UNIVERSIDADE ESTADUAL DE CAMPINAS
FACULDADE DE CIÊNCIAS MÉDICAS

MARCELA MARIA SIMÕES CORA

Consenso Delphi para o diagnóstico de restrição de crescimento fetal e desfechos perinatais adversos: revisão sistemática e recomendações para a prática clínica

CAMPINAS
2023

MARCELA MARIA SIMÕES CORA

Consenso Delphi para o diagnóstico de restrição de crescimento fetal e desfechos perinatais adversos: revisão sistemática e recomendações para a prática clínica

Dissertação apresentada à Faculdade de Ciências Médicas da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do título de Mestra em Ciências, área de concentração Qualificação dos Processos Assistenciais.

ORIENTADOR: PROF. DR. JOÃO RENATO BENNINI JUNIOR

ESTE TRABALHO CORRESPONDE À VERSÃO
FINAL DA DISSERTAÇÃO DEFENDIDA PELA
ALUNA MARCELA MARIA SIMÕES CORA, E ORIENTADA
PELO PROF. DR. JOÃO RENATO BENNINI JUNIOR.

CAMPINAS
2023

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca da Faculdade de Ciências Médicas
Maristella Soares dos Santos - CRB 8/8402

Si51c

Simões Cora, Marcela Maria, 1991-
Consenso Delphi para o diagnóstico de restrição de crescimento fetal e desfechos perinatais adversos : revisão sistemática e recomendações para a prática clínica / Marcela Maria Simões Cora. – Campinas, SP : [s.n.], 2023.

Orientador: João Renato Bennini Júnior.
Dissertação (mestrado profissional) – Universidade Estadual de Campinas, Faculdade de Ciências Médicas.

1. Restrição de crescimento fetal. 2. Insuficiência placentária. 3. Diagnóstico. 4. Desfechos perinatais. 5. Revisão sistemática. 6. Guia de prática clínica. I. Bennini Júnior, João Renato, 1978-. II. Universidade Estadual de Campinas. Faculdade de Ciências Médicas. III. Título.

Informações Complementares

Título em outro idioma: Delphi consensus for the diagnosis of fetal growth restriction and adverse perinatal outcomes : systematic review and recommendations for clinical practice
Palavras-chave em inglês:
Fetal growth restriction
Placental insufficiency
Diagnosis
Perinatal outcomes
Systematic review
Practice guideline
Área de concentração: Qualificação dos Processos Assistenciais
Titulação: Mestra em Ciências
Banca examinadora:
João Renato Bennini Júnior [Orientador]
Carlos Roberto Silveira Correa
Alessandra Cristina Marcolin
Data de defesa: 27-03-2023
Programa de Pós-Graduação: Ciência Aplicada à Qualificação Médica

Identificação e informações acadêmicas do(a) aluno(a)
- ORCID do autor: https://orcid.org/0000-0003-2806-9686
- Currículo Lattes do autor: http://lattes.cnpq.br/2430975375617539

COMISSÃO EXAMINADORA DA DEFESA DE MESTRADO

ALUNA: Marcela Maria Simões Cora

ORIENTADOR: Prof Dr João Renato Bennini Junior

MEMBROS TITULARES:

1. **Prof. Dr. João Renato Bennini Junior**
2. **Prof. Dr. Carlos Roberto Silveira Correa**
3. **Profa. Dra. Alessandra Cristina Marcolin**

Programa de Pós-Graduação em Qualificação dos Processos Assistenciais da Faculdade de Ciências Médicas da Universidade Estadual de Campinas.

A ata de defesa com as respectivas assinaturas dos membros encontra-se no SIGA/Sistema de Fluxo de Dissertação/Tese e na Secretaria do Programa da FCM.

Data de Defesa: 27/03/2023

À minha avó Marilene Rodrigues de Jesus,

que desde muito pequena reconhecia o poder do conhecimento

e soube, ao seu modo, passar seus valores para muitas gerações.

Inspiração de filhos, sobrinhos, netos e bisnetos.

AGRADECIMENTOS

Quando iniciei este trabalho eu sabia do aprendizado acadêmico que viria, mas não imaginava o tamanho da transformação pela qual passaria. Foram anos de aprimoramento profissional e pessoal, e ao longo deste caminho pude contar com pessoas que não só me incentivaram mas foram inspiração e a quem eu quero deixar os meus agradecimentos.

Quero agradecer ao meu orientador Professor Dr João Renato Bennini Junior, por tamanha disponibilidade para me ajudar até mesmo nos horários menos convencionais. Sempre que achei que não conseguiria, ele trouxe motivação de sobra pra continuarmos e irmos além.

À minha família que em todos os meus passos foi e é suporte que me impulsiona. Em especial agradeço à minha filha Luísa Maria Simões Cora, que nasceu durante a realização deste trabalho e, mesmo sem entender, é a maior razão de tudo o que eu faço desde então. Quero agradecer meu marido Caio Fiori Cora, que mesmo sendo da área contábil entende muito deste trabalho após inúmeras "aulas", desabafos, preparações e opiniões.

Não posso deixar de mencionar os meus pais, Antonio de Oliveira e Simões Filho e Maria Lucia de Jesus e Simões, que há muitos anos me deram a preciosidade de um norte e uma base sólida. Se cheguei a qualquer lugar foi porque tive exemplo, estímulo e muito carinho.

Às minhas irmãs, Juliana Carolina Simões e Fernanda Pessorruso, as quais, cada qual com a sua maneira, me ensinam a seguir por caminhos mais alegres e seguros, depois de tê-los aberto por suas próprias experiências.

À minha colega de profissão, de residência e desta aventura, Marcela Oliveira Carniello, que tanto no trabalho quanto na vida acadêmica é uma grande inspiração.

O início deste trabalho com a elaboração da estratégia de busca não teria sido possível sem o apoio da bibliotecária Ana Paula de Morais e Oliveira, a ela também sou muito grata.

Agradeço à Professora Dra Maria Laura Costa do Nascimento, que desde o começo se interessou em me direcionar ao meio acadêmico e sempre está disponível para me ajudar e dar dicas e orientações preciosas para o trabalho e para a vida.

À minha segunda casa UNICAMP, onde vivi 11 anos de transformação e de onde eternamente carregarei o orgulho por ter tido a oportunidade de aprender com os mais excelentes mestres, em um ambiente onde a ciência, o respeito e o comprometimento prosperam.

A todos estes e outros que indiretamente me ajudaram nesta caminhada, o meu muito obrigado!

RESUMO

Introdução: A restrição de crescimento fetal (RCF) é a incapacidade do feto em atingir seu potencial de crescimento determinado geneticamente e afeta cerca de 5% de todas as gestações. Atualmente não existe padrão ouro para o diagnóstico de RCF. Em 2016, uma equipe multicêntrica de especialistas em medicina fetal realizou um estudo baseado no método Delphi para estabelecer um consenso sobre a definição de RCF, dividindo-a em precoce (menos de 32 semanas de idade gestacional) e tardia (32 semanas ou mais) e utilizando uma combinação de parâmetros biométricos fetais e de Dopplervelocimetria materno-fetal. **Objetivos:** revisar a literatura e elaborar recomendações acerca do uso dos critérios propostos pelo consenso Delphi para o diagnóstico RCF. **Métodos:** Revisão sistemática de acordo com o protocolo de revisão sistemática de acurácia e testes diagnósticos do grupo Cochrane que incluiu estudos observacionais analíticos contendo informações sobre a acurácia diagnóstica dos critérios de consenso Delphi para RCF ou dados a partir dos quais esses valores pudessem ser calculados. As pesquisas foram realizadas nas bases PubMed, EMBASE, PMC e Cochrane Library, entre 2010 e 2020. A ferramenta QUADAS-2 foi utilizada para analisar o risco de viés. Os resultados encontrados nos estudos foram sintetizados em uma tabela que contemplou todos os critérios propostos pelo consenso Delphi para o diagnóstico de RCF em relação a desfechos perinatais adversos, com a representação de verdadeiros positivos (VP), falsos positivos (FP), verdadeiros negativos (VN) e falsos negativos (FN), mostrados em forest-plots com seus valores de sensibilidade, especificidade e razões de verossimilhança positiva e negativa com os respectivos intervalos de confiança de 95%. O método GRADE foi utilizado para avaliação da qualidade da evidência e para estruturar as recomendações seguindo o formato do RIGHT Statement. **Resultados:** Foram encontrados 1.521 artigos nas bases de dados pesquisadas que foram reduzidos a 53 estudos após a leitura do título e resumo; após leitura na íntegra, 12 estudos foram selecionados para a coleta de dados. Para responder à pergunta: "Para o diagnóstico de RCF, devemos utilizar os critérios do

consenso Delphi ou PFE > percentil 10 para a IG?” a revisão sistemática encontrou apenas um estudo comparando ambos os testes, com baixa qualidade de evidência. Neste estudo demonstrou-se que ambos os métodos diagnósticos têm precisão semelhante para o diagnóstico de desfechos perinatais adversos. Assim , equiparando os possíveis benefícios e riscos dos resultados, a força da evidência se baseou principalmente nos possíveis custos e preferências e a recomendação de utilizar o consenso Delphi foi considerada fraca. **Conclusão:** Após anos desde a publicação do consenso Delphi para o diagnóstico de RCF, apenas 1 estudo buscou validá-lo e concluiu que ele apresenta uma taxa discretamente maior de detecção de desfechos perinatais adversos quando comparada à definição do PFE <p10. Apesar da baixa qualidade e evidência e força de recomendação fraca, sugerimos que, em locais com profissionais e equipamentos disponíveis, utilizem-se os critérios de consenso Delphi como primeira escolha para o diagnóstico de RCF.

Esta pesquisa não recebeu financiamento específico de nenhuma agência nos setores público, comercial ou sem fins lucrativos. O protocolo de revisão sistemática foi registrado na plataforma PROSPERO sob número: CRD42020204051.

Palavras - chave (DeCS): Restrição de Crescimento Fetal; Insuficiência placentária; Diagnóstico; Desfechos perinatais; Revisão sistemática; Guia de Prática Clínica.

ABSTRACT

Introduction: Fetal growth restriction (FGR) is the failure of the fetus to reach its genetically determined growth potential and is a condition that affects approximately 5% of all pregnancies. Currently, there is no gold standard for diagnosing RCF. In 2016, a multicenter team of fetal medicine specialists conducted a study based on the Delphi method to establish a consensus on the definition of FGR. They divided FGR into early (less than 32 weeks of gestational age) and late (32 weeks or more) and used a combination of fetal biometric and maternal-fetal Doppler velocimetry ultrasound parameters. **Objective:** to review the literature and formulate recommendations regarding the use of the criteria proposed by the Delphi consensus for the diagnosis of RCF. **Methods**: A systematic review was carried out according to the Cochrane Group's systematic review protocol of diagnostic tests and accuracy and included analytical observational studies containing information on the diagnostic accuracy of the Delphi consensus criteria for FGR or data from which these values could be calculated. Searches were performed in PubMed, EMBASE, PMC and Cochrane Library between 2010 and 2020. The QUADAS-2 tool was used to analyze the risk of bias. The results found in the studies were summarized in a table that included all the criteria proposed by the Delphi consensus for the diagnosis of FGR in association to adverse perinatal outcomes, with the representation of true positives (TP), false positives (FP), true negatives (TN) and false negatives (FN), shown in forest-plots with their values of sensitivity, specificity and positive and negative likelihood ratios with the respective 95% confidence intervals. The GRADE method was used to assess the quality of the evidence and to structure the recommendations following the RIGHT Statement format. **Results**: 1521 articles were found in the searched databases, which were reduced to 53 studies after reading the title and abstract; these papers were then read in full, with 12 studies selected for data collection. To answer the question: "For the diagnosis of FGR, should EFW < p1o or Delphi consensus criteria be used?" the systematic review found only one study comparing both tests, with low quality of evidence. In this comparative study, it was demonstrated that both diagnostic methods have similar accuracy for adverse perinatal outcomes. Thus,

considering the possible benefits and risks of the results, the strength of the evidence was mainly based on possible costs and preferences and the recommendation to use the Delphi consensus was considered weak. **Conclusion**: After 6 years since the publication of the Delphi consensus for the diagnosis of FGR, only one study sought to validate it and concluded that it has a slightly higher rate of detection of adverse perinatal outcomes when compared to the definition of EFW <p10. Despite the low quality of evidence and weak strength of recommendation, we suggest that, if professionals and equipment are available, the Delphi consensus criteria be used as the first choice for the diagnosis of FGR.

This study did not receive funding from development agencies in the public, commercial or non-profit sectors. The systematic review protocol was registered on the PROSPERO platform under number: CRD42020204051.

Keywords (DeCS): Fetal Growth Restriction; Placental insufficiency; Diagnosis; Perinatal Outcomes; Systematic review; Practice Guideline.

Ilustrações **Página**

Flowchart of article selection for extracting systematic review data 33

Risk of bias and applicability according to QUADAS. 35

Frequency of assessments of adverse perinatal outcomes among the studies included in the review. 37

Se and Sp in forest plots of each index test related to specific perinatal outcomes. 39

ROC curves for different criteria for the diagnosis of multiple fetal outcomes. 41

DOR results of each index test for the detection of adverse perinatal outcomes presented in forest plots. 43

DOR results of each adverse perinatal outcome in relation to the criteria proposed by the Delphi consensus presented in forest plots. 45

Should Delphi consensus vs EFW < p10 be used to diagnose FGR in singleton pregnancies? GRADE method. 70

Tabelas	**Página**
Characteristics of included studies	55
Delphi consensus criteria for the diagnosis of FGR	58
Criterion or combination of criteria (index test) evaluated by the included studies	59
Se and Sp of each index test related to perinatal outcomes and their respectives confidence intervals.	61
Positive and negative likelihood ratio of each index test related to perinatal outcomes and their respective confidence intervals.	62
Presumed impact of test results	70

Siglas e abreviaturas	**Significado**
ACOG	Colégio Americano de Ginecologia e Obstetrícia (do inglês: *American College of Obstetrician an Gynecolgy*)
AU	Artéria Umbilical
AUC	Área Sob a Curva (do inglês: *Area Under the Curve*)
AUt	Artéria Uterina
CA	Circunferência Abdominal
CMV	Citomegalovírus
DOR	Razão de Chances Diagnóstica (do inglês *Diagnostic Odds Ratio*)
DPA	Desfecho Perinatal Adverso
FDFA/R	Fluxo diastólico final reverso ou ausente na artéria umbilical
HI	Hemorragia intraventricular
IMC	Índice de Massa Corporal
IP	Índice de Pulsatilidade
LES	Lúpus Eritematoso Sistêmico
negLR	Razão de Verossimilhança Negativa (do inglês: *negative likehood ratio*)
OMS	Organização Mundial da Saúde
PFE	Peso Fetal Estimado
PIG	Pequeno para a Idade Gestacional

posLR	Razão de Verossimilhança Positiva (do inglês: *positive likehood ratio*)
RCF	Restrição de Crescimento Fetal
RCP	Relação Cérebro-Placentária
SAAF	Síndrome do Anticorpo Antifosfolípide
SARN	Síndrome da Angústia Respiratória do Recém Nascido
UTI	Unidade de Terapia Intensiva
UTIN	Unidade de Terapia Intensiva Neonatal

Sumário

1. Introdução 17
2. Objetivos 20
    2.1 Objetivos Geral 20
    2.2 Objetivos específicos 20
3. Métodos 21
    3.1. Revisão sistemática 21
  3.1.1 Protocolo e registro 21
  3.1.2. Critérios de elegibilidade 21
  3.1.3. Estratégia de busca 22
  3.1.4. Seleção dos estudos 22
  3.1.5. Extração dos dados 22
  3.1.6. Risco de viés e aplicabilidade 23
  3.1.7. Medidas de desempenho diagnóstico 23
  3.1.8. Síntese dos resultados 23
  3.1.9. Meta-análise 24
    3.2. Formulação das recomendações 25
  3.2.1 Modelo de redação 25
  3.2.2 Definição do foco das recomendações 25
  3.2.3 Cenário e população para as quais as recomendações são viáveis 25
  3.2.4. Definição do grupo de trabalho 25
  3.2.5. Obtenção das evidências e formulação das recomendações 26
4. Resultados 27
    4.1 Revisão sistemática 27
    4.3 Recomendações 63
5. Discussão 74
    5.1 Resumo das Evidências 74
    5.2 Recomendações 75
6. Conclusão 77
7. Referências bibliográficas 78
Apêndice 1: Estratégias de buscas utilizadas em cada base de dados 83
Apêndice 2: Artigos selecionados por título e resumo. 87
Apêndice 3: Artigos excluídos e motivo da exclusão 89
Apêndice 4: Tabela com extração de dados dos artigos selecionados. 90
Anexo 1: Cover Letter da submissão para o American Journal of Obstetrics and Gynecology 102

1. Introdução

A restrição de crescimento fetal (RCF) é definida como a falha do feto em atingir seu potencial de crescimento geneticamente determinado [1], sendo uma condição que afeta cerca de 5% de todas as gestações, com aumento progressivo em sua incidência ao longo das décadas [2]. A etiologia da RCF está associada a vários fatores, que podem ser fetais, placentários ou maternos.

Em relação às etiologias fetais, consideramos alterações cromossômicas, síndromes genéticas, malformações congênitas, erros inatos do metabolismo e infecções intrauterinas, estas últimas associadas a cerca de 10% dos fetos com RCF, principalmente rubéola, toxoplasmose, varicela, herpes zoster e citomegalovírus (CMV) [3,4]. Ainda, sabe-se que uma proporção maior das gestações múltiplas apresenta restrição de crescimento, sendo mais comum nos gêmeos monocoriônicos com transfusão fetal-fetal [5].

Considerando os fatores etiológicos maternos, várias doenças estão associadas, principalmente aquelas que envolvem alterações vasculares, como distúrbios hipertensivos, diabetes mellitus, doenças autoimunes (como a síndrome do anticorpo antifosfolípide - SAAF - e lúpus eritematoso sistêmico - LES) [6-9]. Além disso, o uso de substâncias como álcool, cocaína e tabaco aumentam o risco de RCF, assim como o uso de medicamentos anticonvulsivantes, varfarina, antineoplásicos, antagonistas do ácido fólico, exposição à radiação e vida em grandes altitudes [4].

Embora existam vários componentes etiológicos possíveis para RCF, de todas as condições maternas, as doenças hipertensivas são as condições que mais aumentam o risco de RCF, que nestes casos é cerca de 3 a 4 vezes mais frequente [8,10].

Gestações com alterações na vascularização placentária são sabidamente associadas a piores desfechos maternos e fetais [11] como sofrimento fetal agudo, necessidade de cesariana de urgência, admissão de neonato em unidade de cuidados intensivos (UTI) neonatal, prematuridade iatrogênica, complicações e morte neonatal e óbito

fetal [12-13] A insuficiência placentária pode estar associada ao início da fisiopatologia de condições maternas que se associam à RCF (como os distúrbios hipertensivos incluindo a pré-eclâmpsia e seus espectros de gravidade) ou ainda ser consequência dessas condições, sendo a principal causa de RCF [14].

Determinar quais fetos e/ou gestantes apresentam maiores riscos de desfechos adversos é necessário para buscar o melhor acompanhamento desses casos e decidir o tempo adequado entre a avaliação da vitalidade fetal e o parto, no momento em que os riscos da permanência intrauterina superam aqueles relacionados à prematuridade com base na idade gestacional, etiologia da RCF, grau de comprometimento da vitalidade fetal, experiência e recursos tecnológicos disponíveis.

Apesar da grande incidência da RCF e sua importância clínica no cuidado materno e fetal, o seu diagnóstico ainda não é bem definido e existem definições distintas utilizadas na prática clínica e nos grupos de pesquisa. Mesmo entre grandes instituições de saúde, não há consenso definido: para a Organização Mundial da Saúde (OMS), RCF é quando o peso fetal estimado (PFE) está abaixo do percentil 3 para sua idade gestacional (IG). O Colégio Americano de Obstetrícia e Ginecologia (ACOG), define RCF como peso fetal estimado PFE abaixo do percentil 10 para a IG [15,16].

Em 2016, uma equipe multicêntrica de especialistas internacionais em RCF realizou um estudo fundamentado no método Delphi [17] a fim de estabelecer uma definição para RCF a partir da opinião de especialistas [18]. Consensuou-se RCF precoce como aquela iniciada abaixo de 32 semanas de IG, e tardia, a partir de 32 semanas, uma vez que próximo a esta idade gestacional o crescimento fetal passa a se associar principalmente ao processo de hipertrofia celular, de modo a justificar as diferenças fisiopatológicas entre os fetos restritos precoce e tardios. Os critérios propostos devem ser considerados apenas na ausência de malformações fetais que justifiquem a alteração de crescimento, a fim de detectar, com o proposto pelo consenso Delphi, a restrição de crescimento associada a insuficiência placentária. Assim, a partir desta definição, considera-se RCF precoce o feto que antes de 32 semanas apresente algum dos seguintes critérios: (i) circunferência

abdominal (CA) abaixo do percentil 3 para IG; (ii) PFE abaixo do percentil 3 para a IG; (iii) diástole reversa ou ausente na artéria umbilical (AU); (iv) PFE ou CA abaixo do percentil 10 para IG associado índice de pulsatilidade (IP) da AU ou da artéria uterina (AUt) acima do percentil 95 para IG. Ainda por este consenso, a RCF tardia é definida quando o feto, após 32 semanas de IG, apresente algum dos critérios: (i) CA abaixo do percentil 3 para IG; (ii) PFE abaixo do percentil 3 para IG; (iii) combinação de 2 ou mais achados entre (a) PFE ou CA abaixo do percentil 10 para a IG, (b) redução de 2 ou mais quartis na curva de crescimento do PFE ou da CA, (c) relação cerebroplacentária (RCP) abaixo do percentil 5 para a IG ou IP da AU acima do percentil 95 para IG.

Assim, o consenso Delphi engloba critérios complexos, que por sua vez são definidos com base em uma série de características fetais e placentárias. Considerando que grandes instituições recomendam o uso destes critérios para o diagnóstico de RCF é relevante revisar a literatura e identificar estudos que os tenham validado na prática clínica, a fim de se obter dados agrupados mais consistentes que possam nortear a sua aplicação.

2. Objetivos

2.1 Objetivos Geral

Revisar a literatura e elaborar recomendações acerca do uso dos critérios propostos pelo consenso Delphi para o diagnóstico RCF.

2.2 Objetivos específicos

- Realizar uma revisão sistemática de estudos de validação de teste diagnóstico avaliando o desempenho dos critérios propostos pelo consenso Delphi para o diagnóstico de RCF e detecção de desfechos perinatais adversos.
- Através dos dados obtidos em revisão, realizar o cálculo da sensibilidade, especificidade, valor preditivo positivo, valor preditivo negativo e razão de chances diagnóstica ("*diagnostic odds ratio")* dos critérios propostos pelo consenso Delphi para o diagnóstico de RCF e detecção de desfechos perinatais adversos.
- Elaborar recomendações para o uso na prática clínica dos critérios propostos pelo consenso Delphi para o diagnóstico de RCF baseadas na evidência científica obtida.

3. Métodos

3.1. Revisão sistemática

3.1.1 Protocolo e registro

Esta revisão sistemática foi realizada de acordo com os protocolos de revisões sistemáticas de testes de precisão diagnóstica do grupo Cochrane [19,20]. Uma vez elaborado, o protocolo de pesquisa foi registrado na plataforma PROSPERO (https://www.crd.york.ac.uk/prospero) antes do início das pesquisas (número de registro: CRD42020204051).

3.1.2. Critérios de elegibilidade

Foram analisados estudos observacionais analíticos prospectivos ou retrospectivos dos últimos 10 anos contendo informações sobre a validação diagnóstica dos critérios do consenso Delphi para RCF (sensibilidade, especificidade) ou dados a partir dos quais esses valores pudessem ser calculados. A pergunta da pesquisa foi idealizada de acordo com a modalidade estruturada que envolve: população, teste índice, padrão de referência e condição alvo. População: mulheres com gestações únicas, de fetos sem malformações, distúrbios cromossômicos ou síndromes genéticas, as quais foram submetidas a pelo menos um ultrassom obstétrico durante a gravidez; como testes de índice, foram considerados os critérios escolhidos pelo consenso Delphi [18] para o diagnóstico de RCF (patologia a ser identificada). Uma vez que fetos verdadeiramente restritos apresentam maior risco de desfechos perinatais adversos, como padrão-ouro para o teste diagnóstico (verdadeiros positivos) consideramos a presença de desfecho perinatal adverso. Apenas artigos publicados em inglês foram incluídos. Não foram incluídos artigos não publicados, revisões de literatura, revisões sistemáticas, meta-análises, opiniões de especialistas, cartas ao editor, comentários sobre estudos ou resumos de apresentações em congressos e simpósios.

3.1.3. Estratégia de busca

As buscas foram realizadas nas bases de dados PubMed, EMBASE, PMC e Cochrane Library, incluindo estudos de 2010 a 2020 (último dia de busca em 17 de agosto de 2020). Por se tratar de uma "revisão rápida", não foram incluídos estudos não publicados e os autores não foram contatados durante as buscas.

A seguinte estratégia de busca foi utilizada: ("Fetal Growth Retardation" AND Diagnosis) AND ("Fetal Weight" OR "Ultrasonography, Doppler" OR "Delphi procedure" OR DELPHI OR "Consensus Delphi"). As estratégias para todas as bases de dados eletrônicas, incluindo quaisquer limites usados, são apresentadas no apêndice 1.

3.1.4. Seleção dos estudos

O processo de seleção dos estudos foi dividido em duas fases: inicialmente a partir dos títulos e resumos e seguido da leitura do artigo na íntegra. A seleção por títulos e resumos incluiu estudos que envolveram gestações únicas e fetos sem malformações e que avaliaram os resultados de US obstétrico com Doppler em relação aos desfechos perinatais, contendo informações sobre pelo menos um dos critérios propostos pelo consenso Delphi para o diagnóstico de RCF. A análise dos artigos na íntegra selecionou estudos que apresentassem resultados de acurácia diagnóstica (sensibilidade, especificidade) ou dados a partir dos quais esses valores pudessem ser calculados para um ou mais critérios propostos pelo consenso Delphi para o diagnóstico de RCF.

3.1.5. Extração dos dados

Após a aplicação da estratégia de busca, os artigos foram incluídos na plataforma ENDNOTE (www.myendnoteweb.com) e as primeiras duplicidades identificadas foram excluídas. Os artigos foram então enviados para a plataforma RAYYAN QRI (www.rayyan.qcri.org) [21], onde as novas duplicatas foram identificadas e excluídas. Os artigos foram selecionados por dois revisores [M.M.S.C] e [J.R.B.J.]: na primeira fase pelo título e resumo e na segunda fase pelo texto completo, a partir de formulário desenhado

especificamente para isso, com cegamento entre os revisores. Os conflitos foram resolvidos com base no consenso entre os dois pesquisadores.

Uma vez selecionados os artigos, os dados extraídos incluíram o nome do primeiro autor e ano de publicação, delineamento do estudo, total de participantes, IG mínima para inclusão no estudo, identificação de quais critérios para RCF propostos pelo consenso Delphi foram estudados, número de fetos com PFE abaixo do 10º percentil e 3º percentil, número de fetos alterações no Doppler e identificação dessas alterações, número de fetos que tiveram resultados perinatais adversos e cálculo da especificidade, sensibilidade, razão de verossimilhança positiva e positiva e negativa para cada um dos critérios e resultados avaliados.

3.1.6. Risco de viés e aplicabilidade

Os parâmetros sugeridos pela ferramenta QUADAS-2 [22] foram utilizados para análise do risco de viés em cada estudo incluído na revisão. A análise foi realizada por ambos os pesquisadores, e as diferenças foram resolvidas por consenso.

3.1.7. Medidas de desempenho diagnóstico

As medidas de desempenho calculadas dos testes diagnósticos (para cada critério e desfecho), com seus respectivos intervalos de confiança de 95%, foram as seguintes: sensibilidade, especificidade, razão de verossimilhança positiva e negativa e razão de chances diagnóstica (*diagnostic odds ratio*).

3.1.8. Síntese dos resultados

Os resultados encontrados nos estudos foram resumidos em uma tabela que contemplou todos os critérios propostos pelo consenso Delphi para o diagnóstico de RCF, que foram estudados em relação aos desfechos avaliados: pequeno para a idade gestacional (PIG - peso ao nascer abaixo do percentil 10), admissão no uma unidade de terapia intensiva neonatal (UTIN), morte neonatal, nascimento prematuro, síndrome da angústia respiratória (SARN), Apgar no quinto minuto abaixo de 7, acidose neonatal, óbito

fetal, cesariana de emergência, hemorragia intraventricular (HI) e qualquer desfecho perinatal adverso (DPA).

3.1.9. Meta-análise

Para a análise descritiva, foram calculadas as sensibilidades e as especificidades de cada critério em cada estudo com seus respectivos intervalos de confiança de 95% e apresentados em forest plots. Para determinar a acurácia geral, nós calculamos a Razão de Chances Diagnóstica (do inglês “Diagnostic Odds Ratio” - DOR) usando o modelo de efeitos aleatórios de DerSimonian-Laird, e os forest plots foram construídos representando a metanálise realizada, com seus respectivos intervalos de confiança, agrupando os critérios do consenso Delphi para o diagnóstico de RCF e os desfechos de interesse. A DOR é uma medida de efetividade de um teste diagnóstico; onde DOR = 1 significa nenhum valor diagnóstico, DOR > 1 significa que um teste positivo sugere doença positiva e DOR < 1 significa que um teste negativo sugere doença negativa.

As sensibilidades e as especificidades agrupadas foram obtidas usando o modelo bivariado de Reitsma [23].

Os valores verossimilhança positiva (do inglês positive likehood ratio – posLR) e verossimilhança negativa (do inglês negative likehood ratio – negLR) foram calculados a partir dos dados de sensibilidade e especificidade obtidos dos estudos. A PosLR é definida como “sensibilidade/(1-Especificidade)” e negLR definido por “(1-Especificidade)/Especificidade”; seus valores representam como o resultado do teste muda a probabilidade da doença. Deste modo, posLR na escala de <2, 2 até 5, >5 até 10 e >10 são reconhecidos como um aumento da probabilidade “não significativo”, “pequeno”, “moderado” e “grande”, respectivamente. Para negLR, valores na escala de >0.5, >0,2 até 0,5, 0,1 até 0,2 e < 0,1 representa “não significativa”, “pequena”, “moderada” e “grande” diminuição da probabilidade respectivamente.

A comparação entre as sensibilidades e especificidades obtidas pelo modelo de Reitsma foi realizada por meio de uma análise de metarregressão considerando os critérios como variáveis preditoras dos Logits de sensibilidade e especificidade.

Todas as análises foram realizadas utilizando o pacote “mada”, “meta” e R environment de computação estatística [23].

3.2. Formulação das recomendações

3.2.1 Modelo de redação

As recomendações clínicas foram redigidas de acordo com as orientações do RIGHT Statement [24].

3.2.2 Definição do foco das recomendações

Trata-se de recomendações sobre o uso dos critérios propostos pelo consenso Delphi para o diagnóstico de RCF para detecção de desfechos perinatais adversos em gestações únicas sem malformações fetais.

3.2.3 Cenário e população para as quais as recomendações são viáveis

As presentes recomendações são válidas para a investigação diagnóstica por US quanto à presença de RCF em gestantes, com feto único sem malformações, atendidas pelo serviço de Diagnóstico por Imagem do Hospital da Mulher da UNICAMP.

3.2.4. Definição do grupo de trabalho

O grupo de trabalho foi formado pelos pesquisadores M.M.S.C. e J.R.B. (envolvidos na revisão mencionada previamente), com auxílio da equipe de Diagnóstico por Imagem do Hospital da Mulher da UNICAMP para revisão das recomendações.

3.2.5. Obtenção das evidências e formulação das recomendações

As evidências foram extraídas da revisão sistemática previamente mencionada (item 3.1) e as recomendações foram redigidas seguindo os passos do RIGHT Statement [24] a partir da metodologia e ferramenta GRADE [25,26].

4. Resultados

4.1 Revisão sistemática

Delphi consensus for the diagnosis of fetal growth restriction: systematic review and meta-analysis

Short title: Review: Delphi consensus for FGR diagnosis

PROSPERO registration number: CRD42020204051

Manuscript word count: 3939

A preview of this study was presented on the 26th. São Paulo Congress of Obstetrics and Gynecology that happened online from August 19 to December 12, 2021.

Authors

Marcela Maria Simões-Cora[1], Luís Otávio Sarian[2], João Renato Bennini[1]

[1]Fetal Medicine Program, Department of Gynecology and Obstetrics, Women's Hospital, State University of Campinas

[2]Department of Gynecology and Obstetrics, Women's Hospital, State University of Campinas

Corresponding author

João Renato Bennini, head of the Fetal Medicine Program, Department of Gynecology and Obstetrics, State University of Campinas.

Alexander Fleming Street, 101, Department of Maternal-Fetal Medicine, Unicamp Women's Hospital, State University of Campinas, Brazil

+551935219500

bennini@gmail.com

Funding

The authors declare no potential conflicts of interest with respect to the research, authorship and/or publication of this article.

The authors received no financial support for the research, authorship and/or publication of this article.

Acknowledgments

Ana Paula de Morais e Oliveira, Librarian review professional at the Medical Sciences Faculty's Library, contributed with elaboration of the search strategy and initial management

of the selection of studies. Maria Laura Costa do Nascimento, MD, PhD and Carlos Roberto Silveira Correia MD, PhD, contributed with text reviewing and suggestions for the article.

Abstract

Objective: to perform a diagnostic test accuracy systematic review and meta-analysis of the current data regarding the validation of the Delphi consensus criteria for the diagnosis of Fetal Growth Restriction (FGR). **Data sources:** Searches were performed in PubMed, EMBASE, PMC and Cochrane Library databases, and included studies from 2010 to 2020. Terms used for database searches were: ("Fetal Growth Retardation" AND Diagnosis) AND ("Fetal Weight" OR "Ultrasonography, Doppler" OR "Delphi procedure" OR DELPHI OR "Delphi consensus"). **Study eligibility criteria**: we searched for prospective or retrospective analytical observational studies, containing information on diagnostic validation of Delphi consensus criteria for FGR in singleton pregnancies with no sonographic evidence of structural or genetic abnormalities. **Study appraisal and synthesis methods:** Data was extracted independently by the reviewers and any divergence was evaluated and corrected by a consensus. QUADAS-2 was used to evaluate the quality of the methods. Diagnostic Odds Ratios (DOR) were calculated and forest plots were constructed depicting meta-analysis for the DOR and respective confidence intervals, as well as sensitivity (Se), specificity (Sp), and positive and negative likelihood ratios (LR+ and LR-). **Results:** A total of 12 studies with 4467 fetuses were included. Only one study has focused on validating the Delphi criteria as a whole for the diagnosis of FGR. The most evaluated diagnostic criteria were changes in the pulsatility index (PI) of the umbilical artery (UA) and estimated fetal weight (EFW) below the 3rd percentile. No studies were found that evaluated the fetal growth velocity in quartiles or the elevated PI of the uterine artery (UtA) below 32 weeks of gestation. The Delphi criteria with best DOR performance were: "abdominal circumference (AC) < p3", "EFW < p10 + PI UA > p95" and "A/REDF". **Conclusions**: Our study was the first to systematically review the literature analyzing the criteria proposed by the Delphi consensus for the diagnosis of FGR and its association with adverse perinatal outcomes.

However, more studies should be directed towards the validation of these criteria, especially those that have not yet been analyzed.

Keywords: fetal growth restriction; placental insufficiency; ultrasound; diagnosis; perinatal outcomes; diagnostic test accuracy; review.

Introduction

Fetal growth restriction (FGR) is defined as the failure of the fetus to reach its genetically determined potential [1] and affects about 5% of all pregnancies [2]. The etiologyencompasses fetal, placental and maternal conditions. Placental insufficiency is the most common cause of FGR [3] and its identification is important because pregnancies coursing with placenta-mediated FGR are at higher risk of poor perinatal outcomes [4-8].

In spite of the clinical importance and relatively high incidence of FGR, its diagnostic definition is still under development and the criteria used among different groups remain heterogeneous. In 2016, a multicentre team of international fetal medicine specialists conducted a study based on the Delphi method [9] to establish a consensus on the definition of FGR [10], in the absence of fetal malformations. They divided FGR into early (less than 32 weeks) and late (32 weeks or above) presentations and used a combination of ultrasonographic fetal biometric and maternal-fetal Doppler Velocimetry parameters.

Many international institutions recommend the use of this criteria proposed by the Delphi consensus for the diagnosis of FGR [11,12], however there is a dearth of trials evaluating its performance. In addition, this consensus encompasses associated criteria that may present divergent performances when evaluated separately. Therefore, the effort to validate each criterion before its wide use in clinical practice is justified.

Objectives

To perform a systematic review and meta-analysis of diagnostic accuracy test studies that have evaluated the Delphi consensus as a whole or its individualized criteria to detect adverse perinatal outcomes.

Methods

Protocol and registration

This study was carried out in accordance with Cochrane's directives for systematic reviews of diagnostic accuracy tests [13]. The research protocol was registered in the PROSPERO platform (https://www.crd.york.ac.uk/prospero) prior to the start of searches (Registration number: CRD42020204051; available from: https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42020204051). The final report was presented according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses of Diagnostic Test Accuracy Studies (PRISMA-DTA)[14].

Eligibility criteria

We searched for prospective or retrospective analytical observational studies from the last decade, containing information on diagnostic validation of Delphi consensus criteria for FGR (sensitivity - Se, specificity- Sp; or, if Se/Sp were not reported, those from which data for the calculation of Se and Sp could be retrieved). The research question was elaborated according to a structured modality that involved the population, index test, reference (gold) standard and target condition. Participants (population) were women with single pregnancies, whose fetuses with no sonographic evidence of structural or genetic abnormalities. All participants should have undergone obstetric ultrasound, as for the index test, each criterion encompassed by the Delphi consensus [10] for the diagnosis of early/late FGR was evaluated separately. Thus, the index test was considered "positive" if at least one of the Delphi consensus criteria for FGR was met. True positives were determined by the clinical diagnosis of at least one of the following perinatal results (reference standard): NICU

admission, fifth minute Apgar Score below 7, emergency C-section, neonatal acidosis, preterm birth, respiratory distress syndrome (RDS), small for gestational age (SGA), neonatal death, stillbirth and cerebral intraventricular hemorrhage (IVH). In the absence of a consensual definition, considering that placental insufficiency and its consequent FGR is associated with elevated perinatal morbi-mortality, we defined as the target condition the same parameters used as for the reference standard (poor perinatal outcomes). Only original articles published in English were included. Unpublished articles, literature reviews, systematic reviews, meta-analyses, expert opinion, letters to the editor, comments on studies and abstracts of presentations at congresses and symposia were not included.

Information Sources

Searches were performed in PubMed, EMBASE, PMC and Cochrane Library databases, and included studies from 2010 to 2020 (last day of search was August 17, 2020). Ongoing, unpublished, or in press studies were excluded. Abstracts and conference proceedings were also excluded.

Search strategy

Terms used for database searches were: ("Fetal Growth Retardation" AND Diagnosis) AND ("Fetal Weight" OR "Ultrasonography, Doppler" OR "Delphi procedure" OR DELPHI OR "Delphi consensus"). The complete search strategy is available at https://www.crd.york.ac.uk/PROSPEROFILES/204051_PROTOCOL_20230130.pdf.

Study Selection

The study selection process was divided into two phases: 1) pre-selection based on titles and abstracts and 2) the full article was read and analyzed. The selection was performed by two reviewers, conflicts were solved based on consensus among the researchers. Publications were deemed suitable for inclusion if they reported Se and Sp for at least one of the criteria proposed by the Delphi consensus for the diagnosis of FGR, or if data available

from which Se and Sp could be derived. After conducting the search, articles were inserted in the ENDNOTE platform (www.myendnoteweb.com). Duplicities were identified and excluded. Articles were then inserted into the RAYYAN QRI platform (www.rayyan.qcri.org) [15], and new duplicates were identified and excluded.

Data extraction

Data was extracted independently by the reviewers based on piloted forms and inserted in electronic spreadsheets. After data was collected, they were double checked by the researchers. Any divergence was evaluated and corrected by a consensus.

Once selected, articles were screened for: name of the first author and year of publication, study design, total number of participants, gestational age at the first ultrasound exam for study inclusion, criteria for FGR diagnosis as proposed by the Delphi consensus, adverse perinatal outcomes. The number of fetuses with or without a positive index test or perinatal outcome, as reported in the original article, was also extracted for register and calculation of true positive (TP), true negative (TN), false positive (FP), false negative (FN), Sp, Se, positive (LR+), negative likelihood ratio (LR-) and Diagnostic Odds Ratio (DOR).

Assessment of risk of bias

The QUADAS-2 tool [16] was used to analyze the risk of bias in each study. The analysis was carried out by both researchers, and the differences were resolved by consensus.

Data synthesis

Diagnostic Odds Ratios (DOR) were calculated using the DerSimonian-Laird random effects model and forest plots were constructed depicting meta-analysis for the DOR and respective confidence intervals, subgrouping the analyses based on studies' Delphi diagnostic criteria and outcome of interest.

All analyzes were performed using the “mada” and “meta” [17] libraries for the R environment for statistical computing (R Core Team, 2022) [18].

Results

Study selection

The search strategy yielded 1521 articles from PubMed, EMBASE PMC, and Cochrane Library databases. Manuscripts were then submitted to evaluation by two researchers (M.M.S.C. and J.R.B.) based on studies’ titles and abstracts. After this first screening, 53 articles remained for analysis of the full text. After being read in full, twelve articles (encompassing 4467 fetuses) were selected for data extraction [19-30]. The flowchart of article selection is summarized in Figure 1.

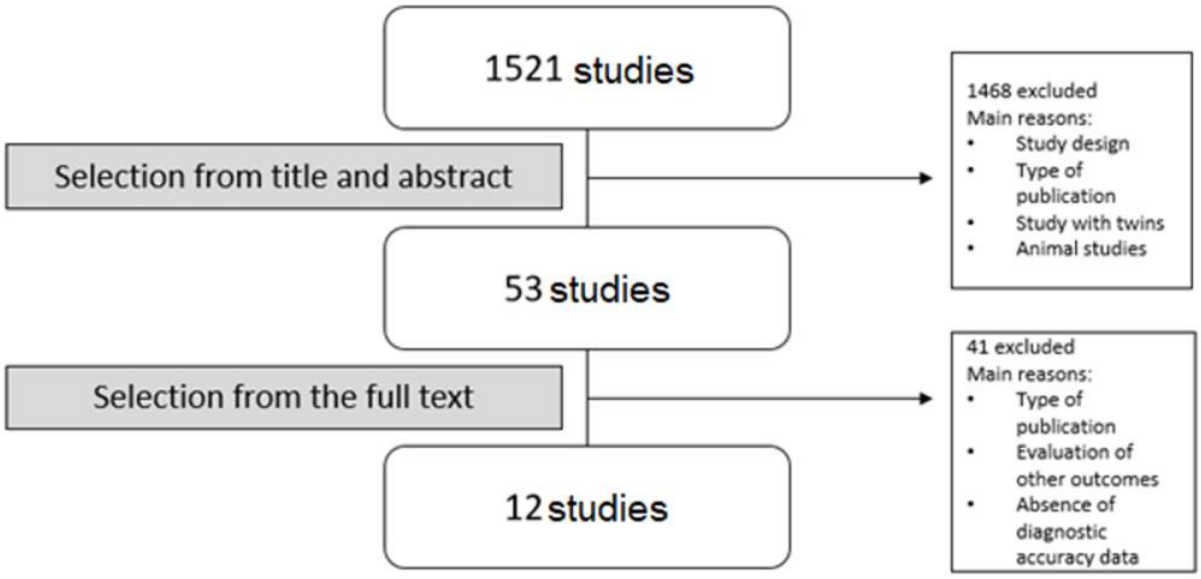


Figure 1: Flowchart of article selection for extracting systematic review data.

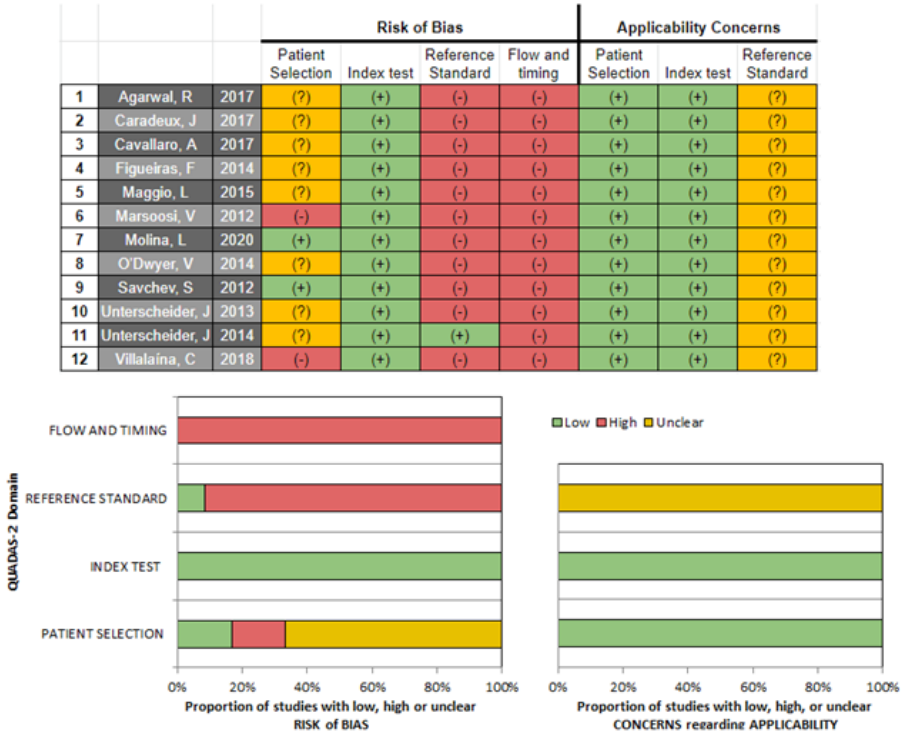
Study characteristics

Key characteristics of the studies included in the systematic review (first author, year of publication, study design, gestational age at the first ultrasound exam for study inclusion,

sample size, formula used to estimate the fetal weight, protocol used for pregnancy termination, Delphi consensus criteria for FGR and adverse perinatal outcomes included for each study) are listed in **Table 1.**

Risk of bias of included studies

The QUADAS method [16] was used to determine the risk of bias and applicability, selection of patients, index test, gold standard and flow and timing of each study (**Figure 2**). According to QUADAS directives, studies' domains represented in red have a high risk of bias, those in yellow have an unclear risk and green a low risk of bias.



			Risk of Bias				Applicability Concerns		
			Patient Selection	Index test	Reference Standard	Flow and timing	Patient Selection	Index test	Reference Standard
1	Agarwal, R	2017	(?)	(+)	(-)	(-)	(+)	(+)	(?)
2	Caradeux, J	2017	(?)	(+)	(-)	(-)	(+)	(+)	(?)
3	Cavallaro, A	2017	(?)	(+)	(-)	(-)	(+)	(+)	(?)
4	Figueiras, F	2014	(?)	(+)	(-)	(-)	(+)	(+)	(?)
5	Maggio, L	2015	(?)	(+)	(-)	(-)	(+)	(+)	(?)
6	Marsoosi, V	2012	(-)	(+)	(-)	(-)	(+)	(+)	(?)
7	Molina, L	2020	(+)	(+)	(-)	(-)	(+)	(+)	(?)
8	O'Dwyer, V	2014	(?)	(+)	(-)	(-)	(+)	(+)	(?)
9	Savchev, S	2012	(+)	(+)	(-)	(-)	(+)	(+)	(?)
10	Unterscheider, J	2013	(?)	(+)	(-)	(-)	(+)	(+)	(?)
11	Unterscheider, J	2014	(?)	(+)	(+)	(-)	(+)	(+)	(?)
12	Villalaína, C	2018	(-)	(+)	(-)	(-)	(+)	(+)	(?)

Figure 2: Risk of bias and applicability according to QUADAS.

Studies that used as inclusion criteria an EFW below the tenth percentile were flagged with an unclear risk of selection bias, since these patients are wisely at high-risk of perinatal morbimortality. The studies that we considered as having low risk of selection bias were those by Molina [22] and Savchev [29], as they included singleton pregnancies, without

evidence of fetal malformations regardless of EFW. Studies by Maroosi [21] and Villalaína [30] were flagged in red for patient selection since they had a retrospective design.

For all studies, ultrasound was performed before birth, thus preceding the perinatal outcomes; hence, the index test was interpreted before the reference standard (perinatal outcome) was known, and for this reason all studies were considered as having a low risk of bias in the index test domain.

We considered FGR as the target condition and defined the reference standard based on the fact that restricted fetuses have higher rates of adverse perinatal outcomes. However, as mentioned before, there is still no gold standard for the definition of FGR. In addition, the results of the US were known to obstetricians and neonatologists; i.e. those professionals who determined the outcome were aware of the index test results, which introduced reference standard bias. Only the study by Unterscheider in 2014 [25] was considered low risk of reference standard bias because the outcomes evaluated were stillbirth and neonatal death; these outcomes are not prone to subjectivity.

In general, the applicability of the reference standard was challenging, since several different perinatal outcomes were evaluated across the studies in this meta analysis. Regarding the flow and timing of the included studies, all studies were considered to bear a high risk of bias since for none of the included studies the interval between the performance of the ultrasounds (index test) and the resolution of the pregnancy (outcomes) was clearly informed or defined. In addition, Marsoosi [21] was flagged in red (high risk of bias regarding flow and timing) because not all live births were submitted to the transfontanellar ultrasound for the diagnosis of IVH.

Synthesis of results

Of the 12 selected studies, seven did not distinguish between early FGR (below 32 weeks of GA) and late FGR (equal to or above 32 weeks of GA). Another five studies included only ultrasound evaluations that started at a gestational age > 32 weeks (Table 1).

Of the 17 criteria proposed by the Delphi consensus for the diagnosis of FGR, only six (alone or in combination) were evaluated in the studies (Table 2); the Delphi criteria or combinations of Delphi criteria evaluated (index tests) are listed in Table 3 and, for clarity, coded from A to L.

The most evaluated Delphi consensus criteria were changes in the PI of the UA and EFW below the 3rd percentile. No studies were found that evaluated the fetal growth velocity in quartiles or the elevated PI of the uterine artery (UtA) below 32 weeks (associated with EFW or AC below the 10th percentile for GA) for the prediction of adverse perinatal outcomes. Only two studies evaluated AC below the 10th percentile for GA associated or not with *Doppler* changes [24,26]. Three studies analyzed CPR associated with EFW below the 10th percentile for GA in fetuses older than 32 weeks [26,27,30].

Regarding the perinatal outcomes evaluated, the included studies provided data on the number of fetuses that were born small for gestational age (SGA - fetal birth weight below the 10th percentile for GA), fetuses that needed to be admitted to NICU, neonatal deaths and stillbirths, premature births, fetuses with RDS, fetuses with a 5-minute Apgar score below 7, neonatal acidosis, emergency cesarean section due to fetal distress, IVH and fetuses that had any perinatal adverse event (APO - may include any of the above outcomes alone or in combination). The outcomes with the highest frequency of assessment between studies were APO, NICU admission, neonatal death and a 5-minute apgar score below 7, as shown in Figure 3.

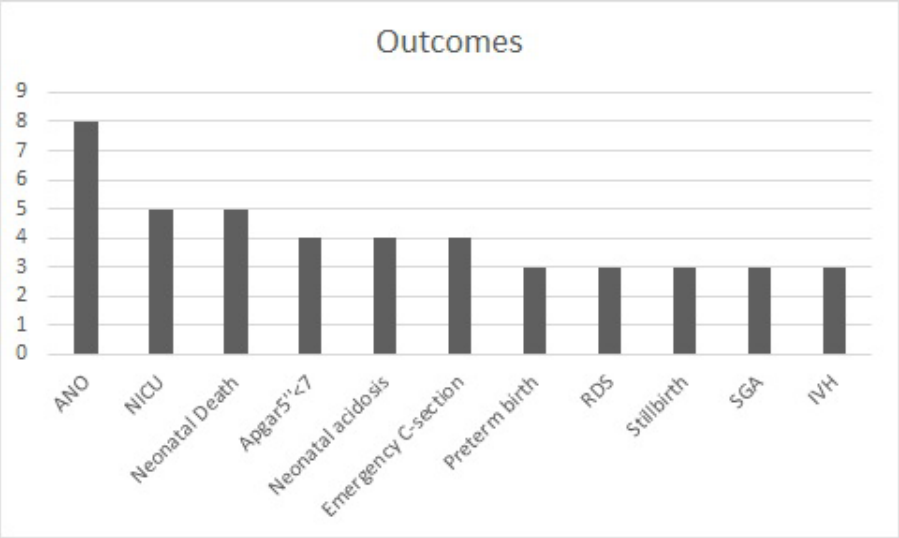


Figure 3: Frequency of assessments of adverse perinatal outcomes among the studies included in the review.

Only one study evaluated the Delphi consensus as a whole, without specifying which criteria were used in each case [22]. In that study, the authors compared the criteria proposed by the Delphi consensus with one of the most common definitions of clinical practice: EFW <p10 only. The authors found that the Se for the current FGR definition (EFW<p10 for GA) for predicting a SGA neonate was higher than that obtained when using the Delphi definition: 41.7% [(95%CI:33.4-50.3%] and 28.8% [95% CI: 21.4-37.1%] respectively. The Se of the Delphi criteria for the prediction of any neonatal outcome (ANO - what they defined as NICU admission, cord pH < 7.10, 5-min Apgar score < 7, RDS, IVH, neonatal seizures or neonatal death) proved to be quite low, but slightly higher than that of the current definition: 10.1% [95% CI: 5.6-16.3%] and 9.3% [95% CI: 5.1-15.5%], respectively. The current definition showed higher discriminatory ability for predicting a SGA neonate (Area under the ROC curve (AUC) =0.69 [95%CI, 0.65–0.73] than did the Delphi criteria (AUC = 0.64 [95% CI 0.60–0.67]; P = 0.001). The AUCs for both definitions for the prediction of composite ANO were poor, despite slightly improved performance using the Delphi consensus definition (AUC = 0.53 [95% CI, 0.50–0.55] *vs.* 0.50 [95% CI, 0.48–0.53]; P = 0.02).

The Se and Sp of the index tests to predict the adverse outcomes are shown in Table 4. The index tests D and I showed a trade-off between Se and Sp, with Se below 60%, but Sp greater than 70% for the evaluated outcomes in all the five studies that made that analysis. Index test J, which combines index tests D and I, yielded Se values ranging from 33% to 100%, Sp ranging from 52% to 84% and a greater variability between studies depending on which outcome is used for the reference standard. As can be seen in Table 4, a very heterogeneous approach to the use of the Delphi criteria (or combinations of Delphi criteria) to predict perinatal outcomes is a hallmark of the current literature status. Forest plots depicting the Se and Sp of each index test to detect adverse outcomes are shown in Figure 4.

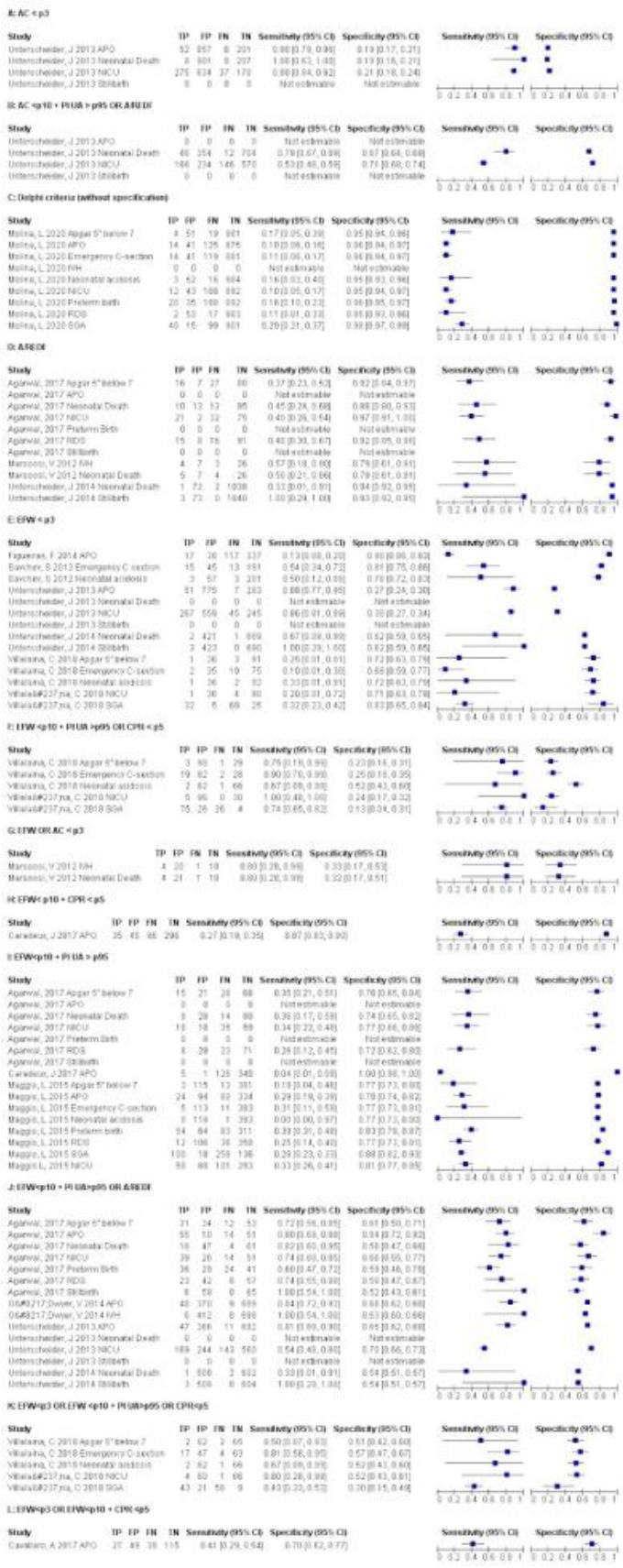


Figure 4: Se and Sp in forest plots of each index test related to specific perinatal outcomes.

The Receiver-operating-characteristics (ROC) curve analisis (figure 5) revealed that criterion D (A/REDF) yields the best trade-offs between sensitivity and specificity for the detection of any perinatal outcome. It is also evident that heterogeneity between studies for criterion D (A/REDF) is relatively low, with a single outlier (Unterscheider, 2014; evaluates A/REDF for stillbirth detection). Heterogeneity between studies for the other criteria is remarkably high; for instance, regarding EFW < p3 (index test E), some studies (e.g. Figueiras and Unterscheider, 2013 evaluating APO as outcome) show reversed performance indicators (Se and Sp).

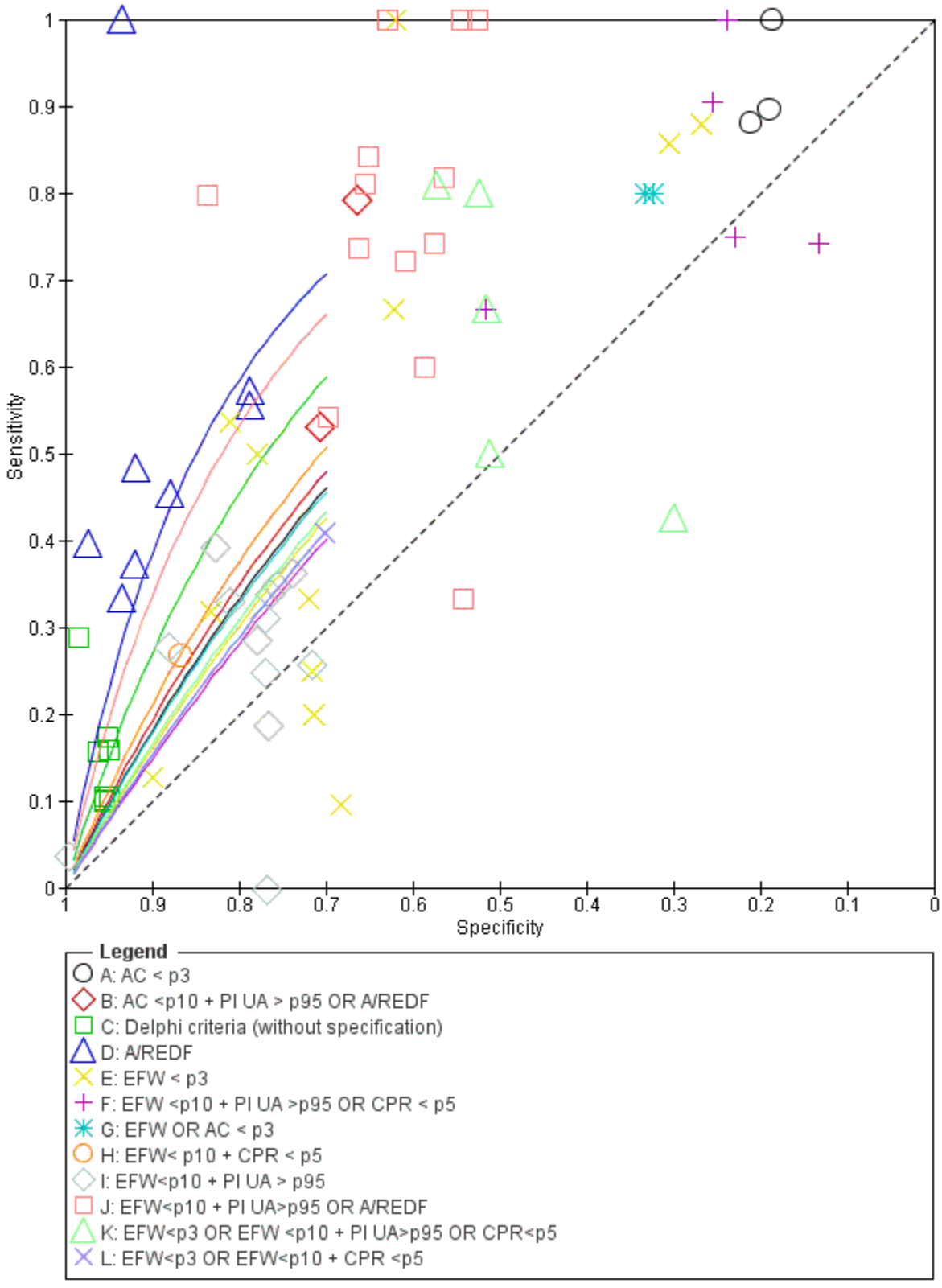


Figure 5: ROC curves for different criteria for the diagnosis of multiple fetal outcomes.

For most index tests, the LR+ showed no or just a slight increase in the probability of adverse perinatal outcomes. Only four index tests showed a large, i.e. LR+ >10 increase in post-test probability related to different outcomes: the presence of any Delphi criteria in the Molina's study [22] increased the chance of a SGA fetus with a LR+ = 15.6 (95%CI = 9.9 to 30.95); the presence of UA A/REDF was associated with a significantly augmented post-test probability of NICU admission, according to one study [19] and stillbirth according to another one [25]; regarding the increase in the chance of APO, the EFW<p10 + PI UA > p95 (index test I) was the only criterion associated with an increase in post-test probability (LR+ = 13.0; 95%CI 1.5 to 110.4) in the study by Caradeux et al [26]. None of the index tests yielded a LR- lower than 0.2, showing zero or a slight decrease of post-test probability, regardless of the outcome being evaluated. The LR + and LR - for each of the index tests as reported by each of the studies included in this analysis are shown in Table 5.

Figure 6 shows the forest plots for index tests' DOR for the detection of adverse perinatal outcomes. Our pooled analyses for index tests A, B, C, D, E, H, I and J resulted in significant DOR values; however, for tests B, C, E and J a high heterogeneity between studies was found. Index test H was evaluated in only one study that evaluated APO. The highest pooled DOR value was obtained for index test D (DOR= 8.0; 95%CI 5.1 to 12.6). For index test D (A/REDF), six outcomes were evaluated (NICU, Neonatal Death, Apgar 5"<7, RDS, IVH and Stillbirth) in three studies included in the analysis [19, 21 and 25], with a total heterogeneity of i2=0%.

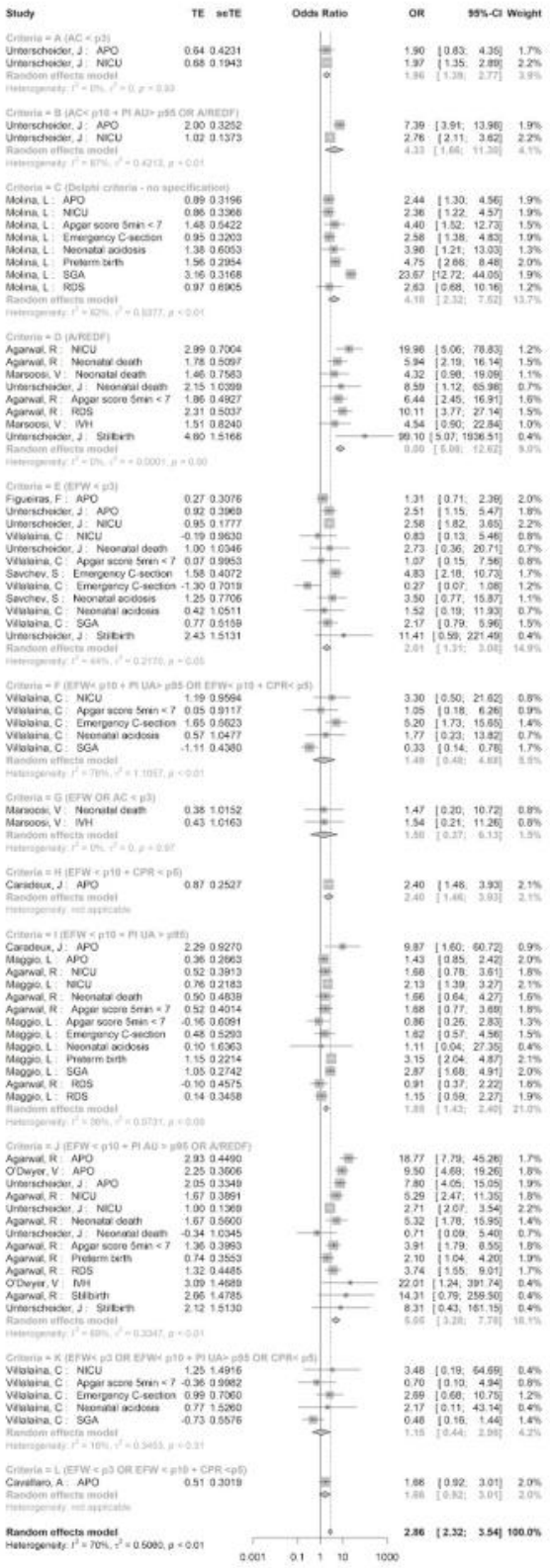


Figure 6: DOR results of each index test for the detection of adverse perinatal outcomes presented in forest plots.

Significant DOR results with low heterogeneity were observed for the detection of the following outcomes (figure 7): NICU, Neonatal Death, Neonatal Acidosis, Preterm birth, IVH and Stilbirth. Significant results, but with high heterogeneity, were observed for the detection of APO, Apgar 5”<7 and Emergency C-section. The meta-analysis showed that the DOR for the presence (positive index test result) of any index test to detect any outcome was 2.86 (95%CI: 2.32-3.54), with a high heterogeneity (70%).

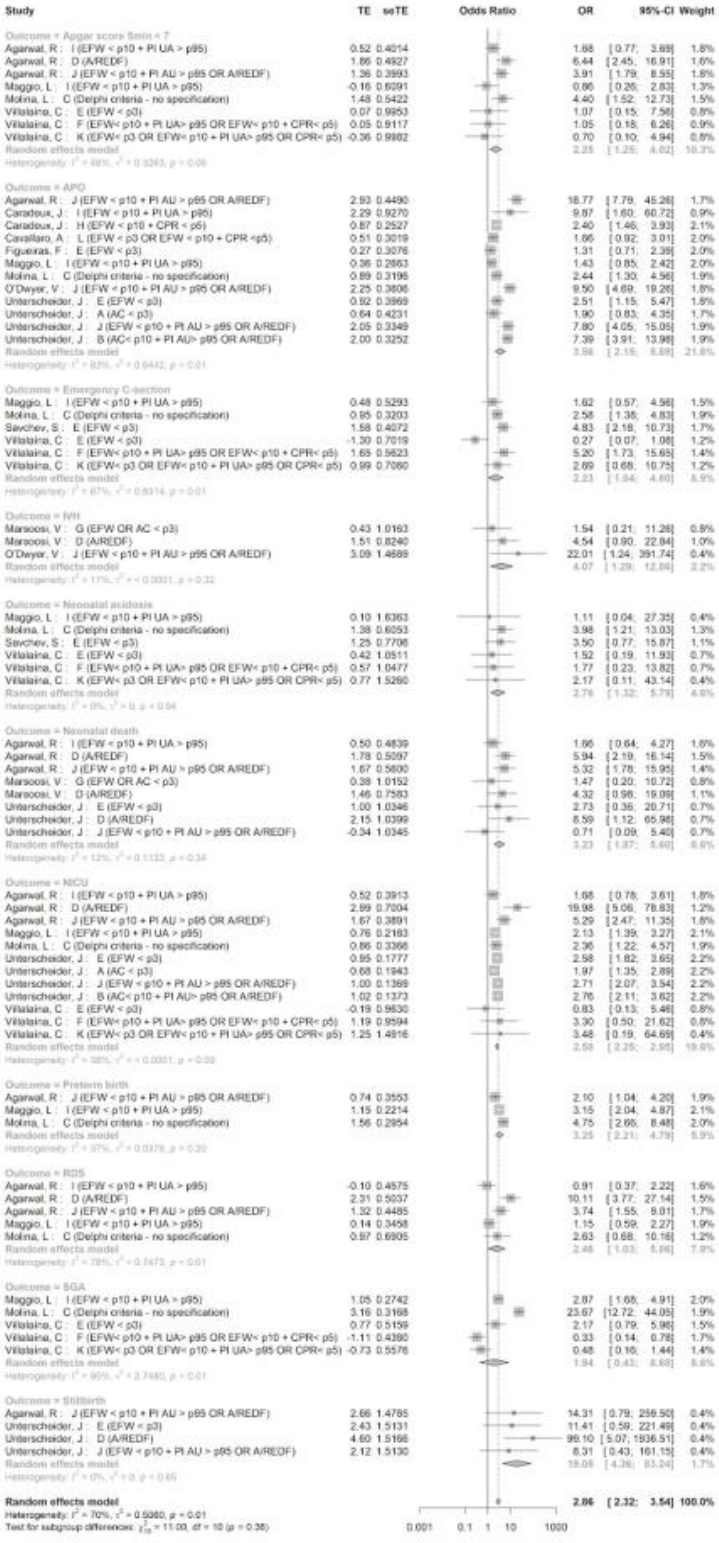


Figure 7: DOR results of each adverse perinatal outcome in relation to the criteria proposed by the Delphi consensus presented in forest plots.

Comment

Principal findings

Since the publication of the Delphi consensus criteria for FGR, only one study has focused on validating the consensus as a whole as a diagnostic test (Molina, 2020). This study showed a slight superiority of the Delphi criteria for the prediction of adverse perinatal outcomes, when compared to the EFW <p10 criterion alone (commonly applied in clinical practice), but both tests showed a poor performance to that diagnosis.

In the last 10 years, the most analyzed diagnostic criteria proposed by the Delphi consensus to the diagnosis of FGR were changes in the PI of the UA and EFW below the 3rd percentile. No studies were found that evaluated the fetal growth velocity in quartiles or the elevated PI of the uterine artery (UtA) below 32 weeks (associated with EFW or AC below the 10th percentile for GA) for the prediction of adverse perinatal outcomes and FGR. This demonstrates an important gap in knowledge about the accuracy of criteria that have already been used in clinical practice.

Although there is no precise definition of which adverse perinatal outcomes actually represent the truly restricted fetus, the most investigated perinatal adverse outcomes according to criteria proposed by the Delphi consensus were ANO, NICU admission, neonatal death and a 5-minute apgar score below 7. The outcomes best associated with the positive findings proposed by the Delphi consensus were: NICU admission, Neonatal Death, Neonatal Acidosis and Preterm birth.

Based on the DOR results, the diagnostic criteria that performed better for predicting adverse perinatal outcomes were: "AC < p3", "EFW < p10 + PI UA > p95" and "A/REDF". The visual analyses of the ROC curve reveal that this criteria yields the best area under the ROC curve compared to all other combinations of criteria and fetal outcomes.

The meta-analysis showed that the DOR for the presence of any Delphi consensus criteria for the diagnosis of FGR to detect any perinatal adverse outcome was 2.86 (95%CI: 2.32-3.54), with a high heterogeneity (70%).

Comparison with Existing Literature

As demonstrated in our review, only one study aimed to evaluate the Delphi consensus for the diagnosis of FGR, so we sought to evaluate studies that had investigated at least one of these criteria for the detection of adverse perinatal outcomes. The purpose of these criteria is to increase diagnostic accuracy in order to reduce the number of false positives and false negatives. Currently, there are renowned institutions that consider a restricted fetus those with EFW below 10t percentile, but from this perspective it is not possible to differentiate a fetus with adequate weight and signs of placental insufficiency, as well as safely affirming the existence of a constitutionally small fetus.

In 2017 Dum et al. performed a review that included 1428 pregnancies and found that among eleven studies evaluated, altered CPR was associated with the presence of adverse perinatal outcomes in eight of them, regardless of birth weight [31]. Nassr and Conde-Agudelo also described this relationship in the population of fetuses with EFW below the 10th percentile [32,33], with evidence that abnormal CPR would increase the risk of adverse outcomes to 49%, while normal CPR would decrease it to 18.2% [33]. In our review, CPR evaluation was studied through four index tests (F, H, K and L). The Delphi criterion "EFW<p10 + CPR<p5" (index test H) to predict APO was examined in only one study (DOR 2.4 ; 95%CI: 1.46-3.93)[26].

Our finding of A/REDF as an important predictor of fetal death is consistent with previous findings. A review of 31 studies including 5909 Doppler assessments found that AEDF yielded a DOR of 3.59 and REDF yielded a DOR of 7.27 for stillbirth[34].

We found that AC below p3 is associated with an increased risk of a composite of adverse perinatal outcomes and NICU admission. The assessment of fetal biometry alone was also previously investigated, with evidence that EFW or AC below p3 was associated with a greater chance of APO (DOR 1.97; 95% CI: 1.33-2.92), NICU admission (DOR 2.87 ; 95 % CI: 1.84-4.47) and stillbirth (DOR 4.26; 95% CI: 1.07-16.93). However, the definition of APO differed across the studies included in this review [35].

Our study revealed a high performance heterogeneity across studies. However, similar discrepant performance has been reported by other studies [36,37], especially when adverse perinatal outcomes are unified to define the reference standard of the diagnostic test. These discrepancies may be ascribed to the different gestational-age thresholds used for patient selection across studies, and to the absence of a standardized definition of conditions included in “APO”. In addition, different protocols for the management of FGR and for pregnancy termination may have also contributed to the variability of performance indicators reported by different authors.

Strengths and Limitations

Our study was the first to systematically review the literature data analyzing the Delphi consensus criteria for the diagnosis of FGR and their association with adverse perinatal outcomes. Our analysis has several strong points: it was an extensive and systematic literature search that was prospectively designed and registered with PROSPERO to reduce the risk of reporting bias. Also, we included a considerable number of 4467 fetuses in our review.

Our analyses suffer from the fact that there is no standardization of the outcomes studied. Some authors have evaluated “any adverse outcome”, but without differentiation between prenatal outcomes (such as emergency C-section or stillbirth) and neonatal outcomes. So we decided to use the definition “any adverse perinatal outcome” (APO) which encompasses the neonatal and prenatal variables evaluated.

Another limitation is that pregnancy follow-up protocols for the cases diagnosed with FGR and information about pregnancies that allowed us to infer the risk of FGR and adverse outcomes were not always mentioned in the included studies.

It is interesting to note that performance analyses exist in the current literature for only 6 out of the 17 existing Delphi criteria. This fact precludes the fulfillment of one of our main objectives, which was to produce a comprehensive review of the performance of Delphi criteria for the diagnosis of the most clinically relevant perinatal outcomes.

Conclusions and implications

Despite the few studies that allow us to validate the Delphi consensus for the diagnosis of FGR, in clinical practice this is the method that has been adopted by many obstetricians and ultrasonographers. Our review showed that the finding of absent or reverse end diastolic flow in the umbilical artery was strongly associated with the development of adverse perinatal outcomes, which may single out this specific criterion (among all those ever studied) as the most important of all Delphi consensus. We found some evidence that severely reduced fetal biometrics (i.e. those below 3rd percentile) and elevations in UA PI also have diagnostic potential for predicting adverse perinatal outcomes.

The validation of Delphi criteria for the diagnosis of FGR is still poorly explored in the current medical literature. More studies should be directed towards including those criteria for which there is still a lack of data. For criteria such as drop in growth velocity, there are no published studies whatsoever. The objective comparison of performance between the Delphi criteria is pending on further observational and validation studies, ideally if all criteria are applied to the same set of patients.

Funding

The authors declare no potential conflicts of interest with respect to the research, authorship and/or publication of this article.

The authors received no financial support for the research, authorship and/or publication of this article.

References

1. Resnik R. Intrauterine growth restriction. *Obstet Gynecol* 2002; **99** : 490-496.

2. Romo A, Carceller R, Tobajas J. Intrauterine growth retardation (IUGR): epidemiology and etiology. *Pediatr Endocrinol* Rev. 2006; **6(3)** : 332-336.

3. Lackman F, Capewell V, Gagnon R, Richardson B. Fetal umbilical cord oxygen values and birth to placental weight ratio in relation to size at birth. *Am J Obstet Gynecol* 2001; **185** : 674-682.

4. Figueiras F, Gratacós E. Update on the Diagnosis and Classification of Fetal Growth Restriction and Proposal of a Stage-Based Management Protocol. *Fetal Diagn Ther* 2014; **36** : 86–98.

5. Baschat AA, Cosmi E, Bilardo CM, Wolf H, Berg C, Rigano S, Germer U, Moyano D, Turan S, Hartung J, Bhide A, Müller T, Bower S, Nicolaides KH, Thilaganathan B, Gembruch U, Ferrazzi E, Hecher K, Galan HL, Harman CR. Predictors of Neonatal Outcome in Early-Onset Placental Dysfunction. *Obstetrics & Gynecology* 2007; **109(2, pt 1)** : 253–261.

6. Francis JH, Permezel M, Davey MA. Perinatal mortality by birthweight centile. *Aust N. Z J Obstet Gynaecol* 2014; **54** : 354-359.

7. Miller SL, Huppi PS, Mallard C. The consequences of fetal growth restriction on brain structure and neurodevelopmental outcome. *J Physiol (Lond)* 2016; **594**: 807-823.

8. The GRIT Study Group. A randomised trial of timed delivery for the compromised preterm fetus: short term outcomes and Bayesian interpretation. *BJOG: An International Journal of Obstetrics and Gynaecology* 2003; **110(1)** : 27–32.

9. Jones J, Hunter D. Consensus methods for medical and health services research. *BMJ*. 1995; **311** : 376–380.

10. Gordijn SJ, Beune IM, Thilaganathan B, Papageorghiou A, Baschat AA, Baker PN, Silver RM, Wynia K, Ganzevoort W. Consensus definition of fetal growth restriction: a Delphi procedure. *Ultrasound Obstet Gynecol* 2016; **48(3)** : 333–339.

11. ISUOG Practice Guidelines: diagnosis and management of small-for-gestational-age fetus and fetal growth restriction. *Ultrasound Obstet Gynecol* 2020; **56** : 298–312.

12. Melamed N, Baschat A, Yinon Y, Athanasiadis A, Mecacci F, Figueras F, Berghella V, Nazareth A, Tahlak M, McIntyre HD, Costa FS , Kihara AB, Hadar E, McAuliffe F, Hanson M, Ma RC, Gooden R, Sheiner E, Kapur A, Divakar H, Ayres-de-Campos D, Hiersch L, Poon LC, Kingdom J, Romero R, Hod M. FIGO (International Federation of Gynecology and Obstetrics) initiative on fetal growth: Best practice advice for screening, diagnosis, and management of fetal growth restriction. I*nt J Gynecol Obstet* 2021; **152(1)** : 3–57.

13. Deeks JJ, Wisniewski S, Davenport C. Chapter 4: Guide to the contents of a Cochrane Diagnostic Test Accuracy Protocol. In: Deeks JJ, Bossuyt PM, Gatsonis C (editors), *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 1.0.0.* The Cochrane Collaboration, 2013. Available on: http://srdta.cochrane.org/.

14. McInnes MDF, Moher D, Thombs BD et al. Preferred Reporting Items for a Systematic Review and Meta-analysis of Diagnostic Test Accuracy Studies: The PRISMA-DTA Statement. JAMA. 2018;**319(4)**:388-39.

15. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan — a web and mobile app for systematic reviews. *Systematic Reviews* 2016; **5** : 210.

16. Whitining PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reistma JB, Leflang MM, Sterne JA, Bossuyt PM. QUADAS-2: A revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011; **155** : 529-536.

17. Doebler P, Holling H. Meta-analysis of diagnostic accuracy with mada. Available from: https://cran.r-project.org/web/packages/mada/vignettes/mada.pdf. Accessed on October 31,2020.

18. . Foundation for Statistical Computing, Vienna, Austria. A language and environment for statistical computing. Available from: https://www.R-project.org/). Accessed on Mai 08, 2022.

19. Agarwal R, Tiwari A, Wadhwa N, Radhakrishnan G, Bhatt S, Batra P. Abnormal umbilical artery Doppler velocimetry and placenta histopathological correlation in fetal growth restriction. *SAJOG* 2017; **23(1)** : 12-16.

20. Maggio L, Dahlke JD, Mendez-Figueroa H, Albright CM, Chauhan SP, Wenstrom KD. Perinatal Outcomes With Normal Compared With Elevated Umbilical Artery Systolic-to-Diastolic Ratios in Fetal Growth Restriction. *Obstetrics & Gynecology* 2015; **125(4)** : 863-869.

21. Marsoosi V, Bahadori F, Esfahani F, Ghasemi-Rad M. The role of doppler indices in predicting intra ventricular hemorrhage and perinatal mortality in fetal growth restriction. *Medical Ultrasonography* 2012; **14(2)** : 125-132.

22. Molina LCG, Odibo L, Zientara S, Običan SG, Rodriguez A, Stout M, Odibo AO. Validation of Delphi procedure consensus criteria for defining fetal growth restriction. *Ultrasound Obstet Gyneco*l 2020; **56** : 61–66.

23. O'Dwyer V, Burke G, Unterscheider J, Daly S, Geary MP, Kennelly MM, McAuliffe FM, O'Donoghue K, Hunter A, Morrison JJ, Dicker P, Tully EC, Malone FD. Defining the residual risk of adverse perinatal outcome in growth-restricted fetuses with normal umbilical artery blood flow. *Am J Obstet Gynecol* 2014; **211** : 421-425.

24. Unterscheider J, Daly S, Geary MP, Kennelly MM, McAuliffe FM, O'Donoghue K, Hunter A, Morrison JJ, Burke G, Dicker P, Tully EC, Malone FD. Optimizing the definition of intrauterine growth restriction: the multicenter prospective PORTO Study. *Am J Obstet Gynecol* 2013; **208** : 291-296.

25. Unterscheider J, O'Donoghue K, Daly S, Geary MP, Kennelly MM, McAuliffe FM, Hunter A, Morrison JJ, Burke G, Dicker P, Tully EC, Malone FD. Fetal growth restriction and the risk of perinatal mortality–case studies from the multicentre PORTO study. *BMC Pregnancy and Childbirth* 2014; **14(63)** : 1-6.

26. Caradeux J, Eixarch E, Mazarico E, Basuki TR, Gratacos E, Figueras F. Longitudinal growth assessment for prediction of adverse perinatal outcome in fetuses suspected to be small-forgestational age. *Ultrasound Obstet Gynecol* 2018; **52:** 325–331.

27. Cavallaro A, Veglia M, Svirko E, Vannuccini S, Volpe G, Impey L; Using fetal abdominal circumference growth velocity in the prediction of adverse outcome in near-term small-for-gestational-age fetuses. *Ultrasound Obstet Gynecol* 2018; **52:** 494–500.

28. Figueiras F, Savchev S, Triunfo S, Crovetto F, Gratacos E. An integrated model with classification criteria to predict small-for-gestational-age fetuses at risk of adverse perinatal outcome. *Ultrasound Obstet Gynecol* 2015; **45** : 279–285.

29. Savchev, S, Figueras F, Cruz-Martinez R, Illa M, Botet F, Gratacos E. Estimated weight centile as a predictor of perinatal outcome in small-for-gestational-age pregnancies with normal fetal and maternal Doppler indices. *Ultrasound Obstet Gynecol* 2012; **39** : 299–303.

30 Villalaína C, Herraiz I, Quezada MS, Gómez-Arriaga PI, Gómez-Montes E, Galindo A. Fetal Biometry and Doppler Study for the Assessment of Perinatal Outcome in Stage I Late-Onset Fetal Growth Restriction. *Fetal Diagn Ther* 2018; DOI: 10.1159/000485124.

31. Dum L, Sherrel H, Kumar S. Review: Systematic review of the utility of the fetal cerebroplacental ratio measured at term for the prediction of adverse perinatal outcome. Placenta, 2017; **54**: 68-75.

32. Nassr AA, Abdelmagied AM, Shazly SAM. Fetal cerebro-placental ratio and adverte perinatal outcome: systematic review and meta-analysis of the association and diagnostic performance. J. Perinat. Med. 2016; 1-7.

33. Conde-Agudelo A, Villar J, Kennedy SH, Papageroghiou AT. Predictive accuracy of cerebroplacental ratio for adverse perinatal and neurodevelopmental outcomes in suspected fetal growth restriction: systematic review and meta-analysis. Ultrasound Obstet Gynecol, 2018; **52**: 430-441.

34. Caradeux J, Martinez-Portilla RJ, Basuki TR, Kiserud T, Figueras F. Risk of fetal death in growth-restricted fetuses with umbilical and/or ductus venosus absent or reversed end-diastolic velocities before 34 weeks of gestation: a systematic review and meta-analysis. American Journal of Obstetrics and Gynecology, 2018; 774-782.

35. Meler E, Martinez-Portilla RJ, Caradeux J, Mazarico E, Gil-Armas C, Boada D, Martinez J, Carrillo P, Camacho M, Figueras F. Severe smallness as predictor of adverse perinatal outcome in suspected late small-for-gestational-age fetuses: systematic review and meta-analysis. Ultrasound Obstet Gynecol, 2022; **60**: 328-337.

36. Ali S, Heuving S, Kawooya MG, Byamugisha J, Grobbee DE, Papageorghiou AT, Klipstein-Grobusch K, Rijken MJ. Prognostic accuracy of antenatal Doppler ultrasound for adverse perinatal outcomes in low-income and middle-income countries: a systematic review. BMJ Open, 2021; **11**:e049799.

37. Martinez-Portilla RJ, Caradeux J, Meler E, Lip-Sosa DL, Sotiriadis A, Figueras F. Third-trimester uterine artery Doppler for prediction of adverse outcome in late small-for-gestational-age fetuses: systematic review and meta-analysis. Ultrasound Obstet Gynecol, 2020; **55**: 575-585.

Table 1: Characteristics of included studies (GA = gestational age; EFW = estimated fetal weight; PI = pulsatility index; UA = umbilical artery; A/REDF = umbilical artery with end diastolic flow absent or reverse; AC = abdominal circumference; IVH = intraventricular hemorrhage; CPR = cerebroplacental relation; RDS = respiratory distress syndrome; APO = any adverse perinatal outcome).

	First Author	Year	Study design	N	Minimum GA	Delphi criterion	Outcomes	Indication of pregnancy resolution	Assessment of EFW
1[19]	Agarwal, R	2017	Prospective	130	24 weeks	• EFW<p10 + PI UA>p95 • A/REDF UA	NICU, neonatal death, preterm birth, RDS, Apgar 5"<7, stillbirth and APO	Management as per hospital protocol (not mentioned).	Not mentioned.
2[20]	Maggio, L	2015	Retrospective	512	17 weeks	• EFW<p10 + PI UA>p95	SGA, emergency C-section, preterm birth, Apgar 5"<7, RDS and APO	Delivery if ARDF/REDF > 34 weeks or a nonreassuring cardiotocografy	Hadlock
3[21]	Marsoosi, V	2012	Prospective	43	23 weeks	• EFW OR AC<p10 + PI UA>p95 • A/REDF UA	IVH and neonatal death	Delivery based on fetal heart frequency and status of mother and fetus	Not mentioned.
4[22]	Molina, L	2020	Prospective	1055	26 weeks	Delphi consensus (as a whole)	IVH, SGA, emergency C-section, neonatal acidosis, NICU, preterm birth, RDS, Apgar 5"<7 and APO	Not prespecified by the study design	Hadlock
5[23]	O'Dwyer, V	2014	Prospective	1116	24 weeks	• EFW<p10 + PI UA>p95 OR A/REDF UA	IVH and APO	Not prespecified by the study design but 34 weeks' gestation when there was absent	Hadlock

								end-diastolic flow in the UA Doppler	
6[24]	Unterschei der, J	20 13	Prospective	1116	24 weeks	• EFW<p10 + PI UA>p95 OR A/REDF UA • AC<p10 + PI UA>p95 OR A/REDF UA • EFW<p3 • AC<p3	NICU, neonatal death, stillbirth and APO	Not prespecified by the study design but 34 weeks' gestation when there was absent end-diastolic flow in the UA Doppler.	Hadlock
7[25]	Unterschei der, J	20 14	Prospective	1116	24 weeks	• EFW<p10 + PI UA>p95 • A/REDF UA • EFW<p3	Neonatal death and stillbirth	Not prespecified by the study design but 34 weeks' gestation when there was absent end-diastolic flow in the UA Doppler.	Hadlock
8[26]	Caradeux, J	20 17	Prospective	472	32 weeks	• EFW<p10 + PI UA>p95 • EFW<p10 + CPR<p5	APO	Management as per hospital protocol (not mentioned).	Hadlock
9[27]	Cavallaro, A	20 17	Prospective	235	36 weeks	• EFW<p3 • EFW<p10 + CPR<p5	APO	Delivery with 37 weeks if EFW<p3, CPR <p5, PI AU >p95, PAPPA < 0,3 or gestational hypertension. EFW p3-p5 delivery with 40 weeks. EFW p5-p10 delivery with 41 weeks.	Hadlock
10[28]	Figueiras, F	20 14	Prospective	509	32 weeks	• EFW<p3	APO	Delivery with 37 weeks if CPR<p5 or MCA < p5 or preeclampsia. Emergency c-section if nonreassuring cardiotocografy or low ph of scalp.	Hadlock

11[29]	Savchev, S	2012	Prospective	264	34 weeks	• EFW<p3	Emergency C-section and neonatal acidosis	Delivey with 37 weeks if EFW < p10, Emergency c-section if signs of fetal suffering.	Hadlock
12[30]	Villalaína, C	2018	Retrospective	131	32 weeks	• EFW<p3 • EFW<p10 + PI UA>p95 OR CPR<05	SGA, emergency C-section, neonatal acidosis, NICU and Apgar 5"<7	Delivery with 37 weeks.	Hadlock

Table 2: Delphi consensus criteria for the diagnosis of FGR (EFW = estimated fetal weight; AC = abdominal circumference; PI = pulsatility index; UA = umbilical artery; UtA = uterine artery; A/REDF = umbilical artery with end diastolic flow absent or reverse; CPR = cerebroplacental ratio).

	Early FGR	Both early and late FGR	Late FGR
Delphi consensus for FGR	UA A/REDF	EFW < p3	EFW < p10 + 2 quartile drop in EFW
	EFW < p10 + UtA PI > p95	AC < p3	EFW < p10 + 2 quartile drop in AC
	CA < p10 + UtA PI > p95	EFW < p10 + UA PI > P95	EFW < p10 + CPR < p5
		AC < p10 + UA PI > P95	CA < p10 + 2 quartile drop in EFW
			CA < p10 + 2 quartile drop in CA
			CA < p10 + CPR < p5
			2 quartile drop in EFW + UA PI > p95
			2 quartile drop in AC + UA PI > p95
			2 quartile drop in EFW + CPR < p5
			2 quartile drop in AC + CPR < p5
	3 criteria	4 criteria	10 criteria

Table 3: Criterion or combination of criteria (index test) evaluated by the included studies (GA = gestational age; EFW = estimated fetal weight; PI = pulsatility index; UA = umbilical artery; A/REDF = umbilical artery with end diastolic flow absent or reverse; AC = abdominal circumference; IVH = intraventricular hemorrhage; CPR = cerebroplacental relation; RDS = respiratory distress syndrome; APO = any adverse perinatal outcome).

Index Test	Criterion/Combination of criteria	Outcomes	Author/study
A	AC < p3	APO and NICU	Unterscheider, J[24]
B	AC <p10 + PI UA > p95 OR A/REDF	APO and NICU	Unterscheider, J[24]
C	Delphi criteria (without specification)	APO, NICU, Apgar5''<7, Emergency C-section, Neonatal Acidosis, Preterm Birth, SGA and RDS.	Molina, L[22]
D	A/REDF	NICU, Neonatal Death, Apgar5"<7, RDS, IVH and Stillbirth	Agarwal, R[19] Marsoosi, V[21] Unterscheider, J[25]
E	EFW < p3	APO, NICU, Neonatal Death, Apgar 5"<7, Emergency C-section, Neonatal Acidosis, SGA and Stillbirth	Unterscheider, J[24,25] Savchev, S[29] Figueiras, F[28] Villalaína, C[30]
F	EFW <p10 + PI UA >p95 OR CPR < p5	NICU, Apgar5"<7, Emergency C-section, Neonatal Acidosis and SGA	Villalaína, C[30]
G	EFW OR AC < p3	Neonatal Death and IVH	Marsoosi, V[21]
H	EFW< p10 + CPR < p5	APO	Caradeux, J[26]
I	EFW<p10 + PI UA > p95	APO, NICU, Neonatal Death, Apgar5"7, Emergency C-section, Neonatal Acidosis, Preterm Birth, SGA and RDS	Caradeux, J[26] Maggio, L[20] Agarwal, R[19]

J	EFW<p10 + PI UA>p95 OR A/REDF	APO, NICU, Neonatal Death, Apgar5"<7, Preterm birth, RDS, IVH and Stillbirth	Agarwal, R[19] O'Dwyer, V[23] Unterscheider, J[24,25]
K	EFW<p3 OR EFW <p10 + PI UA>p95 OR CPR<p5	NICU, Apgar 5"<7, Emergency C-section, Neonatal Acidosis, and SGA	Villalaína, C[30]
L	EFW<p3 OR EFW<p10 + CPR <p5	APO	Cavallaro, A[27]

Table 4: Se and Sp of each index test related to perinatal outcomes and their respectives confidence intervals.

	Index Test	Study	SGA		NICU		Neonatal Death		Preterm birth		RDS		Apgar 5' < 7		Neonatal acidosis		Stillbirth		Emergency Cesarean		IVH		APO	
			Se	Sp	Se	Sp	Se	Sp	Se	Sp	Se	Sp	Se	Sp	Se	Sp	Se	Sp	Se	Sp	Se	Sp	Se	Sp
A	AC < p3	0	-	-	88,14 (84,63-91,51)	21,14 (19,37-24,13)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	89,66 (78,83-96,11)	19 (16,89-21,46)
B	AC < p10 + PI UA > p95 or UA A/REDF	6	-	-	53,21 (47,5-58,85)	70,9 (67,82-74,02)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	79,31 (66,65-88,83)	66,64 (63,81-69,38)
C	Delphi criteria without specification	4	26,8 (21,4-37,1)	98,4 (97,3-99,1)	10 (5,27-16,82)	99,4 (93,85-99,65)	-	-	15,63 (3,81-23,03)	96,22 (94,79-97,36)	10,53 (1,3-33,14)	94,88 (93,56-96,14)	17,39 (4,95-38,78)	95,06 (93,55-96,3)	15,79 (3,38-39,58)	94,98 (93,47-96,23)	-	-	10,53 (5,88-17,03)	95,55 (94,02-96,79)	-	-	10,07 (5,62-16,32)	95,52 (93,99-96,77)
D	UA A/REDF	1	-	-	39,62 (26,45-54)	97,4 (93,33-99,08)	45,45 (24,39-67,79)	87,96 (80,3-93,43)	-	-	48,39 (30,15-66,94)	91,92 (84,7-96,45)	37,21 (22,98-53,27)	91,95 (84,12-96,7)	-	-	-	-	-	-	-	-	-	-
		3	-	-	-	-	55,56 (21,2-86,3)	78,79 (61,09-91,02)	-	-	-	-	-	-	-	-	-	-	-	-	57,14 (18,41-90,1)	78,79 (61,09-91,02)	-	-
		7	-	-	-	-	33,33 (0,82-90,57)	93,51 (91,9-94,89)	-	-	-	-	-	-	-	-	100 (29,24-100)	90,44 (91,82-94,82)	-	-	-	-	-	-
E	EFW < p3	6	-	-	85,56 (61,18-99,28)	30,47 (27,31-33,78)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	87,93 (76,7-95,01)	26,75 (24,1-29,53)
		7	-	-	-	-	66,67 (9,43-99,16)	62,07 (59,14-64,94)	-	-	-	-	-	-	-	-	100 (29,24-100)	61,99 (59,07-64,86)	-	-	-	-	-	-
		10	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	12,69 (7,57-19,53)	89,87 (89,39-90,73)
		11	-	-	-	-	-	-	-	-	-	-	-	-	50 (11,81-88,19)	77,91 (72,34-82,82)	-	-	83,87 (33,87-72,49)	80,93 (75,33-85,74)	-	-	-	-
		12	31,68 (22,78-41,69)	83,33 (65,28-94,36)	20 (0,51-71,64)	71,43 (62,7-79,12)	-	-	-	-	-	-	25 (0,63-80,59)	71,65 (62,99-79,29)	33,33 (0,84-90,57)	71,68 (63,25-79,46)	-	-	9,52 (1,17-30,38)	68,18 (58,62-76,74)	-	-	-	-
F	EFW < p10 + PI UA > p95 or EFW < p10 + CPR < p5	12	42,57 (32,79-52,81)	30 (14,73-49,4)	80 (28,36-99,49)	53,28 (43,3-61,35)	-	-	-	-	-	-	50 (6,76-93,24)	51,18 (42,16-60,15)	66,67 (9,43-99,16)	51,08 (42,57-60,48)	-	-	80,95 (58,09-94,56)	57,27 (47,78-66,66)	-	-	-	-
G	EFW or AC < p3	3	-	-	-	-	60 (26,36-99,49)	32,26 (18,68-51,37)	-	-	-	-	-	-	-	-	-	-	-	-	60 (26,36-99,49)	33,33 (17,29-52,81)		
H	EFW < p10 + CPR < p5	8	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	x	-	-	-	26,72 (19,37-35,15)	86,8 (82,74-90,21)
I	EFW<p10 + PI UA > p95	1	-	-	33,96 (21,52-48,27)	79,62 (65,59-85,52)	36,36 (17,2-59,34)	74,07 (64,75-82,-3)	-	-	25,81 (11,86-44,61)	71,72 (61,78-82,31)	34,68 (21,61-50,93)	75,86 (65,5-84,4)	-	-	-	-	-	-	-	-	-	-
		2	27,93 (23,35-32,89)	88,31 (82,16-92,92)	33,11 (25,68-41,23)	81,16 (76,74-85,07)	-	-	39,42 (31,19-48,12)	82,93 (78,74-86,6)	25 (13,64-39,6)	77,16 (73,06-80,3)	18,75 (4,05-45,65)	76,91 (72,85-80,46)	0 (0-97,5)	76,91 (73,01-80,5)	-	-	31,25 (11,02-58,68)	77,22 (73,27-80,64)	-	-	28,57 (19,24-39,47)	78,04 (73,91-81,67)
		6	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	3,82 (1,25-8,68)	99,71 (98,39-99,99)
J	EFW<p10 + PI UA >p95 or A/REDF	1	-	-	73,58 (59,67-84,74)	66,23 (54,89-76,62)	81,82 (59,72-94,81)	56,48 (48,6-66)	60 (46,54-72,44)	58,57 (46,17-70,23)	74,19 (55,49-88,14)	57,58 (47,23-67,45)	72,09 (56,33-84,67)	60,92 (49,87-71,21)	-	-	100 (54,07-100)	52,42 (43,28-61,46)	-	-	-	-	79,71 (68,31-88,44)	83,61 (71,91-91,65)
		5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	100 (54,57-100)	62,68 (59,96-66,73)	94,21 (72,13-92,52)	85,06 (82,1-87,93)
		6	-	-	54,17 (43,46-59,79)	89,65 (80,34-72,81)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	81,03 (60,59-90,13)	85,41 (82,45-88,27)
		7	-	-	-	-	33,33 (0,84-90,57)	94,23 (91,25-97,2)	-	-	-	-	-	-	-	-	100 (29,24-100)	94,37 (91,29-97,22)	-	-	-	-	-	-
K	EFW < p10 + PI UA > p95 or CPR < p5 or EFW < p3	12	74,26 (64,6-82,44)	13,33 (3,76-30,72)	100 (47,82-100)	23,81 (15,88-32,21)	-	-	-	-	-	-	75 (19,41-99,37)	22,83 (15,86-31,32)	100 (29,24-100)	23,44 (16,41-31,74)	-	-	90,48 (69,62-98,83)	25,49 (17,63-34,65)	-	-	-	-
L	EFW < p3 or EFW < p10 + CPR < p5	9	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	41,43 (29,77-53,83)	70,19 (62,48-77,13)

Table 5: Positive and negative likelihood ratio of each index test related to perinatal outcomes and their respective confidence intervals.

	Index Test	Study	SGA		NICU		Neonatal Death		Preterm birth		RDS		Apgar 5' < 7		Neonatal acidosis		Stillbirth		Emergency Cesarean		IVH		APO	
			LR+	LR-	LR+	LR-	LR+	LR-	LR+	LR-	LR+	LR-	LR+	LR-	LR+	LR-	LR+	LR-	LR+	LR-	LR+	LR-	LR+	LR-
A	AC < p3	6	-	-	1,12 (1,06-1,18)	0,56 (0,40-0,78)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1,11 (1,01-1,21)	0,54 (0,25-1,17)
B	AC < p10 + PI UA > p95 or UA A/REDF	6	-	-	1,83 (1,57-2,12)	0,60 (0,58-0,75)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2,37 (2,09-2,77)	0,31 (0,19-0,52)
C	Delphi criteria without specification	4	15,57 (9,98-30,95)	0,72 (0,65-0,81)	2,17 (1,18-4,01)	0,94 (0,89-1)	-	-	4,14 (2,47-6,94)	0,88 (0,81-0,95)	2,06 (3,54-7,83)	0,94 (0,91-1,1)	3,52 (1,39-8,92)	0,87 (0,72-1,05)	3,15 (1,05-9,19)	0,99 (0,73-1,08)	-	-	2,37 (1,33-4,22)	0,94 (0,88-0,99)	-	-	2,25 (1,26-4,02)	0,94 (0,89-1)
D	UA A/REDF	1	-	-	15,34 (3,73-62,33)	0,62 (0,5-0,77)	3,77 (1,9-7,49)	0,62 (0,42-0,91)	-	-	5,89 (2,81-12,77)	0,56 (0,4-0,79)	4,02 (2,06-10,38)	0,68 (0,54-0,87)	-	-	-	-	-	-	-	-	-	-
		3	-	-	-	-	2,62 (1,09-6,31)	0,96 (0,27-1,2)	-	-	-	-	-	-	-	-	-	-	-	-	2,69 (1,07-6,75)	0,94 (0,23-1,3)	-	-
		7	-	-	-	-	5,14 (1,02-25,86)	0,71 (0,32-1,59)	-	-	-	-	-	-	-	-	15,25 (12,21-19,03)	0	-	-	-	-	-	-
E	EFW < p3	6	-	-	1,23 (1,15-1,31)	0,47 (0,35-0,63)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1,2 (1,08-1,33)	0,45 (0,22-0,91)
		7	-	-	-	-	1,76 (0,79-3,93)	0,94 (0,11-2,66)	-	-	-	-	-	-	-	-	2,63 (2,44-2,84)	0	-	-	-	-	-	-
		10	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1,25 (0,79-2,14)	0,97 (0,9-1,05)
		11	-	-	-	-	-	-	-	-	-	-	-	-	2,26 (0,98-5,2)	0,94 (0,29-1,43)	-	-	2,81 (1,82-4,33)	0,57 (0,26-0,86)	-	-	-	-
		12	1,9 (0,91-4,45)	0,82 (0,67-1,01)	0,7 (0,12-4,13)	1,12 (0,71-1,76)	-	-	-	-	-	-	0,88 (0,16-4,92)	1,05 (0,59-1,86)	1,19 (0,23-6,01)	0,93 (0,41-2,09)	-	-	0,3 (0,09-1,19)	1,33 (1,1-1,6)	-	-	-	-
F	EFW < p10 + PI UA > p95 or EFW < p10 + CPR < p5	12	0,81 (0,44-0,84)	1,31 (1,08-3,39)	1,88 (1,04-2,70)	0,38 (0,07-2,22)	-	-	-	-	-	-	1,22 (0,39-2,77)	0,98 (0,36-2,64)	1,38 (0,61-3,12)	0,85 (0,13-3,23)	-	-	1,83 (1,43-2,58)	0,33 (0,14-0,82)	-	-	-	-
G	EFW or AC < p3	3	-	-	-	-	1,18 (0,72-1,96)	0,62 (0,1-3,85)	-	-	-	-	-	-	-	-	-	-	-	-	1,2 (0,72-1,99)	0,8 (0,1-3,72)		
H	EFW < p10 + CPR < p5	8	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2,02 (1,37-3)	0,84 (0,76-0,94)
I	EFW<p10 + PI UA > p95	1	-	-	1,45 (0,85-2,52)	0,86 (0,69-1,08)	1,4 (0,74-2,66)	0,86 (0,61-1,2)	-	-	0,91 (0,46-1,79)	1,03 (0,81-1,32)	1,44 (0,83-2,51)	0,86 (0,63-2,51)	-	-	-	-	-	-	-	-	-	-
		2	2,89 (1,50-3,8)	0,82 (0,75-0,89)	1,76 (1,29-2,4)	0,82 (0,73-0,93)	-	-	2,31 (1,7-3,13)	0,73 (0,63-0,84)	1,09 (0,85-1,84)	0,97 (0,82-1,15)	0,81 (0,29-2,27)	1,06 (0,83-1,35)	0	1,3 (1,24-1,36)	-	-	1,37 (0,85-2,88)	0,89 (0,64-1,24)	-	-	1,3 (0,89-1,91)	0,92 (0,79-1,08)
		5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	13,02 (1,53-110,36)	0,98 (0,93-1)
J	EFW<p10 + PI UA >p95 or A/REDF	1	-	-	2,16 (1,53-3,1)	0,4 (0,25-0,64)	1,88 (1,4-2,52)	0,32 (0,13-0,79)	1,45 (1,02-2,06)	0,68 (0,47-0,99)	1,75 (1,25-2,38)	0,45 (0,24-0,83)	1,84 (1,34-2,54)	0,49 (0,28-0,76)	-	-	2,1 (1,75-2,53)	0	-	-	-	-	4,86 (2,72-8,68)	0,24 (0,15-0,39)
		5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2,09 (2,5-2,91)	0	2,41 (2,1-2,77)	0,24 (0,13-0,44)
		6	-	-	1,76 (1,54-2,07)	0,66 (0,58-0,75)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2,34 (2,02-2,72)	0,29 (0,17-0,49)
		7	-	-	-	-	0,73 (0,15-3,61)	1,23 (0,55-2,74)	-	-	-	-	-	-	-	-	2,19 (2,05-2,33)	0	-	-	-	-	-	-
K	EFW < p10 + PI UA > p95 or CPR < p5 or EFW < p3	12	0,89 (0,71-1,03)	1,03 (0,73-5,1)	1,31 (1,19-1,45)	0	-	-	-	-	-	-	0,97 (0,55-1,72)	1,09 (0,19-6,16)	1,31 (1,19-1,44)	0	-	-	1,21 (1,02-1,45)	0,37 (0,1-1,45)	-	-	-	-
L	EFW < p3 or EFW < p10 + CPR < p5	9	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-			-	-	1,39 (0,99-2)	0,83 (0,67-1,04)

4.3 Recomendações

Delphi consensus for the diagnosis of Fetal Growth Restriction in singleton pregnancies: recommendations for a tertiary academic brazilian hospital

Authors

Marcela Maria Simões-Cora[1], Luís Otávio Sarian[2], João Renato Bennini[1]

[1]Fetal Medicine Program, Department of Gynecology and Obstetrics, Women's Hospital, State University of Campinas

[2]Department of Gynecology and Obstetrics, Women's Hospital, State University of Campinas

Corresponding author

João Renato Bennini Junior, head of the Fetal Medicine Program, Department of Gynecology and Obstetrics, State University of Campinas.

Alexander Fleming Street, 101, Department of Maternal-Fetal Medicine, Unicamp Women's Hospital, State University of Campinas, Brazil

+551935219500

bennini@unicamp.br

Conflicts of interest: none to disclosure.

Funding

The authors declare no potential conflicts of interest with respect to the research, authorship and/or publication of this article.

Keywords (DeCS): fetal growth restriction; placental insufficiency; ultrasound; diagnosis; perinatal outcomes; diagnostic test accuracy; review.

Executive Summary

For the diagnosis of FGR, should EFW < p10 or Delphi consensus criteria be used?

Both the tests, EFW < p10 and Delphi consensus, have similar accuracy for the diagnosis of adverse perinatal outcomes, which we use to define true positives. In the only study that evaluated the comparison of these two tests, the Delphi consensus had a slightly higher specificity than the EFW<p10, with no significant difference in sensitivity.

Both tests require ultrasound equipment to be performed. For the application of the Delphi consensus, the Doppler Velocimetry function and a qualified professional would be necessary to apply it. An ultrasound examination to assess all possible criteria proposed by the Delphi consensus would take longer than a simple obstetric ultrasound, but not to the point of bringing an important negative impact of the Delphi criteria in terms of costs and patients' preferences.

Based on accuracy, costs and resources, and patients preferences and acceptability, we suggest that places with trained providers and US with Doppler Velocimetry equipment, to use Delphi consensus criteria as a first choice for the diagnosis of FGR and its follow-up.

Quality of evidence: Low (●●○○).

Strength of recommendation: Weak recommendation for using Delphi consensus criteria (↑?)

Abbreviations and acronyms

AC	Abdominal Circumference
ACOG	American College of Obstetricians and Gynecologists
A/REDF	Absent or Reverse end Diastolic Flow
CPR	Cerebroplacental ratio
EFW	Estimated Fetal Weight
FGR	Fetal Growth Restriction
GA	Gestational Age
NICU	Neonatal Intensive Care Unit
PI	Pulsatility Index
UA	Umbilical Artery
US	Ultrasound
UtA	Uterine Artery

Background

Description of the condition

Fetal growth restriction (FGR) is defined as the failure of the fetus to reach its genetically determined growth potential [1] and affects about 5% of all pregnancies, with a progressive increase in its incidence over the decades [2]. The etiology of FGR is associated with several factors, including fetal, placental or maternal conditions. Regarding fetal etiologies, chromosomal defects, genetic syndromes, congenital malformations, congenital errors of metabolism and intrauterine infections should be considered [3,4].

Placental insufficiency, associated or as a consequence of maternal diseases, is the most common cause of FGR [5] and the correct identification of this condition is important in clinical practice because it is associated with poorer perinatal outcomes [6]. Pregnancies affected with FGR due to a placental insufficiency have higher risk of adverse perinatal outcomes [7-10].

The American College of Obstetricians and Gynecologists (ACOG) defines that a restricted fetus is one that has an EFW below the 10th percentile. In 2016, a multicentre team of international fetal medicine specialists conducted a study based on the Delphi method to establish a consensus on the definition of FGR [11]. They divided the FGR into early (<32 weeks) and late (from 32 weeks). So, in the absence of fetal malformations, they considered early FGR in the presence of: (i) fetal abdominal circumference (AC) below the third percentile for gestational age (GA), (ii) estimated fetal weight (EFW) below the third percentile for GA, (iii) absent or reverse umbilical artery (UA) end diastolic flow (A/REDF), (iv) EFW or AC below the tenth percentile for GA and pulsatility index (PI) of the UA or uterine artery (UtA) above the 95th percentile for GA. From this consensus, late FGR is characterized by: (i) fetal AC below the third percentile for GA, (ii) EFW below the third percentile for GA and (iii) the presence of at least two of the following parameters: (a) EFW or AC below the tenth percentile for GA, (b) reduction of more than two quartiles of the CA or

EFW in the growth curve and (c) cerebroplacental relation (CPR) below the fifth percentile for GA or the UA PI above the 95th percentile for GA.

Aims of this guideline

These recommendations aim to guide the choice of an antenatal diagnostic test for FGR in academic hospitals based on the latest and best evidence about diagnostic tests accuracy. Since there is no gold standard for defining FGR, we aim to compare the criteria proposed by the expert consensus (Delphi) with the parameter proposed by ACOG, which is widely used in clinical practice: EFW < p10 for GA.

Target population

These recommendations are suited for singleton pregnant women and fetuses without malformations, chromosomal disorders or fetal genetic syndromes.

End users and settings

These recommendations are intended for obstetricians, radiologists and maternal-fetal medicine specialists to make evidence-based decisions for the choice of a diagnostic test for FGR in the general pregnant population.

Guideline development group

The guideline development group was composed by both the authors: one medical resident in the Obstetrics and Gynecology program and one maternal-fetal medicine specialist, both vinculated to the Hospital Prof Dr Jose Aristodemo Pinotti, a tertiary academic center dedicated to Obstetrics and Gynecology of the State University of Campinas (UNICAMP).

Evidence

Health Care Condition

The recommendations are based on the following question elements:

P	population	Women with single pregnancies and fetuses without malformations, chromosomal disorders or fetal genetic syndromes, who were submitted to evaluation of obstetric ultrasound
I	Index test	Each of the criteria chosen by the Delphi consensus and the current definition of EFW < p10 only
R	Reference standard	Considering that truly restricted fetuses are those that have a higher risk of morbimortality, true positives were determined by the clinical diagnosis of poor perinatal results at the time of delivery and first days of life: NICU admission, fifth minute Apgar below 7, emergency C-section, neonatal acidosis, preterm birth, respiratory distress syndrome (RDS), small for gestational age (SGA), neonatal death, stillbirth and cerebral intraventricular hemorrhage (IVH).
T	Target condition	Fetal Growth Restriction represented by any adverse perinatal outcome (APO)

Systematic Review

We conducted a systematic review according to the Cochrane Diagnostic Test Accuracy (DTA) Reviews [12,13] guidelines to base this recommendation. For a detailed description of the systematic review please refer to topics 3.1 and 4.1 of this document. The clinical recommendations were written according to the guidelines of the RIGHT Statement [14].

Recommendations

For the diagnosis of FGR, should the current definition of EFW<p10 or Delphi consensus criteria be used?

Both the tests, EFW < p10 and Delphi consensus, have similar accuracy for the diagnosis of adverse perinatal outcomes, which we use to define true positives. In the only study that evaluated the comparison of these two tests, the Delphi consensus had a slightly higher specificity than the EFW<p10, with no significant difference in sensitivity.

Both tests require ultrasound equipment to be performed. For the application of the Delphi consensus, the Doppler Velocimetry function and a qualified professional would be necessary to apply it. An ultrasound examination to assess all possible criteria proposed by the Delphi consensus would take longer than a simple obstetric ultrasound, but not to the point of bringing an important negative impact of the Delphi criteria in terms of costs and patients' preferences.

Based on accuracy, costs and resources, and patients preferences and acceptability, we suggest that places with trained providers and US with Doppler Velocimetry equipment, to use Delphi consensus criteria as a first choice for the diagnosis of FGR and its follow-up.

Quality of evidence: Low (●●○○).

Strength of recommendation: Weak recommendation for using Delphi consensus criteria (↑?)

Question: Should Delphi consensus vs. EFW < p10 be used to diagnose FGR in singleton pregnancies ?

Delphi consensus		EFW < p10		Prevalences	5%	13%
Sensitivity	0.10 (95% CI: 0.06 to 0.16)	Sensitivity	0.09 (95% CI: 0.05 to 0.15)			
Specificity	0.95 (95% CI: 0.94 to 0.97)	Specificity	0.92 (95% CI: 0.90 to 0.93)			

Outcome	№ of studies (№ of patients)	Study design	Factors that may decrease certainty of evidence					Effect per 1.000 patients tested				Test accuracy CoE
								pre-test probability of5%		pre-test probability of13%		
			Risk of bias	Indirectness	Inconsistency	Imprecision	Publication bias	Delphi consensus	EFW < p10	Delphi consensus	EFW < p10	
True positives (patients with FGR)	1 studies 139 patients	cohort & case-control type studies	serious	not serious	serious	serious	all plausible residual confounding would suggest spurious effect, while no effect was observed	5 (3 to 8)	5 (3 to 8)	13 (8 to 21)	12 (7 to 20)	⨁⨁◯◯ Low
								0 fewer TP in Delphi consensus		**1 more TP in Delphi consensus**		
False negatives (patients incorrectly classified as not having FGR)								45 (42 to 47)	45 (42 to 47)	117 (109 to 122)	118 (110 to 123)	
								0 fewer FN in Delphi consensus		**1 fewer FN in Delphi consensus**		
True negatives (patients without FGR)	1 studies 916 patients	cohort & case-control type studies	serious	not serious	serious	serious	all plausible residual confounding would suggest spurious effect, while no effect was observed	903 (893 to 922)	874 (855 to 884)	827 (818 to 844)	800 (783 to 809)	⨁⨁◯◯ Low
								29 more TN in Delphi consensus		**27 more TN in Delphi consensus**		
False positives (patients incorrectly classified as having FGR)								47 (28 to 57)	76 (66 to 95)	43 (26 to 52)	70 (61 to 87)	
								29 fewer FP in Delphi consensus		**27 fewer FP in Delphi consensus**		

Figure 1: Should Delphi consensus vs EFW < p10 be used to diagnose FGR in singleton pregnancies? GRADE method.

Evidence to decision process

We used the GRADE approach to assess the quality of evidence [16-18], as described above. Risk of bias was evaluated following the QUADAS-2 assessment described previously in this article. We judged indirectness according to possible consequences of each result for patients and the healthcare service (Table 1).

Table 1. Presumed impact of test results

Test result	True Positive	True Negative	False Positive	False Negative
Presumed impact	Better case management	Safe follow-up in low risk prenatal care	Unnecessary high stress levels for patient and health care providers, higher costs, risk of iatrogenic prematurity	Higher maternal and fetal morbidity and mortality

Review and quality assurance

These recommendations were evaluated at the interdisciplinary team meeting involving obstetricians and sonologists. All suggestions were taken into consideration and addressed before the final text was released.

Funding, declaration and management of interests

These recommendations received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. No potential competing interest was reported by the authors.

Other information

Suggestions for further research

Considering that our review pointed to only one study that analyzed the use of the diagnostic method proposed by the Delphi consensus for the diagnosis of FGR, the authors suggest that studies that seek to validate this method in order to improve it or consolidate it in practice clinic of the different groups of obstetricians.

Limitations

The main limitation of our recommendations is the fact that the review that supported it showed only one study that answered the PIRT question. Still, there is no measurement of the follow-up protocol and resolution of the pregnancy, which may interfere with some of the adverse perinatal outcomes studied.

References

1. Resnik R. Intrauterine growth restriction. *Obstet Gynecol* 2002; **99** : 490-496.

2. Romo A, Carceller R, Tobajas J. Intrauterine growth retardation (IUGR): epidemiology and etiology. *Pediatr Endocrinol* Rev. 2006; **6(3)** : 332-336.

3. Lin CC, Santolaya-Forgas J. Current concepts of fetal growth restriction: part I. Causes, classification, and pathophysiology. Obstet Gynecol 1998; **92**:1044–1055.

4. Wollmann HA. Intrauterine Growth Restriction: Definition and Etiology. Horm Res 1998; **49(2)**:1-6.

5. Lackman F, Capewell V, Gagnon R, Richardson B. Fetal umbilical cord oxygen values and birth to placental weight ratio in relation to size at birth. *Am J Obstet Gynecol* 2001; **185** : 674-682.

6. Figueiras F, Gratacós E. Update on the Diagnosis and Classification of Fetal Growth Restriction and Proposal of a Stage-Based Management Protocol. *Fetal Diagn Ther* 2014; **36** : 86–98.

7. Baschat AA, Cosmi E, Bilardo CM, Wolf H, Berg C, Rigano S, Germer U, Moyano D, Turan S, Hartung J, Bhide A, Müller T, Bower S, Nicolaides KH, Thilaganathan B, Gembruch U, Ferrazzi E, Hecher K, Galan HL, Harman CR. Predictors of Neonatal Outcome in Early-Onset Placental Dysfunction. *Obstetrics & Gynecology* 2007; **109(2, pt 1)** : 253–261.

8. Francis JH, Permezel M, Davey MA. Perinatal mortality by birthweight centile. *Aust N. Z J Obstet Gynaecol* 2014; **54** : 354-359.

9. Miller SL, Huppi PS, Mallard C. The consequences of fetal growth restriction on brain structure and neurodevelopmental outcome. *J Physiol (Lond)* 2016; **594**: 807-823.

10. The GRIT Study Group. A randomised trial of timed delivery for the compromised preterm fetus: short term outcomes and Bayesian interpretation. *BJOG: An International Journal of Obstetrics and Gynaecology* 2003; **110(1)** : 27–32.

11. Gordijn SJ, Beune IM, Thilaganathan B, Papageorghiou A, Baschat AA, Baker PN, Silver RM, Wynia K, Ganzevoort W. Consensus definition of fetal growth restriction: a Delphi procedure. *Ultrasound Obstet Gynecol* 2016; **48(3)** : 333–339.

13. Garritty C, Gartlehner G, Kamel C, King VJ, Nussbaumer-Streit B, Stevens A, Hamel C, Affengruber L. Cochrane Rapid Reviews. Interim Guidance from the Cochrane Rapid Reviews Methods Group 2020.

14. Deeks JJ, Wisniewski S, Davenport C. Chapter 4: Guide to the contents of a Cochrane Diagnostic Test Accuracy Protocol. In: Deeks JJ, Bossuyt PM, Gatsonis C (editors), Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 1.0.0. The Cochrane Collaboration, 2013. Available on: http://srdta.cochrane.org/.

15. Chen Y, Yang K, Marušic A, Qaseem A, Meerpohl JJ, Flottorp S, Akl EA, Schünemann HJ, Chan ES, Falck-Ytter Y, Ahmed F, Barber S, Chen C, Zhang M, Xu B, Tian J, Song F, Shang H, Tang K, Wang Q, Norris SL; RIGHT (Reporting 63 Items for Practice Guidelines in Healthcare) Working Group. A Reporting Tool for Practice Guidelines in Health Care: The RIGHT Statement. Ann Intern Med. 2017 Jan 17;**166(2)**:128-132.

16. Guyatt G, Oxman A, Vist G, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ. 2008;336(7650):924-6.

17. Guyatt G, Oxman A, Kunz R, et al. What is "quality of evidence" and why is it important to clinicians? BMJ. 2008;**336(7651)**:995-8.

18. Schünemann H, Oxman A, Brozek J, et al. Grading quality of evidence and strength of recommendations for diagnostic tests and strategies. BMJ. 2008;**336(7653):**1106-10.

5. Discussão

5.1 Resumo das Evidências

A revisão da literatura apresentou apenas um estudo que se propôs a analisar o método proposto pelo consenso Delphi para o diagnóstico de RCF, a partir dos desfechos perinatais adversos que poderiam estar associados a esta condição. Considerando que a presença de apenas um critério já define o feto como restrito, sem necessidade de contemplar todos os outros critérios também, analisamos cada critério isoladamente e, deste modo, encontramos 12 estudos que avaliaram ao menos um critério do consenso Delphi associado a desfechos perinatais adversos.

Desde a publicação dos critérios de consenso Delphi para RCF, apenas um estudo se concentrou na validação do consenso como um todo como um teste diagnóstico (Molina, 2020 [30]), avaliando os desfechos perinatais adversos que poderiam estar associados a esta condição.

Nos últimos anos, os critérios diagnósticos para RCF, propostos pelo consenso Delphi, mais avaliados foram alterações no IP da AU e PFE abaixo do percentil 3. Não foram encontrados estudos que avaliassem a velocidade de crescimento fetal em quartis ou o IP elevado da AUt abaixo de 32 semanas (associado a PFE ou CA abaixo do percentil 10 para IG) para predizer resultados perinatais adversos e RCF. Os desfechos adversos perinatais mais avaliados em relação aos critérios propostos pelo consenso Delphi foram APO, internação em UTI neonatal, óbito neonatal e Apgar de 5º. minuto abaixo de 7.

O único estudo encontrado que avaliou o consenso Delphi como um todo [30] mostrou uma ligeira superioridade dos critérios Delphi para a predição de desfechos perinatais adversos, quando comparado ao critério de PFE <p10 isoladamente (comumente aplicado na prática clínica). De acordo com a DOR, os achados ultrassonográficos que tiveram melhor desempenho para predizer resultados perinatais adversos foram: "AC < p3", "PFE < p10 + IP AU > p95" e “fluxo diastólico reverso ou ausente na artérias umbilical

(FDFA/R)". Este último, em especial, apresenta a melhor área sob a curva ROC em comparação aos outros testes índice.

A meta-análise mostrou que a DOR para a presença de qualquer critério de consenso Delphi para o diagnóstico de RCF para detectar qualquer desfecho adverso perinatal foi de 2,86 (IC 95%: 2,32-3,54), com alta heterogeneidade (70%). Esta elevada variabilidade entre os estudos já foi apontada em outros estudos que buscaram avaliar a relação dos achados ultrassonográficos com desfechos perinatais adversos [39,40]. Essas diferenças podem ser atribuídas aos diferentes limiares de idade gestacional usados para seleção de pacientes nos estudos e à ausência de uma definição padronizada das condições incluídas no “APO”. Além disso, diferentes protocolos de manejo da RCF e de interrupção da gravidez também podem ter contribuído para a variabilidade dos indicadores de desempenho relatados por diferentes autores.

5.2 Recomendações

A recomendação foi elaborada de acordo com o RIGHT statement [24] e utilizando a ferramenta e metodologia GRADE [25,26], comparando os critérios diagnósticos propostos pelo consenso Delphi para RCF [18] e o critério mais amplamente utilizado na prática clínica do PFE abaixo do percentil 10 para a IG [15,16]. Assim, para responder à pergunta: "Para o diagnóstico de RCF, devemos utilizar os critérios do consenso Delphi ou PFE > percentil 10 para a IG?” foi ponderada a qualidade da evidência a partir dos potenciais efeitos positivos e negativos para a população em que o teste fosse verdadeiramente positivo, verdadeiramente negativo, falsamente positivo e falsamente negativo. Também foram considerados os custos de cada exame (equipamento adequado, profissional capacitado, tempo de exame, considerando US obstétrico simples para medida de peso versus obstétrico com Dopplerfluxometria) e as preferências profissionais.

A revisão sistemática realizada encontrou apenas um estudo comparando ambos os testes, sem especificar o protocolo de seguimento e resolução das gestações, o que acarreta potencial viés na análise dos desfecho e consequentemente baixa qualidade de

evidência. Neste estudo comparativo [30] demonstrou-se que o PFE<p10 e o consenso Delphi têm precisão semelhante para o diagnóstico de desfechos perinatais adversos (aqui utilizados como referência para definição dos casos positivos). No único estudo que avaliou a comparação desses dois testes, o consenso Delphi apresentou especificidade um pouco maior que o PFE<p10, sem diferença significativa na sensibilidade.

Como as medidas de acurácia foram similares para ambos os exames, equiparando os possíveis benefícios e riscos dos resultados, a força da evidência se baseou principalmente nos possíveis custos e preferências e, considerando a heterogeneidade dos recursos técnicos e profissionais dos diferentes serviços de saúde a recomendação de utilizar o consenso Delphi, em detrimento do PFE < p10, para o diagnóstico de RCF foi considerada fraca.

Ambos os testes requerem equipamentos de ultrassom para serem realizados. Para a aplicação do consenso Delphi seria necessária a função Dopplervelocimetria e um profissional habilitado para aplicá-la. Um exame de ultrassom para avaliar todos os possíveis critérios propostos pelo consenso Delphi levaria mais tempo do que um simples ultrassom obstétrico, mas não a ponto de trazer um impacto negativo importante dos critérios Delphi em termos de custos e preferências das pacientes.

Com base na precisão, custos e recursos e nas preferências e aceitabilidade dos pacientes, sugerimos que os locais com provedores treinados e US com equipamento de Dopplervelocimetria usem os critérios de consenso Delphi como primeira escolha para o diagnóstico de RCF e seu acompanhamento.

6. Conclusão

A revisão sistemática encontrou ao todo 12 estudos que apresentaram dados de sensibilidade e especificidade dos critérios propostos pelo consenso Delphi para o diagnóstico de RCF em relação à presença de desfechos perinatais adversos, que caracterizam o feto verdadeiramente restrito. Os critérios Delphi mais avaliados foram alterações no IP da AU e PFE abaixo do percentil 3. Os achados ultrassonográficos que tiveram melhor desempenho para predizer resultados perinatais adversos foram: "CA < p3", "PFE < p10 + IP AU > p95", e "FDFA/R".

Nosso estudo demonstrou que desde a publicação do consenso Delphi para o diagnóstico de RCF apenas um estudo se concentrou em sua validação como um teste diagnóstico. Este estudo comparou os critérios Delphi com a medida isolada do PFE abaixo do percentil 10 e apontou uma discreta superioridade dos critérios Delphi para a predição de desfechos perinatais adversos, apesar de ambos os métodos diagnósticos apresentarem baixa sensibilidade (10,1% versus 9.3%).

Com base na precisão, custos e recursos e nas preferências e aceitabilidade dos pacientes, sugerimos que os locais com provedores treinados e US com equipamento de Doppler Velocimetria usem os critérios de consenso Delphi como primeira escolha para o diagnóstico de RCF e seu acompanhamento.

Qualidade da evidência: Baixa (●●○○).

Força da recomendação: fraca a favor de usar os critérios de consenso Delphi para o diagnóstico de RCF (↑?)

7. Referências bibliográficas

1. Resnik R. Intrauterine growth restriction.Obstet Gynecol 2002; 99:490-6.
2. Romo A, Carceller R, Tobajas J. Intrauterine growth retardation (IUGR): epidemiology and etiology. Pediatr Endocrinol Rev. 2006; 6(3):332-6.
3. Lin CC, Santolaya-Forgas J. Current concepts of fetal growth restriction: part I. Causes, classification, and pathophysiology. Obstet Gynecol 1998; 92:1044–1055.
4. Wollmann HA. Intrauterine Growth Restriction: Definition and Etiology. Horm Res 1998; 49(2):1-6.
5. Townsend R, Khalil A. Fetal Growth Restriction in twins. Best Practice & Research Clinical Obstet Gynaecol 2018; 49: 79-88.
6. Howarth C, et al. Associations of type 1 diabetes mellitus, maternal vascular disease and complications of pregnancy. Diabet Med 2007; 24:1229–34.
7. Yasmeen S, et al. Pregnancy outcomes in women with systemic lupus erythematosus. J Matern Fetal Med 2001; 10:91–6.
8. Allen VM, et al. The effect of hypertensive disorders in pregnancy on small for gestational age and stillbirth: a population based study. BMC Pregnancy Childbirth 2004; 4:17–25.
9. Yasuda M, et al. Prospective studies of the association between anticardiolipin antibody and outcome of pregnancy. Obstet Gynecol 1995; 86:555–9.
10. Rasmussen S, Irgens LM. Fetal growth and body proportion in pre-eclampsia. Obstet Gynecol 2003; 101:575–583.
11. Francis JH, Permezel M, Davey MA. Perinatal mortality by birthweight centile. Aust N. Z J Obstet Gynaecol 2014; 54 : 354-359.
12. Miller SL, Huppi PS, Mallard C. The consequences of fetal growth restriction on brain structure and neurodevelopmental outcome. J Physiol (Lond) 2016; 594: 807-823.

13. The GRIT Study Group. A randomised trial of timed delivery for the compromised preterm fetus: short term outcomes and Bayesian interpretation. BJOG: An International Journal of Obstetrics and Gynaecology 2003; 110(1) : 27–32.
14. Baschat AA, Cosmi E, Bilardo CM, Wolf H, Berg C, Rigano S, Germer U, Moyano D, Turan S, Hartung J, Bhide A, Müller T, Bower S, Nicolaides KH, Thilaganathan B, Gembruch U, Ferrazzi E, Hecher K, Galan HL, Harman CR. Predictors of Neonatal Outcome in Early- Onset Placental Dysfunction. Obstetrics & Gynecology 2007; 109(2, pt 1) : 253–261.
15. American College of Obstetricians and Gynecologists ACOG Practice bulletin no. 134: fetal growth restriction. Obstet Gynecol, 2013; 121:1122-1133.
16. ISUOG Practice Guidelines: diagnosis and management of small-for-gestational-age fetus and fetal growth restriction. Ultrasound Obstet Gynecol 2020; 56 : 298–312.
17. Jones J, Hunter D. Consensus methods for medical and health services research. BMJ. 1995; 311 : 376–380.
18. Gordijn SJ, Beune IM, Thilaganathan B, Papageorghiou A, Baschat AA, Baker PN, Silver RM, Wynia K, Ganzevoort W. Consensus definition of fetal growth restriction: a Delphi procedure. Ultrasound Obstet Gynecol 2016; 48(3) : 333–339.
19. Garritty C, Gartlehner G, Kamel C, King VJ, Nussbaumer-Streit B, Stevens A, Hamel C, Affengruber L. Cochrane Rapid Reviews. Interim Guidance from the Cochrane Rapid Reviews Methods Group 2020.
20. Deeks JJ, Wisniewski S, Davenport C. Chapter 4: Guide to the contents of a Cochrane Diagnostic Test Accuracy Protocol. In: Deeks JJ, Bossuyt PM, Gatsonis C (editors), Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 1.0.0. The Cochrane Collaboration, 2013. Available on: http://srdta.cochrane.org/.
21. Mourad Ouzzani, Hossam Hammady, Zbys Fedorowicz, and Ahmed Elmagarmid. Rayyan — a web and mobile app for systematic reviews. Systematic Reviews (2016) 5:210, DOI: 10.1186/s13643-016-0384-4.

22. Whitining PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reistma JB, Leflang MM, Sterne JA, Bossuyt PM. QUADAS-2: A revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med 2011; 155: 529-536.
23. Doebler P & Holling H. Meta-analysis of diagnostic accuracy with mada. Available from: https://cran.r-project.org/web/packages/mada/vignettes/mada.pdf. Acessed on October 31, 2020.
24. Chen Y, Yang K, Marušic A, Qaseem A, Meerpohl JJ, Flottorp S, Akl EA, Schünemann HJ, Chan ES, Falck-Ytter Y, Ahmed F, Barber S, Chen C, Zhang M, Xu B, Tian J, Song F, Shang H, Tang K, Wang Q, Norris SL; RIGHT (Reporting 63 Items for Practice Guidelines in Healthcare) Working Group. A Reporting Tool for Practice Guidelines in Health Care: The RIGHT Statement. Ann Intern Med. 2017 Jan 17;166(2):128-132. doi: 10.7326/M16-1565. Epub 2016 Nov 22. PMID: 27893062.
25. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, Schünemann HJ; GRADE Working Group. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ. 2008 Apr 26;336(7650):924-6.
26. Guyatt GH, et al. GRADE: going from evidence to recommendations. BMJ, 2008; 336: 1049-1051.
27. Agarwal R, Tiwari A, Wadhwa N, Radhakrishnan G, Bhatt S, Batra P. Abnormal umbilical artery Doppler velocimetry and placenta histopathological correlation in fetal growth restriction. *SAJOG* 2017; 23(1) : 12-16.
28. Maggio L, Dahlke JD, Mendez-Figueroa H, Albright CM, Chauhan SP, Wenstrom KD. Perinatal Outcomes With Normal Compared With Elevated Umbilical Artery Systolic-to-Diastolic Ratios in Fetal Growth Restriction. *Obstetrics & Gynecology* 2015; 125(4) : 863-869.
29. Marsoosi V, Bahadori F, Esfahani F, Ghasemi-Rad M. The role of doppler indices in predicting intra ventricular hemorrhage and perinatal mortality in fetal growth restriction. *Medical Ultrasonography* 2012; 14(2) : 125-132.

30. Molina LCG, Odibo L, Zientara S, Običan SG, Rodriguez A, Stout M, Odibo AO. Validation of Delphi procedure consensus criteria for defining fetal growth restriction. *Ultrasound Obstet Gynecol* 2020; 56 : 61–66.
31. O'Dwyer V, Burke G, Unterscheider J, Daly S, Geary MP, Kennelly MM, McAuliffe FM, O'Donoghue K, Hunter A, Morrison JJ, Dicker P, Tully EC, Malone FD. Defining the residual risk of adverse perinatal outcome in growth-restricted fetuses with normal umbilical artery blood flow. *Am J Obstet Gynecol* 2014; 211 : 421-425.
32. Unterscheider J, Daly S, Geary MP, Kennelly MM, McAuliffe FM, O'Donoghue K, Hunter A, Morrison JJ, Burke G, Dicker P, Tully EC, Malone FD. Optimizing the definition of intrauterine growth restriction: the multicenter prospective PORTO Study. *Am J Obstet Gynecol* 2013; 208 : 291-296.
33. Unterscheider J, O'Donoghue K, Daly S, Geary MP, Kennelly MM, McAuliffe FM, Hunter A, Morrison JJ, Burke G, Dicker P, Tully EC, Malone FD. Fetal growth restriction and the risk of perinatal mortality–case studies from the multicentre PORTO study. *BMC Pregnancy and Childbirth* 2014; 14(63) : 1-6.
34. Caradeux J, Eixarch E, Mazarico E, Basuki TR, Gratacos E, Figueras F. Longitudinal growth assessment for prediction of adverse perinatal outcome in fetuses suspected to be small-forgestational age. *Ultrasound Obstet Gynecol* 2018; 52: 325–331.
35. Cavallaro A, Veglia M, Svirko E, Vannuccini S, Volpe G, Impey L; Using fetal abdominal circumference growth velocity in the prediction of adverse outcome in near-term small-for-gestational-age fetuses. *Ultrasound Obstet Gynecol* 2018; 52: 494–500.
36. Figueiras F, Savchev S, Triunfo S, Crovetto F, Gratacos E. An integrated model with classification criteria to predict small-for-gestational-age fetuses at risk of adverse perinatal outcome. *Ultrasound Obstet Gynecol* 2015; 45 : 279–285.
37. Savchev, S, Figueras F, Cruz-Martinez R, Illa M, Botet F, Gratacos E. Estimated weight centile as a predictor of perinatal outcome in small-for-gestational-age

pregnancies with normal fetal and maternal Doppler indices. *Ultrasound Obstet Gynecol* 2012; 39 : 299–303.

38. Villalaína C, Herraiz I, Quezada MS, Gómez-Arriaga PI, Gómez-Montes E, Galindo A. Fetal Biometry and Doppler Study for the Assessment of Perinatal Outcome in Stage I Late-Onset Fetal Growth Restriction. *Fetal Diagn Ther* 2018; DOI: 10.1159/000485124.

Apêndice 1: Estratégias de buscas utilizadas em cada base de dados

Fonte	Estratégia	Nº de Artigos	Data
PUBMED	(((Fetal Growth Retardation[MeSH Terms]) OR ("Fetal Growth Retardation" OR "Intrauterine Growth Retardation" OR "Growth Retardation, Intrauterine" OR "Intrauterine Growth Restriction" OR "Fetal Growth Restriction"[MeSH Terms])) AND ((Diagnosis[MeSH Terms]) OR (Diagnosis[Title/Abstract] OR Diagnoses[Title/Abstract] OR "Diagnoses[Title/Abstract] AND Examinations"[Title/Abstract] OR "Examinations[Title/Abstract] AND Diagnoses"[Title/Abstract] OR "Postmortem Diagnosis"[Title/Abstract] OR "Diagnoses, Postmortem"[Title/Abstract] OR "Diagnosis, Postmortem"[Title/Abstract] OR "Postmortem Diagnoses"[Title/Abstract] OR "Antemortem Diagnosis"[Title/Abstract] OR "Antemortem Diagnoses"[Title/Abstract] OR "Diagnoses, Antemortem"[Title/Abstract] OR "Diagnosis, Antemortem"[Title/Abstract]))) AND ((("Delphi procedure" OR DELPHI OR "Delphi consensus") OR ((Fetal Weight[MeSH Terms]) OR ("Fetal Weight"[Title/Abstract] OR "Fetal Weights"[Title/Abstract] OR "Weight, Fetal"[Title/Abstract] OR "Weights, Fetal"[Title/Abstract] OR "Body Weight, Fetal"[Title/Abstract] OR "Body Weights, Fetal"[Title/Abstract] OR "Fetal Body Weight"[Title/Abstract] OR "Fetal Body Weights"[Title/Abstract]))) OR ((Ultrasonography, Doppler[MeSH Terms]) OR ("Ultrasonography, Doppler"[Title/Abstract] OR "Doppler Ultrasound"[Title/Abstract] OR "Doppler Ultrasounds"[Title/Abstract] OR "Ultrasound, Doppler"[Title/Abstract] OR "Ultrasounds, Doppler"[Title/Abstract] OR "Doppler Ultrasonography"[Title/Abstract] OR "Doppler Ultrasound Imaging"[Title/Abstract] OR "Doppler Ultrasound Imagings"[Title/Abstract] OR "Imaging, Doppler Ultrasound"[Title/Abstract] OR "Imagings,	861	17/08/2020

	Doppler Ultrasound"[Title/Abstract] OR "Ultrasound Imaging, Doppler"[Title/Abstract] OR "Ultrasound Imagings, Doppler"[Title/Abstract]))) Filters: in the last 10 years, English		
PUBMED	(((Fetal Growth Retardation[MeSH Terms]) OR ("Fetal Growth Retardation" OR "Intrauterine Growth Retardation" OR "Growth Retardation, Intrauterine" OR "Intrauterine Growth Restriction" OR "Fetal Growth Restriction"[MeSH Terms])) AND ((Diagnosis[MeSH Terms]) OR (Diagnosis[Title/Abstract] OR Diagnoses[Title/Abstract] OR "Diagnoses[Title/Abstract] AND Examinations"[Title/Abstract] OR "Examinations[Title/Abstract] AND Diagnoses"[Title/Abstract] OR "Postmortem Diagnosis"[Title/Abstract] OR "Diagnoses, Postmortem"[Title/Abstract] OR "Diagnosis, Postmortem"[Title/Abstract] OR "Postmortem Diagnoses"[Title/Abstract] OR "Antemortem Diagnosis"[Title/Abstract] OR "Antemortem Diagnoses"[Title/Abstract] OR "Diagnoses, Antemortem"[Title/Abstract] OR "Diagnosis, Antemortem"[Title/Abstract]))) AND ((("Delphi procedure" OR DELPHI OR "Delphi consensus") OR ((Fetal Weight[MeSH Terms]) OR ("Fetal Weight"[Title/Abstract] OR "Fetal Weights"[Title/Abstract] OR "Weight, Fetal"[Title/Abstract] OR "Weights, Fetal"[Title/Abstract] OR "Body Weight, Fetal"[Title/Abstract] OR "Body Weights, Fetal"[Title/Abstract] OR "Fetal Body Weight"[Title/Abstract] OR "Fetal Body Weights"[Title/Abstract]))) OR ((Ultrasonography, Doppler[MeSH Terms]) OR ("Ultrasonography, Doppler"[Title/Abstract] OR "Doppler Ultrasound"[Title/Abstract] OR "Doppler Ultrasounds"[Title/Abstract] OR "Ultrasound, Doppler"[Title/Abstract] OR "Ultrasounds, Doppler"[Title/Abstract] OR "Doppler Ultrasonography"[Title/Abstract] OR "Doppler Ultrasound Imaging"[Title/Abstract] OR "Doppler Ultrasound Imagings"[Title/Abstract] OR "Imaging, Doppler	285	17/08/2020

	Ultrasound"[Title/Abstract] OR "Imagings, Doppler Ultrasound"[Title/Abstract] OR "Ultrasound Imaging, Doppler"[Title/Abstract] OR "Ultrasound Imagings, Doppler"[Title/Abstract])))) Filters: in the last 10 years, English		
EMBASE	('intrauterine growth retardation'/exp OR 'intrauterine growth retardation'/syn) AND ('diagnosis'/exp OR 'diagnosis'/syn) AND ('doppler ultrasonography'/exp OR 'doppler ultrasonography'/syn OR 'fetus weight'/exp OR 'fetus weight'/syn OR 'delphi procedure'/exp OR 'delphi procedure' OR 'delphi'/exp OR delphi OR 'delphi consensus') AND (2010:py OR 2011:py OR 2012:py OR 2013:py OR 2014:py OR 2015:py OR 2016:py OR 2017:py OR 2018:py OR 2019:py OR 2020:py) AND [english]/lim AND [embase]/lim NOT ([embase]/lim AND [medline]/lim)	401	17/08/2020

COCHRANE LIBRARY	MeSH descriptor: [Fetal Growth Retardation] explode all trees OR "Fetal Growth Retardation" OR "Intrauterine Growth Retardation" OR "Growth Retardation, Intrauterine" OR "Intrauterine Growth Restriction" OR "Fetal Growth Restriction"):ti,ab,kw AND MeSH descriptor: [Diagnosis] explode all trees OR (Diagnosis OR Diagnoses OR "Diagnoses and Examinations" OR "Examinations and Diagnoses" OR "Postmortem Diagnosis" OR "Diagnoses, Postmortem" OR "Diagnosis, Postmortem" OR "Postmortem Diagnoses" OR "Antemortem Diagnosis" OR "Antemortem Diagnoses" OR "Diagnoses, Antemortem" OR "Diagnosis, Antemortem"):ti,ab,kw AND MeSH descriptor: [Ultrasonography, Doppler] explode all trees OR ("Ultrasonography, Doppler" OR "Doppler Ultrasound" OR "Doppler Ultrasounds" OR "Ultrasound, Doppler" OR "Ultrasounds, Doppler" OR "Doppler Ultrasonography" OR "Doppler Ultrasound Imaging" OR "Doppler Ultrasound Imagings" OR "Imaging, Doppler Ultrasound" OR "Imagings, Doppler Ultrasound" OR "Ultrasound Imaging, Doppler" OR "Ultrasound Imagings, Doppler"):ti,ab,kw OR MeSH descriptor: [Fetal Weight] explode all trees OR ("Fetal Weight" OR "Fetal Weights" OR "Weight, Fetal" OR "Weights, Fetal" OR "Body Weight, Fetal" OR "Body Weights, Fetal" OR "Fetal Body Weight" OR "Fetal Body Weights"):ti,ab,kw OR ("Fetal Weight" OR "Fetal Weights" OR "Weight, Fetal" OR "Weights, Fetal" OR "Body Weight, Fetal" OR "Body Weights, Fetal" OR "Fetal Body Weight" OR "Fetal Body Weights"):ti,ab,kw	60	17/08/2020

Apêndice 2: Artigos selecionados por título e resumo.

Número da revisão	Primeiro autor	Título	Ano	Desenho	Número de pacientes
1	Agarwal, R	Abnormal umbilical artery doppler velocimetry and placental histopathological correlation in fetal growth restriction	2017	Prospectivo	130
2	Akram, H	A comparison of fetal outcome in pregnant women with fetal growth restriction in normal & abnormal umbilical artery doppler	2013	Prospectivo	100
3	Al-Azad, S	Doppler prediction of perinatal outcome in pregnancy induced hypertension and intrauterine growth retardation	2010	Prospectivo	s/dados
4	Ali, A	Comparison of perinatal outcome of growth restricted fetuses with normal and abnormal umbilical artery Doppler waveforms	2014	Prospectivo	100
5	Bardien, N	Placental Insufficiency in Fetuses That Slow in Growth but Are Born Appropriate for Gestational Age: A Prospective Longitudinal Study	2016	Prospectivo	48
6	Barker, E	The role of growth trajectories in classifying fetal growth restriction	2013	Prospectivo	1116
7	Blue, NR	A Comparison of Methods for the Diagnosis of FGR Between the Royal College of Obstetricians and Gynaecologists and the ACOG	2018	Retrospectivo	1704
8	Caradeux, J	Longitudinal growth assessment for prediction of adverse perinatal outcome in fetuses suspected to be small-for-gestational age	2018	Prospectivo	472
9	Cavallaro, A	Using fetal abdominal circumference growth velocity in the prediction of adverse outcome in near-term small-for-gestational-age fetuses	2018	Prospectivo	235
10	Cruz-Martinez, R	Clinical utility of third-trimester UtA Doppler in the prediction of brain hemodynamic deterioration and adverse perinatal outcome in SGA fetuses	2015	Prospectivo	327
11	Del Moral, R	Diagnosis and follow-up of in patients diagnosed with intrau-terine growth restriction in our center	2015	Retrospectivo	62
12	Demirci, O	Maternal and fetal risk factors affecting perinatal mortality in early and late fetal growth restriction	2015	Retrospectivo	271
13	Dahnd, H	Middle cerebral artery Doppler indices better predictor for fetal outcome in IUGR	2011	Prospectivo	121
14	Figueiras, F	An integrated model with classification criteria to predict small-for-gestational-age fetuses at risk of adverse perinatal outcome	2015	Prospectivo	509
15	Filmar, G	Elevated umbilical artery systolic/diastolic ratio in the absence of fetal growth restriction	2013	Retrospectivo	330
16	Flood, K	The role of brain sparing in the prediction of adverse outcomes in intrauterine growth restriction: results of the multicenter PORTO Study	2014	Prospectivo	881
17	Herman, HG	Placental pathology and neonatal outcome in small for gestational age pregnancies with and without abnormal umbilical artery Doppler flow	2018	Prospectivo	158
18	Geerts, L	Placental insufficiency among high-risk pregnancies with a normal umbilical artery resistance index after 32weeks	2016	Prospectivo	210
19	Gomez-Roig, M	Use of placental growth factor and uterine artery doppler PI in pregnancies involving intrauterine FGR or PE to predict perinatal outcomes	2015	Prospectivo	s/dados
20	Griffin, M	Comparison of PlGF and other biomarkers against current US parameters for predicting delivery of SGA infants in women with suspected PE: the PELICAN study	2014	Prospectivo	274
21	Griffin, M	Diagnostic accuracy of placental growth factor and US parameters to predict the SGA infant in women presenting with reduced symphysis-fundus height	2015	Prospectivo	601
22	Kessous, R	Umbilical artery peak systolic velocity measurements for prediction of perinatal outcome among IUGR fetuses	2014	Prospectivo	72
23	Kywe, KL	Uterine artery and umbilical artery doppler in prediction of fetal outcomes in severe preeclampsia	2017	Prospectivo	58
24	Lalkhrawsthangi, K	Study of colour doppler velocimetry in intra-uterine growth retardation	2011	Prospectivo	100
25	Larkin, J	Small for gestational age: The differential mortality when detected versus undetected antenatally	2015	Retrospectivo	125.069

26	Leftwich, HK	Doppler ultrasonography: More than just for intrauterine growth restriction?	2014	Retrospectivo	529
27	Maggio, L	Perinatal outcomes with normal compared with elevated umbilical artery systolic-to-diastolic ratios in fetal growth restriction	2015	Retrospectivo	512
28	Marsoosi, V	The role of Doppler indices in predicting intra ventricular hemorrhage and perinatal mortality in fetal growth restriction	2012	Prospectivo	43
29	Miranda, J	Prediction of fetal growth restriction using estimated fetal weight vs a combined screening model in the third trimester	2017	Prospectivo	1590
30	Molina, L	Validation of the Delphi Procedure Consensus criteria for defining Fetal Growth Restriction (FGR)	2019	Prospectivo	1055
31	Monaghan, C	Perinatal loss at term: role of uteroplacental and fetal Doppler assessment	2018	Retrospectivo	7013
32	Monier, I	Gestational age at diagnosis of early-onset fetal growth restriction and impact on management and survival: a population-based cohort study	2017	Prospectivo	436
33	Monteith, C	Evaluation of normalization of cerebro-placental ratio as a potential predictor for adverse outcome in SGA fetuses	2017	Prospectivo	1116
34	Monteith, C	Is a normalizing CPR a potential predictor for adverse outcome in intrauterine growth restriction: Results of the multicenter PORTO Study	2015	Prospectivo	881
35	O'Dwyer	Defining the residual risk of adverse perinatal outcome in growth-restricted fetuses with normal umbilical artery blood flow	2014	Prospectivo	1116
36	Rial-Crestelo, M	Added value of cerebro-placental ratio and uterine artery Doppler at routine third trimester screening as a predictor of SGA and FGR in non-selected pregnancies	2019	Prospectivo	1030
37	Roma, E	Ultrasound screening for fetal growth restriction at 36 vs 32 weeks' gestation: a randomized trial (ROUTE)	2015	Prospectivo	2586
38	Ropacka-Lesiak, M	Cerebroplacental ratio in prediction of adverse perinatal outcome and fetal heart rate disturbances in uncomplicated pregnancy at 40 weeks and beyond	2015	Retrospectivo	148
39	Roy, A	Perinatal outcome in pregnancies with intra-uterine growth restriction by using umbilical and middle cerebral artery colour Doppler	2012	Prospectivo	50
40	Savchev, S	Estimated weight centile as a predictor of perinatal outcome in small-for-gestational-age pregnancies with normal fetal and maternal Doppler indices	2012	Prospectivo	264
41	Sheth, T	Third-Trimester Fetal Biometry and Neonatal Outcomes in Term and Preterm Deliveries	2016	Retrospectivo	2692
42	Siddiqui, TS	Comparison of perinatal outcome in growth restricted fetuses retaining normal umbilical artery Doppler flow to those with diminished end-diastolic flow	2014	Prospectivo	60
43	Sirico, A	Prediction of adverse perinatal outcome by cerebroplacental ratio adjusted for estimated fetal weight	2018	Retrospectivo	3515
44	Sovio, U	Screening for fetal growth restriction (FGR) using universal third trimester ultrasonography: A prospective cohort study of 3,977 nulliparous women	2015	Prospectivo	3977
45	Sovio, U	Screening for FGR with universal third trimester ultrasonography in nulliparous women in the POP study: a prospective cohort study	2015	Prospectivo	3977
46	Spencer, R	Prediction of perinatal mortality for the mid-trimester intrauterine growth restricted fetus using ultrasound and angiogenic markers	2017	Prospectivo	57
47	Swanzy, L	Neonatal morbidity is increased with the inaccurate diagnosis of fetal growth restriction	2017	Retrospectivo	40.577
48	Triunfo, S	Prediction of delivery of small-for-gestational-age neonates and adverse perinatal outcome by fetoplacental Doppler at 37 weeks' gestation	2017	Prospectivo	946
49	Unterscheider, J	Optimizing the definition of intrauterine growth restriction: The multicenter prospective PORTO study	2013	Prospectivo	1116
50	Unterscheider, J	Fetal growth restriction and the risk of perinatal mortality-case studies from the multicentre PORTO study	2014	Prospectivo	1116
51	Villalaina, C	Fetal Biometry and Doppler Study for the Assessment of Perinatal Outcome in Stage I Late-Onset Fetal Growth Restriction	2018	Retrospectivo	131
52	Vollgraff, HS	Doppler measurements of both umbilical arteries do not improve predictive value for adverse perinatal outcomes in small-for-gestational age fetuses	2018	Prospectivo	124
53	Yamamoto, R	Ultrasonographic prediction of antepartum deterioration of growth-restricted fetuses after late preterm	2018	Retrospectivo	214

Apêndice 3: Artigos excluídos e motivo da exclusão

Número da revisão	Primeiro autor	Motivo da exclusão
2	Akram, H	Não inclui critérios do consenso Delphi.
3	Al-Azad, S	Artigo completo indisponível.
4	Ali, A	Não inclui critérios do consenso Delphi.
5	Bardien, N	Não inclui critérios do consenso Delphi.
6	Barker, E	Artigo completo indisponível.
7	Blue, NR	Desfecho divergente do interesse.
10	Cruz-Martinez, R	Não inclui critérios do consenso Delphi.
11	Del Moral, R	Artigo completo indisponível.
12	Demirci, O	Não inclui critérios do consenso Delphi.
13	Dahnd, H	Não inclui critérios do consenso Delphi.
15	Filmar, G	Não inclui critérios do consenso Delphi.
16	Flood, K	Não inclui critérios do consenso Delphi.
17	Herman, HG	Não inclui critérios do consenso Delphi.
18	Geerts, L	Desfecho divergente do interesse.
19	Gomez-Roig, M	Artigo completo indisponível.
20	Griffin, M	Desfecho e critérios divergentes.
21	Griffin, M	Desfecho divergente do interesse.
22	Kessous, R	Não inclui critérios do consenso Delphi.
23	Kywe, KL	Artigo completo indisponível.
24	Lalkhrawsthangi, K	Artigo completo indisponível.
25	Larkin, J	Artigo completo indisponível.
26	Leftwich, HK	Não inclui critérios do consenso Delphi.
29	Miranda, J	Desfecho divergente do interesse.
31	Monaghan, C	Não é estudo de acurácia diagnóstica.
32	Monier, I	Não é estudo de acurácia diagnóstica.
33	Monteith, C	Não inclui critérios do consenso Delphi.
34	Monteith, C	Artigo completo indisponível.
36	Rial-Crestelo, M	Não inclui critérios do consenso Delphi.
37	Roma, E	Não é estudo de acurácia diagnóstica.
38	Ropacka-Lesiak, M	Não inclui critérios do consenso Delphi.
39	Roy, A	Não inclui critérios do consenso Delphi.
41	Sheth, T	Não é estudo de acurácia diagnóstica.
42	Siddiqui, TS	Não inclui critérios do consenso Delphi.
43	Sirico, A	Não inclui critérios do consenso Delphi.
44	Sovio, U	Desfecho divergente do interesse.
45	Sovio, U	Não inclui critérios do consenso Delphi.
46	Spencer, R	Artigo completo indisponível.
47	Swanzy, L	Não é estudo de acurácia diagnóstica.
48	Triunfo, S	Não é estudo de acurácia diagnóstica.
52	Vollgraff, HS	Não inclui critérios do consenso Delphi.
53	Yamamoto, R	Não inclui critérios do consenso Delphi.

Apêndice 4: Tabela com extração de dados dos artigos selecionados.

Autor	Ano	No. Pacientes	Critério	Desfecho	VP	FP	FN	VN	S	S Mín	S Máx	E	E Mín	E Máx	VPP	VPN	Acurácia
Agarwal, R	2017	130	A/REDF	Neonatal death	10	13	12	95	45,5	36,1	54,8	88,0	74,7	101,3	43,5	88,8	80,8
Agarwal, R	2017	130	A/REDF	RDS	15	8	16	91	48,4	38,5	58,2	91,9	80,8	103,1	65,2	85,0	81,5
Agarwal, R	2017	130	A/REDF	Apgar score 5min < 7	16	7	27	80	37,2	27,1	47,4	92,0	80,8	103,1	69,6	74,8	73,8
Agarwal, R	2017	130	A/REDF	NICU	21	2	32	75	39,6	28,7	50,5	97,4	90,9	103,9	91,3	70,1	73,8
Marsoosi, V	2012	43	A/REDF	Neonatal death	5	7	4	26	55,6	38,6	72,5	78,8	55,7	101,9	41,7	86,7	73,8
Marsoosi, V	2012	43	A/REDF	IVH	4	7	3	26	57,1	40,3	74,0	78,8	54,6	102,9	36,4	89,7	75,0
Unterscheider, J	2014	1116	A/REDF	Stillbirth	3	73	0	1040	100,0	100,0	100,0	93,4	87,9	99,0	3,9	100,0	93,5
Unterscheider, J	2014	1116	A/REDF	Neonatal death	1	72	2	1038	33,3	30,6	36,1	93,5	87,9	99,2	1,4	99,8	93,4

Unterscheider, J	2013	1116	AC < p10 + PI AU > p95 OR A/REDF	APO	46	354	12	704	79,3	76,9	81,8	66,5	61,9	71,2	11,5	98,3	67,2
Unterscheider, J	2013	1116	AC < p10 + PI AU > p95 OR A/REDF	NICU	166	234	146	570	53,2	49,8	56,7	70,9	66,4	75,3	41,5	79,6	65,9
Unterscheider, J	2013	1116	AC > p3	APO	52	857	6	201	89,7	87,8	91,5	19,0	16,4	21,5	5,7	97,1	22,7
Unterscheider, J	2013	1116	AC > p3	NICU	275	634	37	170	88,1	85,9	90,4	21,1	18,5	23,8	30,3	82,1	39,9
Molina, L	2020	1055	Delphi criteria (no specification)	RDS	2	53	17	983	10,5	8,7	12,4	94,9	89,1	100,7	3,6	98,3	93,4
Molina, L	2020	1055	Delphi criteria (no specification)	Neonatal acidosis	3	52	16	984	15,8	13,6	18,0	95,0	89,2	100,8	5,5	98,4	93,6

Molina, L	2020	1055	Delphi criteria (no specification)	Apgar score 5min < 7	4	51	19	981	17,4	15,1	19,7	95,1	89,3	100,8	7,3	98,1	93,4
Molina, L	2020	1055	Delphi criteria (no specification)	NICU	12	43	108	892	10,0	8,1	11,9	95,4	89,9	100,9	21,8	89,2	85,7
Molina, L	2020	1055	Delphi criteria (no specification)	Emergency C-section	14	41	119	881	10,5	8,5	12,5	95,6	90,1	101,0	25,5	88,1	84,8
Molina, L	2020	1055	Delphi criteria (no specification)	Preterm birth	20	35	108	892	15,6	13,3	18,0	96,2	91,2	101,3	36,4	89,2	86,4
Molina, L	2020	1055	Delphi criteria (no specification)	SGA	40	15	99	901	28,8	25,8	31,7	98,4	95,0	101,7	72,7	90,1	89,2

Molina, L	2020	1055	Delphi criteria (no specification)	Neonatal death	0	0	0	0	#NÚM!	#NÚM!	#NÚM!	#NÚM!	#NÚM!	#NÚM!	#NÚM!	#NÚM!	#NÚM!
Molina, L	2020	1055	Delphi criteria (no specification)	IVH	0	0	0	0	#NÚM!	#NÚM!	#NÚM!	#NÚM!	#NÚM!	#NÚM!	#NÚM!	#NÚM!	#NÚM!
Caradeux, J	2017	472	EFW < p10 + CPR < p5	APO	35	45	96	296	26,7	22,0	31,4	86,8	79,4	94,2	43,8	75,5	70,1
Agarwal, R	2017	130	EFW < p10 + PI AU > p95 OR A/REDF	Stillbirth	6	59	0	65	100,0	100,0	100,0	52,4	40,3	64,6	9,2	100,0	54,6
Agarwal, R	2017	130	EFW < p10 + PI AU > p95 OR A/REDF	Neonatal death	18	47	4	61	81,8	74,5	89,1	56,5	44,4	68,5	27,7	93,8	60,8

Agarwal, R	2017	130	EFW < p10 + PI AU > p95 OR A/REDF	RDS	23	42	8	57	74,2	65,6	82,8	57,6	45,6	69,6	35,4	87,7	61,5
Agarwal, R	2017	130	EFW < p10 + PI AU > p95 OR A/REDF	Preterm birth	36	29	24	41	60,0	48,5	71,5	58,6	46,6	70,5	55,4	63,1	59,2
Agarwal, R	2017	130	EFW < p10 + PI AU > p95 OR A/REDF	Apgar score 5min < 7	31	34	12	53	72,1	62,7	81,5	60,9	49,1	72,8	47,7	81,5	64,6
Agarwal, R	2017	130	EFW < p10 + PI AU > p95 OR A/REDF	NICU	39	26	14	51	73,6	63,7	83,4	66,2	54,7	77,7	60,0	78,5	69,2
Agarwal, R	2017	130	EFW < p10 + PI AU > p95 OR A/REDF	APO	55	10	14	51	79,7	69,6	89,8	83,6	74,6	92,6	84,6	78,5	81,5

O'Dwyer, V	2014	1116	EFW < p10 + PI AU > p95 OR A/REDF	IVH	6	412	0	698	100,0	100,0	100,0	62,9	58,3	67,5	1,4	100,0	63,1
O'Dwyer, V	2014	1116	EFW < p10 + PI AU > p95 OR A/REDF	APO	48	370	9	689	84,2	82,0	86,4	65,1	60,5	69,6	11,5	98,7	66,0
Unterscheider, J	2014	1116	EFW < p10 + PI AU > p95 OR A/REDF	Neonatal death	1	508	2	602	33,3	30,6	36,1	54,2	49,9	58,6	0,2	99,7	54,2
Unterscheider, J	2014	1116	EFW < p10 + PI AU > p95 OR A/REDF	Stillbirth	3	509	0	604	100,0	100,0	100,0	54,3	50,0	58,6	0,6	100,0	54,4
Unterscheider, J	2013	1116	EFW < p10 + PI AU > p95 OR A/REDF	APO	47	366	11	692	81,0	78,7	83,4	65,4	60,8	70,0	11,4	98,4	66,2

Unterscheid er, J	2013	1116	EFW < p10 + PI AU > p95 OR A/REDF	NICU	169	244	143	560	54,2	50,7	57,6	69,7	65,2	74,1	40,9	79,7	65,3
Agarwal, R	2017	130	EFW < p10 + PI UA > p95	RDS	8	28	23	71	25,8	17,2	34,4	71,7	57,0	86,4	22,2	75,5	60,8
Agarwal, R	2017	130	EFW < p10 + PI UA > p95	Neonatal death	8	28	14	80	36,4	27,3	45,4	74,1	59,8	88,4	22,2	85,1	67,7
Agarwal, R	2017	130	EFW < p10 + PI UA > p95	Apgar score 5min < 7	15	21	28	66	34,9	24,9	44,9	75,9	61,9	89,8	41,7	70,2	62,3
Agarwal, R	2017	130	EFW < p10 + PI UA > p95	NICU	18	18	35	59	34,0	23,4	44,5	76,6	62,8	90,4	50,0	62,8	59,2
Caradeux, J	2017	472	EFW < p10 + PI UA > p95	APO	5	1	126	340	3,8	1,8	5,9	99,7	95,4	104,0	83,3	73,0	73,1
Maggio, L	2015	512	EFW < p10 + PI UA > p95	Apgar score 5min < 7	3	115	13	381	18,8	15,3	22,2	76,8	69,2	84,4	2,5	96,7	75,0

Maggio, L	2015	512	EFW < p10 + PI UA > p95	Neonatal acidosis	0	118	1	393	0,0	0,0	0,0	76,9	69,3	84,5	0,0	99,7	76,8
Maggio, L	2015	512	EFW < p10 + PI UA > p95	RDS	12	106	36	358	25,0	21,1	28,9	77,2	69,6	84,7	10,2	90,9	72,3
Maggio, L	2015	512	EFW < p10 + PI UA > p95	Emergency C-section	5	113	11	383	31,3	27,2	35,3	77,2	69,6	84,8	4,2	97,2	75,8
Maggio, L	2015	512	EFW < p10 + PI UA > p95	APO	24	94	60	334	28,6	24,3	32,9	78,0	70,6	85,5	20,3	84,8	69,9
Maggio, L	2015	512	EFW < p10 + PI UA > p95	NICU	50	68	101	293	33,1	28,3	38,0	81,2	74,1	88,2	42,4	74,4	67,0
Maggio, L	2015	512	EFW < p10 + PI UA > p95	Preterm birth	54	64	83	311	39,4	34,5	44,4	82,9	76,1	89,7	45,8	78,9	71,3
Maggio, L	2015	512	EFW < p10 + PI UA > p95	SGA	100	18	258	136	27,9	20,8	35,0	88,3	82,5	94,1	84,7	34,5	46,1
Villalaína, C	2018	131	EFW < p10 + PI UA > p95	SGA	43	21	58	9	42,6	24,9	60,3	30,0	18,8	41,2	67,2	13,4	39,7

			OR EFW < p10 + CPR < p5														
Villalaína, C	2018	131	EFW < p10 + PI UA > p95 OR EFW < p10 + CPR < p5	Apgar score 5min < 7	2	62	2	65	50,0	41,3	58,7	51,2	38,9	63,4	3,1	97,0	51,1
Villalaína, C	2018	131	EFW < p10 + PI UA > p95 OR EFW < p10 + CPR < p5	Neonatal acidosis	2	62	1	66	66,7	58,5	74,8	51,6	39,3	63,8	3,1	98,5	51,9
Villalaína, C	2018	131	EFW < p10 + PI UA > p95 OR EFW <	NICU	4	60	1	66	80,0	73,0	87,0	52,4	40,1	64,6	6,3	98,5	53,4

			p10 + CPR < p5														
Villalaína, C	2018	131	EFW < p10 + PI UA > p95 OR EFW < p10 + CPR < p5	Emergency C-section	17	47	4	63	81,0	73,6	88,3	57,3	45,2	69,4	26,6	94,0	61,1
Figueiras, F	2014	509	EFW < p3	APO	17	38	117	337	12,7	9,3	16,1	89,9	81,9	97,8	30,9	74,2	69,5
Savchev, S	2012	264	EFW < p3	Neonatal acidosis	3	57	3	201	50,0	43,9	56,1	77,9	67,4	88,4	5,0	98,5	77,3
Savchev, S	2012	264	EFW < p3	Emergency C-section	15	45	13	191	53,6	47,2	59,9	80,9	71,0	90,9	25,0	93,6	78,0
Unterscheider, J	2013	1116	EFW < p3	APO	51	775	7	283	87,9	86,0	89,9	26,7	23,7	29,8	6,2	97,6	29,9
Unterscheider, J	2013	1116	EFW < p3	NICU	267	559	45	245	85,6	83,1	88,0	30,5	27,3	33,6	32,3	84,5	45,9
Unterscheider, J	2014	1116	EFW < p3	Stillbirth	3	423	0	690	100,0	100,0	100,0	62,0	57,4	66,6	0,7	100,0	62,1

Unterscheider, J	2014	1116	EFW < p3	Neonatal death	2	421	1	689	66,7	63,9	69,4	62,1	57,4	66,7	0,5	99,9	62,1
Villalaína, C	2018	131	EFW < p3	Emergency C-section	2	35	19	75	9,5	4,0	15,0	68,2	53,2	83,2	5,4	79,8	58,8
Villalaína, C	2018	131	EFW < p3	NICU	1	36	4	90	20,0	13,0	27,0	71,4	56,9	86,0	2,7	95,7	69,5
Villalaína, C	2018	131	EFW < p3	Apgar score 5min < 7	1	36	3	91	25,0	17,5	32,5	71,7	57,1	86,2	2,7	96,8	70,2
Villalaína, C	2018	131	EFW < p3	Neonatal acidosis	1	36	2	92	33,3	25,2	41,5	71,9	57,4	86,4	2,7	97,9	71,0
Villalaína, C	2018	131	EFW < p3	SGA	32	5	69	25	31,7	15,0	48,3	83,3	71,3	95,3	86,5	26,6	43,5
Villalaína, C	2018	131	EFW < p3 OR EFW < p10 + PI UA > p95 OR CPR < p5	SGA	75	26	26	4	74,3	58,6	89,9	13,3	6,7	20,0	74,3	13,3	60,3
Villalaína, C	2018	131	EFW < p3 OR EFW < p10 + PI UA > p95 OR CPR < p5	Apgar score 5min < 7	3	98	1	29	75,0	67,5	82,5	22,8	14,6	31,0	3,0	96,7	24,4

Villalaína, C	2018	131	EFW < p3 OR EFW < p10 + PI UA > p95 OR CPR < p5	NICU	5	96	0	30	100,0	100,0	100,0	23,8	15,5	32,1	5,0	100,0	26,7
Villalaína, C	2018	131	EFW < p3 OR EFW < p10 + PI UA > p95 OR CPR < p5	Emergency C-section	19	82	2	28	90,5	85,0	96,0	25,5	17,0	34,0	18,8	93,3	35,9
Cavallaro, A	2017	235	EFW < p3 OR EFW < p10 + RCP <p5	APO	27	48,8884	38,1701	115,1116	41,4	33,9	49,0	70,2	59,9	80,5	35,6	75,1	62,0
Marsoosi, V	2012	43	EFW OR AC < p3	Neonatal death	4	21	1	10	80,0	65,9	94,1	32,3	13,9	50,6	16,0	90,9	38,9
Marsoosi, V	2012	43	EFW OR AC < p3	IVH	4	20	1	10	80,0	65,7	94,3	33,3	14,5	52,2	16,7	90,9	40,0

Anexo 1: Cover Letter da submissão para o American Journal of Obstetrics and Gynecology

Catherine Bradley, MD, MSCE and Roberto Romero, MD, DMedSci

Editors-in-Chief: American Journal of Obstetrics and Gynecology

Campinas, March 11, 2023.

Dear Dr Bradley and Dr Romero,

It is my privilege to submit the manuscript entitled **"Delphi consensus for the diagnosis of fetal growth restriction: systematic review and meta-analysis"** for publication consideration in *American Journal of Obstetrics and Gynecology.*

In recent years, several studies have addressed the technical challenges related to the discrimination of fetuses with fetal growth restriction (FGR). In 2016, a group of specialists, using the Delphi methodology, proposed an amalgam of imaging parameters in an effort to standardize the diagnosis of FGR. Since then, these criteria have been adopted by, among others, the International Federation of Gynecology and Obstetrics (FIGO) and The International Society of Ultrasound in Obstetrics and Gynecology (ISUOG) as the standard for the diagnosis of FGR. However, there is a heterogeneous understanding of the actual performance of the consensus as a whole, because each particular criterion was developed, validated and further tested in clinical practice by distinct study groups and in different populations. In this systematic review and meta-analysis, we revisited the Delphi consensus on the diagnosis of FGR, and revised the entire body of literature pertaining to the consensus as a whole and its individual criterion. In addition, we produced meta-analyses evaluating different scenarios concerning the FGR criteria and postnatal outcomes.

This study is a diagnostic test accuracy systematic review performed according to the PRISMA Diagnostic Test Accuracy protocol. It was registered in the PROSPERO platform (Registration number: CRD42020204051). Since no patient data was directly manipulated in the present study, this review was considered exempt from Institutional Review Board approval. A preview of this study was presented on the 26th. São Paulo Congress of

Obstetrics and Gynecology that happened online from August 19 to December 12, 2021. This manuscript is being submitted solely to *American Journal of Obstetrics and Gynecology* and will not be submitted elsewhere unless a final negative decision is made by the Editors of *American Journal of Obstetrics and Gynecology*. Ana Paula de Morais Oliveira, Maria Laura Costa do Nascimento and Carlos Roberto Silveira Correia are mentioned in the acknowledgements and they previously allowed by email the disclosure of their name.

We hope our manuscript can be of interest to *American Journal of Obstetrics and Gynecology* readers. We are looking forward to receiving your opinion.

Yours sincerely,

João Renato Bennini