004b; Brodin et al., 2015; Broekmans et al., 2014; Broer et al., 2013a; Broer et al., 2013b; Creus et al., 2000; Dzik et al., 2000; Dzik et al., 2008; x; Fábregues et al., 2000; Fang et al., 2015; Fawzy et al., 2002; Gingold et al., 2015; Guo et al., 2014; Hamdine et al., 2015; Heidar et al., 2015; Hendricks et al., 2004; Hendricks et al., 2005a; Hendricks et al., 2005b; Howles et al., 2006; Hsu et al., 2011; Ng et al., 2005a; Ng et al., 2005b; Jayaprakasan et al., 2009a; Jayaprakasan et al., 2009b; Khairy et al., 2008; Kunt et al., 2011; Kwee et al., 2006; Kwee et al., 2007; Kwee et al., 2008; La Cour Freiesleben et al., 2011; La Marca et al., 2011; La Marca et al., 2012; La Marca et al., 2013; Lee et al., 2008; Li et al., 2016; Lorusso et al., 2007; Magnusson et al., 2017; McIlveen et al., 2007; Moon et al., 2016; Mutlu et al., 2013; Muttukrishna et al., 2005; Nelson et al., 2007; Nelson et al., 2015; Oehninger et al., 2015; Olivennes et al., 2009; Peñarrubia et al., 2000; Peñarrubia et al., 2005; Peñarrubia et al., 2010; Polyzos et al., 2013; Popovic-Todorovic et al., 2003; Pouly et al., 2015; Singh et al., 2013; Van rooij et al., 2002; Verberg et al., 2007; Vural et al., 2014; Papaleo et al., 2016 y Erdem et al., 2004). Estos modelos predictivos podían ser simples o multivariados con base en la inclusión de uno o varios marcadores de respuesta ovárica.

Modelos predictivos de gestación tras fecundación in vitro

En este caso, se realizó una revisión sistemática de los modelos predictivos de gestación y de RNV tras FIV publicados en la literatura en los últimos años.

Resultados

Modelos predictivos de respuesta ovárica

La mayor parte de las publicaciones incluidas corresponden a estudios prospectivos (48/66; 72,7%) realizados con base en un número variable de pacientes, si bien en la mayoría se incluyeron 500 o menos (46/66; 69,7%).

A pesar de que históricamente el régimen más utilizado para conseguir la inhibición hipofisaria en ciclos de FIV ha sido el protocolo largo con aGnRH (Daya, 2000), en los últimos años se ha impuesto el protocolo con antagonistas de la GnRH dada la menor duración del tratamiento y la reducción del riesgo de SHO (Al-Inany et al., 2016). Sin embargo, en la mayoría de los estudios incluidos en esta revisión, incluso en los más recientes, la pauta de inhibición hipofisaria utilizada había sido el protocolo largo con aGnRH (48/66) junto con el empleo de FSH recombinante para la estimulación ovárica (34/66; 51,5%). Este dato es importante ya que puede limitar la aplicabilidad clínica actual de los modelos predictivos de respuesta ovárica descritos.

En cuanto a los eventos relacionados con la respuesta ovárica que pretendían predecir los modelos incluidos, en su mayor parte fueron las respuestas ováricas extremas, es decir, la baja respuesta (40/66; 60,6%) y la hiperrespuesta con o sin desarrollo de un SHO (15/66; 22,7%). En este sentido, es importante tener en cuenta las diferentes definiciones de este tipo de eventos existentes en la literatura incluida, hecho que también va a dificultar la aplicación clínica del modelo predictivo. En solo 11 estudios (16,7%) el objetivo final de la predicción fue el número de ovocitos obtenidos.

Respecto al tipo de modelo predictivo, en 30 estudios el modelo se desarrolló con base en la inclusión de un único marcador de respuesta ovárica (modelo simple) y en 36 se realizó con base en varios de estos marcadores (modelo multivariado).

Un aspecto fundamental es valorar la calidad del modelo predictivo desarrollado de acuerdo con las fases expuestas anteriormente (fig. 2). En este sentido, si bien en la mayoría de los estudios revisados se realizó una validación interna del modelo, fundamentalmente a través de la valoración de su capacidad de discriminación mediante el cálculo del AUC_{ROC}, solo en 11 estudios (Akanke et al., 2004; Bancsi et al., 2002; Broekmans et al., 2014; Howles et al., 2006; Lorusso et al., 2007; Nelson et al., 2015; Andersen et al., 2011; Oehninger et al., 2015; Polyzos et al., 2013; Popovic-Todorovic, 2003; Vural et al., 2014) se realizó la calibración del modelo y únicamente en 5 se validó externamente (Broekmans et al., 2014; la Cour Freiesleben et al., 2011; La Marca et al., 2012; Olivennes et al., 2009; Popovic-Todorovic et al., 2003). Por último, tan solo en 2 estudios (Pouly et al., 2015; Magnusson et al., 2017) se analizó el impacto clínico de la aplicación del modelo.

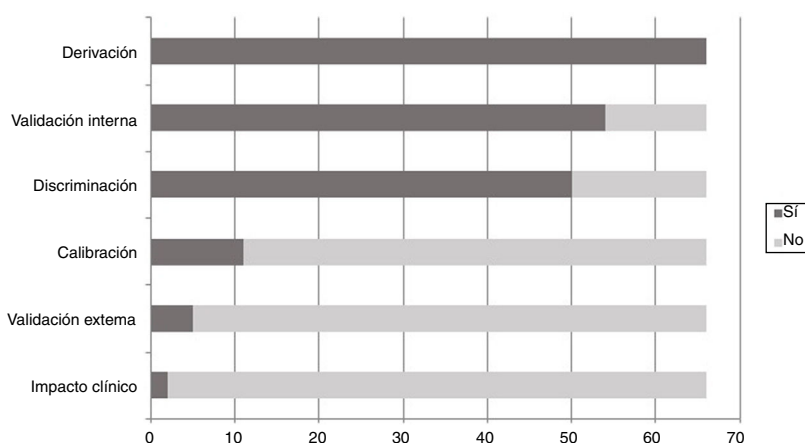


Figura 2 Calidad de los modelos predictivos revisados.

Modelos predictivos de gestación tras fecundación in vitro

Los sistemas de salud británico y holandés han sido los principales responsables del desarrollo de los modelos predictivos al disponer de una gran base de datos por la obligatoriedad del registro de toda la información clínica, de laboratorio y de resultados que emana de cada ciclo de TRA realizado en su territorio. En el caso de Reino Unido, desde 1991 la Autoridad en Fecundación y Embriología Humanas (HFEA) obligaron por ley al registro de todos estos datos y es por ello que el primer modelo predictivo desarrollado fue el «modelo de Templeton» (Templeton et al., 1996) analizando 37.000 ciclos realizados entre 1991 y 1994. El objetivo principal de este primer modelo era predecir la tasa de RNV por ciclo iniciado. Del análisis exhaustivo de estos datos y realizando regresiones univariantes, determinaron que los factores que mayor impacto ejercían sobre los resultados clínicos eran: la edad (objetivando una caída aguda a partir de los 35 años), el tiempo de infertilidad (a mayor tiempo de infertilidad, peor pronóstico), la infertilidad primaria, el número de ciclos previos fallidos y en cuanto a la causa primaria de infertilidad, el factor tubárico (sin especificar el grado de afectación ni la presencia-ausencia de hidrosálpinx).

Leushuis et al. (2009) realizan una revisión sistemática en 2009 sobre los modelos predictivos existentes señalando al modelo de Templeton como el único con una buena capacidad predictiva. Basándose en las conclusiones de esta revisión sistemática, van Loendersloot et al. (2011), aprovechando que en Holanda también existe la obligatoriedad por ley de registrar todos los datos de los ciclos *in vitro*, se plantearon realizar una validación externa del modelo de Templeton con el objeto de valorar si 15 años después de desarrollarse, dados los grandes avances tecnológicos y de conocimiento de la reproducción asistida, y aplicado sobre una población geográficamente diferente, seguía teniendo validez y aplicabilidad. Este grupo estudia 1.537 ciclos de FIV y 1.442 ciclos de ICSI en el Academic Medical Center de Amsterdam. El objetivo principal que se marcan es la tasa de gestación evolutiva y estratifican la población en quintiles de probabilidad de éxito obteniendo una pobre discriminación en la AUC_{ROC} ($C = 0,63$), así como una pobre calibración. Ellos justifican esta pobre capacidad predictiva por disponer

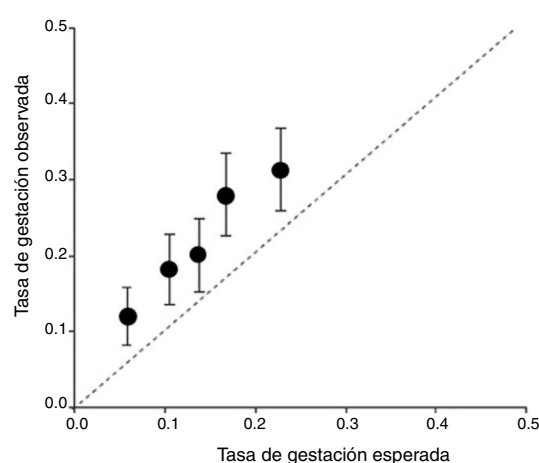


Figura 3 Ejemplo de mala calibración de un modelo predictivo.

de una base de datos mucho más reciente, por la existencia de cambios importantes producidos en ese tiempo en la práctica clínica de laboratorio, que condicionaban la consecución de unas tasas de gestación significativamente superiores, y por un menor tiempo medio de infertilidad previa a la primera consulta.

Por todo ello, al encontrar aplicando el modelo de Templeton una gran desviación entre los valores observados y los valores esperados en tasas de gestación concluyen que el modelo ya no era válido ni operativo en la práctica médica en ese momento. Como consecuencia, ellos desarrollarán su propio modelo predictivo que veremos más adelante (fig. 3).

Otro enfoque diferente lo encontramos en la validación externa del modelo de Templeton que realizan Arvis et al. (2012) intentando demostrar que con unos pequeños ajustes podrían validar la aplicación de dicho modelo 2 décadas después de haberse descrito, a pesar del gran avance y el desarrollo tecnológico producido. El estudio lo realizan sobre 11.208 de FIV-ICSI llevados a cabo en Rennes (Francia) entre 2002 y 2010, y también establecen como objetivo principal la tasa de RNV. Ellos consiguen finalmente validarlo a través de las siguientes fases: en una primera fase demuestran, al igual que van Loendersloot et al. (2011), que el modelo original subestima sistemáticamente, demos-

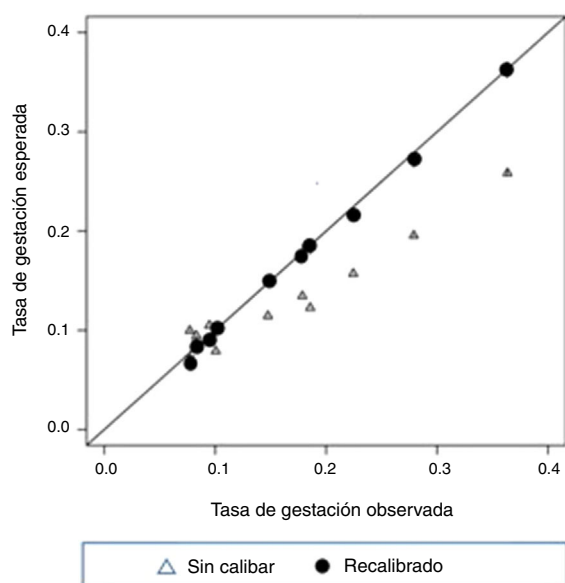


Figura 4 Gráfica de recalibración de un modelo predictivo.

trando tanto la existencia de una pobre discriminación ($C=0,64$) como de una pobre calibración; en una segunda fase, realizando un reajuste con sus propios datos mejoran la discriminación ($C=0,69$) y consiguen una calibración casi perfecta; y finalmente, en una tercera fase ajustando la tendencia lineal en el tiempo hacia una mejora de los resultados y añadiendo 3 nuevos predictores al modelo a partir del análisis de la literatura científica disponible (el tabaquismo, el IMC y los niveles séricos basales de FSH), consiguen mejorar el nivel de discriminación (0,71) manteniendo una calibración casi perfecta entre el valor observado y el esperado (fig. 4).

En 2011, de nuevo con casuística del Reino Unido, se publican los resultados de un nuevo modelo predictivo, el modelo de Nelson-Lawlor o IVFpredict (Nelson y Lawlor, 2011). El objetivo principal de este modelo también era la tasa de RNV. Los autores estudian 144.018 casos registrados por la HFEA entre 2003 y 2007, tratándose prácticamente en el 100% de los casos de transferencias de 2 embriones (DET). Querían utilizar como predictores los mismos del modelo de Templeton actualizados y compararlo con el original. La actualización del modelo de Templeton la realizan a través de 2 mecanismos. En primer lugar, incluyen nuevas variables (todas las causas de esterilidad, gameto propio o donado, tipo de tratamiento hormonal, ciclos FIV e ICSI y número de tratamientos previos) y, en segundo lugar, ajustan la tendencia lineal en el tiempo hacia la mejora en los resultados clínicos tras detectar esa significativa tendencia a la alza desde 2003 hasta 2007 ($p < 0,001$), y lo vuelven a justificar obviamente con el desarrollo tecnológico y con las mejoras en los procedimientos de laboratorio. Después de aplicar estos ajustes al modelo original, constataron que tanto la discriminación como la calibración fueron muy superiores con el nuevo modelo de Nelson-Lawlor que con el original de Templeton en los 10 centiles de probabilidad (fig. 5).

A la vista del alto poder predictivo del nuevo modelo, los autores crean una herramienta novedosa y

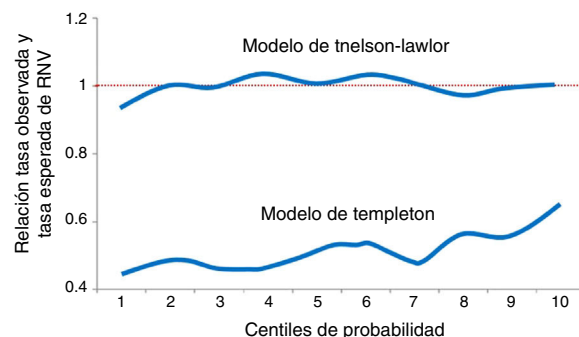


Figura 5 Modelos de Templeton vs. Nelson-Lawlor. Comparación predictiva.

vanguardista, una página web de uso libre a modo de consulta: www.IVFpredict.com.

En 2015, Smith et al. (2015), estudiando 130.960 casos registrados por la HFEA en el Reino Unido entre 2008 y 2010, comparan los modelos de Templeton y de Nelson-Lawlor actualizados mediante coeficientes de regresión multivariable, entre sí y con un modelo basado sencillamente en la edad de la paciente por ser el predictor más potente y el método más ampliamente usado en la práctica clínica habitual. Demostraron que ambos modelos discriminaban mejor que el factor edad de forma aislada y en cuanto a la calibración, constataron que el modelo de Nelson-Lawlor era mejor predictor especialmente en los extremos de probabilidad; es decir los casos de peor y mejor pronóstico. Dicho de otro modo, los casos que no gestaban tenían peor pronóstico al igual que los casos que sí gestaban tenían mejor pronóstico con el modelo de Nelson-Lawlor que con el modelo de Templeton.

Antes de esto, en 2013, van Loendersloot et al. (2013) publican los resultados de su propio modelo predictivo; una vez habían demostrado la «obsolescencia» del modelo de Templeton basándose en los grandes avances en los protocolos clínicos y de laboratorio producidos en la reproducción asistida moderna. Para ello estudian 2.621 casos FIV-ICSI registrados en el Academic Medical Centre de Amsterdam entre 2001 y 2010, según mandato legal al igual que en el Reino Unido con la HFEA y también realizan una validación interna con 515 ciclos del mismo centro registrados entre 2009 y 2011. Como innovación, el modelo permite predecir los resultados clínicos antes del primer ciclo *in vitro* y también después de uno o varios ciclos fallidos.

Seleccionan los predictores del metaanálisis realizado por ellos mismos en el que incluyen variables clínicas disponibles antes de iniciar el ciclo e incluyendo eventualmente datos de la estimulación ovárica y del laboratorio FIV de el/los ciclo/s previos fallidos. Tanto en el desarrollo del modelo como en la validación interna posterior, la discriminación expresada en la AUC_{ROC} fue moderada ($C=0,68$), lo que sería claramente insuficiente para una predicción diagnóstica pero no para una predicción pronóstica como es el caso en todos estos modelos predictivos. La calibración también fue buena en ambas series de estudio especialmente en los 3 primeros quintiles que son lo de mayor relevancia tanto para la paciente como para el médico y los sistemas públicos de salud, por ser los casos de peor pronóstico y en los que

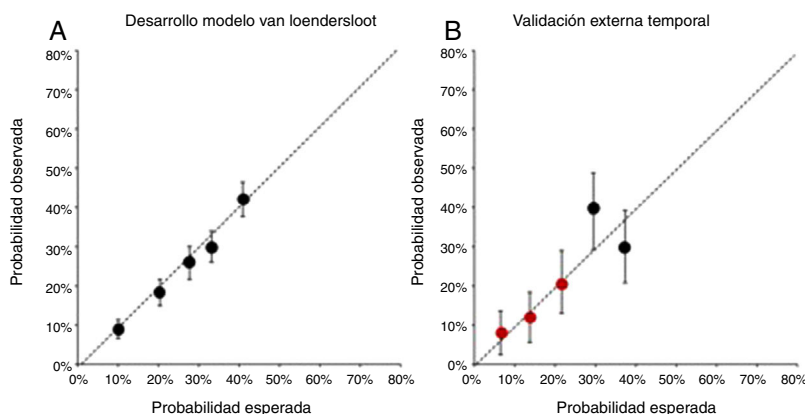


Figura 6 Desarrollo y validación interna del modelo de van Loendersloot. Tomado de [van Loendersloot et al., 2013](#).

hay que decidir si la relación coste-efectividad es razonable ([fig. 6](#)).

Este modelo fue sometido también a una validación externa geográfica con 772 casos registrados entre enero y diciembre del 2013 en el Ospedale Maggiore Policlinico de Milán ([Sarais et al., 2016](#)). Al igual que el original, consigue una calibración casi perfecta tras ajuste de regresión logística en los 5 quintiles de probabilidad.

Más recientemente, fue publicado en 2016 el modelo [Dhillon et al. \(2016\)](#), siendo el primero en incluir predictores como el IMC, la valoración de la reserva ovárica a través del AFC o la etnicidad; teniendo en cuenta, además, factores importantes fruto de los avances tecnológicos y los protocolos de actuación como son la transferencia en estadio de blastocisto o de embrión único (SET). Como buen modelo predictivo solo estudian variables pretratamiento y el desarrollo inicial del modelo se establece analizando 9.915 casos recogidos entre 2008 y 2012 en 12 CARE Clinics del Reino Unido. La validación externa temporal la realizan en 2013 sobre 2.723 ciclos. Tras recalibración de la curva, ellos obtienen un poder predictivo excelente en los 10 centiles de la curva.

Buscando simplificar y facilitar el manejo en la práctica clínica habitual, [La Marca et al. \(2011\)](#) desarrollan un modelo predictivo en el Fertility Centre de Módena (Italia), basado exclusivamente en el factor edad y en la AMH como valoración de reserva ovárica, observando que la peor capacidad predictiva se da en los grupos de peor pronóstico, es decir, en los de mayor edad y menor nivel de AMH, lo que resta mucho valor a este modelo predictivo por ser los casos de mayor interés en la práctica clínica.

Desde un punto de vista práctico, se generaron en Estados Unidos 2 útiles herramientas. La primera fue el desarrollo en 2011 de una aplicación *online* para introducir una serie de parámetros predictores, para asesorar al ginecólogo en cuanto a la tasa de embarazo gemelar esperada tras DET, ayudándole de esta manera a decidir el transfer selectivo de un embrión (eSET) en los casos de buen pronóstico (ForMyOdds.com) ([Jones et al., 2011](#)); la segunda fue la creación de un modelo en 2014 que se incluye en la web de la SART (www.sart.org) para, introduciendo el propio paciente sus datos clínicos, asesorarse de sus expectativas reales en cuanto a tasas de éxito, embarazo gemelar y tasa acumulada en 3 ciclos *in vitro* ([Luke et al., 2014](#)).

Discusión

Es evidente que en medicina existe un interés creciente por la predicción del riesgo individual de los pacientes para desarrollar un determinado problema de salud o para predecir su respuesta a un tratamiento. Este interés ha llevado al desarrollo de modelos predictivos basados en características personales, antecedentes, pruebas diagnósticas y respuesta previa a tratamientos con la finalidad, no de sustituir el papel de los profesionales a la hora de dictaminar sobre un pronóstico o un tratamiento, sino de facilitar la toma de decisiones clínicas.

La predicción de la respuesta ovárica a la estimulación y de los resultados de la FIV en términos de gestación y de RNV es todavía un problema no resuelto en medicina reproductiva. La variabilidad individual a la hora de responder a un TRA hace necesario el mejorar nuestra capacidad predictiva con el fin de reducir las respuestas ováricas subóptimas y extremas, que se asocian a peores tasas de embarazo o a un incremento del riesgo para la salud de la mujer, y de optimizar los resultados en términos de gestación y nacido vivo.

La revisión de los modelos predictivos de respuesta ovárica realizada en el presente trabajo demuestra que la calidad de los modelos publicados en la literatura es baja. En pocos casos se ha realizado una validación externa del modelo y en muy pocas ocasiones se ha llegado a analizar el impacto clínico de la aplicación del modelo. Así, una revisión sistemática de los modelos predictivos en reproducción encontró que, de 29 modelos, solo 8 habían sido validados externamente y en tan solo uno se había llegado a aplicar en la práctica clínica ([Leushuis et al., 2009](#)).

La metodología inadecuada a la hora de desarrollar un modelo predictivo, es decir, el no realizar todas las fases necesarias (derivación, validación interna y externa y análisis del impacto clínico), dificulta e incluso contraindica la aplicación del modelo en la práctica clínica. Para que un modelo predictivo sea ampliamente aceptado es necesaria una evidencia consistente que demuestre que su aplicación conduce a un incremento en el beneficio que obtenga la paciente.

Dado que la FIV puede ser estresante física y emocionalmente, y que no está exenta de riesgos para la salud, las parejas con esterilidad candidatas a una FIV deben ser

bien informadas acerca de las posibilidades de éxito antes de cada ciclo de tratamiento. Desafortunadamente, en este momento, no hay ensayos clínicos controlados aleatorizados que comparen la FIV con la concepción natural. Por lo tanto, la única manera de aconsejar adecuadamente a estas parejas es a través de un pronóstico basado en modelos predictivos.

En conclusión, la calidad de los modelos predictivos de respuesta ovárica publicados hasta el momento es desafortunadamente baja, sobre todo por la falta de una validación externa de los mismos y de un análisis del impacto de su aplicación en la toma de decisiones clínicas. Respecto a los modelos predictivos de gestación tras FIV, existen numerosos trabajos que han informado sobre el desarrollo o la validación de modelos de predicción de gestación tras FIV. Sin embargo, muy pocos han demostrado un buen rendimiento después de su validación externa y aún no se han realizado análisis de impacto para ninguno de estos modelos. En consecuencia, es necesario poder desarrollar y disponer de modelos predictivos realizados con base en una adecuada metodología con la finalidad de optimizar nuestra capacidad de predicción en el campo de la reproducción asistida y alcanzar una verdadera individualización terapéutica y una personalización de la medicina reproductiva. En la actualidad, se está abriendo una puerta que encierra enormes posibilidades predictivas como es el manejo del *big data* y la inteligencia artificial que será cada vez más capaz de analizar estas enormes cantidades de información obteniendo predicciones de forma cada vez más individualizada y precisa. De esta manera, los modelos de predicción podrán fortalecer la toma de decisiones individualizadas y basadas en la evidencia y contribuir al uso racional de estos costosos tratamientos.

Financiación

La presente investigación no ha recibido ayudas específicas provenientes de agencias del sector público, sector comercial o entidades sin ánimo de lucro.

Conflicto de intereses

Los autores declaran no tener ningún conflicto de intereses.

Agradecimientos

A los compañeros del Grupo de Interés de Endocrinología Reproductiva de la Sociedad Española de Fertilidad (Claudio Álvarez, Manuel Álvarez, Elisa Gil, Mónica González, Sonia Lobo, Joaquín Llácer, Elkin Muñoz, Ana Robles, Eric Saucedo).

Bibliografía

- Akande, V.A., Keay, S.D., Hunt, L.P., Mathur, R.S., Jenkins, J.M., Cahill, D.J., 2004. [The practical implications of a raised serum FSH and age on the risk of IVF treatment cancellation due to a poor ovarian response.](#) *Journal of Assisted Reproduction and Genetics* 21, 257–262.
- Al-Azemi, M., Killick, S.R., Duffy, S., Pye, C., Refaat, B., Hill, N., et al., 2011. [Multi-marker assessment of ovarian reserve predicts oocyte yield after ovulation induction.](#) *Human Reproduction* 26, 414–422.
- Al-Inany, H.G., Youssef, M.A., Ayeleke, R.O., Brown, J., Lam, W.S., Broekmans, F.J., 2016. [Gonadotrophin-releasing hormone antagonists for assisted reproductive technology.](#) *Cochrane Database of Systematic Reviews* 4, CD001750.
- Anckaert, E., Smits, J., Schiettecatte, J., Klein, B.M., Arce, J.C., 2012. [The value of anti-Müllerian hormone measurement in the long GnRH agonist protocol: Association with ovarian response and gonadotrophin-dose adjustments.](#) *Human Reproduction* 27, 1829–1839.
- Andersen, A.N., Witjes, H., Gordon, K., Mannaerts, B., Xpect investigators, 2011. [Predictive factors of ovarian response and clinical outcome after IVF/ICSI following a rFSH/GnRH antagonist protocol with or without oral contraceptive pre-treatment.](#) *Human Reproduction* 26, 3413–3423.
- Arce, J.C., la Marca, A., Mirner Klein, B., Nyboe Andersen, A., Fleming, R., 2013. [Antimüllerian hormone in gonadotropin releasing-hormone antagonist cycles: Prediction of ovarian response and cumulative treatment outcome in good-prognosis patients.](#) *Fertility and Sterility* 99, 1644–1653.
- Arvis, P., Leher, P., Guivarch-Levêque, A., 2012. [Simple adaptations to the Templeton model for IVF outcome prediction make it current and clinically useful.](#) *Human Reproduction* 27 (10), 2971–2978.
- Ashrafi, M., Hemat, M., Arabipoor, A., Salman Yazdi, R., Bahman-Abadi, A., Cheraghi, R., 2017. [Predictive values of anti-müllerian hormone, antral follicle count and ovarian response prediction index \(ORPI\) for assisted reproductive technology outcomes.](#) *Journal of Obstetrics and Gynaecology* 37, 82–88.
- Bancsi, L.F., Broekmans, F.J., Eijkemans, M.J., de Jong, F.H., Habbema, J.D., te Velde, E.R., 2002. [Predictors of poor ovarian response in in vitro fertilization: A prospective study comparing basal markers of ovarian reserve.](#) *Fertility and Sterility* 77, 328–336.
- Bancsi, L.F., Broekmans, F.J., Looman, C.W., Habbema, J.D., te Velde, E.R., 2004a. [Impact of repeated antral follicle counts on the prediction of poor ovarian response in women undergoing in vitro fertilization.](#) *Fertility and Sterility* 81, 35–41.
- Bancsi, L.F., Broekmans, F.J., Looman, C.W., Habbema, J.D., te Velde, E.R., 2004b. [Predicting poor ovarian response in IVF: use of repeat basal FSH measurement.](#) *Journal of Reproductive Medicine* 49 (3), 187–194.
- Brodin, T., Hadziosmanovic, N., Berglund, L., Olovsson, M., Holte, J., 2015. [Comparing 4 ovarian reserve markers-associations with ovarian response and live births after assisted reproduction.](#) *Acta Obstetrica et Gynaecologica Scandinavica* 94, 1056–1063.
- Broekmans, F.J., Verweij, P.J., Eijkemans, M.J., Mannaerts, B.M., Witjes, H., 2014. [Prognostic models for high and low ovarian responses in controlled ovarian stimulation using a GnRH antagonist protocol.](#) *Human Reproduction* 29, 1688–1697.
- Broer, S.L., Dölleman, M., van Disseldorp, J., Broeze, K.A., Opmeer, B.C., Bossuyt, P.M., et al., IPD-EXPORT Study Group, 2013a. [Prediction of an excessive response in in vitro fertilization from patient characteristics and ovarian reserve tests and comparison in subgroups: An individual patient data meta-analysis.](#) *Fertility and Sterility* 100, 420–429.
- Broer, S.L., van Disseldorp, J., Broeze, K.A., Dölleman, M., Opmeer, B.C., Bossuyt, P., et al., IMPORT study group, 2013b. [Added value of ovarian reserve testing on patient characteristics in the prediction of ovarian response and ongoing pregnancy: An individual patient data approach.](#) *Human Reproduction Update* 19, 26–36.
- Creus, M., Peñarrubia, J., Fábregues, F., Vidal, E., Carmona, F., Casamitjana, R., et al., 2000. [Day 3 serum inhibin B and FSH and age as predictors of assisted reproduction treatment outcome.](#) *Human Reproduction* 15, 2341–2346.
- Daya, S., 2000. [Gonadotropin releasing hormone agonist protocols for pituitary desensitization in in vitro fertilization and gamete](#)

- intrafallopian transfer cycles. *Cochrane Database and Systematic Reviews* 2, CD001299.
- Dhillon, R.K., McLernon, D.J., Smith, P.P., Fishel, S., Dowell, K., Deeks, J.J., et al., 2016. Predicting the chance of live birth for women undergoing IVF: A novel pretreatment counselling tool. *Human Reproduction* 31 (1), 84–92.
- Dzik, A., Lambert-Messerlian, G., Izzo, V.M., Soares, J.B., Pinotti, J.A., Seifer, D.B., et al., 2008. Anti-müllerian hormone: Correlation of early follicular, ovulatory and midluteal levels with ovarian response and cycle outcome in intracytoplasmic sperm injection patients. *Fertility and Sterility* 89, 1670–1676.
- Erdem, M., Erdem, A., Gursoy, R., Biberoglu, K., 2004. Comparison of basal and clomiphene citrate induced FSH and inhibin B, ovarian volume and antral follicle counts as ovarian reserve tests and predictors of poor ovarian response in IVF. *Journal of Assisted Reproduction and Genetics* 21, 37–45.
- Fábregues, F., Balasch, J., Creus, M., Carmona, F., Puerto, B., Quintó, L., et al., 2000. Ovarian reserve test with human menopausal gonadotropin as a predictor of in vitro fertilization outcome. *Journal of Assisted Reproduction and Genetics* 17, 13–19.
- Fang, T., Su, Z., Wang, L., Yuan, P., Li, R., Ouyang, N., et al., 2015. Predictive value of age-specific FSH levels for IVF-ET outcome in women with normal ovarian function. *Reproductive Biology and Endocrinology* 13, 63.
- Fawzy, M., Lambert, A., Harrison, R.F., Knight, P.G., Groome, N., Hennelly, B., et al., 2002. Day 5 inhibin B levels in a treatment cycle are predictive of IVF outcome. *Human Reproduction* 17, 1535–1543.
- Gingold, J.A., Lee, J.A., Whitehouse, M.C., Rodriguez-Purata, J., Sandler, B., Grunfeld, L., et al., 2015. Maximum basal FSH predicts reproductive outcome better than cycle-specific basal FSH levels: Waiting for a “better” month conveys limited retrieval benefits. *Reproductive Biology and Endocrinology* 13, 91.
- Guo, J., Zhang, Q., Li, Y., Huang, J., Wang, W., Huang, L., et al., 2014. Predictive value of androgens and multivariate model for poor ovarian response. *Reproductive Biomedicine Online* 28, 723–732.
- Hamdine, O., Eijkemans, M.J., Lentjes, E.W., Torrance, H.L., Macklon, N.S., Fauser, B.C., et al., 2015. Ovarian response prediction in GnRH antagonist treatment for IVF using anti-müllerian hormone. *Human Reproduction* 30, 170–178.
- Heidar, Z., Bakhtiyari, M., Mirzamoradi, M., Zadehmodarres, S., Sarfjoo, F.S., Mansournia, M.A., 2015. Prediction of different ovarian responses using anti-müllerian hormone following a long agonist treatment protocol for IVF. *Journal of Endocrinological Investigation* 38, 1007–1015.
- Hendriks, D.J., Broekmans, F.J., Bancsi, L.F., de Jong, F.H., Looman, C.W., te Velde, E.R., 2005a. Repeated clomiphene citrate challenge testing in the prediction of outcome in IVF: A comparison with basal markers for ovarian reserve. *Human Reproduction* 20, 163–169.
- Hendriks, D.J., Broekmans, F.J., Bancsi, L.F., Looman, C.W., de Jong, F.H., te Velde, E.R., 2005b. Single and repeated GnRH agonist stimulation tests compared with basal markers of ovarian reserve in the prediction of outcome in IVF. *Journal of Assisted Reproduction and Genetics* 22, 65–73.
- Hendriks, D.J., Klinkert, E.R., Bancsi, L.F., Looman, C.W., Habema, J.D., te Velde, E.R., et al., 2004. Use of stimulated serum estradiol measurements for the prediction of hyperresponse to ovarian stimulation in in vitro fertilization (IVF). *Journal of Assisted Reproduction and Genetics* 21, 65–72.
- Howles, C.M., Saunders, H., Alam, V., Engrand, P., FSH Treatment Guidelines Clinical Panel, 2006. Predictive factors and a corresponding treatment algorithm for controlled ovarian stimulation in patients treated with recombinant human follicle stimulating hormone (folitropin alfa) during assisted reproduction technology (ART) procedures. An analysis of 1378 patients. *Current Medicine Research Opinion* 22, 907–918.
- Hsu, A., Arny, M., Knee, A.B., Bell, C., Cook, E., Novak, A.L., et al., 2011. Antral follicle count in clinical practice: Analyzing clinical relevance. *Fertility and Sterility* 95, 474–479.
- Ng, E.H., Chan, C.C., Tang, O.S., Ho, P.C., 2005a. Antral follicle count and FSH concentrations after clomiphene citrate challenge test in the prediction of ovarian response during IVF treatment. *Human Reproduction* 20, 1647–1654.
- Ng, E.H., Tang, O.S., Chan, C.C., Ho, P.C., 2005b. Ovarian stromal blood flow in the prediction of ovarian response during in vitro fertilization treatment. *Human Reproduction* 20, 3147–3151.
- Dzik, A., Lambert-Messerlian, G., Izzo, V.M., Soares, J.B., Pinotti, J.A., Seifer, D.B., 2000. Inhibin B response to EFORT is associated with the outcome of oocyte retrieval in the subsequent in vitro fertilization cycle. *Fertility and Sterility* 74, 1114–1117.
- Jayaprakasan, K., Al-Hasie, H., Jayaprakasan, R., Campbell, B., Hopkisson, J., Johnson, I., et al., 2009a. The three-dimensional ultrasonographic ovarian vascularity of women developing poor ovarian response during assisted reproduction treatment and its predictive value. *Fertility and Sterility* 92, 1862–1869.
- Jayaprakasan, K., Jayaprakasan, R., Al-Hasie, H.A., Clewes, J.S., Campbell, B.K., Johnson, I.R., et al., 2009b. Can quantitative three-dimensional power Doppler angiography be used to predict ovarian hyperstimulation syndrome? *Ultrasound in Obstetrics and Gynecology* 33, 583–591.
- Jones, C.A., Christensen, A.L., Salihu, H.M., Carpenter, W., Petrozino, J., Abrams, E., et al., 2011. Prediction of individual probabilities of livebirth and multiple birth events following in vitro fertilization (IVF): A new outcomes counselling tool for IVF providers and patients using HFEA metrics. *Journal of Experimental & Clinical Assisted Reproduction* 8, 3.
- Khairy, M., Clough, A., el-Toukhy, T., Coomarasamy, A., Khalaf, Y., 2008. Antral follicle count at down-regulation and prediction of poor ovarian response. *Reproductive Biomedicine Online* 17, 508–514.
- Kunt, C., Ozaksit, G., Keskin Kurt, R., Cakir Gungor, A.N., Kanat-Pektas, M., Kilic, S., et al., 2011. Anti-müllerian hormone is a better marker than inhibin B, follicle stimulating hormone, estradiol or antral follicle count in predicting the outcome of in vitro fertilization. *Archives of Gynecology and Obstetrics* 283, 1415–1421.
- Kwee, J., Elting, M.E., Schats, R., McDonnell, J., Lambalk, C.B., 2007. Ovarian volume and antral follicle count for the prediction of low and hyper responders with in vitro fertilization. *Reproductive Biology and Endocrinology* 5, 9.
- Kwee, J., Schats, R., McDonnell, J., Schoemaker, J., Lambalk, C.B., 2006. The clomiphene citrate challenge test versus the exogenous follicle-stimulating hormone ovarian reserve test as a single test for identification of low responders and hyperresponders to in vitro fertilization. *Fertility and Sterility* 85, 1714–1722.
- Kwee, J., Schats, R., McDonnell, J., Themmen, A., de Jong, F., Lambalk, C., 2008. Evaluation of anti-müllerian hormone as a test for the prediction of ovarian reserve. *Fertility and Sterility* 90, 737–743.
- La Cour Freiesleben, N., Gerds, T.A., Forman, J.L., Silver, J.D., Nyboe Andersen, A., Popovic-Todorovic, B., 2011. Risk charts to identify low and excessive responders among first-cycle IVF/ICSI standard patients. *Reproductive Biomedicine Online* 22, 50–58.
- La Marca, A., Grisendi, V., Giulini, S., Argento, C., Tirelli, A., Dondi, G., et al., 2013. Individualization of the FSH starting dose in IVF/ICSI cycles using the antral follicle count. *Journal of Ovarian Research* 6, 11.
- La Marca, A., Nelson, S.M., Sighinolfi, G., Manno, M., Baraldi, E., Roli, L., et al., 2011. Anti-müllerian hormone-based prediction

- model for a live birth in assisted reproduction. *Reproductive Biomedicine Online* 22, 341–349.
- La Marca, A., Papaleo, E., Grisendi, V., Argento, C., Giulini, S., Volpe, A., 2012. Development of a nomogram based on markers of ovarian reserve for the individualisation of the follicle-stimulating hormone starting dose in in vitro fertilisation cycles. *British Journal of Obstetrics and Gynecology* 119, 1171–1179.
- La Marca, A., Sunkara, S.K., 2014. Individualization of controlled ovarian stimulation in IVF using ovarian reserve markers: From theory to practice. *Human Reproduction Update* 20 (1), 124–1124.
- Lee, T.H., Liu, C.H., Huang, C.C., Wu, Y.L., Shih, Y.T., Ho, H.N., et al., 2008. Serum anti-müllerian hormone and estradiol levels as predictors of ovarian hyperstimulation syndrome in assisted reproduction technology cycles. *Human Reproduction* 23, 160–167.
- Leushuis, E., van der Steeg, J.W., Steures, P., Bossuyt, P.M., Eijkemans, M.J., van der Veen, F., et al., 2009. Prediction models in reproductive medicine: A critical appraisal. *Human Reproduction Update* 15, 537–552.
- Li, R., Gong, F., Zhu, Y., Fang, W., Yang, J., Liu, J., et al., 2016. Anti-müllerian hormone for prediction of ovarian response in Chinese infertile women undergoing IVF/ICSI cycles: A prospective, multi-centre, observational study. *Reproductive Biomedicine Online* 33, 506–512.
- Lorusso, F., Vicino, M., Lamanna, G., Trerotoli, P., Serio, G., Depalo, R., 2007. Performance of different ovarian reserve markers for predicting the numbers of oocytes retrieved and mature oocytes. *Maturitas* 56, 429–435.
- Luke, B., Brown, M.B., Wantman, E., Jindal, S.K., Racowsky, C., Ball, G.D., 2014. A prediction model for live birth and multiple births within the first three cycles of assisted reproductive technology. *Fertility and Sterility* 102 (3), 744–752.
- Magnusson, Å., Nilsson, L., Oleröd, G., Thurin-Kjellberg, A., Bergh, C., 2017. The addition of anti-müllerian hormone in an algorithm for individualized hormone dosage did not improve the prediction of ovarian response—a randomized, controlled trial. *Human Reproduction* 32, 811–819.
- McIlveen, M., Skull, J.D., Ledger, W.L., 2007. Evaluation of the utility of multiple endocrine and ultrasound measures of ovarian reserve in the prediction of cycle cancellation in a high-risk IVF population. *Human Reproduction* 22, 778–785.
- Moon, K.Y., Kim, H., Lee, J.Y., Lee, J.R., Jee, B.C., Suh, C.S., et al., 2016. Nomogram to predict the number of oocytes retrieved in controlled ovarian stimulation. *Clinical Experimental Reproductive Medicine* 43, 112–118.
- Mutlu, M.F., Erdem, M., Erdem, A., Yildiz, S., Mutlu, I., Arisoy, O., et al., 2013. Antral follicle count determines poor ovarian response better than anti-müllerian hormone but age is the only predictor for live birth in in vitro fertilization cycles. *Journal of Assisted Reproduction and Genetics* 30, 657–665.
- Muttukrishna, S., McGarrigle, H., Wakim, R., Khadum, I., Ranieri, D.M., Serhal, P., 2005. Antral follicle count, anti-müllerian hormone and inhibin B: Predictors of ovarian response in assisted reproductive technology? *British Journal of Obstetrics and Gynaecology* 112, 1384–1390.
- Nelson, S.M., Lawlor, D.A., 2011. Predicting live birth, preterm delivery and low birth weight in infants born from in vitro fertilisation: A prospective study of 144,018 treatment cycles. *PLOS Medicine* 8 (1), e100038.
- Nelson, S.M., Klein, B.M., Arce, J.C., 2015. Comparison of antimüllerian hormone levels and antral follicle count as predictor of ovarian response to controlled ovarian stimulation in good-prognosis patients at individual fertility clinics in 2 multicenter trials. *Fertility and Sterility* 103, 923–930.
- Nelson, S.M., Yates, R.W., Fleming, R., 2007. Serum anti-müllerian hormone and FSH: prediction of live birth and extremes of response in stimulated cycles—implications for individualization of therapy. *Human Reproduction* 22, 2414–2421.
- Oehninger, S., Nelson, S.M., Verweij, P., Stegmann, B.J., 2015. Predictive factors for ovarian response in a corifollitropin alfa/GnRH antagonist protocol for controlled ovarian stimulation in IVF/ICSI cycles. *Reproductive Biology and Endocrinology* 13, 117.
- Olivennes, F., Howles, C.M., Borini, A., Germond, M., Trew, G., Wikland, M., et al., CONSORT study group, 2009. Individualizing FSH dose for assisted reproduction using a novel algorithm: The CONSORT study. *Reproductive Biomedicine Online* 18, 195–204.
- Papaleo, E., Zaffagnini, S., Munaretto, M., Vanni, V.S., Rebonato, G., Grisendi, V., et al., 2016. Clinical application of a nomogram based on age, serum FSH and AMH to select the FSH starting dose in IVF/ICSI cycles: A retrospective two-centres study. *European Journal of Obstetrics & Gynecology and Reproductive Biology* 207, 94–99.
- Peñarrubia, J., Balasch, J., Fábregues, F., Carmona, F., Casamitjana, R., Moreno, V., et al., 2000. Day 5 inhibin B serum concentrations as predictors of assisted reproductive technology outcome in cycles stimulated with gonadotrophin-releasing hormone agonist-gonadotrophin treatment. *Human Reproduction* 15, 1499–1504.
- Peñarrubia, J., Fábregues, F., Manau, D., Creus, M., Carmona, F., Casamitjana, R., et al., 2005. Previous cycle cancellation due to poor follicular development as a predictor of ovarian response in cycles stimulated with gonadotrophin-releasing hormone agonist-gonadotrophin treatment. *Human Reproduction* 20, 622–628.
- Peñarrubia, J., Peralta, S., Fábregues, F., Carmona, F., Casamitjana, R., Balasch, J., 2010. Day-5 inhibin B serum concentrations and antral follicle count as predictors of ovarian response and live birth in assisted reproduction cycles stimulated with gonadotropin after pituitary suppression. *Fertility and Sterility* 94, 2590–2595.
- Peralta, S., Solernou, R., Fabregues, F., Minarro, A., Puerto, B., Penarrubia, J., et al., 2017. Antral follicle count measured after pituitary suppression as predictor of sub-optimal ovarian response and cumulative live birth: A review of 1542 long protocol IVF/ICSI cycles. *Fertility and Sterility* 108, e218.
- Polyzos, N.P., Tournaye, H., Guzman, L., Camus, M., Nelson, S.M., 2013. Predictors of ovarian response in women treated with corifollitropin alfa for in vitro fertilization/intracytoplasmic sperm injection. *Fertility and Sterility* 100, 430–437.
- Popovic-Todorovic, B., Loft, A., Bredkjaer, H.E., Bangsbo, S., Nielsen, I.K., Andersen, A.N., 2003. A prospective randomized clinical trial comparing an individual dose of recombinant FSH based on predictive factors versus a “standard” dose of 150 IU/day in “standard” patients undergoing IVF/ICSI treatment. *Human Reproduction* 18, 2275–2282.
- Popovic-Todorovic, B., Loft, A., Lindhard, A., Bangsbo, S., Andersen, A.M., Andersen, A.N., 2003. A prospective study of predictive factors of ovarian response in “standard” IVF/ICSI patients treated with recombinant FSH. A suggestion for a recombinant FSH dosage nomogram. *Human Reproduction* 18, 781–787.
- Pouly, J.L., Olivennes, F., Massin, N., Celle, M., Caizergues, N., Contard, F., French CONSORT Study Group, 2015. Usability and utility of the CONSORT calculator for FSH starting doses: A prospective observational study. *Reproductive Biomedicine Online* 31, 347–355.
- Rombauts, L., Lambalk, C.B., Schultze-Mosgau, A., van Kuijk, J., Verweij, P., Gates, D., et al., 2015. Intercycle variability of the ovarian response in patients undergoing repeated stimulation with corifollitropin alfa in a gonadotropin-releasing hormone antagonist protocol. *Fertility and Sterility* 104, 884–890.

- Sarais, V., Reschini, M., Busnelli, A., Biancardi, R., Paffoni, A., Somigliana, E., 2016. [Predicting the success of IVF: External validation of the van Loendersloot's model](#). *Human Reproduction*. 31 (6), 1245–1252.
- Scott, R.T., Toner, J.P., Muasher, S.J., Oehninger, S., Robinson, S., Rosenwaks, Z., 1989. [Follicle-stimulating hormone levels on cycle day 3 are predictive of in vitro fertilization outcome](#). *Fertility and Sterility* 51, 651–654.
- Singh, N., Bahadur, A., Malhotra, N., Kalaivani, M., Mittal, S., 2013. [Prospective analysis of ovarian reserve markers as determinant in response to controlled ovarian stimulation in women undergoing IVF cycles in low resource setting in India](#). *Archives of Gynecology and Obstetrics* 288, 697–703.
- Smith, A.D., Tilling, K., Lawlor, D.A., Nelson, S.M. (2015). External validation and calibration of IVFpredict: a national prospective cohort study of 130,960 in vitro fertilisation cycles. *PLoS One*. 10 (4), e0121357.
- Sunkara, S.K., Rittenberg, V., Raine-Fenning, N., Bhattacharya, S., Zamora, J., Coomarasamy, A., 2011. [Association between the number of eggs and live birth in IVF treatment: An analysis of 400 135 treatment cycles](#). *Human Reproduction* 26, 1768–1774.
- Templeton, A., Morris, J.K., Parslow, W., 1996. [Factors that affect outcome of in vitro fertilisation treatment](#). *The Lancet*. 348, 1042–1046.
- Van Loendersloot, L., Repping, S., Bossuyt, P.M., van der Veen, F., van Wely, M., 2014. [Prediction models in in vitro fertilization: Where are we? A mini review](#). *Journal of Advanced Research* 5, 295–301.
- Van Loendersloot, L.L., van Wely, M., Repping, S., van der Veen, F., Bossuyt, P.M., 2011. [Templketon prediction model underestimates IVF success in an external validation](#). *Reproductive Biomedicine Online*. 22, 597–602.
- Van Loendersloot, van Wely, M., Repping, S., Bossuyt, P.M., van der Veen, F., 2013. [Individualized decision-making in IVF: Calculating the chances of pregnancy](#). *Human Reproduction*. 28 (11), 2972–2980.
- Van Rooij, I.A., Broekmans, F.J., te Velde, E.R., Fauser, B.C., Bancsi, L.F., de Jong, F.H., et al., 2002. [Serum anti-müllerian hormone levels: A novel measure of ovarian reserve](#). *Human Reproduction* 17, 3065–3071.
- Vera-Rodriguez M, Chavez, S.L., Rubio, C., Reijo Pera, R.A., Simon, C., 2015. [Prediction model for aneuploidy in early human embryo development revealed by single-cell analysis](#). *Nature Communications* 6, 7601.
- Verberg, M.F., Eijkemans, M.J., Macklon, N.S., Heijnen, E.M., Fauser, B.C., Broekmans, F.J., 2007. [Predictors of low response to mild ovarian stimulation initiated on cycle day 5 for IVF](#). *Human Reproduction* 22, 1919–1924.
- Vural, B., Cakiroglu, Y., Vural, F., Filiz, S., 2014. [Hormonal and functional biomarkers in ovarian response](#). *Archives of Gynecology and Obstetrics* 289, 1355–1361.