

Prueba no invasiva de madurez pulmonar fetal

Revisión sistemática

*Non invasive fetal lung
maturity test*

INFORMES DE EVALUACIÓN DE TECNOLOGÍAS SANITARIAS
AETSA

INFORMES, ESTUDIOS E INVESTIGACIÓN



MINISTERIO
DE SANIDAD, SERVICIOS SOCIALES
E IGUALDAD



RED ESPAÑOLA DE AGENCIAS DE EVALUACIÓN
DE TECNOLOGÍAS Y PRESTACIONES DEL SISTEMA NACIONAL DE SALUD



JUNTA DE ANDALUCÍA
CONSEJERÍA DE SALUD

Prueba no invasiva de madurez pulmonar fetal

Revisión sistemática

*Non invasive fetal lung
maturity test*

INFORMES DE EVALUACIÓN DE TECNOLOGÍAS SANITARIAS
AETSA

Forcada Falcon, Mercedes

Prueba no invasiva de madurez pulmonar. Revisión sistemática.
Mercedes Forcada Falcón, Carolina Carral Sutil, Elena Baños Álvarez,
Ana M^a Carlos Gil, María Piedad Rosario Lozano, Teresa Molina López. —
Sevilla: Agencia de Evaluación de Tecnologías Sanitarias de Andalucía, 2016
58 p; 24 cm. (Colección: Informes, estudios e investigación.
Ministerio de Sanidad, Servicios Sociales e Igualdad. Serie: Informes de
Evaluación de Tecnologías Sanitarias)

ISBN: 978-84-15600-92-3

1. Madurez de los órganos fetales 2. Pulmón / ultrasonografía
3. Embarazo I. Carral Sutil, Carolina II. Baños Álvarez, Elena III. Carlos Gil,
Ana M^a IV. Rosario Lozano, María Piedad V. Molina López, Teresa.
VI. Andalucía. Agencia de Evaluación de Tecnologías Sanitarias VII. España.
Ministerio de Sanidad, Servicios Sociales e Igualdad VIII. España. Ministerio
de Economía y Competitividad.

Autores: Mercedes Forcada-Falcón, Carolina Carral-Sutil, Elena Baños-Álvarez,
Ana M.^a Carlos-Gil, María Piedad Rosario-Lozano, Teresa Molina-López.

Este documento ha sido financiado por el Ministerio de Sanidad, Servicios Sociales
e Igualdad, en el marco del plan anual de trabajo de la Red Española de Agencias de
Evaluación de Tecnologías y Prestaciones del SNS, aprobado en el Pleno del Consejo
Interterritorial de 13 de abril de 2016

Edita: Agencia de Evaluación de Tecnologías Sanitarias de Andalucía
Consejería de Salud

JUNTA DE ANDALUCÍA

Avda. de la Innovación, s/n. Edificio ARENA 1, s/n. Planta baja.
41020 Sevilla
España – Spain

ISBN: 978-84-15600-92-3

NIPO: 680-17-109-X

Para citar este informe:

Forcada-Falcón M, Carral-Sutil C, Baños-Álvarez E, Carlos-Gil AM, Rosario-
Lozano MP, Molina-López T. Prueba no invasiva de madurez pulmonar fetal.
Sevilla: Agencia de Evaluación de Tecnologías Sanitarias de Andalucía. Red
Española de Agencias de Evaluación de Tecnologías Sanitarias y Prestaciones
del SNS; 2017.

Este documento puede ser reproducido en todo o en parte, por cualquier medio,
siempre que se cite explícitamente su procedencia

Prueba no invasiva de madurez pulmonar fetal

Revisión sistemática

*Non invasive fetal lung
maturity test*

INFORMES DE EVALUACIÓN DE TECNOLOGÍAS SANITARIAS
AETSA



MINISTERIO
DE SANIDAD, SERVICIOS SOCIALES
E IGUALDAD



RED ESPAÑOLA DE AGENCIAS DE EVALUACIÓN
DE TECNOLOGÍAS Y PRESTACIONES DEL SISTEMA NACIONAL DE SALUD



JUNTA DE ANDALUCÍA
CONSEJERÍA DE SALUD

Conflicto de interés

Las autoras declaran que no tienen intereses que puedan competir con el interés primario y los objetivos de este informe e influir en su juicio profesional al respecto.

Contribución de las autoras

- Planificación y diseño de la investigación: Elena Baños-Álvarez, Ana M.^a Carlos-Gil, Teresa Molina-López.
- Documentación: María Piedad Rosario-Lozano, Mercedes Forcada-Falcón.
- Obtención de los datos: Mercedes Forcada-Falcón, Carolina Carral-Sutil.
- Análisis y presentación de resultados: Mercedes Forcada-Falcón, Carolina Carral-Sutil, Elena Baños-Álvarez, Ana M.^a Carlos-Gil.
- Elaboración del manuscrito: Mercedes Forcada-Falcón, Carolina Carral-Sutil, Elena Baños-Álvarez, Ana M.^a Carlos-Gil, Teresa Molina-López.

Este manuscrito ha sido leído y aprobado por todas las autoras.

Agradecimientos

Este trabajo se ha beneficiado de forma importante de las aportaciones del Dr. Lucas Cerrillos González, Jefe de Sección de Obstetricia del Hospital de la Mujer del Hospital Universitario Virgen del Rocío (Sevilla); Dr. Rafael Torrejón Cardoso, Director UGC Intercentros Atención Integral a la Mujer, Hospital Universitario Puerta del Mar/Puerto Real (Cádiz) y Dra. Soledad Benot López, Agencia de Evaluación de Tecnologías Sanitarias de Andalucía, Consejería de Salud.

La Agencia de Evaluación de Tecnologías Sanitarias de Andalucía y las autoras agradecen a los revisores de este texto el esfuerzo realizado, su dedicación y sus aportaciones.

Los contenidos del informe son responsabilidad de las autoras.

Índice

Índice	9
Índice de tablas y figuras	11
Puntos clave	13
Key points	15
Justificación	17
Introducción	19
Objetivos	23
Metodología	25
Resultados	29
Aspectos económicos	41
Discusión	43
Referencias	47
Anexos	51

Índice de tablas y figuras

Tabla 1. Condiciones maternas que alteran la madurez pulmonar fetal.....	19
Tabla 2. Métodos bioquímicos diagnósticos de madurez pulmonar	20
Tabla 3. Características de los estudios incluidos	30
Tabla 4. Características de los grupos del estudio Palacio <i>et al.</i> 2012.....	33
Tabla 5. Nivel de evidencia de los estudios de pruebas diagnósticas	38
Tabla 6. Resultados de eficacia en términos de validez diagnóstica en ambos estudios	38
Figura 1: Gráfica sobre riesgo de sesgo y aplicabilidad de los estudios considerados	37
Figura 2: Resumen sobre riesgo de sesgo y aplicabilidad de los estudios considerados	37

Puntos clave

- La prueba no invasiva para la determinación de madurez pulmonar estima el riesgo de morbilidad respiratoria antes del nacimiento mediante un análisis informático que combina hallazgos por ultrasonidos y edad gestacional. Existen dos dispositivos disponibles AQUA y quantusFLM™.
- Se realizó una revisión sistemática para determinar la seguridad, eficacia (validez diagnóstica y precisión) y la efectividad (morbimortalidad del recién nacido pretérmino) de la prueba no invasiva para la determinación de la madurez pulmonar fetal en mujeres embarazadas con riesgo de parto pretérmino.
- Se consultaron las bases de datos referenciales MedLine, EMBASE, *Web of Science* y *Cochrane Library* (hasta abril de 2016). También se consultaron las bases de datos del *Centre for Reviews and Dissemination (CRD)*, el *National Institute for Health and Care Excellence (NICE)*, *International Network of Agencies for Health Technology Assessment (INAHTA)*, el *International Information Network on New and Emerging Health Technologies (EuroScan)*, la Sociedad Española de Neonatología (SENeo) y otras fuentes de información específicas de tecnologías emergentes.
- Se seleccionaron estudios de pruebas diagnósticas que se aplicaron a mujeres embarazadas con riesgo de parto pretérmino. Sólo se incluyeron los estudios que mostraban resultados en términos de seguridad, eficacia y/o efectividad.
- Un investigador extrajo los datos de los artículos. Se evaluó la calidad según la herramienta QUADAS. Para determinar el nivel de evidencia se siguieron los criterios descritos por NICE. Se realizó una síntesis cualitativa de los datos.
- Se identificaron 52 referencias bibliográficas de las que 2 estudios fueron finalmente incluidos para la síntesis. El nivel de evidencia asignado a los estudios fue II/III debido a la posibilidad de riesgo de sesgos moderado/alto. La sensibilidad y especificidad obtenidas para el método de AQUA con el objetivo de predecir madurez pulmonar fetal fue del 95,1 % y del 85,7 %, respectivamente, en comparación con la prueba de referencia (prueba TDx-FLM II en el líquido amniótico). La eficacia de quantusFLM™ para predecir

morbimortalidad respiratoria en los neonatos en términos de sensibilidad, especificidad y precisión diagnóstica fue del 86,2 %, 87 %, 86,8 %, respectivamente, refiriendo valores más bajos en el rango de edad entre 34 + 0 a 39 + 0 semanas de gestación, aun en presencia de mayor tamaño muestral que el grupo de menor edad gestacional. No se facilitaron datos sobre seguridad y efectividad en ninguno de los dos estudios considerados.

- La revisión de la literatura actualizada hasta la fecha nos indica que los resultados son muy preliminares y al no haber resultados publicados de estudios actualmente en marcha, nos orienta a que la prueba está en fase muy inicial. La evidencia identificada no permite establecer indicaciones.

Key points

- Before the birth, non invasive fetal lung maturity test predict the neonatal respiratory morbidity risk with a computer analysis. This analysis combines ultrasound and the gestational age. Two systems are available, AQUA and quantusFLM™.
- A systematic review was performed to determine the safety, efficacy (diagnostic validity and precision) and effectiveness (morbidity and mortality in preterm infants) of quantitative analysis of ultrasound fetal lung texture in pregnant women at risk of preterm delivery.
- Medline, EMBASE, Web of Science and Cochrane Library were consulted (until April 2016). We also searched in the database of The Centre for Reviews and Dissemination (CRD), the National Institute for Health and Care Excellence (NICE), the International Network of Agencies for Health Technology Assessment (INAHTA), International Information Network on New and Emerging Health Technologies (EuroScan), Sociedad Española de Neonatología (SENeo) and specific information sources of emerging technologies.
- Studies about diagnostic test included pregnant woman at risk for preterm delivery. We only considered those studies with results in terms of safety, efficacy and/or effectiveness.
- A researcher extracted data from the articles included. Two researchers evaluated the quality and risk of bias using the QUADAS tool. We determined the level of evidence in accordance with the NICE criteria. We conducted a qualitative data synthesis.
- We identified 52 references without duplicates, of which 2 studies were finally included. The level of evidence was II/III and the risk of bias was moderate/high. To predict lung maturity, when compared with the reference test, the sensitivity and specificity obtained for the AQUA method was 95.1 % and 85.7 %, respectively. The quantusFLM™ diagnostic utility in terms of sensitivity, specificity and diagnostic accuracy was high (86.2 %, 87 %, 86.8 % respectively). The lowest value was in the range of 34 + 0 to 39 + 0 week of gestation, despite having larger sample size. The two studies considered have not provided data on safety and effectiveness.

Justificación

Este informe de evaluación ha sido realizado a petición de la Comisión de Prestaciones, Aseguramiento y Financiación, dependiente del Consejo Interterritorial del Sistema Nacional de Salud, en el marco del plan anual de trabajo de la Red de Agencias.

El Observatorio de Tecnologías Emergentes de la Agencia de Evaluación de Tecnologías Sanitarias de Andalucía (AETSA) detecta, prioriza y evalúa las tecnologías nuevas y emergentes, proporcionando herramientas que anticipan el impacto de éstas y ayudan a la toma de decisiones por parte de la Administración Sanitaria y el personal sanitario.

Su objetivo es evaluar la validez diagnóstica de una prueba no invasiva, basada en las características de textura pulmonar fetal (obtenidas mediante ecografía y analizadas mediante un software específico) combinadas con la edad gestacional. Esta prueba pretende diagnosticar la maduración pulmonar fetal en embarazadas con riesgo de parto pretérmino sin las complicaciones que conlleva la realización de una amniocentesis.

Introducción

La definición de prematuro^{1,2} incluye a todos los recién nacidos con edad gestacional menor de 37 semanas. Según estimaciones de la Organización Mundial de la Salud (OMS), en el mundo nacen 15 millones de niños prematuros cada año. En España, en el año 2014, nacieron 27.992 prematuros (según los datos del Instituto Nacional de Estadística, INE) lo que supone el 6,67 % de todos los partos ocurridos ese año. Aproximadamente un tercio de los partos prematuros se producen por rotura prematura de membranas² antes de la semana 37 y a menudo por razones desconocidas. A veces, el nacimiento prematuro se debe a la indicación de una cesárea electiva^{2,3} realizada con el objetivo de adelantar la fecha prevista de parto. Esto suele deberse a la aparición de situaciones clínicas que precisan este abordaje, ya sea por patología materna o fetal^{4,5}.

En cualquier caso, los recién nacidos prematuros se enfrentan a riesgos inherentes a la propia prematuridad (distrés respiratorio, hemorragia, infección, etc.)² y existe un aumento de la morbilidad asociada sobre todo a patología respiratoria neonatal^{6,7} cuya causa más frecuente es la inmadurez pulmonar fetal⁸.

Dicha inmadurez pulmonar se mantendría incluso más allá de las 37 semanas en algunos neonatos, existiendo estudios^{8,9} que refieren que los neonatos de menos de 39 semanas de gestación presentan una incidencia más alta de morbilidad respiratoria en comparación con los de 39 semanas o más, recomendándose evitar el parto antes de dicha semana¹. La OMS señala que muchos de los recién nacidos prematuros con inmadurez pulmonar que sobreviven presentarán importantes discapacidades a lo largo de su vida, por lo que sería importante valorar la maduración pulmonar previamente al parto con el objetivo de precisar lo más posible su maduración y programar la fecha de parto, permitiendo así la maduración fetal mediante administración de corticoides prenatales¹⁰ y surfactante postnatal¹¹.

La edad gestacional es el predictor más fuerte de inmadurez pulmonar⁸ aunque existen condiciones que pueden modificarla, en la Tabla 1 se describen de forma detallada.

Tabla 1. Condiciones maternas que alteran la madurez pulmonar fetal

Condiciones que aceleran la madurez pulmonar	Condiciones que retrasan la madurez pulmonar
Hipertensión arterial crónica	Diabetes Mellitus descontrolada
Estados hipertensivos inducidos por el embarazo	Isoinmunización Rh
Restricción en el crecimiento	Hernia diafragmática fetal
Ruptura prematura de membranas prolongada	
Tabaquismo, alcoholismo, consumo de cocaína	

Tabla tomada de Ortega *et al.*, 2008¹¹.

Durante el proceso de maduración fetal, los pulmones producen surfactante para evitar el colapso de los alvéolos al final de la espiración⁶. El componente más importante del surfactante alveolar es la lecitina y su síntesis se incrementa hasta las 34 – 36 semanas de gestación, liberándose en parte al líquido amniótico^{6,11}. El déficit de surfactante alveolar se asocia al Síndrome de Distrés Respiratorio neonatal (SDR)^{6,11}, y la maduración pulmonar fetal, por tanto, se traduce en un aumento de la cantidad de surfactante en líquido amniótico en relación con otros componentes que permanecen estables (esfingomielina o albúmina)^{6,11}.

Existen diversas técnicas que miden esta relación entre los componentes del surfactante y los del líquido amniótico, cuyo objetivo es conocer la maduración pulmonar fetal¹¹. Estas técnicas precisan de un procedimiento invasivo, como la amniocentesis⁷, para la toma de muestras que no está exenta de riesgos maternos y fetales. Las más utilizadas¹¹ son:

- Métodos bioquímicos:
 - Cociente lecitina/esfingomielina, considerada la prueba de referencia durante muchos años.
 - Fosfatidil glicerol, desarrollado para mejorar la especificidad de la prueba anterior.
 - Recuento de cuerpos lamelares¹², que mide la concentración de surfactante pulmonar en líquido amniótico.
 - Proporción de surfactante/albúmina en líquido amniótico (TDx-FLM II)¹³ que presenta débil correlación con la madurez pulmonar fetal por lo que actualmente no se encuentra disponible. A continuación, en la Tabla 2 se describen los valores diagnósticos.

Método	Intervalos de referencia	Sensibilidad (%)	Especificidad (%)
L/S ratio	≥ 2: maduro	50 – 100	60 – 97
PG (Amniostat-FLM)	Positivo (≥ 0,5): maduro Negativo (≤ 0,5): inmaduro	83 – 100	47 – 88
S/A ratio	≥ 55 mg/g: maduro 40 – 54 mg/g: indeterminado ≤ 39 mg/g: inmaduro	96 – 100	60 – 84
LBC	> 50.000 / μL: maduro 16 – 49.000/μL: indeterminado < 15.000 /μL: inmaduro	83 – 100	54 – 100

L/S: lecitina/esfingomielina; PG: fosfatidilglicerol; S/A: surfactante/albúmina LBC: conteo cuerpos lamelares; ≥: mayor o igual; >: mayor; <: menor; μL: microlitros.

- Métodos biofísicos: test de Clements¹⁴, basado en la capacidad del surfactante de generar burbujas cuando se agita en presencia de etanol y la absorbancia del líquido amniótico, relacionado con el aumento progresivo de turbidez del líquido amniótico durante la gestación que se puede medir mediante espectrofotometría.

A raíz del auge de las técnicas de imagen, se han desarrollado nuevos métodos diagnósticos de madurez pulmonar fetal cuya ventaja principal frente a los anteriores es que no precisan tomar muestras de líquido amniótico (métodos no invasivos)¹⁵⁻¹⁸. Entre estos se encuentran: medición del tiempo de eyección de la arteria pulmonar fetal mediante Doppler, medición del volumen espiratorio traqueal en fetos pretérminos mediante ecografía, cambios en la maduración de la placenta observados mediante ecografía y cambios ecográficos en el pulmón fetal en relación con el hígado.

En los últimos años, se han desarrollado nuevas técnicas que tratan de evaluar el riesgo de morbilidad respiratoria neonatal en los casos en que se plantea realizar un parto inducido electivo, sobre todo entre las 32 y 39 semanas de gestación. Entre estas técnicas se encuentra la denominada AQUA⁶, un software informático que extrae y procesa imágenes ecográficas seleccionadas de pulmón fetal y que estima las características cuantitativas de textura del tejido.

Posteriormente, se observó que esta técnica presentaba limitaciones a la hora de extraer las características de textura de los tejidos⁷ por lo que el equipo investigador generó un nuevo modelo, utilizando los resultados obtenidos hasta el momento con el método de análisis cuantitativo de AQUA, mejorando la técnica mediante la combinación de algoritmos de aprendizaje, evitando así errores de asociación entre imágenes y textura pulmonar.

Esta nueva prueba se denominó (quantusFML™)⁷ y se basa en el análisis cuantitativo por ultrasonidos de imágenes del pulmón fetal. Este método procesa dichas imágenes y las correlaciona con la edad gestacional a través de un algoritmo, estimando su maduración y el riesgo de morbilidad respiratoria neonatal. La prueba se realiza en dos fases consecutivas⁷.

1. En la primera, se obtiene una ecografía del pulmón fetal (en un plano axial que permita la visualización de las cuatro cámaras cardíacas) que se almacena en formato DICOM y se obtiene la información clínica de la gestación.
2. Posteriormente, esta información se envía a una aplicación web que consta de dos módulos. Uno extrae la información de la imagen y la combina para seleccionar las características de textura pulmonar fetal y el otro es un algoritmo de clasificación que combina las características anteriores con la edad gestacional y estima el riesgo de morbilidad respiratoria neonatal. Para su construcción y validación se han empleado imágenes reales.

Su uso estaría restringido al ámbito hospitalario y necesitaría, para su correcto funcionamiento, de un ecógrafo y un ordenador con acceso a Internet.

La ventaja potencial que podría presentar con respecto a otros métodos similares es que sus resultados son independientes de las condiciones en las que se adquiere la imagen (ecógrafo, iluminación, etc.) y que no precisa de un procedimiento invasivo para su realización. Sin embargo, parece que no existe evidencia clara sobre su validez analítica y diagnóstica.

En relación con el estado de desarrollo de la tecnología, señalar que la prueba no invasiva de maduración pulmonar para determinación del riesgo de morbilidad respiratoria neonatal podría considerarse como una nueva indicación diagnóstica de una tecnología ya existente, como es el uso de ultrasonidos, los cuales ya cuentan con marcado *Conformité Europeen* (CE) y licencia de la *Food and Drug Administration* (FDA).

Objetivos

Los objetivos generales de los informes de síntesis de tecnologías emergentes son:

- Detectar precozmente nuevas tecnologías o cambios en las existentes con impacto potencial sobre el Sistema Sanitario.
- Sintetizar la información disponible sobre las tecnologías detectadas.
- Aportar información actualizada que ayude a la toma de decisiones en los distintos niveles del Sistema Sanitario.

Con este trabajo se pretende dar respuesta a la siguiente pregunta de investigación:

¿Es eficaz (en términos de validez diagnóstica y precisión), segura y efectiva (en términos de disminución de morbi-mortalidad por enfermedad pulmonar del recién nacido) la prueba no invasiva para el diagnóstico de maduración pulmonar fetal en las mujeres embarazadas con riesgo de parto pretérmino?

Los objetivos específicos se centran en:

- Valorar la eficacia de la prueba no invasiva para el diagnóstico de inmadurez pulmonar, en términos de validez diagnóstica (sensibilidad, especificidad, valores predictivos positivos, y negativos) y precisión (concordancia inter e intraobservador).
- Valorar la seguridad de dicha prueba.
- Valorar su efectividad, en términos de disminución de mortalidad y morbilidad de la enfermedad pulmonar en recién nacido.

Metodología

1. Tipo de estudio

Revisión sistemática de la literatura siguiendo las recomendaciones recogidas por la declaración PRISMA¹⁹. Además, para la elaboración del informe se siguió la Guía para la elaboración y adaptación de informes rápidos de evaluación de la Red Española de Agencias de Evaluación de Tecnologías y Prestaciones del SNS²⁰ y el *Quaderns 21 de Salut Pública i Administració de Serveis de Salut*²¹.

2. Búsqueda

La búsqueda se centró en localizar estudios de pruebas diagnósticas. Las siguientes bases de datos referenciales fueron consultadas hasta abril de 2016: MedLine, EMBASE, *Web of Science* y el registro de ensayos clínicos de la Cochrane Library. Además se consultó *International Network of Agencies for Health Technology Assessment (INAHTA)* a través de la base de datos del *Center for Reviews and Dissemination (CRD)*, en el *International Information Network on New and Emerging Health Technologies (EuroScan)*, en la Sociedad Española de Neonatología (SENeo), Asociación Española de Pediatría y en el registro de ensayos clínicos norteamericano *ClinicalTrials.gov* (<http://clinicaltrial.gov/>).

Se realizó una revisión manual en los sitios WEB de agencias no incluidas en INAHTA y de instituciones nacionales e internacionales como el Ministerio de Sanidad, Servicios Sociales e Igualdad, la plataforma de Agencias y Unidades de Evaluación de Tecnologías Sanitarias españolas (AUnETS), Organización Mundial de la Salud (OMS), los *Centers for Disease Control and Prevention (CDC)*, *The Emergency Care Research Institute (ECRI)*, Colaboración Cochrane, *National Institute for Health and Care Excellence (NICE)*, Google Académico, así como una revisión secundaria a partir de las referencias bibliográficas de los artículos recuperados.

La estrategia de búsqueda se muestra en el Anexo 1.

3. Criterios de selección de los artículos recuperados

Para la gestión de los documentos localizados se utilizó el programa *Mendeley Desktop* versión 1.16.1 y se aplicaron los siguientes criterios de inclusión y exclusión para realizar la selección.

Criterios de inclusión

Se han seleccionado aquellos criterios que responden a la pregunta de investigación formulada siguiendo el formato PICO (Población, Intervención, Comparación y Resultados/Outcomes).

- **Población:** Mujeres embarazadas en el tercer trimestre de gestación con riesgo de parto pretérmino (menos de 37 semanas de gestación).
- **Intervención:** Método no invasivo de valoración de la maduración pulmonar fetal.
- **Comparación:**
 - Cualquier otro método, invasivo o no, que permita valorar la maduración pulmonar fetal.
 - Sin prueba de referencia o comparador.
- **Resultados:** Validez diagnóstica (sensibilidad, especificidad y valores predictivos), utilidad (disminución de la morbi-mortalidad respiratoria neonatal) y seguridad del método.

Criterios de exclusión

- Estudios no originales: revisiones narrativas, cartas al director, editoriales, notas de campo.
- *Abstracts* de congresos.
- Estudios que evaluaran exclusivamente validez analítica.

4. Extracción de los datos

Para la extracción de los datos se elaboró una hoja de recogida de variables creada *ad hoc*. Se seleccionaron las variables que aportaran información general de los documentos incluidos para la síntesis de resultados. Estas variables fueron el autor y el año de publicación, el país donde se desarrolló el estudio, los objetivos planteados, las características de los pacientes incluidos, así como datos de la intervención evaluada y el periodo de seguimiento.

Como variables específicas se incluyeron aquellas que evaluaran la eficacia (en términos de validez diagnóstica y precisión), seguridad y efectividad (en términos de disminución de incidencia y morbilidad).

La síntesis de los resultados se realizó de forma cualitativa.

5. Evaluación de la calidad metodológica

Para la valoración del riesgo de sesgo en el establecimiento de la calidad de los artículos originales se siguieron los criterios recomendados por la herramienta QUADAS^{22,23}. Se utilizó el programa *Review Manager* 5.3 para obtener la representación gráfica de los resultados de la evaluación de la calidad. Para asignar el nivel de evidencia a los estudios se utilizó la escala de exactitud diagnóstica de NICE²⁴ (2006). Dicha escala se adjunta en el Anexo 2.

Resultados

1. Descripción de la evidencia disponible

Resultado de la búsqueda

La búsqueda bibliográfica realizada permitió identificar un total de 56 referencias. Tras eliminar duplicados, se revisaron por título y resumen 52, de los cuales 13 fueron seleccionados para su lectura a texto completo. Finalmente se incluyeron 2 estudios en el presente análisis. En el Anexo 3 se resume mediante diagrama de flujo el proceso de selección de los artículos y las causas de exclusión en cada caso. En el Anexo 4 se especifican los artículos excluidos tras su lectura a texto completo y los motivos de exclusión.

Descripción y calidad de los artículos

A continuación se recoge una descripción de los dos estudios incluidos (Tabla 3).

Las características de los estudios incluidos se describen más abajo atendiendo a su diseño, población, intervención y comparador. Además se realiza un análisis de la calidad de la evidencia y de los resultados clínicos.

Diseño de los estudios

Los dos estudios incluidos para síntesis de resultados fueron estudios de pruebas diagnósticas y ambos contenían resultados de validez diagnóstica^{6,7}.

El objetivo de los estudios analizados se centró en determinar la validez diagnóstica de dos métodos no invasivos para predecir la madurez pulmonar (AQUA y quantusFML™) y en uno de ellos se utilizó una prueba de referencia como comparación⁶. Ambos métodos fueron desarrollados por el mismo equipo de investigación. Además, los dos contaron con la aprobación del Comité Ético correspondiente y todas las personas incluidas firmaron un consentimiento informado.

Tabla 3. Características de los estudios incluidos								
Autor y año	Objetivo	Tipo de estudio	Población		Períodos de seguimiento	Intervención	Prueba de referencia	Resultados
			Embarazadas (n)	Semanas gestación				
Palacio <i>et al.</i> ⁶ 2012	Estimar el rendimiento de un método no invasivo para predecir la maduración pulmonar fetal	Prueba diagnóstica	n = 103 Subgrupo A = 52 Subgrupo B = 51	Rango: 24 – 41	48 horas	<ol style="list-style-type: none"> Obtención de ecografía pulmonar fetal y realización de amniocentesis Extracción de características de textura Predicción de madurez pulmonar fetal mediante algoritmo matemático (fase 1 o ensayo) Comparación de la predicción con el resultado de la amniocentesis (fase 2 o validación) Aplicación en dos modelos* 	Ratío surfactante/ albúmina en líquido amniótico	Sensibilidad Especificidad Precisión
Bonet- Carne <i>et al.</i> ⁷ 2015	Estimar el rendimiento de un método no invasivo para predecir la morbilidad respiratoria neonatal	Prueba diagnóstica	n = 150	Rango: 28 – 39	48 horas	<ol style="list-style-type: none"> Realización de ecografía pulmonar fetal Extracción características de textura Combinación datos con edad gestacional Clasificación del riesgo de morbilidad respiratoria neonatal 	ND	Sensibilidad Especificidad VPP VPN Precisión Morbimortalidad respiratoria neonatal

Subgrupo A: consta de 31 casos clasificados como inmadurez y 21 como madurez; Subgrupo B: consta de 31 casos clasificados como inmadurez y 20 como madurez; n: número de pacientes; ND: no datos; VPP: valor predictivo positivo; VPN: valor predictivo negativo
* El primer modelo se ensayó en el subgrupo B y se validó en el A; el segundo modelo se ensayó en el subgrupo A y se validó en el B. Los resultados en ambos estudios consistieron en medir la morbilidad y mortalidad respiratoria neonatal.

Descripción de la población

Los dos estudios incluidos se realizaron en España, en el Departamento de Medicina Materno-Fetal del Hospital Clínic de Barcelona. Los miembros del equipo de investigación en ambos casos pertenecen a este Departamento, al Grupo de Investigación de Medicina Fetal y Perinatal del Institut d'Investigacions Biomèdiques August Pi i Sunyer, al Centro de Investigación Biomédica en Red de Enfermedades Raras y a Transmural Biotech, S.L.

La población de estudio en ambos casos fue heterogénea, con mujeres embarazadas en el tercer trimestre de gestación con riesgo de parto prematuro o indicación de cesárea electiva por causas médicas (metrorragia, rotura prematura de membranas, preeclampsia, retraso del crecimiento intrauterino) dentro de las 48 horas siguientes a la intervención. Los criterios de exclusión en ambos estudios fueron embarazos múltiples y anomalías fetales genéticas o estructurales. En el estudio de Palacio *et al.* 2012⁶ se incluyeron mujeres embarazadas de 24 a 41 semanas de gestación y en el de Bonet-Carné *et al.* 2015⁷, mujeres embarazadas de 28 a 39 semanas de gestación.

La edad media de las pacientes incluidas en cada estudio osciló entre los 31,4 y 32,8 años (con una desviación estándar de $\pm 5,6$ y $\pm 5,7$, respectivamente). Alrededor del 60 % fueron nulíparas. La edad gestacional media en el momento del parto fue de 34,3 semanas (DE $\pm 4,1$) en el estudio de Palacio *et al.* 2012⁶ y de 36,1 semanas (DE $\pm 3,3$) en el de Bonet-Carné *et al.* 2015⁷. Las causas que motivaron el parto prematuro o cesárea electiva, según los casos, fueron: rotura prematura de membranas, preeclampsia o retraso del crecimiento intrauterino, entre las más importantes. El peso al nacer de los neonatos fue alrededor de 2.500 g en ambos estudios. En el 67,4 % de los casos del estudio de Bonet-Carné *et al.* 2015⁷, el parto fue por cesárea; no está disponible este dato en el otro estudio⁶.

En relación con el tamaño muestral, en el estudio de Bonet-Carné *et al.* 2015⁷ se especifica que se estableció en 150 sujetos, al estimar que así se incluirían entre 25 y 35 casos de morbilidad respiratoria neonatal. En el de Palacio *et al.* 2012⁶ no se especifica cómo se llevó a cabo el cálculo del tamaño muestral. En este caso, se optó por incluir a las mujeres embarazadas con indicación de amniocentesis para evaluar la madurez pulmonar fetal durante el periodo de octubre de 2010 a marzo de 2011, incluyéndose también las mujeres embarazadas con indicación de cesárea electiva por razones médicas. El tamaño muestral definitivo fue de 103 sujetos.

En los dos estudios se produjeron pérdidas. En el de Palacio *et al.* 2012⁶ se observó la pérdida de dos mujeres (y sus hijos) incluidos en el estudio sin que se especificaran las causas de la misma. En el de Bonet-Carné *et al.* 2015⁷, se produjo la pérdida de seis casos debido a la mala

calidad de las imágenes obtenidas. En ninguno de los dos artículos se planificó el tratamiento de las pérdidas producidas en el estudio.

Descripción de la intervención y comparadores

La intervención objeto de estudio consiste en un método no invasivo basado en el análisis ultrasonográfico cuantitativo de la textura pulmonar fetal junto a otras variables para predecir el riesgo de morbilidad respiratoria neonatal⁷. Como intervención se consideran dos marcas comerciales, tanto AQUA (*Automatic Quantitative Ultrasound Analysis*)⁶ que sería el prototipo inicialmente desarrollado, como quantusFMLTM, considerado una versión mejorada del anterior.

Tanto AQUA como quantusFMLTM utilizan las imágenes pulmonares fetales obtenidas a través de una ecografía en un corte transversal a nivel de las cuatro cámaras cardíacas fetales. Una vez localizado el plano, se procede a delimitar manualmente el área pulmonar a estudiar mediante un programa específico que cuenta con una herramienta gráfica de usuario. Los ajustes necesarios para obtener imágenes de buena calidad los realiza el profesional que hace las ecografías y se almacenan en formato DICOM para su análisis posterior.

A estas imágenes se les aplica entonces un software diseñado para extraer las características relevantes de textura pulmonar, y esta información se utiliza para determinar descriptores asociados a las imágenes. A través de algoritmos matemáticos se seleccionan descriptores críticos que se combinan para crear y validar un modelo de predicción de maduración pulmonar fetal. El primer modelo desarrollado fue AQUA⁶ y, posteriormente, se desarrolló quantusFMLTM⁷. Este último combina las características de textura pulmonar fetal y la edad gestacional definiendo un algoritmo de clasificación para predecir el riesgo de morbilidad respiratoria neonatal.

En el estudio de Palacio *et al.* 2012⁶ se utilizó como comparador un test de maduración pulmonar fetal que mide la ratio surfactante/albúmina en líquido amniótico (TDx-FLM II). Para mejorar la capacidad de esta prueba para predecir la morbilidad respiratoria neonatal, se utilizaron valores específicos de edad gestacional como puntos de corte que clasificaron los resultados en madurez o inmadurez²⁵. La ecografía y la amniocentesis se realizaron el mismo día en todos los sujetos de la muestra y, en el caso de los partos por cesárea, el líquido amniótico se obtuvo durante la misma, una vez realizada la histerotomía. Las imágenes ultrasonográficas se procesaron en el software AQUA y se identificaron 31 descriptores críticos con los que se desarrolló el modelo. Posteriormente se relacionaron con los resultados de las pruebas de madurez pulmonar

fetal que se realizaron en el líquido amniótico y se clasificaron como imágenes pulmonares maduras o inmaduras.

En el estudio de Bonet-Carné *et al.*⁷ 2015 no se utilizó ninguna intervención como comparador. En este caso, las imágenes ecográficas realizadas se procesaron en el módulo extractor para conseguir las características de textura y se combinaron con los datos clínicos (edad gestacional) de las embarazadas de la muestra. Finalmente, mediante un algoritmo de clasificación se estimó el riesgo de incidencia de morbilidad respiratoria neonatal⁷. Dicho riesgo se comparó con los resultados clínicos neonatales.

Para conseguir su objetivo, se reclutó a todas las embarazadas que cumplían los criterios de selección y se les realizó una ecografía pulmonar fetal, aunque sólo incluyeron en el estudio a aquellas cuyo parto o cesárea ocurrió a las 48 horas de la intervención. En todas ellas se calculó la edad gestacional en base a la longitud cráneo-caudal medida en la ecografía del primer trimestre. Las características maternas y fetales se recogieron posteriormente y se definieron las variables de resultado, en este caso, la mortalidad y morbilidad respiratoria neonatal (definida como síndrome de distrés respiratorio neonatal o taquipnea transitoria del recién nacido). El 36,1 % de las embarazadas en este estudio (52 mujeres) recibieron corticoides prenatales, cuya función es acelerar la maduración fetal, pero las autoras no especifican las características, semanas de gestación o grupo de administración. Cada paso de la intervención fue realizado por distintas personas del equipo investigador que no intervinieron en otras fases, al igual que el análisis de datos.

Los resultados de eficacia en ambos estudios^{6,7} se expresaron en términos de precisión y validez diagnóstica: sensibilidad y especificidad, añadiéndose también los valores predictivos positivo y negativo en el estudio de Bonet-Carné *et al.* 2015⁷.

El valor predictivo de la intervención y su asociación con los resultados de la prueba de referencia en el primer estudio⁶, se probaron mediante una estrategia que consistió en definir dos subgrupos de la muestra creados al azar (A y B) que contenían similar número de fetos maduros e inmaduros en cada uno, según el valor de corte de edad gestacional de la prueba de referencia (Tabla 4). No se especifica en el documento la secuencia de aleatorización que se realizó.

Tabla 4. Características de los grupos del estudio Palacio *et al.* 2012⁶

	Maduros	Inmaduros	Totales
Tamaño muestra	41	62	103
Subgrupo A	21	31	52
Subgrupo B	20	31	51

Se establecieron dos modelos: en el primero, las imágenes del subgrupo B se usaron para probar el algoritmo o intervención y se validaron con las imágenes del subgrupo A, midiéndose el número de aciertos y errores en la clasificación según los resultados de la prueba de referencia en el subgrupo A; en el segundo, las imágenes del subgrupo A se usaron para probar el algoritmo y se validaron con los resultados del subgrupo B. Se calculó la precisión de ambos modelos y la media resultante. Para la validez diagnóstica de la intervención se calculó la sensibilidad y especificidad en relación con la prueba de referencia.

En el estudio de Bonet-Carné *et al.* 2015⁷ los resultados de la intervención se compararon con los resultados clínicos en los recién nacidos, en relación con la presencia o ausencia de morbi-mortalidad respiratoria neonatal, y calcularon la sensibilidad, especificidad y valores predictivos, estableciendo así la validez y precisión de la prueba.

No se especificaron en dichos estudios^{6,7} variables de resultado con respecto a seguridad y tampoco datos sobre disminución de morbi-mortalidad o aumento de supervivencia.

Evaluación de la calidad de la evidencia

La evaluación de los dos documentos incluidos para síntesis de este informe se adjunta en el Anexo 5.

A continuación se describe con detalle la evaluación realizada en relación con los principales riesgos de sesgos (selección de pacientes, prueba de estudio, prueba de referencia y flujo/cronograma de pacientes).

1. Selección de pacientes

Con respecto a la selección de los pacientes realizada, cabe destacar que en ambos estudios se describieron los criterios de inclusión y exclusión propuestos. En el estudio de Bonet-Carné *et al.* 2015⁷ se seleccionaron los casos entre las embarazadas candidatas a parto prematuro o en riesgo en las siguientes 48 horas a la intervención y en el estudio de Palacio *et al.* 2012⁶ se seleccionaron entre las mujeres embarazadas que presentaron indicación de amniocentesis o cesárea electiva.

En el estudio de Palacio *et al.* 2012⁶ el tamaño muestral fue de 103 pacientes y la edad gestacional entre 24 y 41 semanas. En el estudio de Bonet-Carné *et al.* 2015⁷ la muestra fue de 150 pacientes y la edad gestacional osciló entre 28 y 39 semanas.

El reclutamiento realizado a las mujeres embarazadas fue consecutivo y el ámbito de aplicación fue hospitalario, no existiendo exclusiones inapropiadas. Por lo tanto, no parece existir riesgo de sesgo en la selección de los pacientes, aunque no se describe la secuencia de aleatorización en el estudio de Palacio *et al.* 2012⁶.

Aunque ambos estudios mencionaron la existencia de comorbilidades que justificaban el riesgo de parto prematuro, no se indicó si se repartían por igual en todos los grupos o si existían diferencias significativas entre algunos de ellos.

Con la información disponible, los pacientes incluidos y el ámbito de aplicación se ajustan a la pregunta de revisión en relación con la aplicabilidad de la prueba.

2. Prueba de estudio

Los dos estudios describieron adecuadamente la realización e interpretación de la prueba de estudio. En el estudio de Palacio *et al.* 2012⁶ (AQUA) se desarrolló un modelo matemático para evaluar el valor predictivo teórico frente a los resultados de maduración fetal en líquido amniótico. Los autores no especificaron los profesionales que intervinieron en la realización ni en el análisis.

El estudio de Bonet-Carné *et al.* 2015⁷ (quantusFML™) realizó la misma intervención que Palacio *et al.* 2012⁶ aunque con la intención de mejorar el software anterior. Para ello combinaron los datos obtenidos por ecografía con la edad gestacional, teniendo como objetivo, calcular el riesgo (alto o bajo) de morbilidad respiratoria neonatal en embarazadas con riesgo de parto pretérmino. En cada fase del proceso intervinieron diferentes profesionales y las imágenes digitales se sometieron a un proceso de control de calidad.

La validación en el estudio de Palacio *et al.* 2012⁶ se realizó de forma secuencial, en dos subgrupos, aunque no se describe con claridad si la interpretación del método se hizo sin conocer los resultados de la prueba de referencia y al revés. Este aspecto es relevante ya que supone un riesgo de sesgo y podría afectar a los resultados obtenidos. El mismo hecho de clasificar los casos en función de la edad gestacional en ambos estudios^{6,7} podría conllevar sesgos de interpretación si no existe cegamiento, debido a la fuerte correlación entre edad gestacional y maduración pulmonar fetal.

En los dos estudios incluidos^{6,7}, se consideraron valores de corte de edad gestacional para clasificar los casos:

1. En el de Palacio *et al.* 2012⁶, para obtener dos subgrupos aleatorios donde validar la prueba de estudio, aunque no se describe la secuencia de aleatorización.
2. En el de Bonet-Carné *et al.* 2015⁷, para interpretar los resultados.

No se han referido datos sobre los valores de dichos puntos de corte en el primer estudio⁶; en el segundo⁷ el valor de corte (34 semanas) se corresponde con el periodo medio de gestación considerado, de forma que

la proporción de casos en los dos grupos es desigual. Esta clasificación no se corresponde exactamente con los puntos de corte aceptados como prematuridad (37 semanas) por lo que puede ocurrir que no se ajusten a la pregunta de revisión, además al resultar un tamaño muestral muy reducido en el grupo de < de 34 semanas, implicaría limitaciones en su aplicabilidad y en la generalización de resultados.

3. Prueba de referencia

El estudio de Palacio *et al.* 2012⁶ comparó el método AQUA con la ratio surfactante/albúmina mediante luz polarizada (TDx-FLM II). Esta prueba se realizó en el líquido amniótico de todas las pacientes de la muestra, para distinguir entre madurez e inmadurez pulmonar fetal y se establecieron valores de corte específicos, según la edad gestacional. Esta prueba (ratio surfactante/albúmina) presenta alta sensibilidad (96 %) y moderada especificidad (60–84 %). Actualmente, la ratio surfactante/albúmina mediante luz polarizada TDx-FLM II, no está considerada como patrón de referencia y no está disponible desde finales del 2011. Por lo tanto, es posible que la prueba de referencia no hubiera clasificado correctamente los casos.

En el estudio de Bonet-Carné *et al.* 2015⁷ se compararon los resultados del método de estudio con los resultados de morbi-mortalidad neonatal para calcular el riesgo de patología respiratoria neonatal.

En relación con la prueba de referencia se observó riesgo de sesgo alto en la interpretación de la prueba en el estudio de Palacio *et al.* 2012⁶ debido a que, aunque se especifica la secuencia de realización en los dos subgrupos considerados, no identificó si hubo cegamiento, aunque la misma secuencia descrita y los puntos de corte empleados parecen señalar que los resultados de la prueba de referencia podrían conocerse antes que los de la prueba de estudio.

En el otro estudio (Bonet-Carné *et al.* 2015⁷) el riesgo de sesgo fue moderado debido al uso de corticoides prenatales en algunos pacientes que podría afectar a los resultados clínicos (posible factor de confusión). Sin embargo, cabe destacar que tanto la realización de la prueba como la interpretación de los resultados se hicieron a ciegas e independientemente de los resultados clínicos obtenidos.

4. Flujo y cronograma

En el estudio de Palacio *et al.* 2012⁶, tanto las imágenes ecográficas como la obtención del líquido amniótico para realizar la prueba de referencia se obtuvieron el mismo día, en todos los casos. En las cesáreas, se tomó la muestra de líquido amniótico durante la misma y, por tanto, posterior a la ecografía; a las mujeres embarazadas que no se les realizó cesárea, no se especificó cuando se realizó la toma de muestra. En el estudio de

quantusFML™:7 el método se aplicó a todas las pacientes, 48 horas antes de conocer los resultados clínicos neonatales.

En los dos estudios se aplicaron las mismas pruebas a todos los casos incluidos. Sin embargo, no todos se incluyeron en el análisis. En el estudio de Palacio *et al.* 2012⁶ hubo dos pérdidas en el seguimiento y en el de Bonet-Carné *et al.* 2015⁷, seis que fueron debidas a exclusiones por mala calidad de las imágenes obtenidas.

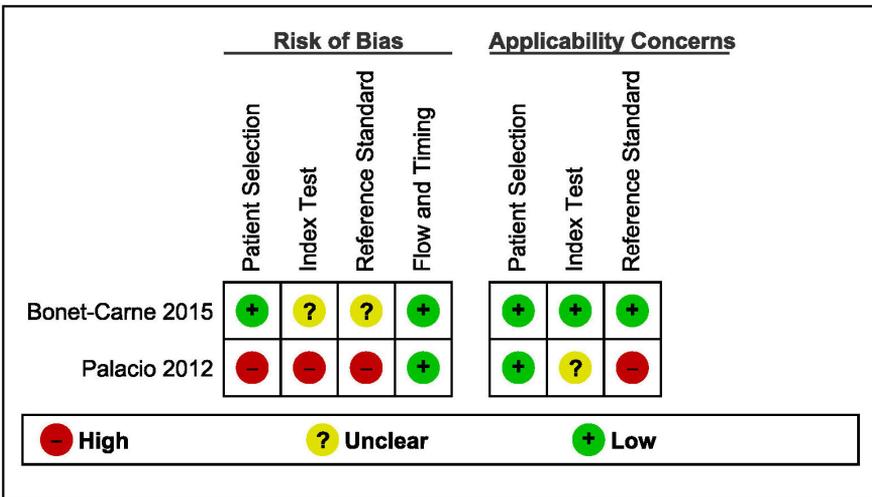
En el seguimiento realizado en el estudio de Palacio *et al.* 2012⁶, la ausencia de descripción de la secuencia de interpretación de resultados entre la prueba de referencia y la de estudio en los dos subgrupos podría conllevar riesgo de sesgo de interpretación.

A continuación se resume gráficamente la valoración del riesgo de sesgo y aplicabilidad de los dos estudios considerados conseguidas mediante la aplicación del programa *Review Manager* 5.3. (Figura 1 y 2 y Tabla 5).

Figura 1: Gráfica sobre riesgo de sesgo y aplicabilidad de los estudios considerados



Figura 2: Resumen sobre riesgo de sesgo y aplicabilidad de los estudios considerados



En relación con el nivel de evidencia asignado, considerando la escala propuesta por NICE en 2006, se obtuvieron los siguientes resultados.

Autor y año	Nivel de evidencia
Palacio <i>et al.</i> ⁶ 2012	III (riesgo de sesgos alto)
Bonet-Carne <i>et al.</i> ⁷ 2015	II (riesgo de sesgos moderado)

2. Resultados clínicos

Eficacia

Los estudios incluidos evaluaron la eficacia de la prueba en términos de sensibilidad, especificidad, valores predictivos y precisión.

Palacio *et al.* 2012⁶ evaluaron la eficacia de la prueba calculando la sensibilidad y especificidad. No presentaron datos sobre valores predictivos ni ofrecieron información suficiente que permitiera su estimación. Los intervalos de confianza estimados para cada medida no fueron facilitados, por lo que no se pudo considerar esta información para conocer ni la significación ni la precisión de los datos. Durante el desarrollo del artículo detallan el cálculo de la precisión, al compararla con la prueba de referencia, siendo en las dos muestras 92,31 % y 88,24 % y considerando que la media de ambas sería el valor de la prueba.

En el estudio de Bonet-Carné *et al.* 2015⁷ calcularon la sensibilidad, especificidad y valores predictivos comparando los resultados de la intervención con los resultados clínicos en los recién nacidos.

A continuación se detallan los resultados obtenidos en ambos estudios.

	Palacio <i>et al.</i> 2012	Bonet-Carné <i>et al.</i> 2015		
	Total (n = 103)	Total (n = 144)	Semanas gestación	
			28 a 33 (n = 38)	34 a 39 (n = 106)
Precisión	90,27 %	86,8 %	92,1 %	84,9 %
Sensibilidad	95,1 %	86,2 %	90,5 %	75,0 %
Especificidad	85,7 %	87,0 %	94,1 %	85,7 %
VPP	-	62,5 %	95,0 %	30,0 %
VPN	-	96,2 %	88,9 %	97,7 %

n: tamaño de la muestra; VPP: valor predictivo positivo; VPN: valor predictivo negativo.

Efectividad

Los estudios incluidos mostraron como resultados de efectividad los siguientes:

En el caso de Palacio *et al.* 2012⁶ los resultados aportados mostraron que el 54,5 % de la muestra ingresó en la Unidad de Cuidados Intensivos (UCI); el 16,8 %, se diagnosticaron de Síndrome Respiratorio Neonatal y el 5 % de ellos presentaron el resultado de muerte.

En cuanto al estudio de Bonet-Carné *et al.* 2015⁷, fueron ingresados en UCI el 36,8 % de la muestra; diagnosticados de morbilidad respiratoria, el 20,1 %; y no ocurrió ninguna muerte neonatal.

En términos de disminución de la mortalidad o supervivencia, en ninguno de los dos estudios se mostraron resultados.

Seguridad

Con respecto a la seguridad, no se declararon efectos adversos directos ni indirectos en relación con la intervención en ninguno de los dos estudios incluidos^{6,7}.

En cualquier caso, la evaluación de la maduración pulmonar mediante métodos no invasivos, ya sea mediante AQUA o quantusFLMTM, está basada en la realización de una ecografía durante el periodo final del embarazo y su seguridad sería similar a la de cualquier prueba de imagen realizada con ecógrafos, estando claramente demostrada en otros estudios²⁶. Su interés radica precisamente en que presentaría menos efectos adversos que la alternativa diagnóstica, que requiere la realización de una amniocentesis¹³.

3. Estudios en marcha localizados

No se han localizado estudios en marcha en las búsquedas bibliográficas realizadas para elaborar este informe.

Aspectos económicos

La búsqueda bibliográfica no proporcionó ningún estudio económico sobre la prueba evaluada, por lo que no se dispone de evidencia sobre aspectos económicos de la misma.

Discusión

La necesidad de valorar la madurez pulmonar fetal en embarazadas con sospecha de sufrimiento fetal en el tercer trimestre obliga actualmente a la realización de técnicas invasivas, como la amniocentesis, para evaluar la cantidad de surfactante pulmonar fetal en líquido amniótico. El objetivo de desarrollar una técnica no invasiva que facilite dicha valoración sin los efectos secundarios y complicaciones maternas y fetales de una prueba invasiva, sería de gran importancia a la hora de tomar decisiones en la práctica clínica diaria, pero también de cara a la seguridad y aceptabilidad por parte del paciente.

La población definida en nuestra pregunta de investigación (mujeres embarazadas en el tercer trimestre con riesgo de parto pretérmino) se corresponde con la población seleccionada en ambos estudios^{6,7}. Con respecto a las pacientes incluidas, además de las señaladas, se incluyeron mujeres embarazadas con indicación de cesárea o parto electivo, no estando disponible información sobre la morbilidad correspondiente a cada indicación y su relación con la toma de decisiones. Todas las mujeres embarazadas incluidas se sometieron a parto o cesárea a las 48 horas de la aplicación de los métodos objeto de estudio.

A todas se les calculó la edad gestacional mediante la longitud cráneo-caudal fetal medida por ecografía en el primer trimestre. En este sentido, nuestra pregunta de investigación se refería a los recién nacidos pretérminos (aquellos que nacen antes de las 37 semanas¹) y, aunque ambos estudios los incluyen, en la muestra incluyen también fetos de mayor edad gestacional (hasta 41 semanas en el estudio de Palacio *et al.* 2012⁶, y 39 semanas en el de Bonet-Carné *et al.* 2015⁷) que los autores justifican porque existen estudios^{8,9} que ponen de manifiesto que existirían fetos inmaduros hasta las 39 semanas de gestación.

Además, no se aportan datos precisos sobre la utilización de corticoides para acelerar la madurez pulmonar fetal en 52 de los casos incluidos en el estudio de Bonet-Carné *et al.* 2015⁷ y tampoco cómo afectó esta circunstancia a los resultados de morbi-mortalidad respiratoria neonatal.

Respecto al tamaño muestral, en el estudio de Palacio *et al.* 2012⁶ no se establece previamente, siendo esta una de las limitaciones para la validación que las propias autoras reconocen. La creación de grupos de diferente edad gestacional en ambos estudios disminuye la posibilidad de un diseño de caso-control, pero los tamaños de cada grupo se reducen aún más por lo que la extrapolación de los resultados en cada uno debe hacerse con precaución.

Ambos estudios aceptan un punto de corte de edad gestacional a las 34 semanas, de forma que en el estudio de Palacio *et al.* 2012⁶ existen 66 casos con edad gestacional menor de 34 semanas y en el de Bonet-Carné *et al.* 2015⁷, 38 casos. En este intervalo de edad gestacional considerado, la inmadurez pulmonar fetal es lo más frecuente y el método estudiado presentaría menos interés para su aplicación en la práctica clínica, ya que en caso de indicación de parto pretérmino o cesárea habría que adoptar medidas para acelerar dicha maduración. Si tenemos en cuenta el grupo de edad gestacional mayor o igual a 34 semanas, encontramos 37 casos en el estudio de Palacio *et al.* 2012⁶ y 106 en el de Bonet-Carné *et al.* 2015⁷, de los cuales los recién nacidos de 37 a 39 semanas (86 casos, y según la definición comúnmente aceptada, considerados a término y, presumiblemente, maduros) lo fueron por parto electivo. En este intervalo de edad es más importante poner de manifiesto la capacidad discriminatoria del método para predecir la maduración pulmonar fetal en aras de poder adoptar medidas que aceleren dicha maduración en caso necesario, por lo que el método tendría aquí más interés. En este sentido, cabe destacar que la sensibilidad y la especificidad del método quantusFML^{TM:7} disminuyen en este grupo con respecto al de menos de 34 semanas. Además, este grupo tiene un tamaño muestral mayor por lo que los resultados podrían ser más robustos, sin embargo, no existen datos sobre los intervalos de confianza establecidos y el error aleatorio aceptado, por lo que dificultaría conocer la precisión y la significación real de los resultados obtenidos.

En relación con el periodo de estudio, en los dos estudios^{6,7} se estableció un periodo de 48 horas entre la realización de los métodos y los resultados clínicos y la prueba de referencia se aplicó a todas las pacientes incluidas, el mismo día que la de estudio en el caso de AQUA⁶. No se especifica en el estudio⁶ si los resultados de la prueba de referencia se conocieron antes que la de estudio o al revés, pero la secuencia de validación de la prueba y la ausencia de cegamiento hace pensar que pudo haber interpretación de los resultados tras conocer los resultados de la otra prueba. En el caso de quantusFML^{TM:7}, el estudio especifica que hubo cegamiento en cada fase del proceso por lo que la aparición de sesgos aquí es más improbable.

En cuanto a la intervención, señalar que los métodos de determinación de madurez pulmonar fetal estudiados son similares en ambos artículos. Tanto AQUA⁶ como quantusFML^{TM:7}, se basan en la aplicación de una serie de algoritmos a las imágenes obtenidas por ecografía de los pulmones fetales para estimar su maduración, aunque los propios investigadores reconocen que el método AQUA no ha conseguido demostrar los efectos esperados⁷.

Ambos utilizan la edad gestacional en su secuencia de aplicación, siendo dicha edad el factor predictor más importante de maduración pulmonar fetal⁶. La asociación de dicho factor con el resultado buscado podría actuar como factor de confusión, por lo que podría sobreestimarse el rendimiento del método.

En relación con la prueba de referencia utilizada, habría que señalar que en el estudio de Palacio *et al.* 2012⁶ se incluye como comparador del método evaluado la prueba TDx-FLM II¹³. Esta prueba está actualmente en desuso y ha sido retirada del mercado por el mismo laboratorio que lo comercializó, por lo que existen dudas en cuanto a la clasificación de las pacientes a las que se les aplicó durante el estudio y que sirvió de base para el ensayo y validación del método AQUA. Además, los puntos de corte utilizados durante el estudio para distribuir los casos en los subgrupos considerados, se basaron en un estudio previo²⁵ realizado con esta prueba de referencia por el mismo equipo investigador que ha desarrollado los dos estudios incluidos para la síntesis de este informe, sin que se detallen o especifique mediante un análisis detallado. Estas circunstancias hacen que los resultados obtenidos deban ser tomados con cautela por la dificultad de extrapolarlos o generalizarlos.

Además, en ninguno de los dos artículos^{6,7} se especifica cómo se trataron las pérdidas. Los resultados se expresaron sobre los casos que finalizaron el seguimiento. Aunque las pérdidas fueron escasas en ambos, no es posible conocer cómo influirían en los resultados, sobre todo en el estudio de Bonet-Carné *et al.* 2015⁷ donde el motivo de exclusión fue la mala calidad de las imágenes. Teniendo en cuenta que estas se tomaron en condiciones ideales por personal entrenado, puede ser una información pertinente para la aplicación en la práctica clínica, donde las imágenes no se toman en condiciones ideales ni por equipos entrenados.

En cuanto a las limitaciones cabe destacar que los resultados de una revisión sistemática se basan en los estudios recuperados tras la realización de una búsqueda bibliográfica. La inclusión solamente de estudios publicados en las principales bases de datos referenciales pudo favorecer el sesgo de publicación (no se incluyó la literatura gris).

Por último señalar que las autoras de los estudios incluidos en este informe incidieron en la importancia de la intervención evaluada para valorar la madurez pulmonar aunque no aportan información sobre el efecto beneficioso de la misma o la disminución de morbi-mortalidad esperada.

El propio equipo de investigación aconseja realizar más estudios de carácter multicéntrico con amplias muestras, cuyos resultados puedan esclarecer las indicaciones de este tipo de métodos de valoración de la maduración pulmonar fetal sobre todo en situaciones clínicas límites donde el parto podría posponerse.

Referencias

1. Kinney, M.V., Howson, C.P, McDougall, L., & Lawn, J.E. Resumen ejecutivo de *Nacidos Demasiado Pronto: Informe de Acción Global sobre Nacimientos Prematuros*. March of Dimes, PMNCH, Save the Children, Organización Mundial de la Salud. 2012
2. Sharp GC, Stock SJ, Norman JE. Fetal assessment methods for improving neonatal and maternal outcomes in preterm prelabour rupture of membranes. *Cochrane Database Syst Rev*. 2014(10):CD010209.
3. Laban M, Mansour GM, Elsafty MS, Hassanin AS, EzzElarab SS. Prediction of neonatal respiratory distress syndrome in term pregnancies by assessment of fetal lung volume and pulmonary artery resistance index. *Int J Gynaecol Obstet*. 2015 Mar;128(3):246-50.
4. Beck AP, Araujo Junior E, Leslie AT, Camano L, Moron AF. Assessment of fetal lung maturity by ultrasound: objective study using gray-scale histogram. *J Matern Fetal Neonatal Med*. 2015 Apr;28(6):617-22.
5. Done E, Debeer A, Gucciardo L, Van Mieghem T, Lewi P, Devlieger R, et al. Prediction of neonatal respiratory function and pulmonary hypertension in fetuses with isolated congenital diaphragmatic hernia in the fetal endoscopic tracheal occlusion era: a single-center study. *Fetal Diagn Ther*. 2015;37(1):24-32.
6. Palacio M, Cobo T, Martínez-Terrón M, Ratta GA, Bonet-Carne E, Amat-Roldan I, et al. Performance of an automatic quantitative ultrasound analysis of the fetal lung to predict fetal lung maturity. *Am J Obstet Gynecol*. 2012;207(6):504 e1-5.
7. Bonet-Carne E, Palacio M, Cobo T, Perez-Moreno A, Lopez M, Piraquive JP, et al. Quantitative ultrasound texture analysis of fetal lungs to predict neonatal respiratory morbidity. *Ultrasound Obstet Gynecol*. 2015;45(4):427-33.
8. Cobo T, Bonet-Carne E, Martínez-Terrón M, Perez-Moreno A, Elías N, Luque J, et al. Feasibility and reproducibility of fetal lung texture analysis by automatic quantitative ultrasound analysis and correlation with gestational age. *Fetal Diagn Ther*. 2012;31(4):230-6.

9. Albinagorta Olórtegu R, Miranda Tapia M. Actualización en la evaluación de la madurez pulmonar fetal por ultrasonido. *Rev Peru Ginecol Obstet.* 2015;61:433-8.
10. Brownfoot FC, Gagliardi DI, Bain E, Middleton P, Crowther CA. Different corticosteroids and regimens for accelerating fetal lung maturation for women at risk of preterm birth. *Cochrane Database Syst Rev.* 2013(8):CD006764.
11. Ortega Berumen, SL. Estandarización de pruebas de maduración pulmonar fetal en el hospital de especialidades del niño y la mujer [tesis doctoral]. México: Facultad de Medicina, Universidad Autónoma de Querétaro; 2008.
12. Besnard AE, Wirjosoekarto SAM, Broeze KA, Opmeer BC, Mol BWJ. Lecithin/sphingomyelin ratio and lamellar body count for fetal lung maturity: a meta-analysis. *Eur J Obstet Gynecol Reprod Biol.* 2013;169(2):177-83.
13. Fantz CR, Powell C, Karon B, Parvin CA, Hankins K, Dayal M, et al. Assessment of the diagnostic accuracy of the TDx-FLM II to predict fetal lung maturity. *Clin Chem.* 2002 May;48(5):761-5.
14. López Pérez ZG. Correlación diagnóstica de maduración pulmonar fetal con test de Clements [tesis doctoral]. Guatemala: Escuela de estudios de Postgrado, Universidad de San Carlos de Guatemala; 2014.
15. Alfirevic Z, Stampalija T, Gyte GM. Fetal and umbilical Doppler ultrasound in high-risk pregnancies. *Cochrane Database Syst Rev.* 2013(11):CD007529.
16. Jiménez Castillejo K, Ruiz Lopez Y, Reyna Villasmil E, Torres Cepeda D, Santos Bolívar J, Aragón Charry J, et al. Doppler de las arterias pulmonares fetales posterior al uso de betametasona para la maduración pulmonar. *Av en Biomed.* 2014;3(3):122-8.
17. García Noguera M. Patrón de los movimientos respiratorios en el feto sano con crecimiento normal. *Salus.* 2012;16(2):23-31.
18. Britto IS, Tedesco GD, Herbst SR, Bussamra LC, de Andrade FM, Araujo Junior E, et al. New anatomical landmarks to study the relationship between fetal lung area and thoracic circumference by three-dimensional ultrasonography. *J Matern Fetal Neonatal Med.* 2012;25(10):1927-32.
19. Urrutia G, Bonfill X. PRISMA: una propuesta para mejorar la publicación de revisiones sistemáticas y metaanálisis. *Med Clin (Barc).* 2010;135(11):507-11.

20. Puñal-Riobóo J, Baños Álvarez E, Varela Lema L, Castillo Muñoz MA, Atienza Merino G, Ubago Pérez R, Triñanes Pego Y, Molina López T y López García M en representación del Grupo de trabajo de la Guía para la elaboración y adaptación de informes rápidos de evaluación de tecnologías sanitarias. Red Española de Agencias de Evaluación de Tecnologías Sanitarias y Prestaciones del SNS. Agencia Gallega para la Gestión del Conocimiento en Salud. Unidad de Asesoramiento Científico-técnico, avalia-t; 2016.
21. Latour J. El diagnóstico. Quaderns de salut pública i administració de serveis de salut, 21. Valencia: Escola Valenciana d'Estudis per a la Salut; 2003.
22. Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol.* 2003;3:25.
23. González Rodríguez MP, Velarde Mayol C. Listas guía de comprobación de estudios sobre pruebas diagnósticas incluidos en las revisiones sistemáticas: declaración QUADAS. *Evid Pediatr.* 2012;8(1):20.
24. National Institute for Health and Care Excellence (NICE). Developing NICE guidelines: the manual [Internet]. London: NICE; 2009. [citado 20 mayo 2016]]URL: <https://www.nice.org.uk/process/pmg20/chapter/updating-guidelines#scheduling-updates>
25. Bennasar M, Figueras F, Palacio M, Bellart J, Casals E, Figueras J, et al. Gestational Age-Specific Cutoff Levels of TDx-FLM II for the Prediction of Neonatal Respiratory Distress Syndrome. *Fetal Diagn Ther.* 2009;25(4):392–6.
26. CADTH. Non-urgent prenatal ultrasound: safety and guidelines. Ottawa: Canadian Agency for Drugs and Technologies in Health (CADTH). Rapid Response: Summary of Abstracts. 2013.

Anexos

Anexo 1. Estrategias de búsqueda

MEDLINE

1. Pregnancy.ti,ab.
2. Pregnancy/
3. 1 OR 2
4. (ultrasound AND texture).ti,ab.
5. (fetal OR neonatal OR maturity).ti,ab.
6. 4 AND 5
7. Lung/us [Ultrasonography]
8. Fetal development/ OR fetal organ maturity/
9. 7 AND 8
10. 6 OR 9
11. 3 AND 10
12. limit 11 to yr="2012 - 2016"

EMBASE

1. pregnancy:ab,ti
2. `pregnancy'/exp
3. ultrasound: ab,ti AND quantitative:ab,ti AND texture: ab,ti
4. fetal:ab,ti OR neonatal:ab,ti OR maturity:ab,ti
5. #3 AND #4
6. `ultrasonography':ab,ti OR `ecography':ab,ti
7. `ecography'/de
8. #6 OR #7
9. `fetus lung'/mj
10. #8 AND #9
11. #1 OR #2
12. #10 AND #11
13. #5 OR #12
14. #13 AND [2012— 2016]/py
15. quantus*
16. #14 OR #15

WOS

1. TS= (pregnancy) AND TS=(fetal lung) AND TS=(ultrasonography)
2. TS= (ultrasound AND quantitative texture)
3. TS= (quantus)
4. TS= (pregnancy AND fetal lung AND ultrasonography) 2012-2016
5. TS= (ultrasound AND quantitative texture) 2012-2016
6. # 5 AND # 6

En búsqueda inicial exploratoria, se consultó la base de datos de tesis doctorales *Open Access Theses and Dissertations* utilizando los siguientes términos, (ultrasound) AND (fetal OR lung) AND (quantitative), así como la base de datos GUPEA (<http://gupea.ub.gu.se/handle/2077>).

Anexo 2: Escala de nivel de evidencia

Niveles de evidencia para estudios de exactitud diagnóstica (NICE 2006) ²⁴		
Niveles de evidencia	Tipo de evidencia	Probabilidad de sesgo
Ia	Revisión sistemática con homogeneidad (a) de estudios nivel 1 (b)	Baja
Ib	Estudios nivel 1 (b)	Baja
II	Estudios nivel 2 (c) Revisión sistemática de estudios nivel 2	Moderada
III	Estudios nivel 3 (d) Revisión sistemática de estudios nivel 3	Alta
IV	Consenso de expertos o recomendaciones basadas en fisiología y estudios básicos	Alta

(a) Homogeneidad significa que no hay o solo son menores variaciones en la dirección y grado de los resultados entre los estudios incluidos.

(b) Estudios nivel 1:

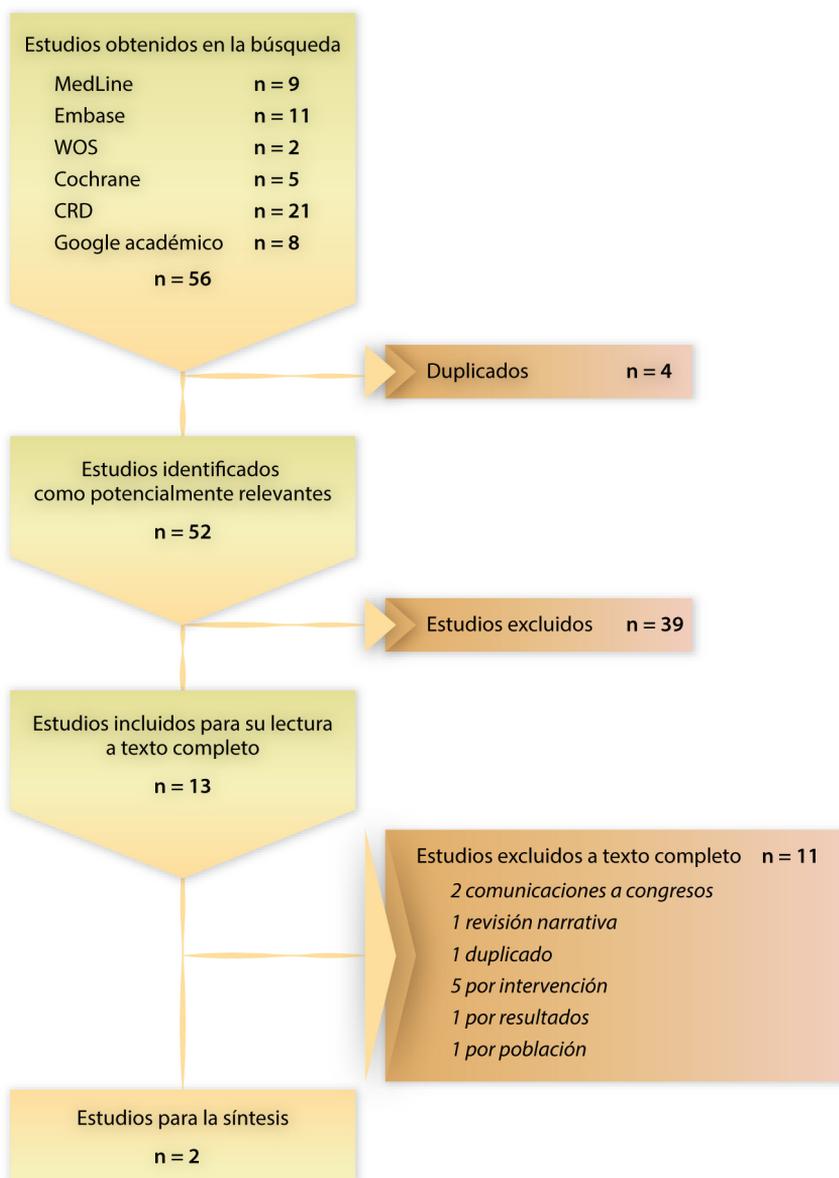
- Usan comparación enmascarada de la prueba con un patrón de referencia validado.
- En una muestra de pacientes que reflejan la población a la que se aplicará el test.

(c) Estudios nivel 2 son aquellos que tiene solo uno de los siguientes:

- Población restringida (la muestra no refleja la población a la que se aplicará el test).
- Utilizan un patrón de referencia de baja calidad (definido como aquel en que el test está incluido en la referencia o donde el test afecta a la referencia).
- La comparación entre el test y la referencia no es enmascarada.
- Estudios de casos y controles.

(d) Estudios nivel 3 son aquellos que tienen dos o más de las características especificadas para los estudios de nivel 2.

Anexo 3. Diagrama de flujo



Anexo 4: Artículos excluidos y motivo de exclusión

Artículos excluidos y motivos de exclusión	
Autor y año	Motivo de exclusión
Beck <i>et al.</i> 2015	Intervención no adecuada
Bonet- Carne <i>et al.</i> 2014	Tipo de estudio. Comunicación a congreso
Bonet-Carne <i>et al.</i> 2014	Duplicado (incorpora la información de los artículos incluidos en el informe)
Cobo <i>et al.</i> 2012	Resultados no adecuados (validez analítica)
García <i>et al.</i> 2012	Intervención no adecuada
Laban <i>et al.</i> 2015	Intervención no adecuada
Albinagorta Olórtegui <i>et al.</i> 2015	Tipo de estudio. Revisión no sistemática
Ortega <i>et al.</i> 2008	Intervención no adecuada
Palacio <i>et al.</i> 2012	Tipo de estudio. Comunicación a congreso
Pérez <i>et al.</i> 2014	Intervención no adecuada
Sharp <i>et al.</i> 2014	Intervención no adecuada

Anexo 5: Evaluación según la herramienta QUADAS-1 y QUADAS-2

Evaluación de calidad mediante QUADAS 1		
	Palacio <i>et al.</i> 2012	Bonet-Carné <i>et al.</i> 2015
¿Es representativa la muestra?	Dudoso	Sí
¿Se describen bien los criterios de selección?	Sí	Sí
¿Clasifica bien el patrón de referencia?	Dudoso	Sí, utilizan los resultados clínicos como comparador.
Si el patrón y la prueba no se aplican simultáneamente, ¿es razonable pensar que la enfermedad no haya evolucionado en ese tiempo?	Sí	Sí
¿Sesgo de verificación parcial?	No	No
¿Sesgo de verificación diferencial?	No	No
¿Es el patrón de referencia independiente de la prueba? (p. ej., la prueba no forma parte del patrón)	Sí	Sí
Detalle de la ejecución de la prueba	Sí	Sí
Detalle de la ejecución del estándar	Sí	No disponible
¿Se interpretó la prueba sin conocer el resultado del patrón?	Dudoso	Sí
¿Se interpretó el patrón sin conocer el resultado de la prueba?	Dudoso	Dudoso
¿La información clínica disponible en la interpretación de la prueba es la misma que habrá cuando se use en la práctica?	Dudoso	Dudoso
¿Se informa de los resultados no interpretables?	No	No
¿Se explican las pérdidas?	No	Sí

Evaluación del riesgo de sesgo (QUADAS-2)			
		Palacio, M. 2012	Bonet-Carné, E. 2015
Selección de pacientes			
Preguntas orientativas	¿Se incluyó muestra de pacientes consecutiva o aleatoria?	Sí	Sí
	¿Se evitó un diseño de casos y controles?	Sí	Sí
	¿Se evitó en el estudio que hubiese exclusiones inapropiadas?	Sí	Sí
Riesgo de sesgo	¿Hay sesgos en la selección de pacientes?	Sí. No se describe secuencia de aleatorización	No
Aplicabilidad	¿Hay dudas de que los pacientes incluidos y el lugar no se ajusten a la pregunta de la revisión?	No	No
Prueba de estudio			
Preguntas orientativas	¿Se realizó la interpretación de la prueba en estudio sin conocer resultados de la prueba de referencia?	No	Sí
	Si se utilizó un punto de corte ¿este se especificó previamente?	Sí	Sí
Riesgo de sesgo	¿Podría haber sesgos en la realización e interpretación de la prueba?	Sí	Sí. al combinar el algoritmo de clasificación con la edad gestacional
Aplicabilidad	¿Hay dudas de que los pacientes incluidos y el lugar no se ajusten a la pregunta de la revisión?	No	No
Prueba de referencia			
Preguntas orientativas	¿El estándar de referencia clasifica correctamente la enfermedad estudiada?	El comparador no está considerado como gold-estándar	Sí. Se utiliza como comparador el resultado clínico
	¿Se procedió a interpretar los resultados de estándar de referencia sin conocer los resultados de la prueba en estudio?	Sí, aunque se aporta escasa información	No disponible
Riesgo de sesgo	¿Podría haber sesgos en la realización e interpretación de la prueba?	Sí, al realizar consecutivamente ensayo y validación de la prueba sobre los mismos grupos	Sí. No se facilita información sobre la utilización de corticoides prenatales
Aplicabilidad	¿Existen dudas acerca de que la enfermedad en estudio definida por la prueba de referencia difiera de la pregunta de revisión?	Sí	No
Flujo y cronograma			
Preguntas orientativas	¿El intervalo entre la prueba de estudio y la prueba de referencia fue adecuado?	Sí, la prueba de estudio y el comparador se realizaron en el mismo día, sin identificar si antes o después del mismo	Se utiliza como comparador el resultado clínico al nacimiento
	¿A todos los pacientes se les realizó la misma prueba de referencia?	Sí	Se utiliza como comparador el resultado clínico al nacimiento
	¿Se incluyeron todos los pacientes en el análisis?	No (hubo 2 pérdidas en el seguimiento)	No (se excluyeron 6 imágenes de mala calidad)
Riesgo de sesgo	¿El flujo de seguimiento del paciente podría haber producido algún sesgo?	No	No

