

**Estimación del peso fetal ecográfico: Valoración
de diferentes métodos de cálculo en mujeres
sanas del área mediterránea en el
tercer trimestre de embarazo.
Una revisión sistemática.**

Trabajo de Fin de Grado

Fecha de última actualización: 28 de Marzo de 2022

Grado en Medicina

Servicio de Ginecología y Obstetricia del Hospital Universitario de La Plana

Realizado por: **Claudia Arambul Sorribes**

Dirección: **Eva María Moya Artuñedo**

TRABAJO DE FIN DE GRADO (TFG) - MEDICINA

EL/LA PROFESOR/A TUTOR/A hace constar su **VISTO BUENO** para la Defensa Pública del Trabajo de Fin de Grado y **CERTIFICA** que el/la estudiante lo ha desarrollado a lo largo de 6 créditos ECTS (150 horas)

TÍTULO del TFG: Estimación del peso fetal ecográfico: Valoración de diferentes métodos de cálculo en mujeres sanas del área Mediterránea en el tercer trimestre de embarazo. Una revisión sistemática.

ALUMNO/A: Claudia Arambul Sorribes

DNI: 53728882P

PROFESOR/A TUTOR/A: Eva María Moya Artuñedo

Fdo (Tutor/a):

RESUMEN

Antecedentes: La estimación del peso fetal es importante para el seguimiento del embarazo por estar relacionado con el desarrollo fetal. No existe un método directo para medirlo y múltiples sistemas ecográficos se desarrollaron para estimarlo; pero el peso puede variar dependiendo del método elegido. **Objetivos:** valorar el método ecográfico más preciso en la estimación del peso fetal en fetos grandes y adecuados para su edad gestacional en mujeres sanas embarazadas del área Mediterránea durante el tercer trimestre de embarazo. **Estrategia de búsqueda:** se revisaron UpToDate, guías de práctica clínica e instituciones de referencia, y se realizó la búsqueda bibliográfica en Pubmed, Web of Science y Scopus. **Criterios de selección:** estudios de tests diagnósticos observacionales publicados entre 2016 y 2022 que evaluaron el peso fetal mediante ecografía en mujeres sanas del área Mediterránea. **Extracción y análisis de datos:** se extrajeron los datos. Se evaluó la calidad metodológica de los estudios mediante QUDADAS-2 y se realizó una descripción y síntesis de los resultados. **Resultados:** Se incluyeron 9 artículos (64546 participantes); seis artículos valoraron el método Hadlock, dos compararon Intergrowth-21 con otros métodos y, uno evaluó el peso fetal estimado con la medición ecográfica seriada. **Limitaciones:** Solo un revisor realizó la búsqueda y el análisis de los artículos y cinco artículos tuvieron riesgo de sesgo en algún ítem de los evaluados. **Conclusiones:** la fórmula Hadlock junto con características materno-fetales es el método más preciso para estimar el peso fetal en mujeres del área mediterránea.

Palabras clave: “peso fetal” “peso al nacimiento”, “macrosomía”, “ultrasonografía” y “tercer trimestre de embarazo.

ABSTRACT

Background: The estimation of fetal weight is important to monitor pregnancies because it is related to fetal development. There is no direct method to measure it and multiple ultrasound systems have been developed to estimate it; although the weight may vary depending on the chosen method. **Objectives:** To evaluate the most accurate ultrasound method to estimate fetal weight in large for and appropriate for their gestational age fetuses, in healthy pregnant women in the Mediterranean area during the third trimester of their pregnancies. **Search strategy:** UpToDate, clinical practice guidelines and reference institutions were reviewed, and the bibliographic search was carried out through Pubmed, Web of Science and Scopus. **Selection criteria:** Observational diagnostic test studies published between 2016 and 2022 that evaluated fetal weight by ultrasound in healthy women in the Mediterranean area. **Data extraction and analysis:** We extracted the data. The methodological quality of the studies was assessed using QUDADAS-2, and a description and synthesis of the results were made. **Results:** 9 articles (64,546 participants) were included; six articles evaluated the Hadlock formula, two compared the Intergrowth-21 method with other formulas, and one article evaluated single fetal weight estimation with serial ultrasound measurements. **Limitations:** Only one reviewer performed the search and analysis of the articles. Five articles had a risk of bias in some of the evaluated items. **Conclusions:** The Hadlock formula together with maternal-fetal characteristics is the most accurate method to estimate fetal weight in healthy women in the Mediterranean area.

Keywords: “fetal weight”, “birth weight”, “macrosomia”, “ultrasonography” and “third trimester of pregnancy”.

EXTENDED SUMMARY

Background: Fetal weight is directly related to fetal development and well-being; it could vary with maternal age and height, ethnicity and fetal sex. An inappropriate fetal weight increases the risk of complications; therefore, its estimation is useful to guide decisions about antenatal follow-up, the moment of delivery, and therapeutic interventions that may follow. Presently, there is no direct method available to quantify it; multiple ultrasound formulas have been developed to estimate it, but the weight can vary depending on the method chosen. Currently, the most widely used formula is the Hadlock formula (based on biparietal diameter, head circumference, abdominal circumference and femur length), which predicts birth weight within 10% in 80% of pregnancies. There are reviews on the use of ultrasound in the third trimester but, they are not focused on the method that best estimates fetal weight during the third trimester in women in the Mediterranean area.

Objectives: to assess the most accurate ultrasound method to estimate fetal weight in large for and appropriate for gestational age in healthy pregnant women in the Mediterranean area during the third trimester of pregnancy.

Methods: The methodology of this systematic review follows the recommendations of the PRISMA 2020 guideline. The bibliographic search began in November 2021 and ended on January 11, 2022. Synthesis publications (UpToDate), clinical practice guidelines and reference institutions of Obstetrics and Gynecology were consulted. The search included studies from PubMed, Web of Science and Scopus. The article selection criteria were: observational diagnostic tests studies published between 2016 and 2022, studies that included healthy pregnant women in the Mediterranean area in the third trimester of gestation with large for or appropriate for gestational age single fetus, and articles whose main method of evaluating fetal weight was the ultrasound. Two screenings were carried out in order to select the articles that met the inclusion criteria and finally, a total of 9 articles were selected, 2 of which came from the manual review of the bibliographic references. The results were analyzed, described and synthesized. The risk of bias of the selected articles was evaluated using the QUADAS-2 tool.

Results: Four prospective cohort studies, one retrospective cohort study, two retrospective case-controls, and two cross-sectional studies were included; with a total of 64546 participants. Cohort studies focused on comparing EFW (estimated fetal weight) using the Hadlock method with other methods, except for one study that compared EFW with fetal growth; case-control studies evaluated parameters that influenced the accuracy of the EFW; and the cross-sectional studies were based on comparing the EFW obtained by the Intergrowth-21 method with other methods. The estimation of the cross-sectional fetal weight was found to be superior to the serial ultrasound measurement. The Hadlock method estimated fetal weight more accurately than the Sciosiosa, the modified Sciosiosa and the 3D ultrasound formulas; the results being similar with respect to the WHO charts. The Intergrowth-21 method overestimated large for gestational age fetuses and underestimated fetuses weighing below the 10th percentile. The addition of maternal-fetal characteristics to the Hadlock method was found to be superior in estimating fetal weight compared to the EFW obtained by the Hadlock formula.

Discussion: Our study shows the accuracy of the Hadlock formula in estimating fetal weight and the great heterogeneity of methods that exist to estimate it. Limitations: three articles turned out to have a high risk of bias in the index and reference test, for interpreting the results knowing those of the opposite test; two studies presented a high risk of bias in the selection of individuals because they were case-control studies; and only one study scored high risk of bias in flow and time for not clearly describing the reference test. On the other hand, only one researcher carried out the search and screening of the articles and no studies were found that evaluated all the available methods for estimating fetal weight in the same population.

Conclusions: 1) The best ultrasound method to estimate fetal weight, both in fetuses with large for and appropriate for gestational age in healthy women in the Mediterranean area, is the Hadlock method. Adding parameters such as maternal and fetal characteristics (maternal height or fetal sex) to this method increase the accuracy of the EFW. 2) EFW by transverse ultrasound measurement is more cost-effective than serial estimation. 3) The addition of 3D ultrasound parameters to the Hadlock formula and maternal-fetal characteristics may be promising to estimate fetal weight more reliably, but more research is required.

ABREVIATURAS

- AEG: Adecuado para la edad gestacional y sexo
- GEG: Grande para la edad gestacional y sexo
- PEG: Pequeño para la edad gestacional y sexo
- PFE: Peso fetal estimado
- PFEp: Peso fetal estimado preciso
- PFEnp: Peso fetal estimado no preciso
- PFEh: Peso fetal estimado mediante el método Hadlock
- PFA: Peso fetal actual
- IG: Proyecto Intergrowth-21
- OMS: Organización Mundial de la Salud
- CA: Circunferencia abdominal
- CC: Circunferencia cefálica
- DBP: Diámetro biparietal
- LF: Longitud del fémur
- IMC: Índice de masa corporal
- FSVU: Flujo sanguíneo de la vena umbilical
- ABC: Área bajo la curva
- GAP: Proyección ajustada a la gestación

ÍNDICE

1. INTRODUCCIÓN	8
2. OBJETIVOS	10
3. PREGUNTAS CLÍNICAS	10
4. MÉTODOS	11
4.1. FUENTES DE INFORMACIÓN	11
4.1.1. RECURSOS DE SÍNTESIS	11
4.1.2. GUÍAS DE PRÁCTICA CLÍNICA	11
4.1.3. INSTITUCIONES DE REFERENCIA	12
4.1.4. BASES DE DATOS	12
4.1.5. BÚSQUEDA MANUAL	13
4.1.6. LITERATURA NO PUBLICADA	13
4.2. CRITERIOS DE SELECCIÓN DE ESTUDIOS	14
4.3. EXTRACCIÓN Y ANÁLISIS DE LOS DATOS	14
5. RESULTADOS	15
5.1. DESCRIPCIÓN DE LOS ESTUDIOS.....	15
5.1.1. RESULTADOS DE LA BÚSQUEDA	15
5.1.2. ESTUDIOS EXCLUIDOS.....	16
5.1.3. CARACTERÍSTICAS Y SÍNTESIS DE LOS ESTUDIOS INCLUIDOS.....	16
5.2. RIESGO DE SESGO DE LOS ESTUDIOS INCLUIDOS.....	27
6. DISCUSIÓN	28
7. CONCLUSIONES	31
8. DECLARACIONES DE INTERÉS	31
9. BIBLIOGRAFÍA	32
10. ANEXOS	38

1. INTRODUCCIÓN

El peso fetal está directamente relacionado con el desarrollo y bienestar fetal. Es un dato importante en el seguimiento del embarazo. Su estimación es de utilidad para orientar decisiones sobre la frecuencia del seguimiento antenatal, el momento del parto y las intervenciones terapéuticas después del mismo [1].

Se considera un feto adecuado para su edad gestacional (AEG) aquel que tiene un peso entre el percentil 10 y 90 para su edad gestacional y sexo; un feto macrosómico o grande para la edad gestacional (GEG) es el que tiene un peso fetal estimado (PFE) mayor del percentil 97 para su edad y sexo o mayor de 4000 gramos al nacimiento; un feto pequeño para la edad gestacional (PEG) se define como un PFE entre el percentil 3 y 10 con estudio de redistribución vascular mediante Doppler normal [2].

El peso fetal podría variar de acuerdo con diferentes variables tales como la edad y altura materna, etnia, sexo, posición y adiposidad fetal; así como, enfermedades materno-fetales. Un peso fetal no adecuado aumenta el riesgo de complicaciones, tanto maternas como fetales. En el caso de feto macrosómico, son mayores las complicaciones fetales como la distocia de hombros, asfixia, hipoglucemia, policitemia o la muerte fetal; y maternas como el trabajo de parto prolongado, parto instrumentado, cesárea, hemorragia postparto o ruptura uterina; la obesidad pre-embarazo y la diabetes gestacional se han relacionado directamente con un feto GEG [3]. Por otra parte, los resultados perinatales adversos de un feto PEG incluyen hipoglucemia, policitemia, asfixia, parto prematuro e incluso muerte neonatal [4]. No solo es importante diagnosticar correctamente a los fetos GEG y PEG para prevenir y tratar las complicaciones clínicas, sino también para evitar las repercusiones negativas a nivel emocional, que desencadenaría un diagnóstico erróneo.

Múltiples sistemas se han utilizado para estimar el peso fetal pero actualmente no se dispone de un método directo para cuantificarlo. En 1970 se introdujo la ecografía bidimensional (2D) en la práctica obstétrica y el peso fetal empezó a ser objeto de investigación [5]. Se desarrollaron decenas de fórmulas para estimar el peso fetal pero ninguna lo estima con total fiabilidad [6]; actualmente, la fórmula ecográfica 2D más

utilizada es la de Hadlock (basada en el diámetro biparietal (DBP), la circunferencia cefálica (CC), la circunferencia abdominal (CA) y la longitud del fémur (LF) fetales, medidos de forma estandarizada [7,8]). Esta fórmula predice el peso al nacer dentro del 10% en la gran mayoría (80%) de los embarazos [9]. No obstante, la variedad de fórmulas desarrolladas determina la ausencia de un método totalmente preciso para estimar el peso fetal, y genera dudas sobre su validez por variar el resultado dependiendo del método elegido [10].

Con la finalidad de crear un modelo estándar para estimar el peso fetal, nacieron dos proyectos a nivel internacional, con participantes de diferentes países: el estudio Intergrowth-21st (IG) [11,12] y las tablas de crecimiento fetal de la Organización Mundial de la Salud (OMS). Aunque ambos se diseñaron de forma similar y permiten calcular el percentil de peso fetal según la edad gestacional, se diferencian por obtener conclusiones contradictorias en referencia al crecimiento fetal: el proyecto IG asumió que el crecimiento fetal no variaba, mientras que el estudio de la OMS demostró que existían diferencias significativas entre el crecimiento fetal de las diferentes poblaciones [13].

Por otra parte, existen estudios que comparan la fórmula Hadlock y el proyecto IG con conclusiones controvertidas: algunos [9,14–16] afirman que el primero es más efectivo que el segundo, otros afirman lo contrario [17], recomiendan el uso de gráficas locales o customizadas según las características antropométricas de la población [18], o directamente, no recomiendan el uso de IG [19].

Existen revisiones sobre el uso de la ecografía en el tercer trimestre [20,21] pero no enfocadas al método que mejor estima el peso fetal durante el tercer trimestre en mujeres del área mediterránea. Por tanto, creemos necesaria una revisión sistemática sobre el mejor método para estimar el peso fetal en la población del área mediterránea, dada su importancia clínica y repercusión en la morbilidad y mortalidad perinatal; y así, poder actualizar las guías de práctica clínica de nuestro Hospital. Asimismo, esta revisión se basa en un estudio actual que se está llevando a cabo sobre nuevas formas de estimación de peso fetal, en el Servicio de Ginecología y Obstetricia del Hospital de La Plana (Vila-Real, Castellón); en este estudio, Almenar et al. revisan el tema hasta el año 2015 [22].

2. OBJETIVOS

El objetivo general de esta revisión sistemática es:

- Valorar el método ecográfico más preciso en la estimación del peso fetal, en fetos AEG en mujeres sanas embarazadas del área mediterránea durante el tercer trimestre de embarazo.

El objetivo específico es:

- Valorar el método ecográfico más preciso en la estimación del peso fetal, en fetos con factores de riesgo macrosómicos en mujeres embarazadas del área mediterránea durante el tercer trimestre de embarazo.

3. PREGUNTAS CLÍNICAS

Para la realización de esta revisión sistemática nos planteamos las siguientes preguntas referentes a la bibliografía:

- ¿Existen publicaciones de síntesis y/o revisiones sistemáticas sobre el mejor método para estimar el peso fetal durante el embarazo? ¿Y durante el embarazo en mujeres del área mediterránea? En caso de que existiesen:
 - ¿Estas publicaciones están actualizadas?
 - ¿Están basadas en la mejor evidencia disponible?
 - ¿Son aplicables a nuestra población diana?
- ¿Disponemos de guías de práctica clínica y/o protocolos sobre el control del peso fetal y su manejo asistencial en caso de alteración? En caso de que hubiesen:
 - ¿Están actualizados?
 - ¿Están basados en la mejor evidencia disponible?
 - ¿Son aplicables a nuestra población diana?
 - ¿Nos facilitan notoriamente el control y el manejo del peso fetal durante la gestación?

4. MÉTODOS

La metodología de esta revisión sistemática sigue las directrices de la guía PRISMA 2020 [23] para revisiones sistemáticas.

4.1. FUENTES DE INFORMACIÓN

Para la elaboración de esta revisión sistemática se realizó una búsqueda bibliográfica que se inició en noviembre de 2021 y finalizó el 11 de enero del 2022. Se consultaron publicaciones de síntesis (Up To Date), guías de práctica clínica e instituciones de referencia de Obstetricia y Ginecología. La búsqueda incluyó estudios de tres bases de datos distintas (PubMed, Web of Science y Scopus). Se establecieron las palabras clave mediante HonSelect: “peso fetal”, “peso al nacimiento”, “macrosomía”, “ultrasonografía” y “tercer trimestre de embarazo”. A continuación, se explican detalladamente las fuentes de información bibliográficas utilizadas, los criterios de búsqueda y los artículos que finalmente se utilizaron como fuente bibliográfica.

4.1.1. Recursos de síntesis

En UpToDate se introdujo en el buscador “ultrasonography fetal weight”. Tras la lectura de los artículos obtenidos, se seleccionaron los que quedan referenciados en la bibliografía [3,24] por tratar el objetivo principal de esta revisión.

4.1.2. Guías de práctica clínica

Se consultaron las siguientes guías de práctica clínica con el objetivo de conocer las recomendaciones:

- National Institute for Health and Clinical Excellence (NICE): Antenatal care [25].
- New Zealand Guidelines Group (NZGG): New Zealand Obstetric Ultrasound Guidelines [26].
- Sistema Nacional de Salud, Ministerio de Sanidad, Servicios Sociales e Igualdad: Guía de Práctica Clínica de atención en el embarazo y el puerperio [27].

4.1.3. Instituciones de referencia

También se consultaron las páginas web de instituciones de referencia de Obstetricia y Ginecología con el fin de conocer las recomendaciones a nivel de la práctica clínica:

- American Collage of Obstetricians and Gynecologists (ACOG): Macrosomia [28], Perinatal Care [29]
- Royal College of Obstetricians and Gynecologists (RCOG): Shoulder Dystocia [30], The Investigation and Management of the Small-or-Gestational-Age Fetus [31]
- International Society of Ultrasound in Obstetrics and Gynecology (ISUOG): Ultrasound assessment of fetal biometry and growth [32].
- International Federation of Gynaecology and Obstetrics (FIGO): FIGO position paper on reference charts for fetal growth and size at birth: Which one to use? [33]
- Sociedad Española de Ginecología y Obstetricia (SEGO): Guía de la exploración ecográfica del III trimestre [34]
- Hospital Clínic de Barcelona, el Hospital Sant Joan de Déu y la Universitat de Barcelona: Screening ecográfico [35], Macrosomía [36], Defectos del crecimiento fetal [37].

4.1.4. Bases de datos

Inicialmente, para realizar la búsqueda en Pubmed se identificaron los términos MeSH (Medical Subject Headings) mediante el catálogo de términos HONselect y se combinaron los términos MeSH con el texto libre correspondiente. Posteriormente, se combinaron los términos de significado similar con el operador booleano OR y estos, con el booleano AND. En Web of Science y Scopus se emplearon los mismos vocablos combinando los términos de la misma forma. La descripción de la búsqueda en cada base de datos se detalla en la tabla 1.

Bases de datos	Descripción de la búsqueda	Número de artículos
PubMed	(((((("fetal weight"[MeSH Terms]) OR ("fetal weight"[Title/Abstract]) OR ("birth weight"[MeSH Terms])) OR ("birth weight"[Title/Abstract])) OR ("fetal macrosomia"[MeSH Terms])) OR ("macrosomia"[Title/Abstract])) OR ("Customized fetal	773

	growth"[Title/Abstract])) AND (((("Ultrasonography, Prenatal"[MeSH Terms]) OR ("ultrasound"[Title/Abstract]) OR (ultrasonography[Title/Abstract]) OR (ultrasonics[Title/Abstract]))) AND (("pregnancy trimester, third"[MeSH Terms]) OR ("third trimester"[Title/Abstract]))	
Web of science	((TS=("fetal weight")) OR TS=("birth weight")) OR TS=("fetal macrosomia")) OR TS=("customized fetal growth") AND (((TS=("Ultrasonography, Prenatal")) OR TS=("ultrasound")) OR TS=("ultrasonography")) OR TS=("ultrasonics")) AND (TS=("pregnancy trimester, third")) OR TS=("third trimester")	475
Scopus	TITLE-ABS-KEY("fetal weight" OR "birth weight" OR "fetal macrosomia" OR "customized fetal growth") AND TITLE-ABS-KEY("ultrasonography, prenatal" OR "ultrasound" OR "ultrasonography" OR "ultrasonics") AND TITLE-ABS-KEY("pregnancy trimester, third" OR "third trimester")	1088

Tabla 1: Estrategia de búsqueda bibliográfica.

Realizada la búsqueda inicial, se decidió filtrar los resultados por fecha de publicación porque en el Servicio de Ginecología y Obstetricia del Hospital de la Plana (Vila-Real, Castellón) se está realizando un estudio [22] con una revisión de la literatura hasta el año 2015. Por tanto, el número de estudios publicados entre 2016 y 2022 (hasta de 11 enero 2022) fueron: 318 en Pubmed, 224 en Web of Science y 458 en Scopus. Los documentos se almacenaron en el programa informático de gestión de referencias bibliográficas Mendeley con el fin de organizarlos y eliminar los duplicados (498). De los 502 artículos restantes, se revisó título y resumen y se eliminaron 424 por no cumplir los criterios de inclusión comentados en el apartado 4.2 (Criterios de selección de estudios). A continuación, se revisó el texto completo de los 78 documentos y se incluyeron los 7 que cumplían con los criterios de inclusión [38–44].

4.1.5. Búsqueda manual

Para completar la selección de los estudios, se obtuvieron mediante búsqueda manual de las referencias bibliográficas 2 artículos [45,46].

4.1.6. Literatura no publicada

Se consultó el proyecto no publicado que se está realizando en el Servicio de Obstetricia y Ginecología en el Hospital de La Plana [22].

4.2. CRITERIOS DE SELECCIÓN DE ESTUDIOS

Los criterios de elegibilidad y exclusión que se siguieron para elegir los artículos de esta revisión se presentan a continuación.

Criterios de inclusión:

- Estudios de test diagnósticos observacionales.
- Artículos publicados entre 2016 y 2022.
- Estudios que incluyeron mujeres embarazadas sanas del área mediterránea en el tercer trimestre de gestación con feto único normo o macrosómico.
- Artículos cuyo método principal en evaluar el peso fetal fue la ecografía.

Criterios de exclusión:

- Estudios no observacionales.
- Artículos publicados fuera del periodo entre 2016 y 2022.
- Estudios que incluyeron mujeres embarazadas con patología materna y/o fetal, de área diferente a la mediterránea o con fetos múltiples o feto PEG.
- Artículos científicos que no evaluaron la ecografía en la estimación del peso fetal.

4.3. EXTRACCIÓN Y ANÁLISIS DE LOS DATOS

Posteriormente al proceso de selección de artículos de las bases de datos, se procedió a extraer la información de los estudios seleccionados. Se obtuvieron los siguientes datos: autor y año de publicación, objetivos, país donde se efectuó el estudio (del área mediterránea), tipo de estudio observacional (cohorte, caso/control o transversal), número de participantes incluidos, prueba índice, prueba control, variable resultado y análisis estadístico.

El riesgo de sesgo de los artículos seleccionados se evaluó mediante una herramienta para revisiones sistemáticas de estudios de precisión diagnóstica: QUADAS-2 [47,48]. Esta herramienta permite evaluar los posibles sesgos en cuatro dominios: la selección de

pacientes, la prueba índice, el estándar de referencia y el flujo y tiempo. Cada dominio es evaluado en términos de su riesgo de sesgo (bajo, alto o incierto) y los primeros tres dominios también se evalúan por su aplicabilidad (preocupación sobre la aplicabilidad baja, alta o incierta) [49].

5. RESULTADOS

5.1. DESCRIPCIÓN DE LOS ESTUDIOS

5.1.1. Resultados de la búsqueda

La Figura 1 muestra el diagrama de flujo del proceso de selección, basado en la guía PRISMA 2020 [23]. De 2336 artículos que se obtuvieron en la búsqueda inicial, se conservaron 502 después de filtrar por fecha de publicación (años 2015-2022) y eliminar duplicados (n= 498). Tras realizar un primer cribado por lectura de título y resumen y un segundo cribado por lectura de texto completo, se identificaron 7 artículos que cumplían los criterios de inclusión. Posteriormente, se añadieron 2 estudios a partir de la revisión manual de las referencias bibliográficas de los artículos incluidos previamente. Por tanto, se incluyeron un total de 9 estudios para síntesis cualitativa.

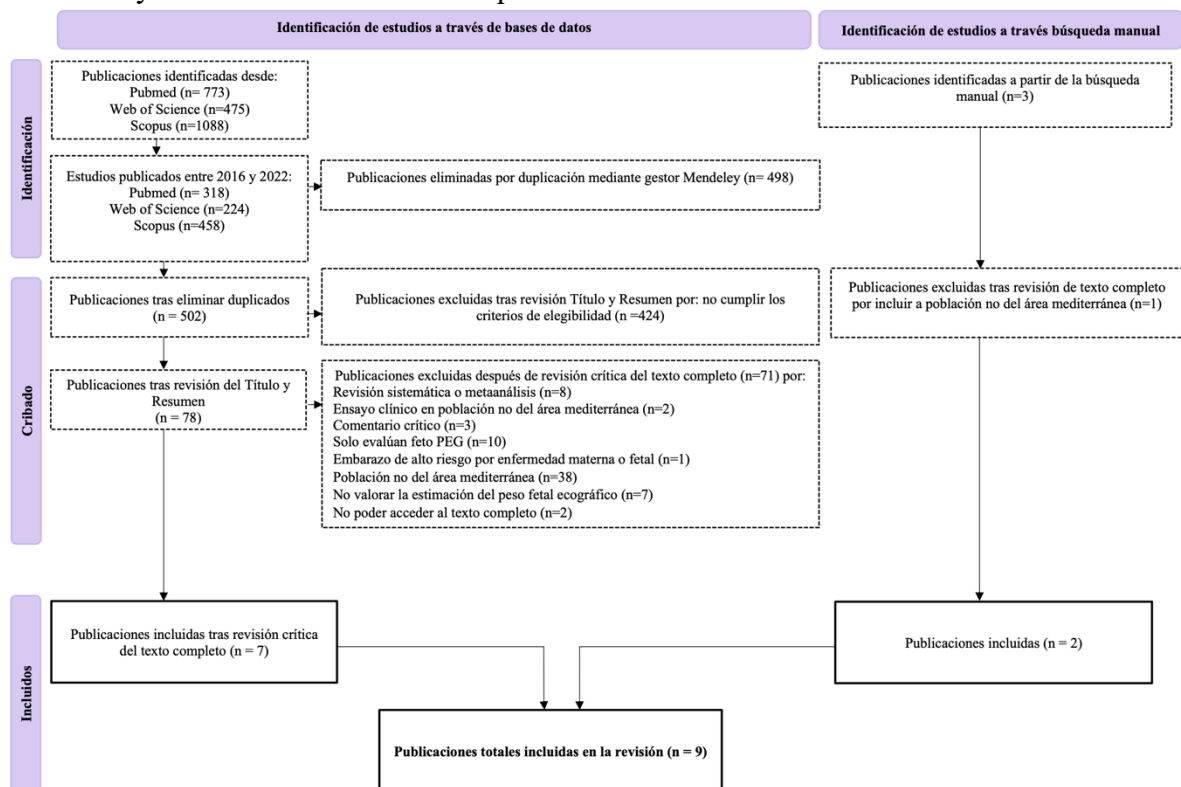


Figura 1: Diagrama de flujo.

5.1.2. Estudios excluidos

Los estudios excluidos en el segundo cribado (tras lectura crítica de los artículos) y la razón por la que fueron excluidos se presentan en el ANEXO 1. La gran mayoría de artículos fueron excluidos por incluir participantes de fuera del área mediterránea (n=38) y solo fetos PEG (n=10).

5.1.3. Características y síntesis de los estudios incluidos

Las características de los artículos incluidos en la revisión se presentan en la tabla 2 (año, país, objetivo, tipo de artículo, número y características de los participantes) y 3 (prueba índice, prueba control, variable resultado y análisis estadístico) y, los principales resultados en la tabla 4.

Las fechas de publicación de los artículos incluidos en la revisión sistemática comprendieron los años: 2017 [42,46], 2018 [38], 2019 [44], 2020 [41,45], 2021 [39,43] y 2022 [40], y fueron desarrollados en diferentes países del área mediterránea: España [38,44], Francia [40,46], Italia [43,45], Israel [39,41] y Egipto [42].

Se encontraron 4 estudios de cohortes prospectivos [38,42–44], 1 cohorte retrospectivo [40], 2 casos y controles retrospectivos [39,41] y 2 estudios transversales retrospectivos [45,46].

El total de participantes que se incluyeron en nuestra revisión fue de 64546. En concreto, 2696 gestantes participaron en la evaluación ecográfica del crecimiento longitudinal con respecto a una evaluación transversal [38]; 531 mujeres fueron evaluadas para valorar la fiabilidad del PFE según Hadlock [41]; 41261 gestantes participaron en la comparación del PFE mediante Hadlock con respecto a Hadlock y las características maternas [39]; en 2152 gestantes se comparó el PFE mediante la fórmula Hadlock con las características maternas y el flujo de la vena umbilical obtenido por Doppler con el PFE mediante Hadlock y las características maternas y, con el PFE solo con la fórmula Hadlock [43]; en 1224 mujeres se comparó el PFE ajustado a la gestación y Hadlock con el PFE ajustado a la gestación y peso real al nacer [44]; 4858 gestantes participaron en la comparación de los gráficos de crecimiento locales de Francia con IG [46]; en 4427 mujeres se

compararon las gráficas de la OMS con IG, Hadlock, Salomon y la estimación del peso fetal según el Colegio Francés de Ecografía Fetal (CFEF) [40]; en 50 gestantes se compararon las fórmulas M3, M6, STT1(Sciosiosa) y STT2 (Sciosiosa modificada) con Hadlock [42] y finalmente, en 7347 gestantes se compararon las gráficas Nicolini, Paladini, OMS e IG [45].

Las gestantes incluidas presentaban una edad gestacional entre 11 y 41 semanas gestacionales. En particular, 21 ± 2 y 32 ± 2 semanas [38], ≥ 33 semanas [40]; entre 24-41 semanas [41]; ≥ 37 semana [42]; 35-37 semanas [43]; 37-41 semanas [45]; 11-14, 21-24 y 30-34 semanas [46] y 38-40 semanas [44].

En la mayoría de los artículos [39–45] se estimó directamente el peso fetal mediante las fórmulas correspondientes o indirectamente, mediante los parámetros biométricos [46]; solamente en uno [38] se midió la CA y su velocidad de crecimiento.

En todos los artículos incluidos, el control de factores de confusión durante el diseño se realizó, sobre todo, pero no únicamente, mediante los criterios de exclusión de los individuos a estudio. Durante el análisis, el control de factores de confusión se realizó en la mayoría de los artículos [38,39,41,43–46] mediante análisis multivariable y en la minoría, mediante análisis de sensibilidad [40].

En cuanto a los hallazgos obtenidos, Caradeux et al.[38] concluyeron que el tamaño (CA) tenía un mejor desempeño predictivo que el crecimiento (velocidad CA); es decir, que la evaluación longitudinal era inferior a una única evaluación transversal a las 32 semanas de gestación para la predicción del PFE, en concreto GEG.

Los estudios de cohortes se centraron en comparar el PFE mediante el método Hadlock con otros métodos. Según Hocquette et al.[40] el PFE en fetos con medidas por debajo del percentil 3 y 10 (p10) fue similar mediante los gráficos de la OMS y Hadlock y, difirió en los fetos con un PFE por encima del percentil 90 (p90) y 97 siendo en este caso, más precisos los gráficos de la OMS frente a Hadlock; IG identificó bajas proporciones de fetos con medidas por debajo del p10 y altas proporciones de fetos con medidas por encima del p90; es decir, subestimó los fetos PEG y sobreestimó los fetos GEG. Por otra parte, Mohsen et al.[42] afirmaron que Hadlock funcionó mejor que las fórmulas M3,

M6, Sciosiosa y Sciosiosa modificada; pero en concreto, el método de volumen fraccional de extremidades (M6) se aproximó a Hadlock en cuanto a precisión y fiabilidad. Por otro lado, Rizzo et al.[43] comprobaron que la adición del parámetro ecográfico FSVU al método Hadlock y a las características maternas mejoró significativamente la predicción de los fetos macrosómicos. Finalmente, Vila-Candel et al.[44] concluyeron que el PFE ajustado por características maternas fue más preciso que el PFE utilizando solamente medidas ecográficas en fetos con un peso inferior a 2500 g y en fetos con peso dentro de los rangos de la normalidad; sin embargo, tanto la ecografía como el método combinado resultaron ser igualmente precisos para predecir la macrosomía fetal.

Los estudios de casos y controles valoraron parámetros que influían en la precisión del PFE. Hendin et al.[39] detectaron en mujeres de más de 41 semanas de gestación y en fetos macrosómicos un PFE no preciso (PFEnp) y Krispin et al.[41] relacionaron la presentación cefálica, la columna vertebral lateral y la placenta anterior con un PFE más preciso.

Finalmente, los estudios transversales se basaron en comparar el PFE según IG con otros métodos. Stampalija et al.[45] mostraron que IG subestimaba los fetos con PFE inferior al p10, mientras que tanto las tablas de crecimiento de la OMS como IG sobreestimaban los fetos por encima del p90, aunque el efecto fue mayor para este último. Por otra parte, Stirnemann et al.[46] concluyeron que IG fue más preciso en estimar las medidas CC, CA y LF que los gráficos locales franceses.

Autor	Año	País	Objetivos	Tipo de estudio	Individuos incluidos	Características de los participantes incluidos
Caradeux et al.	2018	España	Comparar la evaluación del crecimiento longitudinal en el segundo y tercer trimestre con la evaluación transversal en el tercer trimestre en la predicción del tamaño para la edad gestacional (GEG) y la macrosomía.	Cohorte prospectivo	2696	Embarazos únicos consecutivos no seleccionados durante la exploración del segundo (21 ± 2 semanas) y tercer trimestre (32 ± 2 semanas). Los criterios de exclusión fueron: diabetes mellitus preexistente, anomalías cromosómicas confirmadas mediante enfoques genéticos (cariotipo convencional o array-CGH), defectos estructurales sospechados en el momento de las exploraciones de rutina y confirmados después del nacimiento, evidencia de infección fetal confirmada mediante examen microbiológico en sangre materna y/o líquido amniótico, y sospecha de restricción temprana del crecimiento fetal (peso fetal estimado por ecografía inferior al p10, según los estándares locales) o preeclampsia antes de la semana 32.
Hendin et al.	2021	Israel	Examinar los determinantes maternos y fetales asociados con un PFE poco fiable (PFEnp)	Casos y controles retrospectivo	41261	Todas las mujeres con embarazos a término de feto único que dieron a luz dentro de las 2 semanas posteriores al PFE por ecografía. Los criterios de exclusión fueron gestaciones múltiples y anomalías fetales conocidas.
Hocquette et al.	2022	Francia	Evaluar la aplicabilidad de las tablas de crecimiento fetal de la OMS para la CA, la LF y el PFE en las ecografías del segundo y tercer trimestre en una cohorte francesa de nacimiento.	Cohorte retrospectivo	4427	Participantes de una cohorte de nacimiento ELFE (un estudio observacional basado en la población en 349 unidades de maternidad de Francia). La población estuvo constituida por nacidos vivos únicos después de las 33 semanas de gestación ocurridos durante 25 días en 2011. Otros criterios de inclusión fueron: madres de mayores de 18 años, padres que entendían francés, inglés, árabe o turco y padres que vivieron y planearon quedarse en Francia por al menos tres años y dieron su consentimiento informado por escrito.
Krispin et al.	2020	Israel	Asignar parámetros asociados con desviaciones significativas en el peso fetal estimado ecográfico y evaluar los resultados del trabajo de parto en tales circunstancias.	Casos y controles retrospectivo	531	Mujeres con una gestación única entre las 24-41 semanas de gestación que se sometieron a la evaluación del peso fetal ecográfica dentro de una semana antes del parto en un único centro médico terciario de Israel. Se incluyeron las mujeres que experimentaron un parto con un nacido vivo dentro de la semana posterior a la estimación del peso fetal mediante ecografía. Se excluyeron las gestaciones múltiples, malformaciones fetales conocidas y muertes fetales intrauterinas. El grupo de estudio estuvo compuesto por todos los embarazos en los que el PFE difería en al menos un 20% por debajo o por encima del peso real al nacer. El grupo de control incluyó embarazos emparejados en los que el PFE ecográfico fue preciso y difería hasta en un 1% por debajo o por encima del peso real al nacer.
Mohsen et al.	2017	Egipto	Evaluar dos nuevas fórmulas de ultrasonido 3D (M3 y M6) y dos nuevas 2D (STT1 y STT2) para la estimar el peso fetal, compararlas con	Cohorte prospectivo	50	Cualquier paciente embarazada a partir de 37 semanas de gestación, que acudiera a nuestro departamento para una ecografía obstétrica durante su último tercer trimestre fue incluida en el estudio solo si su parto se produjo dentro de los cinco días posteriores a la ecografía. Criterios de inclusión: se incluyeron embarazos únicos, dentro de los cinco días

			la fórmula Hadlock y comparar su estimación con el peso fetal real.			posteriores al parto y embarazos normales o RCIU. Criterios de exclusión: embarazos múltiples, DIU, pacientes en estado crítico (por la necesidad de una ecografía rápida corta), embarazos de menos de 37 semanas de gestación, excepto trabajo de parto prematuro.
Rizzo et al.	2021	Italia	Evaluar el rendimiento diagnóstico de los Doppler venoso y arterial fetales para predecir la macrosomía en el tercer trimestre de embarazo y construir un modelo de predicción multiparamétrico que incluya las características del embarazo, la ecografía y el Doppler, capaz de predecir con precisión la macrosomía.	Cohorte prospectivo	2152	Embarazos únicos sometidos a evaluación por ultrasonido a las 35-37 semanas de gestación. Criterios de inclusión: correcta datación del embarazo desde el primer trimestre, ausencia de anomalías cromosómicas o estructurales fetales y parto en nuestra unidad.
Stampalija et al.	2020	Italia	Evaluar la práctica clínica a nivel nacional entre los centros de referencia de tercer nivel en el uso de tablas de crecimiento biométrico fetal y, evaluar el impacto en la detección del crecimiento fetal de las tablas de crecimiento nacionales e internacionales.	Transversal retrospectivo	7347	Embarazos únicos sin complicaciones de 37-41 semanas con resultado perinatal sin incidentes e información disponible sobre paridad, altura materna y peso. Embarazos complicados por defectos estructurales o cromosómicos fetales. anomalías, enfermedades maternas pre-existentes o el desarrollo de complicaciones obstétricas fueron excluidos. Sólo las gestantes con IMC antes del embarazo de 17-30kg/m ² se seleccionaron.
Stirneman et al.	2017	Francia	Evaluar las diferencias potenciales en el tamaño fetal entre la población francesa y la población internacional del proyecto IG y medir el impacto del cambio a las referencias IG para el tamaño fetal.	Transversal retrospectivo	4858	Embarazos de bajo riesgo reclutados prospectivamente dentro de la red del Colegio Nacional Francés de Ecografía Fetal durante un período de 6 semanas. Se realizó una selección adicional en base a los criterios del proyecto IG. Se incluyeron mujeres con bajo riesgo de resultados adversos maternos y perinatales. Así, sólo las mujeres mayores o iguales a 18 años y menores o iguales a 35 años, que tuvieran un IMC mayor o igual a 18,5kg/ m ² y menor a 30kg/m ² , sin antecedentes obstétricos o ginecológicos clínicamente relevantes, que iniciaron el control obstétrico antes de la semana 14 de gestación.
Vila-Candel et al.	2019	España	Desarrollar y evaluar una ecuación basada en parámetros clínicos maternos y biometría ecográfica del tercer trimestre (método combinado), y compararla con el PFE mediante la fórmula Hadlock.	Cohorte prospectivo	1224	Se incluyeron mujeres con una primera cita prenatal y ultrasonidos entre las semanas 5 y 12 de embarazo, embarazo de feto único sin anomalías fetales y nacimiento entre las semanas 38 y 40. Se excluyeron del estudio los embarazos complicados por polihidramnios, hipotiroidismo, preeclampsia, diabetes gestacional y oligohidramnios.

Tabla 2: Características de los artículos incluidos: Año, país, objetivo, tipo de artículo, número y características de los participantes.

Autor	Prueba índice	Prueba control	Variable resultado	Análisis estadístico
Caradeux et al.	Evaluación ecográfica transversal de la CA a las 32 ± 2 semanas de gestación.	La evaluación del crecimiento longitudinal se realizó mediante el cálculo de la velocidad z y el crecimiento condicional	CA y velocidad de crecimiento de CA, GEG y macrosomía	Se realizaron la prueba t de Student y la χ^2 de Pearson (o prueba exacta de Fisher) para las comparaciones univariadas entre grupos de variables cuantitativas y cualitativas, respectivamente. Las mediciones de CA se transformaron a valores z (valor z de CA) según los estándares IG. La asociación entre crecimiento longitudinal y GEG/macrosomía se evaluó mediante regresión logística, donde el modelo basal de comparación fue la evaluación transversal al tercer trimestre. Es decir, el análisis de regresión logística se realizó para evaluar si la combinación de una puntuación z CA alta de 32 semanas (> percentil 90) y una velocidad z CA alta de 22 a 32 semanas se asoció con un mayor riesgo de GEG o macrosomía. En este modelo de regresión logística se obtuvieron: valor z CA, valor z CA + valor z CA velocidad y, valor z CA + CA condicional (el CA condicional permite cuantificar una medición de CA observada en el tercer trimestre dado el valor esperado de una medición del segundo trimestre). Los modelos se compararon evaluando la mejora en su Nagelkerke R2 como una medida de bondad de ajuste. El rendimiento de la predicción para GEG y macrosomía se determinó mediante el análisis de la curva característica operativa del receptor (ROC) y estas curvas ROC pareadas se compararon mediante el método De Long.
Hendin et al.	PFE según Hadlock y las características maternas	PFE según Hadlock	PFE	Las características de las mujeres se describieron como proporciones para las variables categóricas y como media y desviación estándar para las variables continuas. La significación se evaluó mediante la prueba de Chi-cuadrado y la prueba exacta de Fisher para variables categóricas. Se utilizó la prueba t de Student para el análisis de variables continuas con distribución normal y la prueba U de Mann-Whitney para el análisis de variables continuas con distribución sesgada. La cohorte del estudio se estratificó aún más en categorías de peso al nacer para delinear el papel del peso fetal en la precisión del PFE. Los resultados del estudio se presentaron como razones de probabilidad (OR) e intervalos de confianza (IC) del 95 %. Un valor bilateral < 0,05 indicó significación estadística. Se realizaron análisis estadísticos multivariados.
Hocquette et al.	Gráficas de crecimiento fetal de la OMS	IG, Hadlock, Salomon, y CFEF (solo PFE)	CA, LF y PFE	Primero se describieron las características maternas, del embarazo y de los recién nacidos y se compararon las características de los recién nacidos incluidos en el estudio con los excluidos por falta de información de ultrasonido usando Chi-cuadrado o la prueba t de Student, respectivamente. Luego se calcularon las proporciones de segundo y tercer trimestre CA, LF y PFE y los percentiles (< percentil 3, <percentil 10, >percentil 90 y >percentil 97) según cada gráfico en el conjunto de la muestra y en la submuestra de embarazos de bajo riesgo. Se realizaron todos los análisis en la totalidad de la muestra y en una submuestra de embarazos de bajo riesgo. Finalmente, se llevó a cabo un análisis de sensibilidad para evaluar el impacto de los casos excluidos (que no se pudieron calcular por la falta de parámetros biométricos).

Krispin et al.	Hadlock	-	PFE, presentación fetal, ubicación de la placenta y posición de la columna.	Las variables continuas se compararon utilizando el modelo lineal general. Se utilizaron las pruebas de Chi-cuadrado y exacta de Fisher para las variables categóricas, según correspondiera. Las diferencias se consideraron significativas cuando el valor p fue inferior a 0,05. Se utilizó un análisis de regresión logística multivariable para ajustar los resultados a posibles factores de confusión. Las imprecisiones del PFE se ajustaron a los días desde el ultrasonido hasta el nacimiento y el peso al nacer
Mohsen et al.	M3, M6, STT1, STT2	Hadlock	PFE	Primero se realizó un análisis descriptivo de la muestra y se evaluaron los métodos de estimación del peso al nacer de la siguiente manera: (a) Correlación con PFA (peso fetal actual) utilizando el coeficiente de correlación de Pearson. (b) Diferencia entre PFA y PFE. (c) Error sistemático entre PFA y PFE. (d) Error aleatorio. (e) Error porcentual absoluto. (f) Fracciones de estimaciones dentro del 5%, 10% y 15% del PFA. El error sistemático se calculó de la siguiente manera: $[(PFE-PFA)/PFA] \times 100$ y esto reflejó la desviación sistemática o regular firmada del modelo del PFA y es una medida de precisión del modelo. El error aleatorio se calculó como la desviación estándar del error sistemático multiplicado por 100. Esto reflejó la desviación aleatoria (no sistemática) de PFA y es una medida de la precisión del modelo. Los datos numéricos se mostraron como media \pm desviación estándar, mientras que los datos no paramétricos se mostraron por la mediana. Se utilizó una prueba T de muestra para comparar los errores sistemáticos de todos los modelos. Se utilizó la prueba T para muestras pareadas para comparar los nuevos métodos de estimación del peso fetal con el método MH (Hadlock). Se consideró un resultado significativo si valor $p < 0,05$.
Rizzo et al.	Hadlock + características maternas y Hadlock + parámetro Doppler FSVU	Hadlock	PFE	Las variables biométricas, los datos de estudio Doppler y el PFE se expresaron como valores z. Los datos se expresaron como mediana y rango intercuartílico o número (n) y porcentaje (%). Las variables continuas se compararon mediante la prueba U de Mann-Whitney, mientras que las categóricas mediante Chi-cuadrado. Se realizó un análisis de regresión logística multivariante para explorar la fuerza de la asociación entre las características maternas y del embarazo, el PFE, el flujo sanguíneo de la vena umbilical (FSVU) y la macrosomía. Los resultados se informaron como razones de probabilidad ajustadas (ORa) con sus intervalos de confianza (IC) del 95%. La colinealidad se probó mediante la correlación de Pearson. Finalmente, se llevó a cabo el análisis de la curva de características operativas del receptor (ROC) para dilucidar el rendimiento diagnóstico de FSVU y las características maternas y ultrasonográficas en la predicción de macrosomía y el área bajo la curva (ABC) calculada.
Stampalija et al.	Nicolini, Paladini, IG y OMS	-	CC, CA, LF y PFE	Las características de la población estuvieron representadas por mediana y rango intercuartílico (RIC) o por frecuencia y porcentaje. Para evaluar el impacto clínico de las tablas de crecimiento examinadas, se calcularon para cada parámetro biométrico y semana gestacional el porcentaje de fetos cuyas medidas biométricas estaban por debajo y por encima p10 y p90 de cada gráfico de crecimiento. Se aplicaron regresiones por cuantiles para estimar la probabilidad de cada medición ecográfica por semana gestacional para los percentiles correspondientes, con el fin de poder comparar gráficamente el percentil en curvas de la población de estudio con las tablas de crecimiento evaluadas.

Stirneman et al.	Gráficos locales	IG	CC, CA, LF	El impacto de cambiar a las referencias del IG se cuantificó comparando el valor z tanto en el IG como en las referencias derivadas localmente para CC, CA y LF. Específicamente, las diferencias estandarizadas fueron las diferencias entre la media obtenida en la muestra francesa y la media de los datos del IG agrupados, expresados como unidades de la desviación estándar (DE). Por lo tanto, las diferencias estandarizadas fueron similares a un valor z. Los valores de diferencia estandarizados de menos de 0,5 se especificaron previamente como adecuados para reflejar que la población francesa podría haber sido incluida en el conjunto de datos agrupados utilizados para construir los estándares internacionales
Vila-Candel et al.	Proyección ajustada por gestación (GAP) y PFE mediante Hadlock (PFEh).	PFE obtenido por regresión multivariante (PFEr) utilizando variables antropométricas, demográficas, ultrasonográficas y obstétrico-neonatales.	PFE	Datos estadísticos descriptivos: media y desviación estándar (variables continuas), o mediana y rango intercuartílico (variables categóricas). Normalidad: Kolmogorov-Smirnov (variables continuas); prueba no paramétrica (variables no normales). En el análisis bivariado se estudiaron las correlaciones entre las variables dependientes (peso al nacer) e independientes recogidas mediante la prueba de la t de Student para comparar medias de variables cuantitativas. Todas las variables con significación estadística ($p < 0,05$) y valor clínico se incluyeron en un análisis multivariante para identificar la ecuación de predicción de peso al nacer más precisa (PFEp). La precisión de los diferentes métodos (PFEh/ PFEr) para estimar el peso fetal luego se evaluó calculando el coeficiente R de correlación de Pearson entre los pesos fetales estimados obtenidos usando cada ecuación y el peso real al nacer. Se usó el Coeficiente de Correlación Intraclase (CCI) para evaluar el grado de concordancia de ambos métodos con el peso neonatal real.

Tabla 3: Características de los artículos incluidos: Prueba índice, prueba control, variable resultado y análisis estadístico.

Autor	Principales resultados
Caradeux et al.	<p>Con respecto a la predicción de GEG, las áreas bajo la curva (ABC) para el valor z CA de 32 semanas, la velocidad z de CA y la CA condicional fueron 0,7841 (intervalo de confianza [IC] del 95 % 0,7472–0,8209), 0,6106 (IC del 95 %, 0,5645–0,6567) y 0,5533 (IC 95 %, 0,5072–0,5995), respectivamente. Tanto las ABC de la velocidad z de CA ($p < 0,001$) como las de CA condicional ($p < 0,001$) fueron significativamente más bajas que las ABC del valor z de CA transversales. Al comparar ambas evaluaciones longitudinales, la velocidad CA z se desempeñó significativamente mejor que la CA condicional ($p < 0,001$). Con respecto a la predicción de macrosomía, las ABC para el valor z de CA de 32 semanas, la velocidad z de CA y la CA condicional fueron 0,7524 (IC del 95 %, 0,7168–0,7879), 0,6082 (IC del 95 %, 0,5637–0,6528) y 0,5548 (IC 95 %, 0,5102–0,5995), respectivamente. Tanto las ABC de la velocidad z de CA ($p < 0,001$) como las de CA condicional ($p < 0,001$) fueron significativamente más bajas que las ABC del valor z de CA transversales. Al comparar ambas evaluaciones longitudinales, la velocidad z CA se desempeñó significativamente mejor que la CA condicional ($p < 0,001$). En términos de predicción, el ABC para la combinación del valor z CA alta de 32 semanas más velocidad z CA alta de 22 a 32 semanas no fue significativamente mejor que un modelo que solo consideró el tamaño (valor z CA alto).</p>
Hendin et al.	<p>En total, 41261 partos cumplieron los criterios de inclusión durante el período de estudio. De ellos, 1721 (4,17%) tenían PFEnp (estimación de peso fetal no preciso). El mayor IMC materno y el aumento de peso se asociaron positivamente con un PFEnp. La proporción de las gestantes que dieron a luz a una edad gestacional < 39 semanas fue menor en el grupo PFEnp mientras que la proporción de mujeres con $> 41^{0/7}$ semanas de gestación al momento del parto fue mayor en el grupo PFEnp: 12274 (31%) frente a 437 (25,4 %), OR 0,75, IC 95 % (0,67–0,84), $P < 0,001$ y 397 (23,1 %) vs. 7010 (17,7 %), OR 1,39, IC 95 % (1,24–1,56), $P < 0,001$, respectivamente. Un mayor peso fetal al nacimiento se asoció con un PFEnp (media de 3554 gramos frente a 3289), $p < 0,001$. El intervalo desde PFE hasta el parto fue mayor en el grupo PFEnp en comparación con el grupo PFEp (4,6 frente a 3,3, respectivamente, $p < 0,001$). Después del análisis de regresión multivariable, se encontró que la macrosomía se asoció positivamente con PFEnp. Es posible que el peso al nacer represente un factor de confusión en estudios previos que abordaron la estimación del peso fetal.</p>
Hocquette et al.	<p>La población de estudio incluyó 7747 nacidos vivos únicos, incluidos 4427 (57,1%) embarazos que eran de bajo riesgo. La edad gestacional media en la ecografía fue de 22,4 y 32,5 semanas, y el PFE medio fue de 539 y 2008 g. en el segundo y tercer trimestre, respectivamente. La edad gestacional media al nacer fue de 39,7 semanas y el peso medio al nacer fue de 3337 g. Las comparaciones entre la muestra del estudio y los nacimientos excluidos debido a la falta de datos en las ecografías mostraron mayor edad gestacional al nacer (39,7 vs. 39,6 semanas de gestación, $P < 0,001$) en la muestra de análisis. En la muestra total, las proporciones de nacimientos con CA por debajo del percentil 3 y 10 fueron 2,1 y 8,2% en la ecografía del tercer trimestre según tablas de la OMS. Para LF, estos valores fueron 1,8 y 8,2% para el tercer trimestre. En cuanto al PFE en el tercer trimestre, las proporciones por debajo del percentil 3 y 10 fueron 2,1 y 9,4%, y 2,2 y 9,2%, según tablas de la OMS y Hadlock, respectivamente. Las proporciones según las tablas Intergrowth e IG2 estaban por debajo 1,7 y 5,3% para cada umbral, para todos los parámetros y trimestres. Los análisis para la muestra de bajo riesgo produjeron proporciones ligeramente más bajas por debajo del percentil 10 para todos los gráficos; la mayor diferencia de 1,6% fue para LF por debajo del percentil 10 en el tercer trimestre según Hadlock. En cuanto al percentil 90 y 97: en el total de la muestra, estas proporciones para CA fueron 10,2 y 3,4% según las tablas de la OMS en la ecografía del segundo trimestre y 8,9, 2,4 y 9,6, 3,3% según las tablas de la OMS y Hadlock en la ecografía del tercer trimestre, respectivamente. Para la LF, estos valores fueron 9,5 y 3,5% de los fetos según la OMS en el segundo trimestre y 12,7 y 4,7% en el tercer trimestre. En el segundo trimestre del embarazo, las proporciones de fetos con PPE por encima del percentil 90 y 97 según la OMS fueron superiores a lo esperado (16,3 y 5,9%), pero más cercanas a los valores esperados en el tercer trimestre (11,1 y 3,8%). Los gráficos Intergrowth e IG2 arrojaron constantemente proporciones sobre los percentiles 90 y 97 que eran más del doble de lo esperado. Los análisis con embarazos de bajo riesgo arrojaron proporciones ligeramente más bajas (menos del 0,8 % de diferencia). En los análisis de sensibilidad, los resultados fueron muy similares.</p>

Krispin et al.	177 mujeres presentaron un PFEnp (de al menos un 20% de desviación del peso real al nacer) y se consideraron elegibles para formar parte del grupo de estudio; estos se compararon con 354 mujeres que presentaron un PFEp. La mediana de la edad gestacional al momento del parto (37,0 frente a 38,0 semanas, $p < 0,001$), la mediana del peso fetal estimado (2591 frente a 3198 g, $p < 0,001$) y la mediana del peso al nacer (2916 frente a 3158 g, $p = 0,001$) fueron todos menor en el grupo de PFEnp en comparación con el grupo de PFEp. Los parámetros ecográficos diferían significativamente entre los grupos. Las mujeres en el grupo de PFEnp presentaron tasas más altas de presentación podálica (20,34 vs. 9,89 %, $p = 0,001$), tasas más bajas de placenta anterior (44,6 vs. 53,67 %, $p = 0,002$) y posición lateral de la columna (80,23 vs. 88,42%, $p < 0,021$), en comparación con el grupo de PFEp. Se utilizó un análisis multivariado para controlar los posibles factores de confusión; se encontró que la posición no lateral de la columna era un factor de riesgo independiente de imprecisiones en PFE [OR = 2,505, IC del 95 %: 1,159–5,416, $p = 0,0196$]. El PFE fue 607 g menor en el grupo de estudio ($p < 0,001$). Se observaron tasas más altas de ingreso en la unidad de cuidados intensivos neonatales (UCIN) (20,34 frente a 11,02 %, $p = 0,003$) y complicaciones respiratorias neonatales (7,34 frente a 3,39 %, $p = 0,042$) en el subgrupo PFEnp. El análisis multivariante demostró que la imprecisión del PFE fue un factor de riesgo independiente para el ingreso en la UCIN (OR = 2,4, IC del 95 %: 1,034–5,577, $p = 0,041$).
Mohsen et al.	La edad gestacional de los fetos, en el momento de la evaluación, osciló entre 34 y 41 semanas (mediana de 38 semanas), según FUM (fecha de última menstruación) y de 33 a 40 semanas (mediana de 37 semanas) según ultrasonido. La mayoría de los fetos (43 fetos) estaban en posición cefálica, solo siete fetos estaban en presentación podálica o transverso/oblicuo. El peso al nacer (PFA) osciló entre 1970 y 3815 g con un media \pm DE de $3064,6 \pm 454,8$ g y mediana de 3100 g. Siete pacientes (14%) dieron a luz bebés con bajo peso al nacer, es decir < 2500 g, mientras que ninguno de los pacientes presentaron macrosomía (0%). Todos los PFE obtenidos mediante los modelos estudiados se correlacionaron con PFA con un nivel de significación $p = 0,01$. Fórmula M3: Media \pm DE $2831,7 \pm 405,4$, Diferencias con PA (media \pm DE) $232,9 \pm 229,7$, correlación de Pearson R: 0,864, $p50,0005$. Fórmula M6: Media \pm DE $2913,7 \pm 449,8$, Diferencias con PA (media \pm DE) $150,9 \pm 186,6$, correlación de Pearson R: 0,915, $p50,0005$. FórmulaSTT1: Media \pm DE $3354,8 \pm 328$, Diferencias con PA (media \pm DE) $290,2 \pm 338,9$, correlación de Pearson R: 0,669, $p50,0005$. FórmulaSTT2: Media \pm DE $3455,6 \pm 410,6$, Diferencias con PA (media \pm DE) 391 ± 229 57, correlación de Pearson R: 0,865, $p50,0005$. Fórmula Hadlock: Media \pm DE $3121,6 \pm 438,2$, Diferencias con PA (media \pm DE) $\pm 215,4$ NA, correlación de Pearson R: R: 0,884, $p50,0005$. PA fetal: $3064,6 \pm 454,8$ (Todas las unidades en gramos). Los errores sistemáticos para M3 y M6, en el estudio, fue 7.2% y 4.8% respectivamente. Los errores aleatorios, por otro lado, fueron 7,6% y 6,2% para M3 y M6, respectivamente. En comparación, Hadlock tenía un error sistemático más bajo (2,3%) y un comparable error aleatorio (7,7%) en nuestro estudio. De manera similar, los errores porcentuales absolutos de todos los modelos, excepto el modelo M6, fueron significativamente más altos que Hadlock. Los porcentajes de PFE dentro del 5% y 10% de los PFA también fueron más altos mediante la fórmula MH llegando al 48% y 86%, respectivamente; frente al 40% y 78%, respectivamente, para el Modelo 6 de volumen de muslo fraccionado. La fórmula original y modificada de Scioscia revelaron errores sistemáticos y aleatorios mucho más altos que el volumen fraccional de las extremidades y Hadlock.
Rizzo et al.	La macrosomía fetal se vio en un 9,8% (IC 95% 8,1-11,1; 211/2152) de los embarazos, mientras que el peso al nacer > 4000 g se registró en el 7,7% (IC 95% 6,6-8,9; 165/2152) recién nacidos. Los valores z del CC ($-0,02$ vs $0,59$ $p < 0,0001$), DBP ($0,03$ vs $0,57$ $p < 0,0001$), CA ($0,02$ vs $0,86$ $p < 0,0001$), LF ($0,04$ vs $0,39$ $p < 0,0001$) y PFE ($0,09$ vs $1,23$ $p < 0,0001$) fueron mayores en fetos macrosómicos en comparación con fetos con PFA con percentil igual o inferior a 90. Los valores z de FSVU/CA fueron mayores ($-0,03$ vs $0,88$ $p < 0,0001$) en los fetos que presentaban macrosomía, mientras que no hubo diferencia en cualquiera de los otros parámetros Doppler explorados entre los dos grupos. En el análisis de regresión logística multivariable FSVU (aOR1.84 95% IC 1.54-2.14) se asoció de forma independiente con macrosomía. Al explorar la precisión diagnóstica de los diferentes parámetros, el valor z del PFE tuvo un ABC de 0,750 (IC 95% 0,63-0,81) y de 0,801 (IC 95% 0,66-0,86) solo y en asociación con características maternas para la predicción de macrosomía, respectivamente. La adición de FSVU a este modelo mejoró significativamente la predicción de la macrosomía fetal proporcionada por la madre y parámetros ultrasonográficos, con un ABC de 0,892 (IC 95% 0,76-0,91)

Stampalija et al.	Al aplicar las graficas IG, el porcentaje de fetos con CC, CA y LF por debajo del p10 fue 3,9%, 3,6% y 2,3%, y por encima del p90 fueron 29,9%, 32,5% y 46%, respectivamente. Los porcentajes de CC, CA y LF al aplicar los gráficos de crecimiento de la OMS fueron por debajo del p10 6,3%, 7,2% y 5,3%, y por encima del p90 22,8%, 21,3% y 31,9%, respectivamente. En cuanto a las gráficas Nicolini, la mitad de los centros (n=6) las utilizaban; no obstante, solo proporcionan datos sobre DBP, CC y CA por lo que para la LF y el PFE tenían que utilizar otras gráficas. Por esto, la mayor variabilidad se observó en la LF y PFE. En general, las tablas de crecimiento IG identificaron la menor proporción de fetos por debajo del p10 y la mayor proporción de fetos por encima del p90.
Stirneman et al.	Los percentiles estimados de la subpoblación de bajo riesgo utilizando la misma metodología estadística que el estudio IG (es decir, el mismo conjunto de potencias polinómicas fraccionarias) se comparan visualmente con el estándar IG. Los dos conjuntos de percentiles demostraron una clara similitud entre las dos poblaciones, lo que sugirió que el tamaño de la CC fetal en la población francesa era comparable al observado en IG. Esto se confirmó aún más al analizar las diferencias estandarizadas de CC, que permanecieron cerca de 0 y dentro de $\pm 0,5$ cuando se calcularon en diferentes ventanas de edad gestacional a lo largo del embarazo. Para evaluar el impacto de cambiar de gráficos locales a los gráficos IG, también se estimaron los valores z individuales para CA, LF y CC en la muestra en función de los dos gráficos diferentes. En cualquier distribución de valores z, la desviación estándar (SD) esperada era 1; luego se comparó la SD real del uso de IG y gráficos locales. Se encontró que los gráficos de IG dieron una SD de 1,01, 0,99 y 1,00 para CC, CA y LF respectivamente, y los gráficos locales 0,75, 0,75 y 0,78; por lo tanto, los gráficos IG estaban más cerca de la distribución de población esperada que los gráficos locales. Esto se confirmó ya que dos valores de SD de 15 comparaciones estaban por encima del límite pre-especificado para IG frente a 9 valores utilizando los gráficos locales. Igualmente, el estándar IG identificó una proporción de CC por debajo del percentil 3y 10 más cerca del valor esperado que los gráficos locales. Las proporciones de CA por debajo del percentil 3 y 10 según cada gráfico era comparables; sin embargo, las proporciones de CA por encima de los percentiles 90 y 97 mostraron diferencias significativas, ambas alejadas del valor esperado. Finalmente, la proporción de LF estuvo más cerca del valor esperado usando los gráficos locales que el IG, mientras que la proporción de LF más grandes fue sobreestimada en el IG y subestimada en los gráficos locales.
Vila-Candel et al.	La edad gestacional media al momento del parto fue $39,14 \pm 1,5$ semanas (rango 35-40) y el peso medio real al nacer de la población de estudio fue de $3254 \pm 448,4$ g. Cuarenta y nueve (4,0 %) tuvieron un peso al nacer <2500 g, 1118 (91,3 %) pesaron entre 2500 y 3999 g y 57 (4,7 %) pesaron >4000 g. La diferencia del PFEh con el PFA fue de $187,4g \pm 361,4$ (IC 95%: 167,1-207,7), mientras que, para el PFEr, la diferencia fue de $-0,68g \pm 315,6$ (IC 95%: -18,3- 17,1), siendo las diferencias estadísticamente significativas respectivamente (prueba t pareada $p<0,001$; $p<0,001$). El error de valor relativo medio para PFEr fue menor que para PFEh ($0,97 \% \pm 10,1$ vs. $6,67 \% \pm 12,1$), y las diferencias fueron estadísticamente significativas mediante la prueba t pareada ($5,7 \% \pm 2,0$, IC del 95 %: 5,4- 6,0, $p<0,001$). La proporción de pesos estimados por ecografía dentro de $<10\%$ del peso real al nacer fue significativamente menor que la obtenida con el método combinado (65,3% vs 82,7%) siendo la diferencia significativa (Test de McNemar, $p<0,001$). Hubo diferencias estadísticamente significativas entre los métodos de estimación del peso por categorías de peso al nacer. Para los recién nacidos con <2500 g de peso al nacer, la proporción de estimaciones dentro de $<10\%$ del peso real al nacer para el PFEr fue significativamente mayor que para el PFEh (20,4 % frente a 16,3 %, $p=0,005$). Para los fetos con peso normal al nacer (2500-3999 g), el método combinado fue significativamente mayor que el PFAh (84,5 vs. 65,7%, $p<0,001$). Por último, en los recién nacidos macrosómicos (>4000 g), la proporción de estimaciones dentro de $<10\%$ del peso real al nacer para el método combinado fue menor que para el método de ultrasonido (100% frente a 98,2 %), sin embargo, las diferencias no fueron estadísticamente significativas ($p=0,238$), lo que sugiere que tanto la ecografía como el método clínico son igualmente precisos para predecir macrosomía fetal.

Tabla 4: Principales resultados de los artículos incluidos en la revisión.

5.2. RIESGO DE SESGO DE LOS ESTUDIOS INCLUIDOS

La evaluación de la calidad metodológica de cada uno de los estudios incluido en la revisión se detalla en el ANEXO 2. En la Tabla 5 y en la Figura 2 se resumen los resultados de la evaluación QUADAS-2.

Tres artículos [38,45,46] resultaron tener un alto riesgo de sesgo en la prueba índice y la de referencia, sobre todo, por interpretar los resultados conociendo los de la prueba contraria. Dos estudios [39,41] presentaron un alto riesgo de sesgo en la selección de los individuos por ser estudios de casos y controles. Y tan solo un estudio [45] obtuvo un alto riesgo de sesgo en flujo y tiempo, por no describir claramente la prueba de referencia.

En cuanto a la preocupación sobre la aplicabilidad de los resultados, fue baja para todos los artículos en la selección de los pacientes y la prueba índice. En cambio, la preocupación sobre la aplicabilidad de la prueba de referencia fue alta en un artículo [45] por no evaluar ninguna y solo estudiar la prueba índice.

Estudios incluidos	Probabilidad de sesgos				Preocupación sobre la aplicabilidad de los resultados		
	Selección de los individuos	Prueba índice	Prueba de referencia	Flujo y tiempos	Selección de los pacientes	Prueba índice	Prueba de referencia
Caradeux et al.	Baja	Alta	Alta	Baja	Baja	Baja	Baja
Hendin et al.	Alta	Incierta	Incierta	Baja	Baja	Baja	Incierta
Hocquette et al.	Baja	Incierta	Incierta	Baja	Baja	Baja	Baja
Krispin et al.	Alta	Incierta	Incierta	Incierta	Baja	Baja	Incierta
Mohsen et al.	Baja	Baja	Baja	Baja	Baja	Baja	Baja
Rizzo et al.	Baja	Incierta	Incierta	Baja	Baja	Baja	Incierta
Stampalija et al.	Baja	Alta	Alta	Alta	Baja	Baja	Alta
Stirnemann et al.	Baja	Alta	Alta	Baja	Baja	Baja	Baja
Vila-Candel et al.	Baja	Incierta	Baja	Baja	Baja	Baja	Baja

Tabla 5: Riesgo de sesgo y aplicabilidad de los estudios incluidos en la revisión según QUADAS-2.

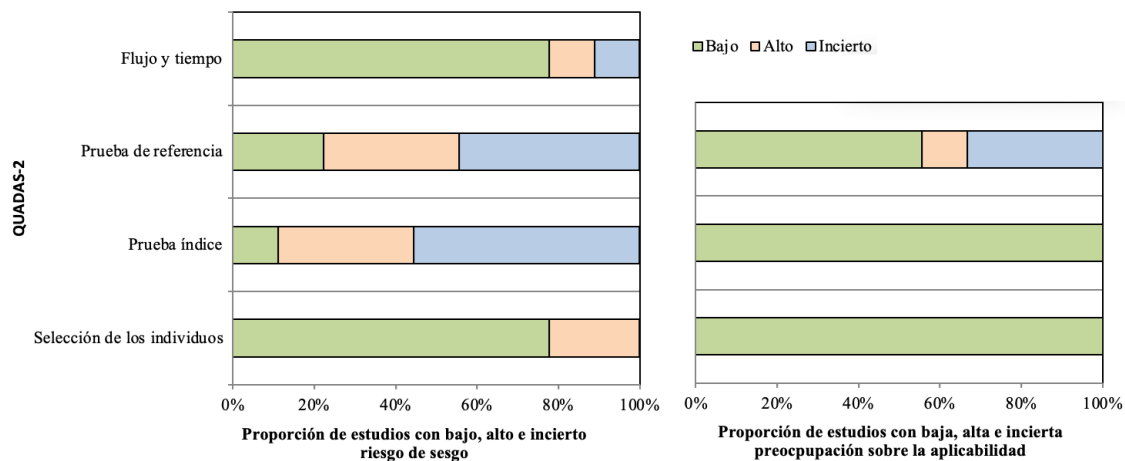


Figura 2: Gráficas según QUADAS-2 de proporción de riesgo de sesgo y preocupación sobre la aplicabilidad de los estudios incluidos en la revisión.

6. DISCUSIÓN

Nuestro estudio muestra que el método ecográfico Hadlock es el que mejor estima el peso fetal. Resulta evidente la gran heterogeneidad de métodos que existen para estimarlo, y que distintas fórmulas pueden informar diferentes percentiles para la misma medición fetal; sin un consenso, esto puede conllevar a errores en la práctica clínica con todas las repercusiones materno-fetales que podría implicar. Las participantes de esta revisión son mujeres embarazadas de bajo riesgo procedentes de diferentes países del área mediterránea.

En general, UpToDate, las guías de práctica clínica y las instituciones obstétricas de referencia a nivel internacional [26,29,31,33] y nacional [34–37] recomiendan calcular el PFE mediante el método Hadlock; sin embargo, la fecha de última actualización de algunas de ellas no es reciente [29,31,35,36]. Existe discrepancia entre las dos guías de ACOG ya que en una [29] se recomienda Hadlock para estimar el peso fetal, pero en otra [28] se afirma que ninguna fórmula basada en biometría por ultrasonido funciona significativamente mejor que otras para la detección de la macrosomía. Las guías NICE [25] tan solo recomiendan el uso de la biometría fetal sin especificar qué fórmula se debería de usar. ISUOG [32] es la única guía que recomienda el uso de las tablas de crecimiento prescriptivas IG. Finalmente, no se han encontrado guías ni publicaciones sobre el manejo del embarazo normal ni de fetos macrosómicos del RCOG; tan solo se

han encontrado guías sobre el manejo de la distocia de hombros como complicación de los fetos GEG y de los fetos PEG [30,31].

En cuanto a los estudios incluidos en la revisión, todos coincidieron en que la ecografía 2D es el método de elección para calcular el PFE; solo en un artículo se tuvo en cuenta la ecografía 3D [42] y en él se afirmó que la ecografía 3D es un método prometedor y con resultados similares en cuanto a precisión que la 2D. Este resultado coincide con otros estudios [50–53] en los que se determinó que la ecografía 3D estima el peso fetal dentro del 10% del peso al nacer. No obstante, Tuuli et al.[54] compararon la ecografía 3D con el método Hadlock y concluyeron que este último proporcionaba un PFE más preciso. Si bien, estudios más recientes concluyeron que el PFE con método 3D se aproxima más al peso fetal al nacer [55–57].

En nuestra revisión, existe una gran variabilidad entre los resultados debido a que los artículos incluidos utilizaron diferentes metodologías para estimar el peso fetal. La gran mayoría de los artículos incluidos compararon Hadlock con otros métodos [39–44]. Sobre ellos, se puede concluir que en mujeres sanas del área mediterránea el PFE mediante Hadlock es más preciso en comparación con los otros; además, si se tienen en cuenta características como la altura materna o el sexo fetal aumenta la fiabilidad del PFE. Estos resultados coinciden con la literatura [33,35,36,58–61] en la que se afirma que la fórmula ecográfica bidimensional Hadlock predice el percentil de peso al nacer con mayor precisión y fiabilidad.

Stirnemann et al. [46] concluyeron que los gráficos IG son más precisos que los gráficos de crecimiento franceses; pero lo contrario afirmó un artículo más reciente [62] en el cual, se recomienda el uso de las gráficas locales francesas para estimar el peso fetal; sin embargo, ambos artículos no tuvieron en cuenta el método Hadlock, por lo que sus conclusiones podrían estar limitadas. Hocquette et al.[40] y Stampalija et al.[45] coincidieron en que las gráficas IG sobreestiman el peso de fetos macrosómicos pero discreparon en cuanto a las tablas de la OMS: el primero afirmó que son más precisas en detectar fetos GEG y el segundo que los sobreestiman. Además, Hocquette et al.[40] es el único de los estudios incluidos en los que se afirma que los gráficos de la OMS son superiores al método Hadlock para estimar el peso fetal en fetos macrosómicos; estos

resultados podrían explicarse por las limitaciones del estudio (pérdidas de seguimiento y de medidas ecográficas).

Por otra parte, solo un estudio incluido en nuestra revisión compara la estimación del peso fetal según una valoración ecográfica transversal con mediciones seriadas longitudinales, afirmando que la primera es superior en cuanto a precisión [38] y coincidiendo con la literatura: la velocidad de crecimiento registrada mediante mediciones seriadas no proporciona información adicional para estimar el peso fetal, en comparación con el último PFE [63–65].

Con respecto a las limitaciones de nuestra revisión sistemática, la comparación de los artículos resultó complicada por utilizar, los autores, diferentes metodologías para obtener y analizar sus datos. Además, no se encontró ningún artículo que evaluase en la misma población todos los métodos disponibles para estimar el peso fetal. Por otra parte, la razón por la que gran parte de artículos fueron excluidos en el segundo cribado, fue que las participantes no procedían del área mediterránea, lo que pudo conllevar a la pérdida de posibles resultados valiosos para nuestra revisión. Pese a que los estudios incluyeron embarazadas en el tercer trimestre de gestación, no todas las mujeres presentaban la misma edad gestacional. Cinco [38,39,41,45,46] de los estudios incluidos tuvieron riesgo de sesgo en algún ítem de los evaluados y en algunos estudios [38,39,43] hubo factores que no fueron examinados, como la experiencia de los ecografistas. Finalmente, solo un investigador realizó la búsqueda y el cribado de los artículos, lo que pudo sesgar los resultados de la revisión y disminuir la validez interna de nuestro estudio.

Con relación a las fortalezas, destaca el gran tamaño muestral (64546 participantes incluidas) y la baja preocupación sobre la aplicabilidad de los resultados, sobre todo en el contexto de la prueba índice y la selección de los individuos.

Con todo esto, podemos realizarnos nuevas preguntas para continuar con la investigación: ¿Aumenta la precisión del PFE si al método Hadlock y las características maternas y fetales, se añaden otros parámetros de ecografía 2D? ¿Y si a estos se añaden parámetros de ecografía 3D? ¿Se podría desarrollar un método para estimar el peso fetal que no dependa de las medidas tomadas por el ecografista? Probablemente, se deberían realizar

más estudios para encontrar un método más preciso que los actuales para estimar el peso fetal.

7. CONCLUSIONES

Las conclusiones de nuestra revisión sistemática son:

- El mejor método ecográfico para estimar el peso fetal, tanto en fetos con peso AEG y GEG en mujeres sanas del área mediterránea, es el método Hadlock. Añadir a este método las características maternas y fetales (altura materna o sexo fetal) aumentan la precisión del PFE.
- El PFE mediante una sola estimación ecográfica cercana de la fecha probable de parto es más rentable que la estimación ecográfica en serie.
- Los métodos IG y OMS son imprecisos para estimar los pesos extremos y pueden conllevar a errores e implicaciones negativas para la práctica clínica.
- La adición de parámetros ecográficos 3D al método Hadlock y las características maternas y fetales puede ser prometedor para estimar el peso fetal con más fiabilidad; no obstante, se requiere más investigación en este campo.

8. DECLARACIONES DE INTERÉS

Los participantes declaran no tener conflicto de intereses ni fuentes de financiación.

9. BIBLIOGRAFÍA

- [1] Mongelli M, Gardosi J. Longitudinal study of fetal growth in subgroups of a low-risk population: Fetal growth in a low-risk population. *Ultrasound Obstet Gynecol* 1995;6:340–4. <https://doi.org/10.1046/j.1469-0705.1995.06050340.x>.
- [2] Battaglia FC, Lubchenco LO. A practical classification of newborn infants by weight and gestational age. *J Pediatr* 1967;71:159–63. [https://doi.org/10.1016/S0022-3476\(67\)80066-0](https://doi.org/10.1016/S0022-3476(67)80066-0).
- [3] Jaques S Abramowicz, Jennifer T Ahn. Fetal macrosomia. UpToDate 2021.
- [4] Divon MY. Fetal growth restriction: Screening and diagnosis. UpToDate 2021.
- [5] Deter RL, Lee W, Yeo L, Erez O, Ramamurthy U, Naik M, et al. Individualized growth assessment: conceptual framework and practical implementation for the evaluation of fetal growth and neonatal growth outcome. *Am J Obstet Gynecol* 2018;218:S656–78. <https://doi.org/10.1016/j.ajog.2017.12.210>.
- [6] Coutinho CM, Melchiorre K, Thilaganathan B. Stillbirth at term: Does size really matter? *Int J Gynecol Obstet* 2020;150:299–305. <https://doi.org/10.1002/ijgo.13229>.
- [7] Hadlock FP, Harrist RB, Sharman RS, Deter RL, Park SK. Estimation of fetal weight with the use of head, body, and femur measurements—A prospective study. *Am J Obstet Gynecol* 1985;151:333–7. [https://doi.org/10.1016/0002-9378\(85\)90298-4](https://doi.org/10.1016/0002-9378(85)90298-4).
- [8] Hadlock FP, Harrist RB, Martinez-Poyer J. In utero analysis of fetal growth: a sonographic weight standard. *Radiology* 1991;181:129–33. <https://doi.org/10.1148/radiology.181.1.1887021>.
- [9] Hammami A, Mazer Zumaeta A, Syngelaki A, Akolekar R, Nicolaides KH. Ultrasonographic estimation of fetal weight: development of new model and assessment of performance of previous models. *Ultrasound Obstet Gynecol* 2018;52:35–43. <https://doi.org/10.1002/uog.19066>.
- [10] Aviram A, Yogev Y, Ashwal E, Hirsch L, Danon D, Hadar E, et al. Different formulas, different thresholds and different performance—the prediction of macrosomia by ultrasound. *J Perinatol* 2017;37:1285–91. <https://doi.org/10.1038/jp.2017.134>.
- [11] Papageorghiou AT, Ohuma EO, Altman DG, Todros T, Ismail LC, Lambert A, et al. International standards for fetal growth based on serial ultrasound measurements: the Fetal Growth Longitudinal Study of the INTERGROWTH-21st Project. *The Lancet* 2014;384:869–79. [https://doi.org/10.1016/S0140-6736\(14\)61490-2](https://doi.org/10.1016/S0140-6736(14)61490-2).

- [12] Stirnemann J, Villar J, Salomon LJ, Ohuma E, Ruyan P, Altman DG, et al. International estimated fetal weight standards of the INTERGROWTH-21st Project: International estimated fetal weight standards. *Ultrasound Obstet Gynecol* 2017;49:478–86. <https://doi.org/10.1002/uog.17347>.
- [13] Kiserud T, Benachi A, Hecher K, Perez RG, Carvalho J, Piaggio G, et al. The World Health Organization fetal growth charts: concept, findings, interpretation, and application. *Am J Obstet Gynecol* 2018;218:S619–29. <https://doi.org/10.1016/j.ajog.2017.12.010>.
- [14] Sovio U, Smith GCS. Comparison of estimated fetal weight percentiles near term for predicting extremes of birthweight percentile. *Am J Obstet Gynecol* 2021;224:292.e1–292.e19. <https://doi.org/10.1016/j.ajog.2020.08.054>.
- [15] Nwabuobi C, Odibo L, Camisasca-Lopina H, Leavitt K, Tuuli M, Odibo AO. Comparing INTERGROWTH-21st Century and Hadlock growth standards to predict small for gestational age and short-term neonatal outcomes. *J Matern-Fetal Neonatal Med Off J Eur Assoc Perinat Med Fed Asia Ocean Perinat Soc Int Soc Perinat Obstet* 2020;33:1906–12. <https://doi.org/10.1080/14767058.2018.1533945>.
- [16] Zhu C, Ren Y-Y, Wu J-N, Zhou Q-J. A Comparison of Prediction of Adverse Perinatal Outcomes between Hadlock and INTERGROWTH-21st Standards at the Third Trimester. *BioMed Res Int* 2019;2019:7698038. <https://doi.org/10.1155/2019/7698038>.
- [17] Kabiri D, Romero R, Gudicha DW, Hernandez-Andrade E, Pacora P, Benschalom-Tirosh N, et al. Prediction of adverse perinatal outcome by fetal biometry: comparison of customized and population-based standards. *Ultrasound Obstet Gynecol* 2020;55:177–88. <https://doi.org/10.1002/uog.20299>.
- [18] Hocquette A, Durox M, Wood R, Klungsoyr K, Szamotulska K, Berrut S, et al. International versus national growth charts for identifying small and large-for-gestational age newborns: A population-based study in 15 European countries. *Lancet Reg Health - Eur* 2021;8:100167. <https://doi.org/10.1016/j.lanepe.2021.100167>.
- [19] Zeitlin J, Vayssière C, Ego A, Goffinet F. More validation is needed before widespread adoption of INTERGROWTH-21st fetal growth reference standards in France: Correspondence. *Ultrasound Obstet Gynecol* 2017;49:547–8. <https://doi.org/10.1002/uog.17423>.
- [20] Grivell RM, Wong L, Bhatia V. Regimens of fetal surveillance for impaired fetal growth. *Cochrane Database Syst Rev* 2009. <https://doi.org/10.1002/14651858.CD007113.pub2>.

- [21] Bricker L, Medley N, Pratt JJ. Routine ultrasound in late pregnancy (after 24 weeks' gestation). *Cochrane Database Syst Rev* 2015. <https://doi.org/10.1002/14651858.CD001451.pub4>.
- [22] Almenar Agustí L. Estimación de peso fetal; valoración de diferentes métodos de cálculo [Tesis doctoral inédita]. Servicio de Obstetricia y Ginecología Hospital de La Plana, Vila-Real, n.d.
- [23] Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. Declaración PRISMA 2020: una guía actualizada para la publicación de revisiones sistemáticas. *Rev Esp Cardiol* 2021;74:790–9. <https://doi.org/10.1016/j.recesp.2021.06.016>.
- [24] Andrew P MacKenzie, Edmund F Funai., Courtney D Stephenson. Prenatal assessment of gestational age, date of delivery, and fetal weight. UpToDate 2021.
- [25] National Institute for Health and Clinical Excellence. Antenatal care. Monitoring fetal growth. NICE 2021:1–95.
- [26] New Zealand GUIDELINES GROUP, Ministry of Health. New Zealand obstetric ultrasound guidelines. 2019.
- [27] Guía de Práctica Clínica de Atención en el Embarazo y Puerperio. GuíaSalud n.d. <https://portal.guiasalud.es/gpc/embarazo-puerperio/> (accessed March 7, 2022).
- [28] American Collage of Obstetricians and Gynecologists. Macrosomia. *Obstet Gynecol* 2020;135:18–35. <https://doi.org/10.1097/AOG.0000000000003606>.
- [29] American Academy of Pediatrics, American College of Obstetricians and Gynecologists, editors. Guidelines for perinatal care. Eighth edition. Elk Grove Village, IL : Washington, DC: American Academy of Pediatrics ; The American College of Obstetricians and Gynecologists; 2017.
- [30] Crofts J, I Montague, TJ Draycott, C Winter, R Fox. Shoulder Dystocia (Green-top Guideline No. 42). *R Coll Obstet Amp Gynaecol* n.d. <https://www.rcog.org.uk/en/guidelines-research-services/guidelines/gtg42/> (accessed March 14, 2022).
- [31] SC Robson, WL Martin, RK Morris. Small-for-Gestational-Age Fetus, Investigation and Management (Green-top Guideline No. 31). *R Coll Obstet Amp Gynaecol* n.d. <https://www.rcog.org.uk/en/guidelines-research-services/guidelines/gtg31/> (accessed March 14, 2022).
- [32] Salomon L j., Alfirevic Z, Da Silva Costa F, Deter R l., Figueras F, Ghi T, et al. ISUOG Practice Guidelines: ultrasound assessment of fetal biometry and growth.

- Ultrasound Obstet Gynecol 2019;53:715–23. <https://doi.org/10.1002/uog.20272>.
- [33] Visser GHA, Nicholson WK, Barnea ER, Ramasauskaite D, Nassar AH, FIGO Safe Motherhood, Newborn Health Committee. FIGO position paper on reference charts for fetal growth and size at birth: Which one to use? *Int J Gynaecol Obstet Off Organ Int Fed Gynaecol Obstet* 2021;152:148–51. <https://doi.org/10.1002/ijgo.13500>.
- [34] Ramírez JA, Navarro BP, Alvarado EA, Bueno JAS, Ruiz H, González CB. Guía de la exploración ecográfica del III trimestre 2020 n.d.:42.
- [35] Bennasar M, Borobio V, Puerto, B. Screening ecográfico fetal. Hospital Clínic, Hospital Sant Joan de Déu, Universitat de Barcelona 2016.
- [36] Molinet C, Masoller N, Boguñà J. Macrosomía. Hospital Clínic, Hospital Sant Joan de Déu, Universitat de Barcelona 2018.
- [37] Figueras F, Gómez L, Eizarch E, Paules C, Mazarico A, Pérez M, et al. Defectos del crecimiento fetal 2019.
- [38] Caradeux J, Eixarch E, Mazarico E, Basuki TR, Gratacós E, Figueras F. Second-to Third-Trimester Longitudinal Growth Assessment for the Prediction of Largeness for Gestational Age and Macrosomia in an Unselected Population. *Fetal Diagn Ther* 2018;43:284–90. <https://doi.org/10.1159/000477460>.
- [39] Hendin N, Levin G, Tsur A, Ilan H, Rottenstreich A, Meyer R. Factors Associated with More Than 500 Grams Inaccuracy in Sonographic Fetal Weight Estimation. *Isr Med Assoc J IMAJ* 2021;23:43–7.
- [40] Hocquette A, Zeitlin J, Heude B, Ego A, Charles M-A, Monier I. World Health Organization fetal growth charts applied in a French birth cohort. *J Gynecol Obstet Hum Reprod* 2022;51:102308. <https://doi.org/10.1016/j.jogoh.2021.102308>.
- [41] Krispin E, Dreyfuss E, Fischer O, Wiznitzer A, Hadar E, Bardin R. Significant deviations in sonographic fetal weight estimation: causes and implications. *Arch Gynecol Obstet* 2020;302:1339–44. <https://doi.org/10.1007/s00404-020-05732-x>.
- [42] Mohsen LA, Amin MF. 3D and 2D ultrasound-based fetal weight estimation: a single center experience. *J Matern Fetal Neonatal Med* 2017;30:818–25. <https://doi.org/10.1080/14767058.2016.1187125>.
- [43] Rizzo G, Mappa I, Bitsadze V, Khizroeva J, Makatsarya A, D'Antonio F. The added value of umbilical vein flow in predicting fetal macrosomia at 36 weeks of gestation: A prospective cohort study. *Acta Obstet Gynecol Scand* 2021;100:900–7. <https://doi.org/10.1111/aogs.14047>.
- [44] Vila-Candel R, Soriano-Vidal FJ, Castro-Sánchez E. Third trimester ultrasound

scan combined with a clinical method for accurate birthweight prediction at term: a cohort study in Spain. *Rev Colomb Obstet Ginecol* 2019;70:27. <https://doi.org/10.18597/rcog.3201>.

[45] Stampalija T, Ghi T, Rosolen V, Rizzo G, Ferrazzi EM, Prefumo F, et al. Current use and performance of the different fetal growth charts in the Italian population. *Eur J Obstet Gynecol Reprod Biol* 2020;252:323–9. <https://doi.org/10.1016/j.ejogrb.2020.06.059>.

[46] Stirnemann JJ, Fries N, Bessis R, Fontanges M, Mangione R, Salomon LJ. Implementing the INTERGROWTH-21st fetal growth standards in France: a ‘flash study’ of the College Français d’Echographie Foetale (CFEF): Implementing INTERGROWTH-21st fetal size charts in France. *Ultrasound Obstet Gynecol* 2017;49:487–92. <https://doi.org/10.1002/uog.17223>.

[47] Whiting PF, Rutjes AWS, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011;155:529–36. <https://doi.org/10.7326/0003-4819-155-8-201110180-00009>.

[48] Bristol U of. QUADAS-2 n.d. <https://www.bristol.ac.uk/population-health-sciences/projects/quadas/quadas-2/> (accessed March 3, 2022).

[49] Ciapponi A. QUADAS-2: instrumento para la evaluación de la calidad de estudios de precisión diagnóstica. *Evid Actual En Práctica Ambulatoria* 2015;18. <https://doi.org/10.51987/evidencia.v18i1.6341>.

[50] Gibson KS, Stetzer B, Catalano PM, Myers SA. Comparison of 2- and 3-Dimensional Sonography for Estimation of Birth Weight and Neonatal Adiposity in the Setting of Suspected Fetal Macrosomia. *J Ultrasound Med Off J Am Inst Ultrasound Med* 2016;35:1123–9. <https://doi.org/10.7863/ultra.15.06106>.

[51] Lee W, Comstock CH, Kirk JS, Smith RS, Monck JW, Deenadayalu R, et al. Birthweight prediction by three-dimensional ultrasonographic volumes of the fetal thigh and abdomen. *J Ultrasound Med Off J Am Inst Ultrasound Med* 1997;16:799–805. <https://doi.org/10.7863/jum.1997.16.12.799>.

[52] Liang RI, Chang FM, Yao BL, Chang CH, Yu CH, Ko HC. Predicting birth weight by fetal upper-arm volume with use of three-dimensional ultrasonography. *Am J Obstet Gynecol* 1997;177:632–8. [https://doi.org/10.1016/s0002-9378\(97\)70157-1](https://doi.org/10.1016/s0002-9378(97)70157-1).

[53] Song TB, Moore TR, Lee JI, Kim YH, Kim EK. Fetal weight prediction by thigh volume measurement with three-dimensional ultrasonography. *Obstet Gynecol*

2000;96:157–61. [https://doi.org/10.1016/s0029-7844\(00\)00884-x](https://doi.org/10.1016/s0029-7844(00)00884-x).

[54] Tuuli MG, Kapalka K, Macones GA, Cahill AG. Three-Versus Two-Dimensional Sonographic Biometry for Predicting Birth Weight and Macrosomia in Diabetic Pregnancies. *J Ultrasound Med Off J Am Inst Ultrasound Med* 2016;35:1925–30. <https://doi.org/10.7863/ultra.15.08032>.

[55] Sharma KA, Das D, Dadhwal V, Deka D, Singhal S, Vanamail P. Two-dimensional fetal biometry versus three-dimensional fractional thigh volume for ultrasonographic prediction of birthweight. *Int J Gynaecol Obstet Off Organ Int Fed Gynaecol Obstet* 2019;145:47–53. <https://doi.org/10.1002/ijgo.12770>.

[56] Lee W, Mack LM, Sangi-Haghpeykar H, Gandhi R, Wu Q, Kang L, et al. Fetal Weight Estimation Using Automated Fractional Limb Volume With 2-Dimensional Size Parameters: A Multicenter Study. *J Ultrasound Med Off J Am Inst Ultrasound Med* 2020;39:1317–24. <https://doi.org/10.1002/jum.15224>.

[57] Wu X, Niu Z, Xu Z, Jiang Y, Zhang Y, Meng H, et al. Fetal weight estimation by automated three-dimensional limb volume model in late third trimester compared to two-dimensional model: a cross-sectional prospective observational study. *BMC Pregnancy Childbirth* 2021;21:365. <https://doi.org/10.1186/s12884-021-03830-5>.

[58] Blue NR, Savabi M, Beddow ME, Katukuri VR, Fritts CM, Izquierdo LA, et al. The Hadlock Method Is Superior to Newer Methods for the Prediction of the Birth Weight Percentile. *J Ultrasound Med Off J Am Inst Ultrasound Med* 2019;38:587–96. <https://doi.org/10.1002/jum.14725>.

[59] Marien M, Perron S, Bergeron A-M, Singbo N, Demers S. Comparison of the Accuracy of INTERGROWTH-21 and Hadlock Ultrasound Formulae for Fetal Weight Prediction. *J Obstet Gynaecol Can JOGC J Obstet Gynecol Can JOGC* 2021;43:1254–9. <https://doi.org/10.1016/j.jogc.2021.03.015>.

[60] Di Lorenzo G, Monasta L, Ceccarello M, Cecotti V, D'Ottavio G. Third trimester abdominal circumference, estimated fetal weight and uterine artery doppler for the identification of newborns small and large for gestational age. *Eur J Obstet Gynecol Reprod Biol* 2013;166:133–8. <https://doi.org/10.1016/j.ejogrb.2012.10.010>.

[61] Frick AP, Syngelaki A, Zheng M, Poon LC, Nicolaides KH. Prediction of large-for-gestational-age neonates: screening by maternal factors and biomarkers in the three trimesters of pregnancy. *Ultrasound Obstet Gynecol Off J Int Soc Ultrasound Obstet Gynecol* 2016;47:332–9. <https://doi.org/10.1002/uog.15780>.

[62] Verspyck E, Senat MV, Subtil D, Vayssiere C, Comité de pilotage du groupe des

RPC “Courbes de croissance” du CNGOF. [Which fetal growth curve reference should be now chosen for our country?]. *Gynecol Obstet Fertil Senol* 2021;49:801–2. <https://doi.org/10.1016/j.gofs.2021.05.001>.

[63] Larsen T, Greisen G, Petersen S. Prediction of birth weight by ultrasound-estimated fetal weight: a comparison between single and repeated estimates. *Eur J Obstet Gynecol Reprod Biol* 1995;60:37–40. [https://doi.org/10.1016/0028-2243\(95\)02079-9](https://doi.org/10.1016/0028-2243(95)02079-9).

[64] Hedriana HL, Moore TR. A comparison of single versus multiple growth ultrasonographic examinations in predicting birth weight. *Am J Obstet Gynecol* 1994;170:1600–4; discussion 1604-1606. [https://doi.org/10.1016/s0002-9378\(94\)70329-9](https://doi.org/10.1016/s0002-9378(94)70329-9).

[65] Tarca AL, Hernandez-Andrade E, Ahn H, Garcia M, Xu Z, Korzeniewski SJ, et al. Single and Serial Fetal Biometry to Detect Preterm and Term Small- and Large-for-Gestational-Age Neonates: A Longitudinal Cohort Study. *PloS One* 2016;11:e0164161. <https://doi.org/10.1371/journal.pone.0164161>.

10.ANEXOS

ANEXO 1: Artículos excluidos en el segundo cribado y razón de exclusión:

Artículos excluidos	Razón de exclusión
Agrawal C, Agrawal KK, Gandhi S. Assessment of fetal growth using the ratio of the transverse cerebellar diameter to abdominal circumference. <i>Int J Gynecol Obstet</i> [Internet]. 2016;135(1):33–7. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84977594202&doi=10.1016%2Fj.ijgo.2016.03.025&partnerID=40&md5=2f7f50a1308148843f7b2955187172de	Selección de embarazadas con fetos PEG
Akolekar R, Ciobanu A, Zingler E, Syngelaki A, Nicolaides KH. Routine assessment of cerebroplacental ratio at 35-37 weeks' gestation in the prediction of adverse perinatal outcome. <i>Am J Obstet Gynecol</i> . 2019 Jul;221(1):65.e1-65.e18.	Participantes del estudio de Reino Unido
Andreasen LA, Tabor A, Nørgaard LN, Ringsted C, Sandager P, Rosthøj S, et al. Multicenter randomized trial exploring effects of simulation-based ultrasound training on obstetricians' diagnostic accuracy: value for experienced operators. <i>Ultrasound Obstet Gynecol Off J Int Soc Ultrasound Obstet Gynecol</i> . 2020 Apr;55(4):523–9.	Participantes del estudio de Dinamarca
Ashimi Balogun O, Sibai BM, Pedroza C, Blackwell SC, Barrett TL, Chauhan SP. Serial Third- Trimester Ultrasonography Compared With Routine Care in Uncomplicated Pregnancies: A Randomized Controlled Trial. <i>Obstet Gynecol</i> . 2018 Dec;132(6):1358–67.	Tipo de artículo: ensayo clínico aleatorizado
Azagidi A, Ibitoye B, Makinde O, Idowu B, Aderibigbe A. Fetal gestational age determination using ultrasound placental thickness. <i>J Med Ultrasound</i> [Internet]. 2020;28(1):17–23. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85082763353&doi=10.4103%2FJMU.JMU_127_18&partnerID=40&md5=261716e1be0d6b350cc3f2820d590e48	Participantes del estudio de Nigeria
Blitz MJ, Rochelson B, Stork LB, Augustine S, Greenberg M, Sison CP, et al. Effect of Maternal Body Mass Index and Amniotic Fluid Index on the Accuracy of Sonographic Estimation of Fetal Weight in Late Gestation. <i>Am J Perinatol</i> . 2018 Nov;35(13):1235–40.	Participantes del estudio de Estados Unidos
Blue NR, Savabi M, Beddow ME, Katukuri VR, Fritts CM, Izquierdo LA, et al. The Hadlock Method Is Superior to Newer Methods for the Prediction of the Birth Weight Percentile. <i>J ultrasound. Med Off J Am Inst Ultrasound Med</i> . 2019 Mar;38(3):587–96.	Participantes del estudio de Estados Unidos
Boutin A, Joseph KS. Contextual performance of diagnostics for fetal surveillance. <i>BJOG An Int J Obstet Gynaecol</i> [Internet]. 2021;128(2):237. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85086447889&doi=10.1111%2F1471-0528.16334&partnerID=40&md5=b131be190a53c863c431187b919d65c2	Tipo de artículo: comentario
Cheng YKY, Leung TY, Lao TTH, Chan YM, Sahota DS. Impact of replacing Chinese ethnicity- specific fetal biometry charts with the INTERGROWTH-21st standard. <i>BJOG An Int J Obstet Gynaecol</i> [Internet]. 2016 Sep;123:48–55. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84987624259&doi=10.1111%2F1471-0528.14008&partnerID=40&md5=87a66e7fe170358fdcacd46b7ce65266	Participantes del estudio de China
Ciobanu A, Anthoulakis C, Syngelaki A, Akolekar R, Nicolaides KH. Prediction of small-for- gestational-age neonates at 35-37 weeks' gestation: contribution of maternal factors and growth velocity between 32 and 36 weeks. <i>Ultrasound Obstet Gynecol Off J Int Soc Ultrasound Obstet Gynecol</i> . 2019 May;53(5):630–7.	Selección de embarazadas con fetos PEG

Ciobanu A, Formuso C, Syngelaki A, Akolekar R, Nicolaides KH. Prediction of small-for-gestational-age neonates at 35-37 weeks' gestation: contribution of maternal factors and growth velocity between 20 and 36 weeks. <i>Ultrasound Obstet Gynecol Off J Int Soc Ultrasound Obstet Gynecol</i> . 2019 Apr;53(4):488–95.	Selección de embarazadas con fetos PEG
Ciobanu A, Khan N, Syngelaki A, Akolekar R, Nicolaides KH. Routine ultrasound at 32 vs 36 weeks' gestation: prediction of small-for-gestational-age neonates. <i>Ultrasound Obstet Gynecol Off J Int Soc Ultrasound Obstet Gynecol</i> . 2019 Jun;53(6):761–8.	Selección de embarazadas con fetos PEG
Coutinho CM, Melchiorre K, Thilaganathan B. Stillbirth at term: Does size really matter? <i>Int J Gynaecol Obstet Off organ Int Fed Gynaecol Obstet</i> . 2020 Sep;150(3):299–305.	Tipo de artículo: revisión
Deter RL, Lee W, Dicker P, Tully EC, Cody F, Malone FD, et al. Third-trimester growth diversity in small fetuses classified as appropriate-for-gestational age or small-for-gestational age at birth. <i>Ultrasound Obstet Gynecol Off J Int Soc Ultrasound Obstet Gynecol</i> . 2021 Dec;58(6):882–91.	No evalúa la estimación del peso fetal mediante ecografía
Deter RL, Lee W, Kingdom JCP, Romero R. Fetal growth pathology score: a novel ultrasound parameter for individualized assessment of third trimester growth abnormalities. <i>J Matern neonatal Med Off J Eur Assoc Perinat Med Fed Asia Ocean Perinat Soc Int Soc Perinat Obstet</i> . 2018 Apr;31(7):866–76.	Selección de embarazadas con fetos PEG
Deter RL, Lee W, Sangi-Hagheykar H, Tarca AL, Li J, Yeo L, et al. Personalized third-trimester fetal growth evaluation: comparisons of individualized growth assessment, percentile line and conditional probability methods. <i>J Matern & NEONATAL Med</i> . 2016 Jan;29(2):177–85.	Participantes del estudio de Estados Unidos
Deter RL, Lee W, Yeo L, Erez O, Ramamurthy U, Naik M, et al. Individualized growth assessment: conceptual framework and practical implementation for the evaluation of fetal growth and neonatal growth outcome. <i>Am J Obstet Gynecol</i> . 2018 Feb;218(2S):S656–78.	Tipo de artículo: revisión
Di Mascio D, Rizzo G, Buca D, D'Amico A, Leombroni M, Tinari S, et al. Comparison between cerebroplacental ratio and umbilicocerebral ratio in predicting adverse perinatal outcome at term. <i>Eur J Obstet Gynecol Reprod Biol</i> . 2020 Sep;252:439–43.	No evalúa la estimación del peso fetal mediante ecografía
Drukker L, Droste R, Chatelain P, Noble JA, Papageorghiou AT. Expected-value bias in routine third-trimester growth scans. <i>Ultrasound Obstet Gynecol Off J Int Soc Ultrasound Obstet Gynecol</i> . 2020 Mar;55(3):375–82.	Participantes del estudio de Reino Unido
Dubinsky TJ, Sonneborn R. Trouble With the Curve Pearls and Pitfalls in the Evaluation of Fetal Growth. <i>J ULTRASOUND Med</i> . 2020 Sep;39(9):1839–46.	Selección de embarazadas con fetos PEG
Dyer E, Chudleigh T. Peer review of third trimester abdominal circumference measurements. <i>ULTRASOUND</i> . 2021 May;29(2):83–91.	Selección de embarazadas con fetos PEG
Erkamp JS, Voerman E, Steegers EAP, Mulders AGMGJ, Reiss IKM, Duijts L, et al. Second and third trimester fetal ultrasound population screening for risks of preterm birth and small-size and large-size for gestational age at birth: a population-based prospective cohort study Fetal ultrasound screening for common adverse birth outcome. <i>BMC Med</i> . 2020 Apr;18(1):63.	Participantes del estudio de Países Bajos
Eslamian ZL, Zarean E, Moshfeghi M, Heidari Z. Evaluation of the predictive value of fetal Doppler ultrasound for neonatal outcome from the 36th week of pregnancy. <i>J Res Med Sci</i> . 2018 Feb;23:13.	No evalúa la estimación del peso fetal mediante ecografía
Flatley C, Kumar S. Is the fetal cerebroplacental ratio better than the estimated fetal weight in predicting adverse perinatal outcomes in a low risk cohort? <i>J Matern neonatal Med Off J Eur Assoc Perinat Med Fed Asia Ocean Perinat Soc Int Soc Perinat Obstet</i> . 2019 Jul;32(14):2380–6.	Participantes del estudio de Australia

Galán Arévalo MS, Mahillo-Fernández I, Esteban LM, Andeyro-García M, Piñeiro Pérez R, Saénz de Pipaón M, et al. Perinatal adverse effects in newborns with estimated loss of weight percentile between the third trimester ultrasound and delivery. The growin study. J Clin Med [Internet]. 2021 Oct;10(20). Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85116637342&doi=10.3390%2Fjcm10204643&partnerID=40&md5=b86955dea38ca93afd5ff3a45a6d0e51	No evalúa la estimación del peso fetal mediante ecografía
Galjaard S, Ameye L, Lees CC, Pexsters A, Bourne T, Timmerman D, et al. Sex differences in fetal growth and immediate birth outcomes in a low-risk Caucasian population. Biol Sex Differ. 2019 Sep;10(1):48.	Participantes del estudio de Países Bajos
Gembicki M, Offerman DR, Weichert J. Semiautomatic Assessment of Fetal Fractional Limb Volume for Weight Prediction in Clinical Praxis: How Does It Perform in Routine Use? J ultrasound Med Off J Am Inst Ultrasound Med. 2021 Apr;	Participantes del estudio de Alemania
Gjessing HK, Grøttum P, Økland I, Eik-Nes SH. Fetal size monitoring and birth-weight prediction: a new population-based approach. Ultrasound Obstet Gynecol Off J Int Soc Ultrasound Obstet Gynecol. 2017 Apr;49(4):500–7.	Participantes del estudio de Noruega
Goetzinger KR, Odibo AO. Disorders of fetal growth [Internet]. Evidence-based Obstetrics and Gynecology. 2018. 451–465 p. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85101481662&doi=10.1002%2F9781119072980.ch43&partnerID=40&md5=200a7ed0fc0574d990450f10f2da5721	Tipo de artículo: comentario
Grantz KL, Kim S, Grobman WA, Newman R, Owen J, Skupski D, et al. Fetal growth velocity: the NICHD fetal growth studies. Am J Obstet Gynecol. 2018 Sep;219(3):285.e1-285.e36.	Participantes del estudio de Estados Unidos
Hua X, Shen M, Reddy UM, Buck Louis G, Souza JP, Gülmözoglu AM, et al. Comparison of the INTERGROWTH-21st, National Institute of Child Health and Human Development, and WHO fetal growth standards. Int J Gynaecol Obstet Off organ Int Fed Gynaecol Obstet. 2018 Nov;143(2):156–63.	Participantes del estudio de Estados Unidos
Khan N, Ciobanu A, Karampitsakos T, Akolekar R, Nicolaides KH. Prediction of large-for-gestational-age neonate by routine third-trimester ultrasound. Ultrasound Obstet Gynecol Off J Int Soc Ultrasound Obstet Gynecol. 2019 Sep;54(3):326–33.	Participantes del estudio de Reino Unido
Kiserud T, Benachi A, Hecher K, Perez RG, Carvalho J, Piaggio G, et al. The World Health Organization fetal growth charts: concept, findings, interpretation, and application. Am J Obstet Gynecol. 2018 Feb;218(2S):S619–29.	Tipo de artículo: revisión
Lee K-S, Kim HY, Lee SJ, Kwon SO, Na S, Hwang HS, et al. Prediction of newborn's body mass index using nationwide multicenter ultrasound data: a machine-learning study. BMC Pregnancy Childbirth. 2021 Mar;21(1):172.	Participantes del estudio de Corea del Sud
Mack LM, Kim SY, Lee S, Sangi-Haghpeykar H, Lee W. Automated Fractional Limb Volume Measurements Improve the Precision of Birth Weight Predictions in Late Third-Trimester Fetuses. J ultrasound Med Off J Am Inst Ultrasound Med. 2017 Aug;36(8):1649–55.	Participantes del estudio de Estados Unidos
Marien M, Perron S, Bergeron A-M, Singbo N, Demers S. Comparison of the Accuracy of INTERGROWTH-21 and Hadlock Ultrasound Formulae for Fetal Weight Prediction. J Obstet Gynaecol Canada JOGC = J d'obstetrique Gynecol du Canada JOGC. 2021 Nov;43(11):1254–9.	Participantes del estudio de América del Norte
Maruotti GM, Saccone G, Martinelli P. Third trimester ultrasound soft-tissue measurements accurately predicts macrosomia. J Matern neonatal Med Off J Eur Assoc Perinat Med Fed Asia Ocean Perinat Soc Int Soc Perinat Obstet. 2017 Apr;30(8):972–6.	Tipo de artículo: revisión

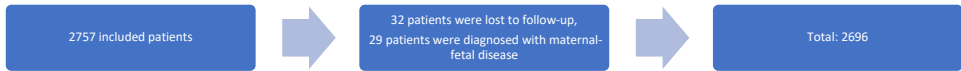
Mazzone E, Dall'Asta A, Kiener AJO, Carpano MG, Suprani A, Ghi T, et al. Prediction of fetal macrosomia using two-dimensional and three-dimensional ultrasound. <i>Eur J Obstet Gynecol Reprod Biol.</i> 2019 Dec;243:26–31.	Pacientes incluidos de alto riesgo (con enfermedad materno-fetal)
Miranda J, Triunfo S, Rodriguez-Lopez M, Sairanen M, Kouru H, Parra-Saavedra M, et al. Performance of third-trimester combined screening model for prediction of adverse perinatal outcome. <i>Ultrasound Obstet Gynecol Off J Int Soc Ultrasound Obstet Gynecol.</i> 2017 Sep;50(3):353–60.	No evalúa la estimación del peso fetal mediante ecografía
Monier I, Ego A, Benachi A, Hocquette A, Blondel B, Goffinet F, et al. Comparison of the performance of estimated fetal weight charts for the detection of small- and large-for-gestational age newborns with adverse outcomes: a French population-based study. <i>BJOG.</i> 2021 Nov;	Falta de acceso al artículo
Monier I, Ego A, Benachi A, Hocquette A, Blondel B, Goffinet F, et al. Unisex vs sex-specific estimated fetal weight charts for fetal growth monitoring: a population-based study. <i>Am J Obstet Gynecol MFM</i> [Internet]. 2022 Nov;4(1):100527. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85120179112&doi=10.1016%2Fj.ajogmf.2021.100527&partnerID=40&md5=171df6efd1470fd2ea5208715c6abd37	Selección de embarazadas con fetos PEG
Morales-Roselló J, Loscalzo G, Gallego A, Jakaitė V, Perales-Marín A. Which is the best ultrasound parameter for the prediction of adverse perinatal outcome within 1 day of delivery? <i>J Matern neonatal Med Off J Eur Assoc Perinat Med Fed Asia Ocean Perinat Soc Int Soc Perinat Obstet.</i> 2021 Oct;1–9.	Falta de acceso al artículo
Norris T, Johnson W, Petherick E, Cameron N, Oddie S, Johnson S, et al. Investigating the relationship between fetal growth and academic attainment: secondary analysis of the Born in Bradford (BiB) cohort. <i>Int J Epidemiol.</i> 2018 Oct;47(5):1475–84.	Participantes del estudio de Reino Unido
Parikh LI, Iqbal SN, Jelin AC, Overcash RT, Tefera E, Fries MH. Third trimester ultrasound for fetal macrosomia: optimal timing and institutional specific accuracy. <i>J Matern neonatal Med Off J Eur Assoc Perinat Med Fed Asia Ocean Perinat Soc Int Soc Perinat Obstet.</i> 2019 Apr;32(8):1337–41.	Participantes del estudio de Estados Unidos
Peixoto AB, Rodrigues da Cunha Caldas TM, Godoy Silva TA, Silva Gomes Caetano MS, Martins 6. WP, Martins Santana EF, et al. Assessment of ultrasound and Doppler parameters in the third trimester of pregnancy as predictors of adverse perinatal outcome in unselected pregnancies. <i>Ginekol Pol.</i> 2016;87(7):510–5.	Participantes del estudio de Brasil
Pugh SJ, Hinkle SN, Kim S, Albert PS, Newman R, Grobman WA, et al. Combined Influence of Gestational Weight Gain and Estimated Fetal Weight on Risk Assessment for Small- or Large-for- Gestational-Age Birth Weight: A Prospective Cohort Study. <i>J ultrasound Med Off J Am Inst Ultrasound Med.</i> 2018 Apr;37(4):935–40.	Participantes del estudio de Estados Unidos
Ray CL, Grangé G. Routine third trimester ultrasound in low risk pregnancy confers no benefit!: AGAINST: Arguments for a routine third trimester ultrasound: what the meta-analysis does not show! <i>BJOG An Int J Obstet Gynaecol</i> [Internet]. 2016;123(7):1122. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84978416170&doi=10.1111%2F1471-0528.13939&partnerID=40&md5=19b4433fa90981eca73f4014e6c3f479	Tipo de artículo: metaanálisis
Roeckner JT, Odibo L, Odibo AO. The value of fetal growth biometry velocities to predict large for gestational age (LGA) infants. <i>J Matern neonatal Med Off J Eur Assoc Perinat Med Fed Asia Ocean Perinat Soc Int Soc Perinat Obstet.</i> 2020 Jun;1–6.	Participantes del estudio de Estados Unidos
Ross MG. Ultrasound estimated fetal weight. <i>Am J Obstet Gynecol</i> [Internet]. 2017;217(6):709–10. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.085030635354&doi=10.1016%2Fj.ajog.2017.08.104&partnerID=40&md5=b5d5a2079878a5c1fbbf2eab6676f0a7	Tipo de artículo: comentario

Sabir F, Saeed S, Akhter N, Akram S, Salam R. Agreement between clinical and ultrasonographic estimation of abnormal birth weight of term fetus. <i>Pakistan J Med Heal Sci</i> [Internet]. 2018;12(4):1378–80. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85062584983&partnerID=40&md5=cee4eb1f8df12c80d8917de43f6dbd3e	Participantes del estudio de Pakistan
Sekar R, Khatun M, Barrett HL, Duncombe G. A prospective pilot study in assessing the accuracy of ultrasound estimated fetal weight prior to delivery. <i>Aust N Z J Obstet Gynaecol</i> . 2016 Feb;56(1):49–53	Tipo de artículo: ensayo clínico aleatorio
Self A, Papageorghiou AT. Ultrasound Diagnosis of the Small and Large Fetus. <i>Obstet Gynecol Clin North Am</i> [Internet]. 2021;48(2):339–57. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85105469092&doi=10.1016%2Fj.ogc.2021.03.003&partnerID=40&md5=ec302086d0558bbfc3bb100650399135	Tipo de artículo: revisión
Sharma KA, Das D, Dadhwal V, Deka D, Singhal S, Vanamail P. Two-dimensional fetal biometry versus three-dimensional fractional thigh volume for ultrasonographic prediction of birthweight. <i>Int J Gynaecol Obstet Off organ Int Fed Gynaecol Obstet</i> . 2019 Apr;145(1):47–53.	Participantes del estudio de India
Sheth T, Glantz JC. Third-Trimester Fetal Biometry and Neonatal Outcomes in Term and Preterm Deliveries. <i>J ultrasound Med Off J Am Inst Ultrasound Med</i> . 2016 Jan;35(1):103–10.	Participantes del estudio de Estados Unidos
Simcox LE, Myers JE, Cole TJ, Johnstone ED. Fractional fetal thigh volume in the prediction of normal and abnormal fetal growth during the third trimester of pregnancy. <i>Am J Obstet Gynecol</i> . 2017 Oct;217(4):453.e1-453.e12.	Selección de embarazadas con fetos PEG
Sirico A, Diemert A, Glosemeyer P, Hecher K. Prediction of adverse perinatal outcome by cerebroplacental ratio adjusted for estimated fetal weight. <i>Ultrasound Obstet Gynecol Off J Int Soc Ultrasound Obstet Gynecol</i> . 2018 Mar;51(3):381–6.	Participantes del estudio de Alemania
Sirico A, Diemert A, Glosemeyer P, Hecher K. Third Trimester Umbilical Artery Doppler in Low- Risk Pregnancies and its Correlation to Estimated Fetal Weight and Birthweight. <i>Ultraschall Med</i> . 2021 Jun;42(3):285–90.	Participantes del estudio de Alemania
Sousa-Santos RF, Miguelote RF, Cruz-Correia RJ, Santos CC, Bernardes JFMAL. Development of a birthweight standard and comparison with currently used standards. What is a 10th centile? <i>Eur J Obstet Gynecol Reprod Biol</i> . 2016 Nov;206:184–93.	No evalúa la estimación del peso fetal mediante ecografía
Sovio U, Moraitis AA, Wong HS, Smith GCS. Universal vs selective ultrasonography to screen for large-for-gestational-age infants and associated morbidity. <i>Ultrasound Obstet Gynecol Off J Int Soc Ultrasound Obstet Gynecol</i> . 2018 Jun;51(6):783–91.	Participantes del estudio de Reino Unido
Sovio U, Smith GCS. Comparison of estimated fetal weight percentiles near term for predicting extremes of birthweight percentile. <i>Am J Obstet Gynecol</i> . 2021 Mar;224(3):292.e1-292.e19.	Participantes del estudio de Reino Unido
Stephens K, Moraitis A, Smith GCS. Routine Third Trimester Sonogram: Friend or Foe. <i>Obstet Gynecol Clin North Am</i> . 2021 Jun;48(2):359–69.	Tipo de artículo: revisión
Tarca AL, Hernandez-Andrade E, Ahn H, Garcia M, Xu Z, Korzeniewski SJ, et al. Single and Serial Fetal Biometry to Detect Preterm and Term Small- and Large-for-Gestational-Age Neonates: A Longitudinal Cohort Study. <i>PLoS One</i> . 2016;11(11):e0164161.	Participantes del estudio de Estados Unidos
Valiño N, Giunta G, Gallo DM, Akolekar R, Nicolaides KH. Biophysical and biochemical markers 30-34 weeks' gestation in the prediction of adverse perinatal outcome. <i>Ultrasound Obstet Gynecol Off J Int Soc Ultrasound Obstet Gynecol</i> . 2016 Feb;47(2):194–202.	Participantes del estudio de Reino Unido

Valiño N, Giunta G, Gallo DM, Akolekar R, Nicolaides KH. Biophysical and biochemical markers 35-37 weeks' gestation in the prediction of adverse perinatal outcome. <i>Ultrasound Obstet Gynecol Off J Int Soc Ultrasound Obstet Gynecol</i> . 2016 Feb;47(2):203–9.	Participantes del estudio de Reino Unido
Wastlund D, Moraitis AA, Thornton JG, Sanders J, White IR, Brocklehurst P, et al. The cost- effectiveness of universal late-pregnancy screening for macrosomia in nulliparous women: a decision analysis. <i>BJOG</i> . 2019 Sep;126(10):1243–50.	Participantes del estudio de Reino Unido
Williams M, Turner S, Butler E, Gardosi J. Fetal growth surveillance - Current guidelines, practices and challenges. <i>Ultrasound</i> . 2018 May;26(2):69–79.	Tipo de artículo: revisión
Wu X, Niu Z, Xu Z, Jiang Y, Zhang Y, Meng H, et al. Fetal weight estimation by automated three- dimensional limb volume model in late third trimester compared to two-dimensional model: a cross- sectional prospective observational study. <i>BMC Pregnancy Childbirth</i> . 2021 May;21(1):365.	Participantes del estudio de China
Yee LM, Grobman WA. Relationship Between Third-Trimester Sonographic Estimate of Fetal Weight and Mode of Delivery. <i>J ultrasound Med Off J Am Inst Ultrasound Med</i> . 2016 Apr;35(4):701–6.	No evalúa la estimación del peso fetal mediante ecografía
Zhou J, Xiong Y, Ren Y, Zhang Y, Li X, Yan Y. Three-dimensional power Doppler ultrasonography indicates that increased placental blood perfusion during the third trimester is associated with the risk of macrosomia at birth. <i>J Clin Ultrasound</i> . 2021 Jan;49(1):12–9.	Participantes del estudio de China
Zhu C, Ren Y-Y, Wu J-N, Zhou Q-J. A Comparison of Prediction of Adverse Perinatal Outcomes between Hadlock and INTERGROWTH-21st Standards at the Third Trimester. <i>Biomed Res Int</i> . 2019;2019:7698038.	Participantes del estudio de China
Zohav E, Zohav E, Rabinovich M, Shenhav S, Ovadia YS, Anteby EY, et al. Local cerebroplacental ratio reference ranges are better predictors for adverse delivery outcomes in normal weight fetuses during pregnancy. <i>J Matern Neonatal Med [Internet]</i> . 2021;34(21):3475–80. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85075419542&doi=10.1080%2F14767058.2019.1685968&partnerID=40&md5=d7816701cdb939ddb1b2a6e7d790e60d	Selección de embarazadas con fetos PEG

ANEXO 2: Evaluación de la calidad de los estudios con la herramienta QUADAS-2.

Second- to Third-Trimester Longitudinal Growth Assessment for the Prediction of Largeness for Gestational Age and Macrosomia in an Unselected Population (Caradeux et al.)

- **Phase 1:** State the review question:
 - o Patients: 2696 unselected singleton pregnancies scanned at 21 ± 2 and 32 ± 2 weeks was created.
 - o Index test(s): cross-sectional evaluation at the third trimester
 - o Reference standard and target condition: second- to third- trimester longitudinal growth assessment, by means of AC growth velocity
- **Phase 2:** Draw a flow diagram for the primary study: Participants:


```
graph LR; A[2757 included patients] --> B[32 patients were lost to follow-up, 29 patients were diagnosed with maternal-fetal disease]; B --> C[Total: 2696]
```
- **Phase 3:** Risk of bias and applicability judgments. QUADAS-2 is structured so that 4 key domains are each rated in terms of the risk of bias and the concern regarding applicability to the research question (as defined above). Each key domain has a set of signalling questions to help reach the judgments regarding *bias and applicability*.

DOMAIN 1: PATIENT SELECTION

A. Risk of Bias

Describe methods of patient selection: Between January 2007 and December 2012, a prospective cohort of unselected consecutive singleton pregnancies was constructed, each referred to the Department of Maternal-Fetal Medicine in the Hospital Clinic of Barcelona for second- (21 ± 2 weeks) and third- trimester scan (32 ± 2 weeks). Exclusion criteria were: preexisting diabetes mellitus, chromosomal anomalies confirmed by genetic approaches (conventional karyotype or array-CGH), structural defects suspected at the time of routine scans and confirmed postnatally, evidence of fetal infection confirmed by microbiological examination in maternal blood and/or amniotic fluid, and suspected early fetal growth restriction (ultrasound estimated fetal weight <10th centile, according to local standards [18]) or preeclampsia before 32 weeks

- ❖ Was a consecutive or random sample of patients enrolled? **Yes/No/Unclear**
- ❖ Was a case-control design avoided? **Yes/No/Unclear**
- ❖ Did the study avoid inappropriate exclusions? **Yes/No/Unclear**

Could the selection of patients have introduced bias? **RISK: LOW/HIGH/UNCLEAR**

B. Concerns regarding applicability

Describe included patients: 2696 unselected singleton pregnancies scanned at 21 ± 2 and 32 ± 2 weeks was created. A total of 188 (7%) newborns qualified for LGA and 182 (6.8%) for macrosomia (LGA).

Is there concern that the included patients do not match the review question?
CONCERN: LOW/HIGH/UNCLEAR

DOMAIN 2: INDEX TEST(S)

If more than one index test was used, please complete for each test.

A. Risk of Bias

Describe the index test and how it was conducted and interpreted: Cross-sectional evaluation at the third trimester. At 21 ± 2 and 32 ± 2 weeks of gestation, ultrasound scan examinations were performed using a General Electric Voluson E8 (GE Medical Systems, Zipf, Austria) ultrasound machine equipped with a 6–2 MHz linear curved-array transducer by certified ultra-sonographers. At both scans, AC was measured at a transverse section of the abdomen at the level of the portal sinus and stomach, with ellipse placing at the outer surface.

- ❖ Were the index test results interpreted without knowledge of the results of the reference standard? Yes/No/Unclear
- ❖ If a threshold was used, was it pre-specified? Yes/No/Unclear

Could the conduct or interpretation of the index test RISK: LOW /HIGH/UNCLEAR have introduced bias?

B. Concerns regarding applicability

Is there concern that the index test, its conduct, or interpretation differ from the review question? CONCERN: LOW /HIGH/UNCLEAR

DOMAIN 3: REFERENCE STANDARD

A. Risk of Bias

Describe the reference standard and how it was conducted and interpreted: Second- to third- trimester longitudinal growth assessment, by means of AC (Abdominal Circunference) growth velocity. At 21 ± 2 and 32 ± 2 weeks of gestation, ultrasound scan examinations were performed using a General Electric Voluson E8 (GE Medical Systems, Zipf, Austria) ultrasound machine equipped with a 6–2 MHz linear curved-array transducer by certified ultra-sonographers. At both scans, AC was measured at a transverse section of the abdomen at the level of the portal sinus and stomach, with ellipse placing at the outer surface.

- ❖ Is the reference standard likely to correctly classify the target condition? Yes/No/Unclear
- ❖ Were the reference standard results interpreted without knowledge of the results of the index test? Yes/No/Unclear

Could the reference standard, its conduct, or its interpretation have introduced bias? RISK: LOW /HIGH/UNCLEAR

B. Concerns regarding applicability

Is there concern that the target condition as defined by the reference standard does not match the review question? CONCERN: LOW /HIGH/UNCLEAR

DOMAIN 4: FLOW AND TIMING

A. Risk of Bias

Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2x2 table (refer to flow diagram): Exclusion criteria were: preexisting diabetes mellitus, chromosomal anomalies confirmed by genetic approaches (conventional karyotype or array-CGH), structural defects suspected at the time of routine scans and confirmed postnatally, evidence of fetal infection confirmed by microbiological examination in maternal blood and/or amniotic fluid, and suspected early fetal growth restriction (ultrasound estimated fetal weight <10th centile, according to local standards [18]) or preeclampsia before 32 weeks .

Describe the time interval and any interventions between index test(s) and reference standard: At 21 ± 2 and 32 ± 2 weeks of gestation, ultrasound scan examinations were performed.

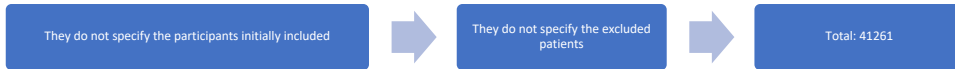
- ❖ Was there an appropriate interval between index test(s) and reference standard? Yes/No/Unclear
- ❖ Did all patients receive a reference standard? Yes/No/Unclear
- ❖ Did patients receive the same reference standard? Yes/No/Unclear
- ❖ Were all patients included in the analysis? Yes/No/Unclear

Could the patient flow have introduced bias?

RISK: LOW /HIGH/UNCLEAR

Factors associated with more than 500 grams inaccuracy in sonographic fetal weight estimation (Hendin et al.)

- **Phase 1:** State the review question:
 - o Patients: women carrying singleton term gestations who delivered within 2 weeks after undergoing a sonographic fetal weight estimation.
 - o Index test(s): Hadlock formula for uEFW (unreliable estimated fetal weight).
 - o Reference standard and target condition: Hadlock formula for aEFW (accurate estimated fetal weight).
- **Phase 2:** Draw a flow diagram for the primary study: Participants:



- **Phase 3:** Risk of bias and applicability judgments. QUADAS-2 is structured so that 4 key domains are each rated in terms of the risk of bias and the concern regarding applicability to the research question (as defined above). Each key domain has a set of signalling questions to help reach the judgments regarding *bias and applicability*.

DOMAIN 1: PATIENT SELECTION

A. Risk of Bias

Describe methods of patient selection: This study was conducted at a single, tertiary medical center between 2011 and 2019. Unreliable EFW was defined as > 500 grams discordance between it and the actual birth weight. We allocated the study cohort into two groups: unreliable EFW (cases) and accurate EFW (controls).

- ❖ Was a consecutive or random sample of patients enrolled? Yes/No/**Unclear**
- ❖ Was a case-control design avoided? Yes/**No**/Unclear
- ❖ Did the study avoid inappropriate exclusions? Yes/No/**Unclear**

Could the selection of patients have introduced bias? RISK: LOW/HIGH/UNCLEAR

B. Concerns regarding applicability

Describe included patients: Overall, 41261 deliveries met the inclusion criteria during the study period. The study cohort comprised all women carrying singleton term gestations who delivered within 2 weeks after undergoing a sonographic fetal weight estimation. The study was conducted from 2011 to 2019. Exclusion criteria were multiple gestations and known fetal anomalies. We used the Hadlock A formula for EFW as it is widely used and has been shown to produce the smallest systematic mean errors on a normal fetal population. We allocated the study cohort into two groups: uEFW (cases) and accurate EFW (controls).

Is there concern that the included patients do not match the review question?

CONCERN: LOW/HIGH/UNCLEAR

DOMAIN 2: INDEX TEST(S)

If more than one index test was used, please complete for each test.

A. Risk of Bias

Describe the index test and how it was conducted and interpreted: We used the Hadlock A formula for EFW as it is widely used and has been shown to produce the smallest systematic mean errors on a normal fetal population. We defined uEFW as a difference of more than 500 grams from actual birth weight. We chose 500 grams as cutoff due to previous works attributing this cutoff a clinical significance, as errors of this magnitude could have a potential impact on delivery mode planning. The unreliable EFW (uEFW) are the cases of our study.

- ❖ Were the index test results interpreted without knowledge of the results of the reference standard? Yes/No/Unclear
- ❖ If a threshold was used, was it pre-specified? Yes/No/Unclear

Could the conduct or interpretation of the index test **RISK: LOW /HIGH/UNCLEAR** have introduced bias?

B. Concerns regarding applicability

Is there concern that the index test, its conduct, or interpretation differ from the review question? **CONCERN: LOW /HIGH/UNCLEAR**

DOMAIN 3: REFERENCE STANDARD

A. Risk of Bias

Describe the reference standard and how it was conducted and interpreted: We used the Hadlock A formula for EFW. The accurate EFW are the controls of our study.

- ❖ Is the reference standard likely to correctly classify the target condition? Yes/No/Unclear
- ❖ Were the reference standard results interpreted without knowledge of the results of the index test? Yes/No/Unclear

Could the reference standard, its conduct, or its interpretation have introduced bias? **RISK: LOW /HIGH/UNCLEAR**

B. Concerns regarding applicability

Is there concern that the target condition as defined by the reference standard does not match the review question? **CONCERN: LOW /HIGH/UNCLEAR**

DOMAIN 4: FLOW AND TIMING

A. Risk of Bias

Describe any patients who did not receive the index test(s) and/or reference standard or whowere excluded from the 2x2 table (refer to flow diagram): Exclusion criteria were multiple gestations and known fetal anomalies.

Describe the time interval and any interventions between index test(s) and reference standard: it does not appear in the article.

- ❖ Was there an appropriate interval between index test(s) and reference standard? Yes/No/**Unclear**
- ❖ Did all patients receive a reference standard? **Yes**/No/Unclear
- ❖ Did patients receive the same reference standard? **Yes**/No/Unclear
- ❖ Were all patients included in the analysis? **Yes**/No/Unclear

Could the patient flow have introduced bias?

RISK: LOW /HIGH/UNCLEAR

World Health Organization fetal growth charts applied in a French birth cohort (Hocquette et al.)

- **Phase 1:** State the review question:
 - Patients: Data came from the ELFE birth cohort, an observational population-based study in 349 maternity units randomly selected from the 544 in metropolitan France. The population was constituted of singleton or twin live births after 33 weeks' gestation.
 - Index test(s): Biometric measurements and EFW were translated into percentiles according to the WHO chart.
 - Reference standard and target condition: Biometric measurements and EFW were translated into percentiles according to Intergrowth, Hadlock et al., Salomon et al. and CFEF (EFW only).

- **Phase 2:** Draw a flow diagram for the primary study: Participants:



- **Phase 3:** Risk of bias and applicability judgments. QUADAS-2 is structured so that 4 key domains are each rated in terms of the risk of bias and the concern regarding applicability to the research question (as defined above). Each key domain has a set of signalling questions to help reach the judgments regarding *bias and applicability*.

DOMAIN 1: PATIENT SELECTION

A. Risk of Bias

Describe methods of patient selection: Data came from the ELFE birth cohort, an observational population-based study in 349 maternity units randomly selected from the 544 in metropolitan France.

- ❖ Was a consecutive or random sample of patients enrolled? **Yes/No/Unclear**
- ❖ Was a case-control design avoided? **Yes/No/Unclear**
- ❖ Did the study avoid inappropriate exclusions? **Yes/No/Unclear**

Could the selection of patients have introduced bias? **RISK: LOW/HIGH/UNCLEAR**

B. Concerns regarding applicability

Describe included patients: The population was constituted of singleton or twin live births after 33 weeks' gestation occurring during 25 days in 2011. Other inclusion criteria were: mothers aged 18 years and over, parents who understood French, English, Arabic or Turkish and parents who lived and planned to stay in France for at least three years and gave written informed consent (N = 18,329).

Is there concern that the included patients do not match the review question?

CONCERN: LOW/HIGH/UNCLEAR

DOMAIN 2: INDEX TEST(S)

If more than one index test was used, please complete for each test.

A. Risk of Bias

Describe the index test and how it was conducted and interpreted: Biometric measurements and EFW were translated into percentiles according to the WHO chart. Our main outcomes were FL, AC and EFW. The WHO and Intergrowth charts are international prescriptive standards built from samples of low-risk singleton pregnancies both published in 2017.

❖ Were the index test results interpreted without knowledge of the results of the reference standard? Yes/No/Unclear

❖ If a threshold was used, was it pre-specified? Yes/No/Unclear

Could the conduct or interpretation of the index test have introduced bias? RISK: LOW /HIGH/UNCLEAR

B. Concerns regarding applicability

Is there concern that the index test, its conduct, or interpretation differ from the review question? CONCERN: LOW/HIGH/UNCLEAR

DOMAIN 3: REFERENCE STANDARD

A. Risk of Bias

Describe the reference standard and how it was conducted and interpreted: Biometric measurements and EFW were translated into percentiles according to Intergrowth, Hadlock et al., Salomon et al. and CFEF (EFW only). Our main outcomes were FL, AC and EFW. The WHO and Intergrowth charts are international prescriptive standards built from samples of low-risk singleton pregnancies both published in 2017. Fetal biometric (1984) and EFW (1991) charts for Hadlock are reference charts based on two samples of 392 and 361 singletons from mothers with no history of diabetes and hypertension in one single center in the US. Salomon's charts are fetal biometric and EFW references from 19,647 and 18,959 routine ultrasounds respectively in two samples of unselected pregnancies published in 2006 and 2007. Finally, the CFEF's EFW chart was based on 33,143 routine ultrasounds in a sample of unselected pregnancies in seven French towns and was published in 2014.

❖ Is the reference standard likely to correctly classify the target condition? Yes/No/Unclear

❖ Were the reference standard results interpreted without knowledge of the results of the index test? Yes/No/Unclear

Could the reference standard, its conduct, or its interpretation have introduced bias? RISK: LOW /HIGH/UNCLEAR

B. Concerns regarding applicability

Is there concern that the target condition as defined by the reference standard does not match the review question? CONCERN: LOW /HIGH/UNCLEAR

DOMAIN 4: FLOW AND TIMING

A. Risk of Bias

Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2x2 table (refer to flow diagram): We excluded multiple births (n = 574) because of their specific growth patterns. We also excluded births with an ultrasound outside the gestational age ranges recommended in France for at least one ultrasound (n = 1118) because we aimed to include screening ultrasounds only and not follow-up visits to monitor growth. We also sought to select fetuses with all data at both time points as the aim was to compare between measures and trimesters. Fetuses were excluded when all ultrasound information was missing for one or both of the routine ultrasounds (n = 4637) which can occur when the results from ultrasounds are not retranscribed in records at the maternity hospital or if women did not have one of the two ultrasounds. Finally, fetuses were excluded when transverse abdominal diameter had been collected without AC (n = 3819), and when gestational age at the ultrasound or one ultrasound measure (FL or AC but transverse abdominal diameter not collected either) was missing (n = 434).

Describe the time interval and any interventions between index test(s) and reference standard: At the same time

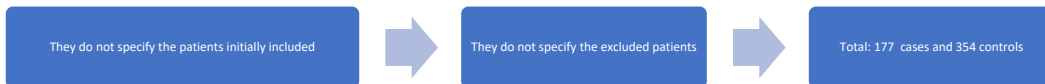
- ❖ Was there an appropriate interval between index test(s) and reference standard? **Yes/No/Unclear**
- ❖ Did all patients receive a reference standard? **Yes/No/Unclear**
- ❖ Did patients receive the same reference standard? **Yes/No/Unclear**
- ❖ Were all patients included in the analysis? **Yes/No/Unclear**

Could the patient flow have introduced bias?

RISK: LOW /HIGH/UNCLEAR

Significant deviations in sonographic fetal weight estimation: causes and implications (Krispin et al.)

- **Phase 1:** State the review question:
 - o Patients: All women who experienced a live birth delivery within 1 week following sonographic EFW were included.
 - o Index test(s): Hadlock
 - o Reference standard and target condition: -
- **Phase 2:** Draw a flow diagram for the primary study:



- **Phase 3:** Risk of bias and applicability judgments. QUADAS-2 is structured so that 4 key domains are each rated in terms of the risk of bias and the concern regarding applicability to the research question (as defined above). Each key domain has a set of signalling questions to help reach the judgments regarding *bias and applicability*.

DOMAIN 1: PATIENT SELECTION

A. Risk of Bias

Describe methods of patient selection: We collected data from all pregnant women who underwent sonographic EFW and delivered in our university-affiliated tertiary medical center between 2012 and 2018.

- ❖ Was a consecutive or random sample of patients enrolled? Yes/No/Unclear
- ❖ Was a case-control design avoided? Yes/No/Unclear
- ❖ Did the study avoid inappropriate exclusions? Yes/No/Unclear

Could the selection of patients have introduced bias? RISK: LOW/HIGH/UNCLEAR

B. Concerns regarding applicability

Describe included patients: All women who experienced a live birth delivery within 1 week following sonographic EFW were included. Multiple gestations, known fetal malformations and intrauterine fetal deaths were excluded. The study group was comprised of all pregnancies in which the sonographic EFW differed by at least 20% below or above actual birth weight. These were matched in a 1:2 ratio to controls, according to: maternal age, gravidity, parity and gestational age at delivery. The control group included matched pregnancies in which sonographic EFW was accurate and differed by up to 1% below or above actual birth weight.

Is there concern that the included patients do not match the review question?

CONCERN: LOW/HIGH/UNCLEAR

DOMAIN 2: INDEX TEST(S)

If more than one index test was used, please complete for each test.

A. Risk of Bias

Describe the index test and how it was conducted and interpreted: All fetal biometry evaluations were performed in our ultra- sound unit solely by senior obstetricians/gynecologists specialized in obstetric imaging or by skilled ultrasound technicians. All evaluations were documented by images, clips and written report. EFW were calculated following measurements of bi-parietal diameter (BPD), head circum- ference (HC), abdominal circumference (AC) and femur length (FL) according to the Hadlock equation and in accordance with international guidelines.

❖ Were the index test results interpreted without knowledge of the results of the reference standard? Yes/No/Unclear

❖ If a threshold was used, was it pre-specified? Yes/No/Unclear

Could the conduct or interpretation of the index test RISK: LOW /HIGH/UNCLEAR have introduced bias?

B. Concerns regarding applicability

Is there concern that the index test, its conduct, or interpretation differ from the review question? CONCERN: LOW /HIGH/UNCLEAR

DOMAIN 3: REFERENCE STANDARD

A. Risk of Bias

Describe the reference standard and how it was conducted and interpreted: -

❖ Is the reference standard likely to correctly classify the target condition? Yes/No/Unclear

❖ Were the reference standard results interpreted without knowledge of the results of the index test? Yes/No/Unclear

Could the reference standard, its conduct, or its interpretation have introduced bias? RISK: LOW /HIGH/UNCLEAR

B. Concerns regarding applicability

Is there concern that the target condition as defined by the reference standard does not match the review question? CONCERN: LOW /HIGH/UNCLEAR

DOMAIN 4: FLOW AND TIMING

A. Risk of Bias

Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2x2 table (refer to flow diagram): .Multiple gestations, known fetal malformations and intrauterine fetal deaths were excluded.

Describe the time interval and any interventions between index test(s) and reference standard: Unclear

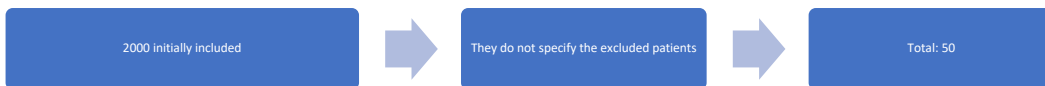
- ❖ Was there an appropriate interval between index test(s) and reference standard? Yes/No/**Unclear**
- ❖ Did all patients receive a reference standard? Yes/No/**Unclear**
- ❖ Did patients receive the same reference standard? Yes/No/**Unclear**
- ❖ Were all patients included in the analysis? **Yes**/No/Unclear

Could the patient flow have introduced bias?

RISK: LOW /HIGH/UNCLEAR****

3D and 2D ultrasound-based fetal weight estimation: a single center experience (Mohsen et al.)

- **Phase 1:** State the review question:
 - o Patients: Fifty pregnant females were included. Inclusion criteria: singleton pregnancy, within five days of delivery and normal or IUGR pregnancy.
 - o Index test(s): two new 3D (M3, M6)and two new 2D (STT1, STT2) ultrasound formulae for fetal weight estimation against the modified Hadlock formula and compare their estimation to the actual fetal weight.
 - o Reference standard and target condition: Hadlock
- **Phase 2:** Draw a flow diagram for the primary study: Participants:



- **Phase 3:** Risk of bias and applicability judgments. QUADAS-2 is structured so that 4 key domains are each rated in terms of the risk of bias and the concern regarding applicability to the research question (as defined above). Each key domain has a set of signalling questions to help reach the judgments regarding *bias and applicability*.

DOMAIN 1: PATIENT SELECTION

A. Risk of Bias

Describe methods of patient selection: Any pregnant patient presenting to our department for an obstetric ultrasound scan during her late third trimester (>37 weeks) was evaluated and included in the study only if her delivery occurred within five days after the ultrasound study.

<ul style="list-style-type: none"> ❖ Was a consecutive or random sample of patients enrolled? ❖ Was a case-control design avoided? ❖ Did the study avoid inappropriate exclusions? 	<p><u>Yes</u>/No/Unclear</p> <p><u>Yes</u>/No/Unclear</p> <p><u>Yes</u>/No/Unclear</p>
---	--

Could the selection of patients have introduced bias? RISK: LOW/HIGH/UNCLEAR

B. Concerns regarding applicability

Describe included patients: Any pregnant patient presenting to our department for an obstetric ultrasound scan during her late third trimester (>37 weeks) was evaluated and included in the study only if her delivery occurred within five days after the ultrasound study. Fifty pregnant females were included in our study. Inclusion criteria: singleton pregnancy, within five days of delivery, and normal or IUGR pregnancy were included (because we were aiming to evaluate fetal weight estimation in clinical practice). Exclusion criteria: multiple pregnancies, IUFD, patients in critical condition (because of the need for a short quick scan), pregnancies less than 37 weeks gestation, except if in preterm labor. One preterm and one IUGR patients were included in this study.

Is there concern that the included patients do not match the review question?

CONCERN: LOW/HIGH/UNCLEAR

DOMAIN 2: INDEX TEST(S)

If more than one index test was used, please complete for each test.

A. Risk of Bias

Describe the index test and how it was conducted and interpreted: two new 3D (M3, M6) and two new 2D (STT1, STT2) ultrasound formulae for fetal weight estimation against the modified Hadlock formula and compare their estimation to the actual fetal weight. All patients had a basic fetal growth evaluation, i.e. measurement of BPD, AC and FL. Then a routine evaluation of fetal well-being was performed, i.e. cranial, cardiac, abdominal and spinal evaluation, calculation of heart rate and Doppler assessment of umbilical arteries and middle cerebral arteries. Then 3D evaluation of the fetal thigh and arm was added. Images were taken with the femur and humerus in longitudinal view. Image depth and magnification were adjusted such that the thigh and arm volumes of interest fill at least two thirds of the display. The transducer focal zone was adjusted to have the femoral or humeral shaft in its center. Each volume was acquired from a sagittal sweep of the limb with a sweep angle of 90. Estimated fetal weight (EFW) was calculated according to the fractional thigh and arm volumes, mid-thigh soft tissue thickness and MH formula (MH) as previously described. Two formulae were used for the fractional thigh and arm volumes; these were Model 3 (M3) and Model 6 (M6) in the study by Lee et al. In his study, these provided the least percentage and random errors. The same names were used in our study, to avoid confusion. The original Scioscia formula and a modified one were also used for comparison. These will be given short names of STT 1 and STT 2, respectively.

- ❖ Were the index test results interpreted without knowledge of the results of the reference standard? Yes/No/Unclear
- ❖ If a threshold was used, was it pre-specified? Yes/No/Unclear

Could the conduct or interpretation of the index test RISK: LOW /HIGH/UNCLEAR have introduced bias?

B. Concerns regarding applicability

Is there concern that the index test, its conduct, or interpretation differ from the review question? CONCERN: LOW /HIGH/UNCLEAR

DOMAIN 3: REFERENCE STANDARD

A. Risk of Bias

Describe the reference standard and how it was conducted and interpreted: Hadlock method. All patients had a basic fetal growth evaluation, i.e. measurement of BPD, AC and FL. Then a routine evaluation of fetal well-being was performed, i.e. cranial, cardiac, abdominal and spinal evaluation, calculation of heart rate and Doppler assessment of umbilical arteries and middle cerebral arteries.

- ❖ Is the reference standard likely to correctly classify the target condition? Yes/No/Unclear
- ❖ Were the reference standard results interpreted without knowledge of the results of the index test? Yes/No/Unclear

Could the reference standard, its conduct, or its interpretation have introduced bias? RISK: LOW /HIGH/UNCLEAR

B. Concerns regarding applicability

Is there concern that the target condition as defined by the reference standard does not match the review question? CONCERN: LOW /HIGH/UNCLEAR

DOMAIN 4: FLOW AND TIMING

A. Risk of Bias

Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2x2 table (refer to flow diagram): Unclear

Describe the time interval and any interventions between index test(s) and reference standard: At the same time

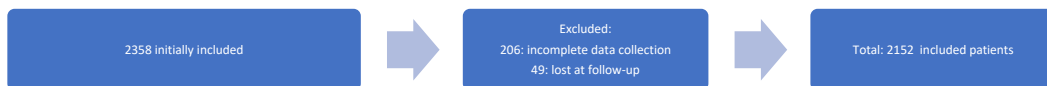
- ❖ Was there an appropriate interval between index test(s) and reference standard? Yes/No/Unclear
- ❖ Did all patients receive a reference standard? Yes/No/Unclear
- ❖ Did patients receive the same reference standard? Yes/No/Unclear
- ❖ Were all patients included in the analysis? Yes/No/Unclear

Could the patient flow have introduced bias?

RISK: LOW /HIGH/UNCLEAR

The added value of umbilical vein flow in predicting fetal macrosomia at 36 weeks of gestation: a prospective cohort study (Rizzo et al.)

- **Phase 1:** State the review question:
 - o Patients: 2156 singleton pregnancies scheduled for routine ultrasound assessment at 36 weeks of gestation.
 - o Index test(s): Hadlock, maternal variables and UVBF
 - o Reference standard and target condition: Hadlock
- **Phase 2:** Draw a flow diagram for the primary study



- **Phase 3:** Risk of bias and applicability judgments. QUADAS-2 is structured so that 4 key domains are each rated in terms of the risk of bias and the concern regarding applicability to the research question (as defined above). Each key domain has a set of signalling questions to help reach the judgments regarding *bias and applicability*.

DOMAIN 1: PATIENT SELECTION

A. Risk of Bias

Describe methods of patient selection: This was a prospective cohort study including consecutive singleton pregnancies undergoing ultrasound assessment at 35-37 weeks of gestation in a dedicated antenatal clinic in the Department of Obstetrics and Gynecology, Università di Roma Tor Vergata between January 2016 and January 2020.

- ❖ Was a consecutive or random sample of patients enrolled? Yes/No/Unclear
- ❖ Was a case-control design avoided? Yes/No/Unclear
- ❖ Did the study avoid inappropriate exclusions? Yes/No/Unclear

Could the selection of patients have introduced bias? **RISK: LOW/HIGH/UNCLEAR**

B. Concerns regarding applicability

Describe included patients: singleton pregnancies undergoing ultrasound assessment at 35-37 weeks of gestation. Inclusion criteria were first trimester pregnancy dating by crown-rump length at 11-14 weeks of gestations, absence of fetal chromosomal or structural anomalies and delivery in our unit.

Is there concern that the included patients do not match the review question?

CONCERN: LOW/HIGH/UNCLEAR

DOMAIN 2: INDEX TEST(S)

If more than one index test was used, please complete for each test.

A. Risk of Bias

Describe the index test and how it was conducted and interpreted: All women underwent detailed ultrasound assessment including biometrics, maternal and fetal Doppler. Recordings were performed using a Samsung W80 or Hera W10 ultrasound devices equipped with a 1-8 MHz volumetric probe. All the recordings were performed by two of the Authors. Fetal biometric measurements included biparietal diameter, head circumference, abdominal circumference (AC) and femur length, obtained following International Society of Ultrasound in Obstetrics & Gynecology recommendations, while EFW was computed using Hadlock IV formula. Doppler assessment included pulsatility index of the uterine, umbilical artery and middle cerebral arteries and cerebroplacental ratio, defined as a ratio between the middle cerebral artery and umbilical artery pulsatility index according to a previously published methodology (21,22). Umbilical vein (UV) diameter and time-averaged maximum velocity were measured in UV intra-abdominal portion by real time and Doppler ultrasonography. A semi-automatic measurement software was used to obtain UV diameter values. The UV absolute value of flow corrected for fetal size (UVBF/AC).

❖ Were the index test results interpreted without knowledge of the results of the reference standard? **Yes/No/Unclear**

❖ If a threshold was used, was it pre-specified? **Yes/No/Unclear**

Could the conduct or interpretation of the index test **RISK: LOW /HIGH/UNCLEAR** have introduced bias?

B. Concerns regarding applicability

Is there concern that the index test, its conduct, or interpretation differ from the review question? **CONCERN: LOW /HIGH/UNCLEAR**

DOMAIN 3: REFERENCE STANDARD

A. Risk of Bias

Describe the reference standard and how it was conducted and interpreted: EFW based in Hadlock formula.

❖ Is the reference standard likely to correctly classify the target condition? **Yes/No/Unclear**

❖ Were the reference standard results interpreted without knowledge of the results of the index test? **Yes/No/Unclear**

Could the reference standard, its conduct, or its interpretation have introduced bias? **RISK: LOW /HIGH/UNCLEAR**

B. Concerns regarding applicability

Is there concern that the target condition as defined by the reference standard does not match the review question? **CONCERN: LOW /HIGH/UNCLEAR**

DOMAIN 4: FLOW AND TIMING

A. Risk of Bias

Describe any patients who did not receive the index test(s) and/or reference standard or whowere excluded from the 2x2 table (refer to flow diagram): During the study period 2358 pregnancies met the inclusion criteria; out of these, 206 were excluded for the incomplete data collection (n= 157) and lost at follow-up (n=49) leaving 2152 pregnancies available for the analysis ,

Describe the time interval and any interventions between index test(s) and reference standard: At the same time

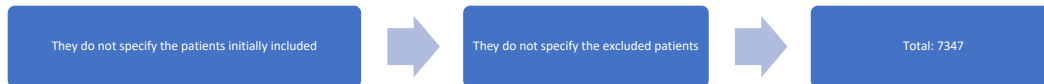
- ❖ Was there an appropriate interval between index test(s) and reference standard? Yes/No/Unclear
- ❖ Did all patients receive a reference standard? Yes/No/Unclear
- ❖ Did patients receive the same reference standard? Yes/No/Unclear
- ❖ Were all patients included in the analysis? Yes/No/Unclear

Could the patient flow have introduced bias?

RISK: LOW /HIGH/UNCLEAR

Current use and performance of the different fetal growth charts in the Italian population (Stampalija et al.)

- **Phase 1:** State the review question:
 - Patients: Uncomplicated singleton pregnancies delivered at 37-41 weeks with uneventful perinatal outcome
 - Index test(s): Fetal growth charts examined: Nicolini, Paladini, WHO and IG.
 - Reference standard and target condition: -
- **Phase 2:** Draw a flow diagram for the primary study



- **Phase 3:** Risk of bias and applicability judgments. QUADAS-2 is structured so that 4 key domains are each rated in terms of the risk of bias and the concern regarding applicability to the research question (as defined above). Each key domain has a set of signalling questions to help reach the judgments regarding *bias and applicability*.

DOMAIN 1: PATIENT SELECTION

A. Risk of Bias

Describe methods of patient selection: A questionnaire was sent to 14 Italian tertiary referral centers to explore biometric reference growth charts used in clinical practice. In order to investigate which fetal growth charts are in use at national level, a survey was performed involving 14 Italian tertiary referral centers. A short questionnaire was prepared to explore the biometric and Doppler charts used in clinical practice. The questionnaire was e-mailed to the lead of each Unit. In the case of no response, up to two e-mail reminders were sent at three-week intervals. Frequency and percentage distribution of adopted fetal growth charts were analyzed.

- ❖ Was a consecutive or random sample of patients enrolled? **Yes/No/Unclear**
- ❖ Was a case-control design avoided? **Yes/No/Unclear**
- ❖ Did the study avoid inappropriate exclusions? **Yes/No/Unclear**

Could the selection of patients have introduced bias? RISK: LOW/HIGH/UNCLEAR

B. Concerns regarding applicability

Describe included patients (prior testing, presentation, intended use of index test and setting): Uncomplicated singleton pregnancies delivered at 37-41 weeks with uneventful perinatal outcome and available information on parity, maternal height, weight and race were included. Pregnancies complicated by fetal structural or chromosomal anomalies, pre-existing maternal disease or development of obstetrics complications were excluded. Only mothers with a pre-pregnancy body mass index (BMI) in the range of 17-30 were selected. The pregnancies were dated based on first trimester ultrasound measurement of crown rump length. Sonographic measurements were obtained during routine screening ultrasound examinations, recommended at 19-21 and 30-34 weeks of gestation according to national guidelines.¹⁰ A smaller number of fetal biometric data were collected out of the screening time interval when ultrasound was performed as a complement of clinical examination.

Is there concern that the included patients do not match the review question?

CONCERN: LOW/HIGH/UNCLEAR

DOMAIN 2: INDEX TEST(S)

If more than one index test was used, please complete for each test.

A. Risk of Bias (for all tests)

Describe the index test and how it was conducted and interpreted:

- The Nicolini growth charts were the result of a prospective cross-sectional multicentric (n=8) study conducted in Northern Italy in the years 1984-1985 on uncomplicated pregnancies (n=1426).⁴ The standardization was performed together with the evaluation of intra- and inter-observer reproducibility. Reference ranges from 12 to 41 gestational weeks for biparietal diameter (BPD), head circumference (HC) and abdominal circumference (AC) were calculated.
- The Paladini growth charts were based on a prospective cross-sectional study conducted in three referral centers for prenatal diagnosis on 626 fetuses in early 2000ies.⁵ Exclusion criteria were applied for all maternal and/or fetal conditions that could possibly affect fetal growth. Reference ranges from 16 to 40 weeks of gestation were computed for BPD, HC, AC, femur (FL), tibia, humerus, ulna and radius length.
- The IG-21st prescriptive standards were derived from an international prospective longitudinal multicentric multiethnic study (8 countries) on 4321 pregnancies in years 2009-2014.⁷ Strict selection criteria were applied in order to include only low-risk pregnancies of healthy and well-nourished mothers from medium-high socioeconomic status.⁹ Strict methodological criteria were applied for the biometric measurements that were blinded. Growth standards from 14 to 40 weeks of gestation were obtained for BPD, HC, AC, FL and EFW.
- The WHO fetal growth charts are based on a prospective longitudinal multicentric multiethnic study on 1387 women conducted in 10 countries in years 2009-2014.⁸ Participants had no known health, environmental, nutritional or socio-economic constraints. Fetal, maternal and neonatal clinical conditions were retained, thus providing reference curves from 14 to 40 weeks of gestation for BPD, HC, AC, FL and EFW.

❖ Were the index test results interpreted without knowledge of the results of the reference standard? Yes/**No**/Unclear

❖ If a threshold was used, was it pre-specified? **Yes**/No/Unclear

Could the conduct or interpretation of the index test **RISK: LOW /HIGH/UNCLEAR** have introduced bias?

B. Concerns regarding applicability

Is there concern that the index test, its conduct, or interpretation differ from the review question?

CONCERN:
LOW /HIGH/UNCLEAR

DOMAIN 3: REFERENCE STANDARD

A. Risk of Bias

Describe the reference standard and how it was conducted and interpreted: -

❖ Is the reference standard likely to correctly classify the target condition? Yes/**No**/Unclear

❖ Were the reference standard results interpreted without knowledge of the results of the index test? Yes/**No**/Unclear

Could the reference standard, its conduct, or its interpretation have introduced bias? **RISK: LOW /HIGH/UNCLEAR**

B. Concerns regarding applicability

Is there concern that the target condition as defined by the reference standard does not match the review question?

CONCERN:
LOW /HIGH/UNCLEAR

DOMAIN 4: FLOW AND TIMING

A. Risk of Bias

Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2x2 table (refer to flow diagram): There are no information in the study.

Describe the time interval and any interventions between index test(s) and reference standard: There are no reference standard.

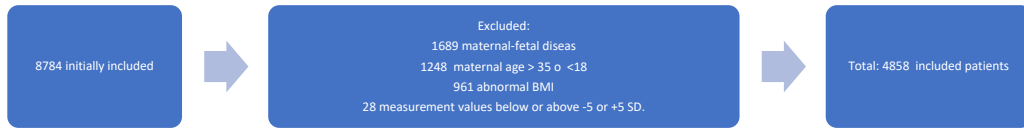
- ❖ Was there an appropriate interval between index test(s) and reference standard? Yes/No/**Unclear**
- ❖ Did all patients receive a reference standard? Yes/**No**/Unclear
- ❖ Did patients receive the same reference standard? Yes/**No**/Unclear
- ❖ Were all patients included in the analysis? **Yes**/No/Unclear

Could the patient flow have introduced bias?

RISK: LOW /HIGH**/UNCLEAR**

Implementing the Intergrowth-21st fetal growth standards in France: a « flash » study of the college français d'echographie foetale (CFEF). (Stirnemann et al.)

- **Phase 1:** State the review question:
 - o Patients (setting, intended use of index test, presentation, prior testing):
 - o Index test(s): french local growth charts
 - o Reference standard and target condition: IG
- **Phase 2:** Draw a flow diagram for the primary study



- **Phase 3:** Risk of bias and applicability judgments. QUADAS-2 is structured so that 4 key domains are each rated in terms of the risk of bias and the concern regarding applicability to the research question (as defined above). Each key domain has a set of signalling questions to help reach the judgments regarding *bias and applicability*.

DOMAIN 1: PATIENT SELECTION

A. Risk of Bias

Describe methods of patient selection: Pregnant women contributed with a single measure and were prospectively and consecutively included over a fixed study period of six weeks. the FGLS participants were first selected at the level of a geographic area and then at the individual level within each study site.

❖ Was a consecutive or random sample of patients enrolled?	<u>Yes</u> /No/Unclear
❖ Was a case-control design avoided?	<u>Yes</u> /No/Unclear
❖ Did the study avoid inappropriate exclusions?	<u>Yes</u> /No/Unclear

Could the selection of patients have introduced bias? RISK: LOW/HIGH/UNCLEAR

B. Concerns regarding applicability

Describe included patients: Pregnants with singleton pregnancies without congenital malformations and be scanned between September 8th 2014 and October 18th 2014.. Within this dataset, a subsample of women was selected who met, as closely as possible, the strict inclusion criteria of the Fetal Growth Longitudinal Study (FGLS) of the IG project. For the individual level criteria, we similarly included women who were at low risk of adverse maternal (e.g. absence of smoking, diabetes, hypertension, renal disease, maternal symptoms of preeclampsia or other medical conditions associated with fetal growth disturbances) and perinatal outcomes (patients referred for suspicion of fetal anomaly, history of fetal anomaly or IUGR). Thus, only women older than or equal to 18 years and younger than or equal to 35 years, who had a body mass index (BMI) greater than or equal to 18.5 kg/m2 and less than 30 kg/m2, with no clinically relevant obstetric or gynaecological history, who initiated antenatal care at less than 14 weeks' gestation were used in the subsequent analysis. We also a priori excluded measurements below or above -5 and +5 SD respectively.

Is there concern that the included patients do not match the review question?
CONCERN: LOW/HIGH/UNCLEAR

DOMAIN 2: INDEX TEST(S)

If more than one index test was used, please complete for each test.

A. Risk of Bias

Describe the index test and how it was conducted and interpreted: All scans were performed based on a routine indication. In France, 3 scans are planned for all pregnancies at 11-14 weeks, 21-24 weeks and 30-34 weeks as part of routine follow-up. To assess the similarity of size in French fetuses we decided to use head circumference (HC), the main fat-free skeletal measure of linear growth for comparison. First, the charts obtained from fitting an "IG-like" fractional polynomial regression model (i.e. the same functional form defined by the set of fractional polynomial powers) to our newly collected data were visually compared to those from the IG Standards. Similarities in the French and IG populations were analysed based on standardized differences, at various gestational age windows, for the fetal HC. Standardized differences are therefore similar to a Z score. Standardized difference values of less than 0.5 was pre-specified as adequate for reflecting that the French population could have been included in the pooled dataset used to construct the international standards.

❖ Were the index test results interpreted without knowledge of the results of the reference standard? Yes/No/Unclear

❖ If a threshold was used, was it pre-specified? Yes/No/Unclear

Could the conduct or interpretation of the index test RISK: LOW /HIGH/UNCLEAR have introduced bias?

B. Concerns regarding applicability

Is there concern that the index test, its conduct, or interpretation differ from the review question?

CONCERN:
LOW /HIGH/UNCLEAR

DOMAIN 3: REFERENCE STANDARD

A. Risk of Bias

Describe the reference standard and how it was conducted and interpreted: Similarities in the French and IG populations were analysed based on standardized differences, at various gestational age windows, for the fetal HC. To assess the similarity of size in French fetuses we decided to use head circumference (HC), the main fat-free skeletal measure of linear growth for comparison. Standardized differences are therefore similar to a Z score. Standardized difference values of less than 0.5 was pre-specified as adequate for reflecting that the French population could have been included in the pooled dataset used to construct the international standards.

❖ Is the reference standard likely to correctly classify the target condition? Yes/No/Unclear

❖ Were the reference standard results interpreted without knowledge of the results of the index test? Yes/No/Unclear

Could the reference standard, its conduct, or its interpretation have introduced bias? RISK: LOW /HIGH/UNCLEAR

B. Concerns regarding applicability

Is there concern that the target condition as defined by the reference standard does not match the review question?

CONCERN:
LOW /HIGH/UNCLEAR

DOMAIN 4: FLOW AND TIMING

A. Risk of Bias

Describe any patients who did not receive the index test(s) and/or reference standard or whowere excluded from the 2x2 table (refer to flow diagram): There were 8784 scans and of these 1689 were excluded because they were indicated (maternal conditions (n=1,090), maternal symptoms (311), or due to previous history or suspected fetal anomaly or IUGR (599)). Of the 7095 routinely performed scans, 1248 women/examinations (14.2%) were excluded because of maternal age >35 (n=1,220) or <18 (n=28), plus 961 (10.9%) because of an abnormal BMI (n=193 and 768 <18.5 and >30 respectively). Lastly, 28 examinations were excluded because of measurement values below or above -5 or +5 SD

Describe the time interval and any interventions between index test(s) and reference standard: At the same time.

- ❖ Was there an appropriate interval between index test(s) and reference standard? **Yes/No/Unclear**
- ❖ Did all patients receive a reference standard? **Yes/No/Unclear**
- ❖ Did patients receive the same reference standard? **Yes/No/Unclear**
- ❖ Were all patients included in the analysis? **Yes/No/Unclear**

Could the patient flow have introduced bias?

RISK: LOW /HIGH/UNCLEAR

Third trimester ultrasound scan combined with a clinical method for accurate birthweight prediction at term: a cohort study in Spain (Vila-Candel et al)

- **Phase 1:** State the review question:
 - o Patients: A total of 1,224 women with singleton pregnancies who had undergone foetal ultrasound scanning (USS) at 34 weeks were recruited.
 - o Index test(s): Third trimester ultrasound biometry (combined method).
 - o Reference standard and target condition: Hadlock
- **Phase 2:** Draw a flow diagram for the primary study



- **Phase 3:** Risk of bias and applicability judgments. QUADAS-2 is structured so that 4 key domains are each rated in terms of the risk of bias and the concern regarding applicability to the research question (as defined above). Each key domain has a set of signalling questions to help reach the judgments regarding *bias and applicability*.

DOMAIN 1: PATIENT SELECTION

A. Risk of Bias

Describe methods of patient selection: We performed a prospective cohort study in women seen at La Ribera University Hospital (LRUH) in Valencia (Spain) for pregnancy follow-up and delivery between January and December 2016. During the study period, 2,017 women consented to be included in the study. A representative sample size was calculated. The null hypothesis was that there were no statistically significant differences in birthweight accuracy between the clinical and sonographic methods. Assuming an expected difference between both estimates of weight (EFWm and EFWa) greater than or equal to 200 g to consider statistically significant differences, accepting an alpha risk of 0.05 and beta of 0.2 in a two-sided test, a common standard deviation of 450, and a drop-out rate of 30%, then the necessary sample would be 39 participants within each group. However, all pregnant women who agreed to participate during the study period (one year) were included.

- ❖ Was a consecutive or random sample of patients enrolled? **Yes/No/Unclear**
- ❖ Was a case-control design avoided? **Yes/No/Unclear**
- ❖ Did the study avoid inappropriate exclusions? **Yes/No/Unclear**

Could the selection of patients have introduced bias? RISK: LOW/HIGH/UNCLEAR

B. Concerns regarding applicability

Describe included patients: Women with a first prenatal appointment and USS between 5 and 12 weeks of pregnancy, single-foetus pregnancy with no foetal abnormalities, and birth between 38 and 40 weeks.

**Is there concern that the included patients do not match the review question?
CONCERN: LOW/HIGH/UNCLEAR**

DOMAIN 2: INDEX TEST(S)

If more than one index test was used, please complete for each test.

A. Risk of Bias

Describe the index test and how it was conducted and interpreted: Ultrasound estimated foetal weight (EFW) was calculated using the Hadlock 2 formula (8,20), using four foetal indices: Biparietal diameter (BPD), femur length (FL) and abdominal circum-ference (AC), recorded at 34 weeks at LRUH (Log 10 weight = $1,326 - 0.00326 AC \times FL + 0.0107 HC + 0.0438AC + 0.158 \times FL$). Ultrasound estimated foetal weight (EFW) was calculated using the Hadlock 2 formula (8,20), using four foetal indices: Biparietal diameter (BPD), femur length (FL) and abdominal circum-ference (AC), recorded at 34 weeks at LRUH (Log 10 weight = $1,326 - 0.00326 AC \times FL + 0.0107 HC + 0.0438AC + 0.158 \times FL$). The ultrasound estimations (EFWa) were each compared to the actual birthweight. Finally, methods were compared: Estimated foetal weight by GAP method (EFW_a) (USS at third trimester with the Hadlock 2 formula adjusted by gestational age at birth and foetal gender), and estimated foetal weight with the combined clinical and GAP method (EF- Wmr) (multivariate lineal regression method) were calculated to compare with the actual birthweight.

❖ Were the index test results interpreted without knowledge of the results of the reference standard? Yes/No/Unclear

❖ If a threshold was used, was it pre-specified? Yes/No/Unclear

Could the conduct or interpretation of the index test RISK: LOW /HIGH/UNCLEAR have introduced bias?

B. Concerns regarding applicability

Is there concern that the index test, its conduct, or CONCERN: LOW /HIGH/UNCLEAR interpretation differ from the review question?

DOMAIN 3: REFERENCE STANDARD

A. Risk of Bias

Describe the reference standard and how it was conducted and interpreted: The accuracy of the different methods (EFW_a/ EFW_{mr}) for estimating foetal weight was then evaluated by calculating the Pearson correlation R coefficient between the estimated foetal weights obtained using each equation and actual birthweight. Intraclass Correlation Coefficient (ICC) was used to evaluate the degree of agreement of both methods with the actual neonatal weight. Absolute error was defined as the absolute value of EFW_a//EFW_{mr} minus the actual birthweight, and the relative error value as the absolute EFW_a//EFW_{mr} error divided by the birthweight multiplied by 100.

❖ Is the reference standard likely to correctly classify the target condition? Yes/No/Unclear

❖ Were the reference standard results interpreted without knowledge of the results of the index test? Yes/No/Unclear

Could the reference standard, its conduct, or its interpretation have introduced bias? RISK: LOW /HIGH/UNCLEAR

B. Concerns regarding applicability

Is there concern that the target condition as defined by the reference standard does not match the review question? CONCERN: LOW /HIGH/UNCLEAR

DOMAIN 4: FLOW AND TIMING

A. Risk of Bias

Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2x2 table (refer to flow diagram): Of the initial 2,017 pregnant women delivered at the LRUH during the study period, 458 women (22.7%) did not meet the inclusion criteria: 36 (7.9%) twin pregnancies, 27 (5.9%) <35 weeks, 334 (72.9%) >40 weeks at birth, 22 (4.8%) polyhydramnios, 34 (7.4%) oligohydramnios, and five (1.0%) foetal deformities. Missing data from antenatal maternity records resulted in 335 women (16.6%) subsequently excluded from the final analysis: 108 (32.2%) without documented third trimester US, and 227 (67.8%) without SFH recorded.

Describe the time interval and any interventions between index test(s) and reference standard: At the same time

- ❖ Was there an appropriate interval between index test(s) and reference standard? Yes/No/Unclear
- ❖ Did all patients receive a reference standard? Yes/No/Unclear
- ❖ Did patients receive the same reference standard? Yes/No/Unclear
- ❖ Were all patients included in the analysis? Yes/No/Unclear

Could the patient flow have introduced bias?

RISK: LOW /HIGH/UNCLEAR