PONTIFÍCIA UNIVERSIDADE CATÓLICA DE CAMPINAS

AMANDA GABRIELLE CENZI

COMPARAÇÃO DOS MODELOS DE AVALIAÇÃO ULTRASSONOGRÁFICA IOTA E O-RADS NA ACUIDADE DIAGNÓSTICA DE MASSAS ANEXIAIS MALIGNAS: UMA REVISÃO SISTEMÁTICA COM METANÁLISE

CAMPINAS

2025

PONTIFÍCIA UNIVERSIDADE CATÓLICA DE CAMPINAS
ESCOLA DE CIÊNCIAS DA VIDA
PROGRAMA DE PÓS-GRADUAÇÃO *STRICTO SENSU* EM CIÊNCIAS DA SAÚDE

AMANDA GABRIELLE CENZI

COMPARAÇÃO DOS MODELOS DE AVALIAÇÃO ULTRASSONOGRÁFICA IOTA E O-RADS NA ACUIDADE DIAGNÓSTICA DE MASSAS ANEXIAIS MALIGNAS: UMA REVISÃO SISTEMÁTICA COM METANÁLISE

Dissertação apresentada ao Programa de Pós-Graduação *Stricto Sensu* em Ciências da Saúde da Escola de Ciências da Vida, da Pontifícia Católica de Campinas, como exigência para obtenção do título de Mestre em Ciências da Saúde.

Orientador: Prof. Dr. José Alexandre Mendonça

.

CAMPINAS
2025

Sistema de Bibliotecas e Informação - SBI
Gerador de fichas catalográficas da Universidade PUC-Campinas
Dados fornecidos pelo(a) autor(a).

Cenzi, Amanda Gabrielle

C395c Comparação dos modelos de avaliação ultrassonográfica IOTA e O-RADS na acuidade diagnóstica de massas anexiais malignas: uma revisão sistemática com metanálise / Amanda Gabrielle Cenzi. - Campinas: PUC-Campinas, 2025.

82

Orientador: José Alexandre Mendonça.

Dissertação (Mestrado em Ciências da Saúde) - Programa de Pós-Graduação Stricto Sensu, Escola de Ciências da Vida, Pontifícia Universidade Católica de Campinas, Campinas, 2025.
Inclui bibliografia.

1. Ultrassonografia. 2. IOTA. 3. O-RADS. I. Mendonça, José Alexandre . II. Pontifícia Universidade Católica de Campinas. Escola de Ciências da Vida. Programa de Pós-Graduação Stricto Sensu. III. Título.

PONTIFÍCIA UNIVERSIDADE CATÓLICA DE CAMPINAS
ESCOLA DE CIÊNCIAS DA VIDA
PROGRAMA DE PÓS-GRADUAÇÃO *STRICTO SENSU* EM CIÊNCIAS DA SAÚDE

AMANDA GABRIELLE CENZI

COMPARAÇÃO DOS MODELOS DE AVALIAÇÃO ULTRASSONOGRÁFICA IOTA E O-RADS NA ACUIDADE DIAGNÓSTICA DE MASSAS ANEXIAIS MALIGNAS: UMA REVISÃO SISTEMÁTICA COM METANÁLISE

Dissertação defendida e aprovada em 02 de dezembro de 2025 pela Comissão Examinadora

Prof. Dr. José Alexandre Mendonça
Orientador da Dissertação e Presidente da Comissão Examinadora
Pontifícia Universidade Católica de Campinas (PUC-Campinas)

Prof. Dr. Rodrigo Menezes Jales
Universidade Estadual de Campinas (UNICAMP)

Profa. Dra. Alessandra Gambero
Pontifícia Universidade Católica de Campinas (PUC-Campinas)

CAMPINAS
2025

DEDICO este trabalho aos colegas ultrassonografistas, cujos olhos enxergam através – e além – das imagens em escalas de cinza; e aos colegas ginecologistas, que escutam os silêncios do corpo feminino e o acolhem em seus territórios mais íntimos.

AGRADECIMENTOS

A Deus, minha eterna gratidão pela vida, por todas as oportunidades que me foram dadas e por Sua presença constante.

Agradeço aos meus amados pais, Mara e José Renato, por me permitirem conhecer e vivenciar o amor infinito; por serem o alicerce da minha vida, sinônimo de carinho, doação, afeto e verdade; por confiarem em mim, pelo apoio incondicional e pela compreensão sincera diante das minhas ausências.

Ao meu companheiro e melhor amigo, Roberto, por sempre enxergar em mim a minha melhor versão, mesmo antes que eu pudesse reconhecê-la; por me incentivar, me escutar e estar ao meu lado ao longo dessa caminhada.

Ao meu orientador, Prof. Dr. José Alexandre Mendonça, pela dedicação, paciência, confiança e pelos conselhos valiosos; pelo conhecimento generosamente compartilhado e por ter me mostrado, com sensibilidade e firmeza, que tudo seria possível.

Ao Dr. Carlos Menossi por ter acreditado em mim e no meu potencial acadêmico desde o início.

Às minhas queridas amigas, Luciana Peres, Tamiris e Maria Carolina, por me ouvirem com tanto carinho, por acolherem meus desabafos e por estarem presentes nos momentos mais desafiadores dessa jornada.

Ao Dr. Estevão, pelos ensinamentos e incentivos.

Aos professores do curso de Mestrado em Ciências da Saúde da PUC-Campinas, pelo compromisso com a formação de seus alunos e pelas inestimáveis contribuições, que enriqueceram minha trajetória com diferentes olhares e saberes.

Aos meus colegas de pós-graduação, pela amizade e parceria.

Aos ilustres membros da banca de qualificação e defesa, meu sincero agradecimento por aceitarem o convite e pelo precioso tempo dedicado à leitura atenta deste trabalho. Sua presença é uma honra, e sua notável experiência é de grande valor para o enriquecimento da dissertação e para meu desenvolvimento acadêmico.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) – Código de Financiamento 001.

"Se vi mais longe, foi por estar sobre os ombros de gigantes"

Isaac Newton

(1643 – 1727)

RESUMO

Introdução: O diagnóstico etiológico de massas anexiais é crucial para a conduta adequada. Embora a ultrassonografia pélvica seja o exame complementar de escolha inicial, a sua eficácia pode ser otimizada pela padronização de critérios. Modelos estruturados como o IOTA (*International Ovarian Tumor Analysis*) e o O-RADS (*Ovarian-Adnexal Reporting and Data System*) foram desenvolvidos para aprimorar a reprodutibilidade e a segurança na estratificação de risco de malignidade. **Objetivos:** Comparar a acurácia diagnóstica do sistema O-RADS em relação aos modelos IOTA *Simple Rules* (IOTA-SR) e IOTA *ADNEX* na triagem de massas anexiais malignas por ultrassonografia. **Métodos:** Foi conduzida uma revisão sistemática com metanálise de acurácia diagnóstica, registrada no PROSPERO (CRD42024581159). A busca bibliográfica abrangeu as bases PubMed, Embase, Ovid, LILACS e Cochrane Library, incluindo estudos longitudinais com confirmação histopatológica ou seguimento clínico como padrão de referência. A síntese estatística utilizou um modelo bivariado de efeitos aleatórios, e o risco de viés foi avaliado pelas ferramentas QUADAS-2 e QUADAS-C. A certeza de evidência foi avaliada e classificada pelo GRADE. No total, 21 estudos foram incluídos na análise qualitativa e 16 na análise quantitativa. **Resultados:** Os resultados demonstraram que o O-RADS apresentou sensibilidade estatisticamente superior ao IOTA-SR (p = 0,024) e ao *ADNEX* (p = 0,010), o que se traduz em aproximadamente oito diagnósticos adicionais de malignidade por 1.000 pacientes avaliadas. A especificidade mostrou-se estatisticamente semelhante entre os modelos (p > 0,10). Em termos de acurácia sumária bivariada, o O-RADS obteve sensibilidade de 96,1% (IC95%: 94,4–97,4) e especificidade de 73,1% (IC95%: 61,9–82,0). Em comparação, o IOTA *Simple Rules* (SR) apresentou sensibilidade de 92,8% (IC95%: 89,6–95,1) e especificidade de 79,8% (IC95%: 69,7–87,2). O desempenho diagnóstico global, avaliado pela AUC e DOR, foi elevado para todos: o O-RADS, o IOTA SR e o IOTA *ADNEX* alcançaram AUCs de 0,958, 0,941 e 0,966, e DORs sumários de 80,1, 57,4 e 84,5, respectivamente. A avaliação do risco de viés (QUADAS-2 e QUADAS-C) indicou baixo risco de viés em 94% das análises. Pelo sistema GRADE, a certeza da evidência foi classificada como moderada para a sensibilidade e alta para a especificidade. **Conclusão:** Conclui-se que, embora todos os modelos demonstrem alta acurácia, o sistema O-RADS apresentou o melhor desempenho estatístico na triagem de massas anexiais malignas comparado ao IOTA-SR e IOTA *ADNEX*. Essa vantagem se manifesta na maior sensibilidade, crucial para a segurança oncológica ao minimizar a taxa de falsos negativos.

Palavras-Chave: IOTA. O-RADS. Neoplasias anexiais. Metanálise. Acurácia Diagnóstica.

ABSTRACT

Introduction: The etiologic diagnosis of adnexal masses is crucial for appropriate management. Although pelvic ultrasonography is the initial complementary imaging modality of choice, its efficacy can be optimized by the standardization of criteria. Structured models such as IOTA (International Ovarian Tumor Analysis) and O-RADS (Ovarian-Adnexal Reporting and Data System) were developed to enhance reproducibility and safety in malignancy risk stratification. **Objectives:** To compare the diagnostic accuracy of the O-RADS system relative to the IOTA Simple Rules (IOTA-SR) and IOTA ADNEX models in the triage of malignant adnexal masses by ultrasonography. **Methods:** A systematic review with diagnostic accuracy meta-analysis was conducted, registered in PROSPERO (CRD42024581159). The literature search covered the PubMed, Embase, Ovid, LILACS, and Cochrane Library databases, including longitudinal studies with histopathological confirmation or clinical follow-up as the reference standard. Statistical synthesis utilized a bivariate random-effects model, and the risk of bias was assessed using the QUADAS-2 and QUADAS-C tools. The certainty of evidence was assessed and rated using the GRADE approach. In total, 21 studies were included in the qualitative analysis and 16 in the quantitative analysis. **Results:** The results demonstrated that O-RADS showed a sensitivity statistically superior to IOTA-SR ($p = 0.024$) and ADNEX ($p = 0.010$), which translates to approximately eight additional diagnoses of malignancy per 1,000 evaluated patients. Specificity was statistically similar among the models ($p > 0.10$). In terms of bivariate summary accuracy, O-RADS achieved a sensitivity of 96.1% (95%CI: 94.4–97.4) and a specificity of 73.1% (95%CI: 61.9–82.0). In comparison, IOTA Simple Rules (SR) presented a sensitivity of 92.8% (95%CI: 89.6–95.1) and a specificity of 79.8% (95%CI: 69.7–87.2). Global diagnostic performance, assessed by AUC and DOR, was high for all models: O-RADS, IOTA SR, and IOTA ADNEX achieved AUCs of 0.958, 0.941, and 0.966, and summary DORs of 80.1, 57.4, and 84.5, respectively. The risk of bias assessment (QUADAS-2 and QUADAS-C) indicated a low risk of bias in 94% of the analyses. By the GRADE system, the certainty of evidence was classified as moderate for sensitivity and high for specificity. **Conclusion:** It is concluded that, although all models demonstrate high accuracy, the O-RADS system showed the best statistical performance in the triage of malignant adnexal masses compared to IOTA-SR and IOTA ADNEX. This advantage is manifested in the higher sensitivity, which is crucial for oncologic safety by minimizing the false-negative rate.

Keywords: IOTA. O-RADS. Adnexal Neoplasms. Meta-Analysis. Diagnostic Accuracy.

LISTA DE TABELAS

Tabela 1- Estratégia de Busca nas bases de dados eletrônicas ---------- 31

Tabela 2 – Características dos estudos incluídos na Revisão Sistemática --------- 35

Tabela 3 – Estudos elegíveis para a síntese quantitativa (metanálise) ---------- 38

Tabela 4 – Proporção e taxa de detecção de massas malignas segundo os diferentes modelos avaliados ---------- 40

Tabela 5 – Nível de evidência pelo GRADE ---------- 57

Tabela 6 – Interpretação resumida das etapas de seleção, número de estudos e situação final---------- 58

LISTA DE FIGURAS

Figura 1- Fluxograma de seleção dos estudos pela metodologia PRISMA -------- 34

Figura 2 – *Forest plots* de sensibilidade e especificidade do O-RADS ---------- 41

Figura 3 – *Forest plot* de razão de chances diagnóstica (DOR) do O-RADS. ----- 42

Figura 4 – Curva SROC do O-RADS obtida a partir do modelo bivariado de Reitsma ---------- 43

Figura 5 – *Forest plots* de sensibilidade e especificidade do IOTA-SR---------- 44

Figura 6 – *Forest plot* de razão de chances diagnóstica (DOR) do IOTA-SR------ 45

Figura 7 – Curva SROC do IOTA SR obtida a partir do modelo bivariado de Reitsma ---------- 46

Figura 8 – *Forest plots* de sensibilidade e especificidade do IOTA *ADNEX*-------- 47

Figura 9 – *Forest plot* de razão de chances diagnóstica (DOR) do IOTA *ADNEX* 48

Figura 10 – Curva SROC do IOTA *ADNEX* obtida a partir do modelo univariado 49

Figura 11 – *Forest plots* de sensibilidade e especificidade do IOTA 2 etapas + IOTA SRR---------- 50

Figura 12 – SAUC combinando O-RADS e IOTA SR ---------- 51

Figura 13 – SAUC combinando O-RADS e IOTA *ADNEX* ---------- 52

Figura 14 – Risco de viés e aplicabilidade: revisão do julgamento dos autores--- 54

Figura 15 – Gráfico de barras agrupados da avaliação de risco de viés para os estudos incluídos na metanálise ---------- 55

Figura 16 – Fluxo de inclusão e análise dos estudos na revisão sistemática e na metanálise.---------- 58

LISTA DE ABREVIATURAS E SIGLAS

ADNEX	*Assessment of Different Neoplasias in the Adnexa*
ACR	*American College of Radiology* (Colégio Americano de Radiologia)
AUC	*Area Under the Curve* (área sob a curva)
CA-125	*Cancer Antigen 125* (Antígeno Carboidrato 125)
CEUS	*Contrast-Enhanced Ultrasound* (ultrassom com contraste)
D1	domínio 1
D2	domínio 2
D3	domínio 3
D4	domínio 4
DOR	*Diagnostic Odds Ratio* (razão de chances diagnóstica)
D.P.	Desvio padrão
EUA	Estados Unidos da América
FN	*False negative* (falso negativo)
FP	*False positivo* (falso positivo)
GE	*General Electric HealthCare*
GI-RADS	*Gynecologic Imaging Reporting and Data System*
GRADE	*Grading of Recommendations, Assessment, Development and Evaluation*
GRADEpro GDT	*GRADEpro Guideline Development Tool*
HE4	*Human Epididymis Protein 4* (Proteína do Epidídimo Humano)
I	Indeterminada
I^2	Índice de Inconsistência (Heterogeneidade)
IA	Inteligência artificial
IC95%	Intervalo de Confiança de 95%
ln(DOR)	*Log diagnostic odds ratio* (logaritmo natural do DOR)
INCA	Instituto Nacional de câncer
IOTA	*International Ovarian Tumor Analysis*
LILACS	Literatura Latino-Americana e do Caribe em Ciências da Saúde
M	Maligna
M1	modelo 1
M2	modelo 2
MA	massas anexiais
MEDLINE	*Medical Literature Analysis and Retrieval System Online*
MeSH	*Medical Subject Headings* (Descritores em Ciências da Saúde)
min	minuto
n	número
N/A	*not available*
O-RADS	*Ovarian-Adnexal Reporting and Data System*
O-RADS MRI	O-RADS por Ressonância Magnética
PBE	Prática Baseada em Evidências

PICO	Acrônimo para*: Patient or Problem, Intervention, Comparasion, Outcomes*
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-*Analyses*
PRISMA-DTA	*Preferred Reporting Items for Systematic Reviews and Meta-Analyses of Diagnostic Test Accuracy Studies.*
PROSPERO	*International Prospective Register of Systematic Reviews*
QUADAS-2	*Quality Assessment of Diagnostic Accuracy Studies-2*
QUADAS-C	*Quality Assessment of Diagnostic Accuracy Studies-comparative*
RM	Ressonância Nuclear magnética
RMI	*Risk of Malignancy Index*
ROMA	*Risk of Ovarian Malignancy Algorithm*
SAUC	*Summary Area Under the Curve* (Área Sob a Curva Sumária)
SR	*Simple Rules*
SROC	*Summary Receiver Operating Characteristic Curve* (Curva ROC Sumária)
SRR	*Simple Rules Risk*
TN	*true negative* (veradeiro negativo)
TP	*true positive* (verdadeiro positivo)
US	ultrassonografia
VN	verdadeiro negativo
VP	verdadeiro positivo
v2020	versão 2020 (O-RADS)
v2022	versão 2022 (O-RADS)

SUMÁRIO

1. INTRODUÇÃO ---- 16

1.1. Massas Anexiais ---- 16

1.2. Massas Anexiais e IOTA ---- 19

1.3. Massas Anexiais e O-RADS ---- 23

2. JUSTIFICATIVA ---- 26

3. OBJETIVOS ---- 27

3.1. Objetivo Geral ---- 27

3.2. Objetivos Específicos ---- 27

4. METODOLOGIA ---- 28

4.1. Desenho do estudo ---- 28

4.2. Amostra ---- 28

4.2.1. Tipos de estudos ---- 28

4.2.2. Tipos de participantes ---- 28

4.2.3. Tipos de desfechos ---- 29

4.3. Protocolo e Registro ---- 29

4.4. Critérios de Elegibilidade ---- 29

4.4.1. Critérios de inclusão: ---- 29

4.4.2. Critérios de exclusão: ---- 30

4.5. Busca na Literatura ---- 30

4.6. Seleção dos estudos ---- 31

4.7. Coleta de Dados ---- 32

4.8. Avaliação do risco de viés ---- 32

4.9. Análise dos dados (metodologia estatística) ---- 33

5. RESULTADOS ---- 34

5.1. Seleção dos estudos ---- 34

5.2. Coleta de dados ---- 35

5.3. Metanálise ---- 39

5.3.1. O-RADS ---- 40

5.3.2. IOTA-SR ---- 43

5.3.3. IOTA *ADNEX* ---- 46

5.3.4. IOTA 2 etapas + IOTA SRR ---- 49

5.3.5. Comparação entre O-RADS e IOTA SR ---- 50

5.3.6. Comparação entre O-RADS e IOTA *ADNEX* ---------- 51
5.4. Risco de Viés ---------- 52
5.5. Nível de Evidência ---------- 55
6. DISCUSSÃO ---------- 59
7. CONCLUSÃO ---------- 68
8. REFERÊNCIAS ---------- 69
9. APÊNDICES ---------- 76
10. ANEXOS ---------- 80

1. INTRODUÇÃO

1.1. Massas Anexiais

As massas anexiais são lesões originárias nos ovários, tubas uterinas, ligamentos redondos e resquícios embrionários dos ductos de Wolf, segundo o *American College of Obstetricians and Gynecologists* (ACOG, 2016). São um achado frequente na prática ginecológica e radiológica.

Podem acometer mulheres de todas as idades. As causas são variadas e incluem alterações funcionais, processos inflamatórios e infecciosos, e tumores benignos e malignos (ACOG, 2016). Os diagnósticos diferenciais envolvem condições não ginecológicas, provenientes de estruturas pélvicas adjacentes, ampliando a complexidade da avaliação clínica e por imagem.

A sua abordagem é desafiadora pela multifatoriedade e individualização, a depender tanto da etiologia, quanto de outras variáveis como idade, prole, quadro clínico e extensão do acometimento (ACOG, 2016; Carvalho; Moretti-Marques; Silva Filho, 2020).

São mais frequentes em pacientes na menacme, em que a benignidade também é maior. Massas anexais malignas são mais comuns na pré-adolescência e na pós menopausa (Carvalho; Moretti-Marques; Silva Filho, 2020).

A sua incidência é desconhecida, visto que, na maioria das vezes, a condição é assintomática e, portanto, não diagnosticada, sendo frequente o achado incidental pelo grande emprego atual da ultrassonografia pélvica e transvaginal na prática diária. Em alguns casos podem ocorrer sintomas inespecíficos como dor aguda ou intermitente, aumento do volume abdominal, febre, perda ponderal, distúrbios menstruais, virilização e secreção vaginal (ACOG, 2016; Carvalho; Moretti-Marques; Silva Filho, 2020).

O exame clínico pélvico pode oferecer alguns critérios para diferenciação, mas possui baixa sensibilidade (Carvalho; Moretti-Marques; Silva Filho, 2020), com limitação para avaliação anexial, mesmo em situação ideal, tanto por fatores relacionados à lesão, como pequeno volume, quanto por fatores individuais, como o panículo adiposo.

Em caso de malignidade, o câncer de ovário é a neoplasia de pior prognóstico, com uma taxa de sobrevida de cinco anos em estágio avançado de 51,6% nos anos

de 2015 a 2021, segundo o *National Cancer Institute* (NIH, 2025). É a terceira neoplasia ginecológica mais comum, ficando atrás do câncer do colo do útero e do endométrio. Cerca de 95% das neoplasias ovarianas são derivadas das células epiteliais, e 5% provêm de células germinativas ou estromais (INCA, 2025).

Diante de sua agressividade, o tratamento deve ser iniciado o mais precocemente possível e preferencialmente em centro de referência em oncologia ginecológica, com melhores resultados e, assim, aumento da sobrevida (Andreotti et al., 2020; Pelayo *et al.,* 2023a).

Assim, o diagnóstico etiológico e, principalmente, diferencial entre massas tumorais benignas e malignas é fundamental para a escolha do tratamento adequado, seja ele expectante ou cirúrgico (Sokalska *et al.*, 2009).

Tumores anexais são responsáveis por um número considerável de cirurgias ginecológicas. Estima-se que mais de 200.000 mulheres com massa anexial passam por cirurgia anualmente nos Estados Unidos da América (EUA), embora o câncer de ovário seja encontrado em cerca de apenas 10% dessas pacientes. Estudos populacionais documentaram uma taxa de complicação de 3% a 15% em cirurgias para tumores benignos (Yoeli-Bik *et al*., 2023).

Atualmente não há evidência científica de que o rastreamento do câncer de ovário na população geral traga mais benefícios do que riscos, conforme diretrizes da *US Preventive Services Task Force* (USPSTF, 2018) (Armstrong *et al*., 2021). No entanto, diante de pacientes com massas anexiais diagnosticadas, a avaliação complementar é fundamental. Embora nenhum método isolado ou combinado ofereça acurácia absoluta para o diagnóstico de malignidade, tais ferramentas são indispensáveis para a estratificação de risco. Isso permite diferenciar pacientes com baixo risco — que podem ser manejadas em hospitais gerais por ginecologistas generalistas, com tratamento conservador ou priorizando cirurgias minimamente invasivas — daquelas com alta suspeita de malignidade, que requerem referenciamento para centros especializados em ginecologia oncológica (Carvalho; Moretti-Marques; Silva Filho, 2020).

Um estudo publicado em 2017 concluiu que procedimentos cirúrgicos para lesões benignas podem ser evitados com melhoria na avaliação pré-operatória das mesmas (Andreotti *et al*., 2020), otimizando, assim, a triagem do paciente e reduzindo cirurgias desnecessárias sem deixar de detectar o câncer (Yoeli-Bik *et al.,* 2023).

A ultrassonografia pélvica transvaginal é considerada um exame complementar eficaz para a avaliação da região anexial, sendo um método seguro, por não utilizar radiação e economicamente mais acessível, de acordo com o *Royal College of Obstetricians and Gynaecologists* (RGOG, 2011) (Carvalho; Moretti-Marques; Silva Filho, 2020; Pelayo *et al*., 2023a).

O reconhecimento de padrões de imagem ultrassonográfica por profissional especializado é considerado um dos métodos mais precisos na discriminação entre lesões benignas e malignas (Andreotti *et al*., 2020; Sokalska *et al.*, 2009) e, em conjunto com a história clínica, marcadores tumorais e outros métodos de imagem, como a ressonância nuclear magnética (RM), contribui para o diagnóstico diferencial entre massas anexais (Carvalho; Moretti-Marques; Silva Filho, 2020).

No entanto, sua eficácia pode ser limitada pela sua natureza subjetiva, altamente dependente da experiência do examinador (Liu *et al*., 2024), tornando a classificação ultrassonográfica de massas anexiais como benignas ou malignas um desafio (Pelayo *et al.,* 2023a). Embora ultrassonografistas experientes possam distinguí-las corretamente, os exames ecográficos são frequentemente realizados e interpretados por profissionais com níveis variados de conhecimento, confiança e expertise (Andreotti *et al*., 2020), o que interfere na estratificação de risco e, assim, nos resultados e abordagens dos pacientes (Barra; Jorge; Condé, 2021).

Nesse contexto, diretrizes, sistemas de classificação e modelos de previsão foram desenvolvidos para padronizar a avaliação ultrassonográfica das massas anexais (Guo *et al*., 2022), melhorando os resultados, a segurança e a qualidade na assistência à paciente acometida com esses agravos, com o objetivo de otimizar os resultados do câncer de ovário e, ao mesmo tempo, minimizar procedimentos cirúrgicos desnecessários em pacientes com baixo risco de malignidade, com redução da morbidade cirúrgica e custos relacionados ao procedimento, além da manutenção, quando possível, da competência hormonal ovariana (Andreotti *et al*., 2018, 2020).

1.2. Massas Anexiais e IOTA

O grupo IOTA (*International Ovarian Tumor Analysis*), europeu, na tentativa de homogeneizar e padronizar critérios, facilitando a reprodutibilidade por profissionais com diferentes níveis de experiência, publicou, em 1999, um consenso com os termos, definições e medidas para descrever as características ultrassonográficas das lesões anexiais (Timmerman, D. *et al*., 2000), tanto em parâmetros morfológicos por imagem em modo B, quanto em parâmetros de vascularização e fluxo sanguíneo por imagem com *Doppler* colorido.

As imagens são analisadas em modo B e inicialmente caracterizadas de acordo com sua localização (ovariana ou paraovariana). Devem ser medidas no seu diâmetro longitudinal, anteroposterior e transverso. Então, a análise de sua morfologia envolve a descrição da sua composição e ecogenicidade dominantes; regularidade das paredes internas e externas; presença ou ausência de projeções papilares sólidas internas/vegetações; presença ou ausência de septos internos/loculações; e presença de sombra acústica posterior (Timmerman, D. *et al*., 2000).

A composição de uma lesão pode ser sólida e, neste caso, ela será hiperecoica; líquida, apresentando-se como anecoica se líquido flúido, hipoecoica se houver debris espessos como no caso de mucina, e isoecoica se debris finos e homogêneos; ou heterogênea, se houver componente sólido e líquido (Timmerman, D. *et al.,* 2000).

Em imagens sólidas, analisa-se a regularidade das suas paredes externas. Em imagens líquidas, a avaliação detalhada de suas paredes internas é de extrema importância para a possível caracterização de alguma irregularidade. A presença de tecido sólido com mais de 3 milímetros emergindo da superfície interna da lesão e se projetando para a sua luz define uma projeção sólida, que também deve ser descrita e medida na análise morfológica. Projeções fibrilares sólidas que atravessam completamente a lesão de uma parede à outra formando loculações são definidas como septos, que devem ser referidos juntamente com o número de loculações formadas, podendo ser medidos em sua espessura máxima (Timmerman, D*. et al.,* 2000).

A presença ou ausência do fenômeno de sombra acústica posterior deve estar incluída na descrição, assim como a presença ou ausência de líquido livre na pelve e sua quantificação (Timmerman, D. *et al.,* 2000).

O estudo *Doppler* colorido, corretamente calibrado, também é utilizado qualitativamente para a caracterização de vascularização em lesões sólidas e papilas. *Escores* são atribuídos de acordo com o resultado: *escore* 1 se nenhum fluxo for demonstrado, *escore* 2 se houver fluxo mínimo, *escore* 3 para fluxo moderado e *escore* 4 se vascularização intensa (Timmerman, D. *et al.,* 2000).

Após a devida análise da localização e morfologia, com os descritores preconizados, e a medição da lesão, ela pode ser classificada em uma das seis categorias específicas do léxico: 1. Cisto unilocular; 2. Cisto unilocular sólido; 3. Cisto multilocular; 4. Cisto multilocular sólido; 5. Tumor sólido (80% de sua composição é sólida) e 6. Não classificável (por dificuldade de visualização devido à artefatos como calcificações produtoras de sombra acústica posterior) (Timmerman, D. *et al*., 2000).

Para a validação dos descritores definidos, alguns modelos de análise de regressão logística foram propostos pelo grupo IOTA para a estratificação de risco e, então, estudados em trabalhos retrospectivos e, posteriormente, em trabalhos prospectivos, com resultados semelhantes (Barra; Jorge; Condé, 2021).

Inicialmente, entre 2000 e 2004, foram utilizados dois modelos: o modelo 1 (M1) e o modelo 2 (M2). O M1 contemplava doze variáveis clínicas e ultrassonográficas: idade, antecedente pessoal de neoplasia ovariana, uso de terapia hormonal, dor ao exame físico, presença de ascite, presença de tumor sólido, maior diâmetro da lesão, *escore Doppler*, maior diâmetro da papila, presença de fluxo papilar, irregularidade nas paredes internas e sombra acústica posterior (Timmerman, D. *et al.,* 2005), enquanto o M2 contemplava seis variáveis: idade, presença de ascite, maior diâmetro da papila, presença de fluxo papilar, irregularidade nas paredes internas e sombra acústica posterior, sendo que ambos tiveram performance semelhante na diferenciação entre benignidade e malignidade (Timmerman, D*. et al.,* 2010).

Analisaram a performance dos modelos com a adição do valor do CA-125, uma glicoproteína utilizada como marcador de alguns tumores, sem evidências de melhora significativa na acurácia diagnóstica (Timmerman, D*. et al*., 2007).

Em 2008, o grupo definiu dez variáveis ecográficas para a diferenciação de benignidade e malignidade, aplicadas somente para lesões não fisiológicas: as *Simple Rules* (dez Regra Simples) de IOTA (IOTA-SR), um modelo dicotômico, bem como a sua interpretação (Timmerman, D. *et al.,* 2008).

Cinco critérios estariam relacionados à benignidade (cisto unilocular, papila menor que 7 milímetros, sombra acústica posterior presente, cisto multilocular com paredes lisas e regulares e com diâmetro máximo menor que 100 milímetros, e ausência de vascularização ao *Doppler* colorido (*escore* 1)), e cinco critérios à malignidade (tumor sólido de contorno irregular, ascite, pelo menos quatro papilas, cisto sólido multilocular com diâmetro máximo maior ou igual a 100 milímetros e intensa vascularização ao *Doppler* colorido (*escore* 4)). Na presença de pelo menos um critério de benignidade e nenhum de malignidade, a interpretação é de baixo risco de malignidade (1% a 5%); na presença de pelo menos um critério de malignidade e nenhum de benignidade, a interpretação é de alto risco de malignidade (75% a 90%) e, se houver ambos os critérios ou nenhum, o risco é de 25% a 50% (Timmerman, D. *et al*., 2008).

As *Simple Rules* apresentaram um bom desempenho diagnóstico tanto em centros de oncologia quanto em outros centros não especializados, fundamentando sistemas de manejo clínico (Timmerman, D. *et al.,* 2016). No entanto, não são aplicáveis à totalidade das massas anexiais: cerca de 20-25% dos casos permanecem inconclusivos. Nessas situações, a definição diagnóstica exige a intervenção de variáveis adicionais, como a avaliação subjetiva por um examinador experiente ou a utilização de métodos outros propedêuticos complementares (Andreotti *et al.,* 2020).

Um outro modelo proposto pelo grupo IOTA em 2014 foi o *ADNEX (Assessment of Different Neoplasias in the Adnexa*), modelo de regressão logística politômico que utiliza três variáveis clínicas (idade, centro de referência em oncologia e CA-125 - opcional, mas recomendado) e seis variáveis ultrassonográficas (maior diâmetro da lesão, número de projeções papilares sólidas, maior diâmetro da projeção papilar sólida, presença de mais de 10 lóculos, sombra acústica posterior, ascite) e estratifica os percentuais de risco de cada massa.

O modelo *ADNEX* demonstrou boa discriminação bem entre tumores benignos e malignos e discriminação razoável a excelente entre quatro tipos de malignidade ovariana (*borderline*, tumor primário invasivo maligno estágio I, tumor primário invasivo maligno estágio II-IV e metástase), com potencial de melhorar a triagem e as decisões de tratamento, reduzindo assim a morbidade e a mortalidade associadas à patologia anexial (Andreotti *et al*., 2020; Van Calster *et al*., 2014, 2015).

Para abordar os resultados inconclusivos das *Simple Rules* e fornecer um risco numérico de malignidade para lesões anexiais, um outro modelo foi desenvolvido em 2016 pelo grupo: o *Simple Rules Risk* (SRR), avaliação de risco de regra simples. Ele utiliza as mesmas características ultrassonografias de benignidade e malignidade das *Simple Rules*, que são, então, inseridas em uma calculadora disponível online, a '*Simple Rules Risk calculator'*, juntamente com informações do centro de referência do seguimento, oncológico ou não (Hiett *et al*., 2022).

Uma abordagem alternativa para a estratificação de risco é a estratégia de duas etapas do IOTA. Primeiramente, aplicam-se os descritores simples de benignidade, que não exigem acesso imediato a um computador e, nos casos em que estes não são aplicáveis, prossegue-se a avaliação com o modelo *ADNEX* ou SRR. Quando validada externamente com pacientes tratados cirurgicamente ou com conduta expectante, essa estratégia de duas etapas apresentou desempenho satisfatório (Landolfo *et al.,* 2023; Timmerman, S. *et al.,* 2023).

No entanto, apesar da ampla validação internacional, a aceitação do consenso foi limitada na prática clínica geral nos Estados Unidos da América (EUA) e no Canadá (Andreotti *et al*., 2018), o que pode estar relacionado com a preferência por uma abordagem de reconhecimento de padrões em vez de um modelo matemático (ADNEX), à falta de validação desses modelos nos EUA (Hiett *et al*., 2022), bem como com a ausência de orientação mais detalhada na avaliação de muitas lesões que são quase certamente benignas (Andreotti *et al*., 2020).

Outros sistemas de caracterização e manejo das massas ovarias foram, então, propostos, como o consenso da Sociedade de Radiologistas em Ultrassom, o índice morfológico da Universidade de Kentucky e o GI-RADS (*Gynecologic Imaging Reporting and Data System*) (Andreotti *et al*., 2020), no entanto todos com vieses que não os tornaram populares.

1.3. Massas Anexiais e O-RADS

Diante da necessidade de uma ferramenta de relatório padrão universalmente reconhecida que fosse precisa, útil e que incluísse todos os descritores e definições pertinentes, o Colégio Americano de Radiologia (ACR) publicou, em 2018, o léxico padrão, seguido por um Guideline de Consenso denominado O-RADS US v2020 ("*Ovarian-Adnexal Reporting and Data System*"). Trata-se de um sistema de estratificação e manejo das lesões anexais com o objetivo de diminuir ou eliminar ambiguidades em laudos ultrassonográficos através da limitação da variabilidade na linguagem dos relatórios, de melhorar a qualidade e comunicação entre os médicos e de fornecer uma recomendação de gerenciamento para cada categoria de risco, consistentes na prática clínica (Andreotti *et al.,* 2018).

Os membros do comitê concordaram que os termos descritores de ultrassom desenvolvidos anteriormente pelo grupo IOTA eram os mais robustos baseados em evidências disponíveis na literatura e o aderiram na sua totalidade, ocasionalmente com substituições acordadas para manter a familiaridade entre os usuários (Andreotti *et al.,* 2018).

As diretrizes propostas – colaborativas, multidisciplinares e internacionais – incorporaram as abordagens comuns europeias e norte-americanas e incluíram todas as categorias de risco com as respectivas estratégias de gestão, que anteriormente não foram incluídas em nenhum dos sistemas (Andreotti *et al.,* 2018, 2020). No entanto, não substituem a realização de uma anamnese e exame físico completos e a avaliação da necessidade de testes adicionais (Andreotti *et al*., 2020).

Apesar das melhorias, o O-RADS US v2020 exibiu especificidade relativamente baixa, com meta-análises relatando uma faixa de 75% a 82% (Han; Wen; Hu, 2024; Lee *et al.,* 2023; Liu *et al.,* 2024), o que poderia levar ao tratamento excessivo. Para abordar essas limitações e melhorar o seu desempenho diagnóstico, o O-RADS US v2020 foi atualizado em sua versão O-RADS v2022, incorporando variáveis adicionais, como bilocularidade em lesões císticas e sombreamento em lesões sólidas lisas (Liu *et al.,* 2024; Strachowski *et al*., 2023).

O modelo O-RADS se aplica somente em lesões de etiologia ovarias ou tubária e para aquelas indeterminadas, mas com suspeita dessas origens. Se uma lesão pélvica for claramente identificada como não ovariana ou tubária, o sistema O-RADS

é apropriado apenas para casos de cisto paraovariano ou de inclusão peritoneal; caso contrário, não se aplica (Andreotti *et al.,* 2020).

A estratificação de risco é dividida em seis categorias, com recomendações específicas em cada caso e, na maioria das vezes também categorizadas em pacientes na pré e na pós menopausa (Andreotti *et al*., 2020; Strachowski *et al.,* 2023).

A categoria 0 (O-RADS 0) é aplicada em casos em que a avaliação é considerada incompleta por dificuldades relacionadas à técnica do exame, com recomendação de repetir o estudo ou, em casos específicos, realizar complementação com outro método de imagem, como a Ressonância Nuclear Magnética (Andreotti *et al*., 2020).

O-RADS 1 é uma categoria relevante apenas para paciente na pré menopausa, em que o ovário é normal ou os achados são considerados fisiológicos (folículos ovarianos e corpo lúteo menores que 3 centímetros), não sendo necessários acompanhamentos adicionais (Andreotti *et al.,* 2020; Strachowski *et al*., 2023).

A categoria 2 (O-RADS 2), quase certamente benigna, com risco de malignidade de 1%, compreende a maioria dos cistos uniloculares menores que 10 centímetros (3 a 10 centímetros em pacientes não menopausadas), chamados de cistos simples; cistos uniloculares com ecos ou septos incompletos internos, chamados de cistos não simples; cistos biloculares menores que 10 centímetros (um septo completo – duas loculações); e outros cistos, desde que típicos, como hemorrágicos, dermóides, endometriomas, paraovarianos, de inclusão peritoneal; e a hidrossalpinge.

Geralmente nesta categoria nenhum acompanhamento adicional é necessário. No entanto, alguns subgrupos podem necessitar de conduta adicional (Andreotti *et al.,* 2020). Em pacientes na pré-menopausa com cistos simples entre 5 e 10 centímetros recomenda-se reavaliação em 12 meses para confirmar sua origem funcional ou para reavaliar a presença de anormalidades na sua parede; em pacientes na pós menopausa com cistos simples entre 3 e 10 centímetros ou cistos não simples menores que 3 centímetros, recomenda-se pelo menos 1 ano de acompanhamento para caracterizar a redução do seu tamanho ou sua estabilidade, com consideração de acompanhamento anual por até 5 anos. Para cistos não simples acima de 3 centímetros ou cistos biloculados, é recomendada a reavaliação em 6 meses.

As lesões benignas clássicas pertencentes a este grupo, citadas anteriormente, também são manejadas de maneira específica de acordo com o período de vida da paciente (pré ou pós menopausa) e seu tamanho (Strachowski *et al*., 2023) (Anexo 1).

O-RADS 3 engloba achados com baixo risco de malignidade, de 1 a 10%, como os listados na categoria 2 e maiores que 10 centímetros; cistos uniloculares de qualquer tamanho com paredes irregulares; cistos multiloculares (acima de 3 loculações) menores que 10 centímetros e sem componente sólido; e lesões sólidas lisas, com sombra acústica posterior e não vascularizadas (a partir da categoria 3, o *escore* de cor passa a ser incorporado ao sistema de estratificação de risco) (Andreotti *et al.,* 2020; Strachowski *et al.,* 2023). A maioria das lesões desta categoria são benignas e, por isso, o manejo pode ser feito por um ginecologista generalista, mas com recomendação de considerar a reavaliação da imagem em 6 meses e a avaliação de imagens sólidas por ecografista especialista ou Ressonância Nuclear Magnética, para reduzir o risco de ignorar características mais suspeitas (Strachowski *et al.,* 2023).

A categoria 4 (O-RADS 4), com risco intermediário de malignidade (10 a 50%), inclui os achados listados na categoria 3, mas maiores que 10 centímetros, além de cistos com irregularidade de paredes internas ou septos e com projeções sólidas e *escore* 4 de vascularização, justificando, em seu manejo, a avaliação de um ginecologista oncológico, além de avaliação de imagem por especialista, ainda que o caso seja conduzido por ginecologista generalista. Em paciente na pós menopausa, outros métodos complementares como marcadores séricos podem ser de grande valia (Andreotti *et al*., 2020).

O-RADS 5, em que o risco de malignidade é alto (acima de 50%), engloba lesões, de qualquer tamanho, sólidas com ou sem sombra acústica posterior, sólidas irregulares, e cistos uniloculares com quatro ou mais papilas, cistos bi ou multiloculares com componente sólido e alto *escore* de cor. A presença de ascite e/ou nódulos peritoneais também é incluída nesta categoria, exceto quando há ascite associada a um cisto fisiológico ou quase certamente a uma lesão benigna. Nesta categoria, a recomendação é de acompanhamento por um ginecologista oncologista *(Andreotti et al.*, 2020; Strachowski *et al*., 2023).

Os Anexos 2 e 3 trazem a tabela e o algoritmo de categorias de avaliação, respectivamente, apresentados pelo Colégio Americano de Radiologia (ACR), em sua versão mais atualizada (v2022) (Strachowski *et al.,* 2023).

2. JUSTIFICATIVA

Justifica-se o presente estudo para identificar, entre os modelos IOTA e o sistema O-RADS, aquele que apresenta a melhor acurácia diagnóstica, proporcionando, assim, a menor perda de malignidade.

3. OBJETIVOS

3.1. Objetivo Geral

Realizar uma revisão sistemática da literatura afim de comparar a acurácia diagnóstica dos modelos ultrassonográficos IOTA e O-RADS na classificação de massas anexiais.

3.2. Objetivos Específicos

1. Levantar dados na literatura atual para determinar a acurácia dos modelos ultrassonográficos IOTA e O-RADS no diagnóstico de massas anexiais malignas.
2. Realizar uma metanálise comparativa entre a acurácia dos modelos ultrassonográficos IOTA e O-RADS na detecção de massas anexiais malignas.
3. Avaliar o risco de viés dos estudos incluídos na metanálise por meio das ferramentas QUADAS-2 e QUADAS-C, assegurando a validade interna das comparações.
4. Determinar o nível de evidência e a certeza dos achados utilizando a abordagem GRADE (*Grading of Recommendations, Assessment, Development and Evaluation*), considerando critérios de risco de viés, inconsistência, imprecisão e viés de publicação.

4. METODOLOGIA

4.1. Desenho do estudo

Trata-se de uma revisão sistemática da literatura com metanálise, conduzida conforme as diretrizes do *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy,* versão 2.0 (Deeks *et al.,* 2023).

Para a formulação da pergunta de pesquisa, utilizou-se a estratégia PICO (*Patient or Problem, Intervention, Comparasion, Outcomes),* metodologia prevista pela Prática Baseada em Evidências (PBE), em que:

- **P** (*Patient* / população) refere-se a pacientes com massas anexiais;
- **I** (*Intervention* / intervenção) refere-se à ultrassonografia da pelve, utilizando a análise pelos modelos IOTA e O-RADS;
- **C** (*Comparison* / comparação) refere-se a um grupo controle;
- **O** (*Outcome* / desfecho) refere-se à acurácia diagnóstica.

A pergunta norteadora foi: *Qual a acurácia diagnóstica dos modelos IOTA e O-RADS na classificação ultrassonográfica de massas anexiais?*

4.2. Amostra

4.2.1. Tipos de estudos

Foram incluídos estudos longitudinais observacionais prospectivos ou retrospectivos que compararam a acurácia diagnóstica dos modelos ultrassonográficos de estratificação IOTA e O-RADS para massas anexiais em relação à histopatologia da lesão, nos casos em que houve exérese cirúrgica. O “follow up” ultrassonográfico foi aceito nos casos em que a cirurgia não foi realizada/recomendada.

4.2.2. Tipos de participantes

A população estudada foi composta por mulheres que apresentaram massas anexiais unilateral ou bilateral, diagnosticadas pela ultrassonografia da pelve, por via

abdominal ou transvaginal, classificadas e estratificadas pelos modelos IOTA e O-RADS.

4.2.3. Tipos de desfechos

Os desfechos analisados foram relacionados à acurácia diagnóstica dos modelos utilizados na classificação das imagens. Portanto, a análise da performance diagnóstica foi feita através:

- da sua sensibilidade (desfecho primário)
- da sua especificidade (desfecho primário)
- área sob a curva ROC (AUC)
- razão de chances diagnóstica (DOR)

4.3. Protocolo e Registro

Esta revisão sistemática com metanálise foi conduzida e redigida conforme as diretrizes do *Preferred Reporting Items for Systematic Reviews and Meta-Analyses* (PRISMA 2020) (Page *et al*., 2021) utilizando a sua extensão específica para estudos de acurácia diagnóstica, o PRISMA-DTA (Mcinnes *et al.,* 2018), para o relato transparente dos métodos e resultados (Apêndice 1).

O seu protocolo foi previamente incluído na base Internacional de Registro Prospectivo de Revisões Sistemáticas – PROSPERO (*International Prospective Register of Systematic Reviews*), sob o número de identificação CRD42024581159, disponível em: https://www.crd.york.ac.uk/PROSPERO/view/CRD42024581159.

4.4. Critérios de Elegibilidade

4.4.1. Critérios de inclusão:

- Estudos que utilizaram análise de imagem ultrassonográfica de massas anexais e sua estratificação de risco segundo os modelos IOTA e O-RADS, e que comparam o seu desempenho diagnóstico. Não foram impostas restrições em relação ao equipamento utilizado e em relação aos seus detalhes técnicos.

- Estudos que tenham como padrão de referência o resultado de anatomopatológico (histopatologia) após cirurgia ou seguimento por imagem.
- Estudos datadas a partir do ano de 2018 (ano em que o O-RADS foi publicado) até o ano atual.
- Não foram impostas restrições de idiomas.

4.4.2. Critérios de exclusão:

- Estudos que utilizaram ressonância nuclear magnética como método principal de imagem, ainda que tenham empregado os modelos IOTA e O-RADS para estratificação de risco de massas anexiais.
- Estudos sem a comparação direta entre os modelos IOTA e O-RADS.
- Estudos cujos dados relatados não forem suficientes para a avaliação do desempenho diagnóstico dos modelos, mesmo após tentativa de contato por e-mail com os autores.
- Relatos de casos, erratas e cartas, diretrizes, diretrizes clínicas / terapêuticas (Guidelines), manuscritos apenas com resumo, resumos publicados em congressos, revisões de literatura, revisões sistemáticas ou metanálises.
- Estudos duplicados.
- Estudos considerados de baixa qualidade metodológica.

4.5. Busca na Literatura

A busca de alta sensibilidade, ampla, abrangente e com estratégia pré-definida (palavras-chave e termos *MeSH*) foi realizada entre 1º de janeiro de 2025 e 31 de janeiro de 2025, nas seguintes bases de dados eletrônicas: PubMed (que inclue os registros indexados no MEDLINE – *Medical Literature Analysis and Retrieval System Online -,* principal base da National Library of Medicine), LILACS (*Literatura Latino-Americana e do Caribe em Ciências da Saúde*), Embase, Ovid e Cochrane Library.

A Tabela 1 demonstra o processo de busca realizado:

Tabela 1- Estratégia de Busca nas bases de dados eletrônicas

Bases de Dados	Estratégia de Busca	Número de artigos encontrados
PubMed	(((((((((cancer of ovary[MeSH Terms]) OR (cancer of the ovary[MeSH Terms])) OR (adnexal disease[MeSH Terms])) OR (adnexal diseases[MeSH Terms]))) AND ((((((((((echography[MeSH Terms]) AND (“IOTA”)) OR (“INTERNATIONAL OVARIAN TUMOR ANALYSIS”)) OR (“IOTA SR”)) OR (“SIMPLE RULES”)) OR (“ADNEX”)) OR (“IOTA models”)) AND (“O-RADS”)) OR (“ovarian adnexal reporting and data system”))))) AND ((((cancer early detection[MeSH Terms]) OR (cancer early diagnosis[MeSH Terms])) OR (diagnoses[MeSH Terms]))))	33
Embase	(‘ovary disease’/exp OR ‘ovary tumor’/exp OR ‘ovary cancer’/exp OR ‘adnexa disease’/exp OR ‘adnexal 31âncer (gynecologic)’/exp) AND ‘echography’/exp AND ‘international ovarian tumor analysis’/exp AND ‘gynecologic 31âncer31 reporting and data system’/ exp AND (‘diagnosis’/exp OR ‘tumor diagnosis’/exp OR ‘cancer diagnosis’/exp)	3
Cochrane	(Ovarian neoplasm):ti, ab,kw OR (adnexal disease):ti, ab, kw AND (“Echography”):ti, ab,kw AND (“IOTA”):ti, ab,kw OR (“International ovarian tumor analysis”):ti, ab,kw AND (“O-RADS”):ti,ab,kw OR (“ovarian adnexal reporting and data system”):ti,ab,kw AND (Diagnosis):ti,ab,kw	0
Ovid	(((((((((cancer of ovary[MeSH Terms]) OR (cancer of the ovary[MeSH Terms])) OR (adnexal disease[MeSH Terms])) OR (adnexal diseases[MeSH Terms]))) AND ((((((((((echography[MeSH Terms]) AND (“IOTA”)) OR (“INTERNATIONAL OVARIAN TUMOR ANALYSIS”)) OR (“IOTA SR”)) OR (“SIMPLE RULES”)) OR (“ADNEX”)) OR (“IOTA models”)) AND (“O-RADS”)) OR (“ovarian adnexal reporting and data system”))))) AND ((((31âncer early detection[MeSH Terms]) OR (31âncer early diagnosis[MeSH Terms])) OR (diagnoses[MeSH Terms]))))	0
LILACS	(mh⊛adnexal disease)) OR (mh⊛adnexal diseases)) OR (mh⊛cancer of ovary)) OR (mh⊛cancer of the ovary)) OR (mh⊛ovary neoplasm)) OR (mh⊛ovary neoplasms)) AND (mh⊛echography)) OR (mh: (ultrasonography)) AND (“IOTA”) OR (“international ovarian tumor analysis”) AND (“O-RADS”) OR (“ovarian adnexal reporting and data system”) AND (mh⊛diagnoses)) OR (mh⊛31âncer early detection)) OR (mh⊛31âncer early diagnosis))	0

Fonte: Elaboração própria (2025). **Legenda:** ADNEX = *Assessment of Different Neoplasias in the adnexa*; IOTA = *International Ovarian Tumor Analysis*; LILACS = Literatura Latino-Americana e do Caribe em Ciências da Saúde**;** MeSH = *Medical Subject Headings*; O-RADS = *Ovarian-Adnexal Reporting and Data System*; SR = *Simple Rules*.

Ainda, as referências dos artigos selecionados foram manualmente avaliadas para buscar outros estudos potencialmente elegíveis.

4.6. Seleção dos estudos

Critérios de elegibilidade (inclusão e exclusão), previamente estabelecidos e descritos anteriormente, foram empregados para a seleção pareada dos artigos em duas etapas: primeiramente, com a leitura do título e resumos por dois pesquisadores (AGC e JAM) e de maneira independentes, com seleção de um primeiro grupo. Discordâncias foram resolvidas por consenso. A elegibilidade dos estudos selecionados nesta primeira etapa de triagem foi confirmada com sua leitura na

íntegra, também de forma pareada e independente, e, desses, foram selecionados àqueles que obedeceram aos critérios para esta pesquisa. Todas as discordâncias foram resolvidas por consenso. Neste processo, o gerenciador de referência bibliográfica *Mendeley* foi utilizado para excluir artigos duplicados e um gerenciador de revisão sistemática, o *software on-line Rayyan, foi* utilizado para auxiliar a organização dos artigos elegíveis.

4.7. Coleta de Dados

Um pesquisador (AGC), com conhecimento do tema, coletou os dados dos estudos selecionados utilizando uma planilha de Excel e o segundo pesquisador (JAM) revisou os dados coletados. Discordâncias foram resolvidas por consenso.

Para casos em que os dados analisados não estiveram totalmente elucidados, foi realizada tentativa de contato com os autores por e-mail para maior detalhamento e, para os casos em que não obtivemos respostas, os estudos foram excluídos.

4.8. Avaliação do risco de viés

Os artigos elegíveis foram inicialmente submetidos à análise de risco de viés por dois membros (AGC e JAM), pela ferramenta *Quality Assessment of Diagnostic Accuracy Studies-2* (QUADAS-2) (Whiting *et al.,* 2011), em quatro domínios: seleção de pacientes, teste índice, teste referência ou padrão-ouro, e fluxo e tempo. Em seguida, a ferramenta QUADAS-C (Comparative) (Yang *et al.,* 2021) foi aplicada para avaliar o risco de viés na comparação direta entre os testes diagnósticos (IOTA e O-RADS). Os estudos com baixo risco de viés em pelo menos três domínios do QUADAS-2 e baixo risco de viés na comparação (QUADAS-C) foram classificados como de alta qualidade e, ao contrário, os estudos com alto risco de viés ou risco incerto em pelo menos três domínios e/ou alto risco de viés na comparação foram classificados como de baixa qualidade. Àqueles que fugiram dos critérios descritos anteriormente foram considerados de moderada qualidade.

4.9. Análise dos dados (metodologia estatística)

Para a síntese dos dados de acurácia diagnóstica foi conduzida uma metanálise em três estágios. Inicialmente, foram realizadas análises univariadas utilizando modelos de efeitos aleatórios para estimar separadamente as medidas globais de sensibilidade, especificidade, razão de chances diagnóstica (DOR), área sob a curva ROC (AUC) e respectivos intervalos de confiança de 95%, além de avaliar a heterogeneidade por meio do teste de Cochran e do índice I^2.

Em seguida, foi aplicado o modelo bivariado de efeitos aleatórios descrito por Reitsma et al. (2005) para estimar simultaneamente sensibilidade e especificidade, obtendo-se também as curvas ROC resumidas (SROC – *Summary Receiver Operating Characteristic*).

A comparação entre métodos diagnósticos (que foi metodologicamente possível entre O-RADS e IOTA-SR e O-RADS e IOTA *ADNEX*) foi realizada incluindo o fator “método” como covariável no modelo bivariado. A presença de heterogeneidade entre os estudos foi explorada pelo índice I^2 segundo a abordagem de Zhou & Dendukuri (Zhou; Dendukuri, 2010). A disparidade no número de estudos e, consequentemente, no tamanho amostral total entre dois modelos é inerente à metanálise de dados de pacientes a nível de estudo e não compromete a validade da comparação. O modelo de regressão bivariada empregado é capaz de ajustar as estimativas de efeito para essa diferença amostral, permitindo uma comparação robusta da acurácia relativa (β) entre os testes, conforme as diretrizes metodológicas. Todas as análises foram conduzidas no software R, versão 4.4.2 (R Core Team, 2024), utilizando o pacote mada (Doebler, 2017).

A avaliação de viés de publicação por meio de gráficos de funil ou testes estatísticos associados não foi realizada neste estudo. Isso se deve ao fato de que tais métodos não possuem validação adequada para metanálises de acurácia diagnóstica.

A certeza da evidência foi avaliada seguindo a metodologia GRADE, conforme detalhado no manual do grupo de trabalho (Schünemann *et al.*, 2013). Para a construção das tabelas de perfil de evidência e resumo dos achados (Summary of Findings), utilizou-se o software GRADEpro GDT (McMaster University; Evidence Prime, 2015).

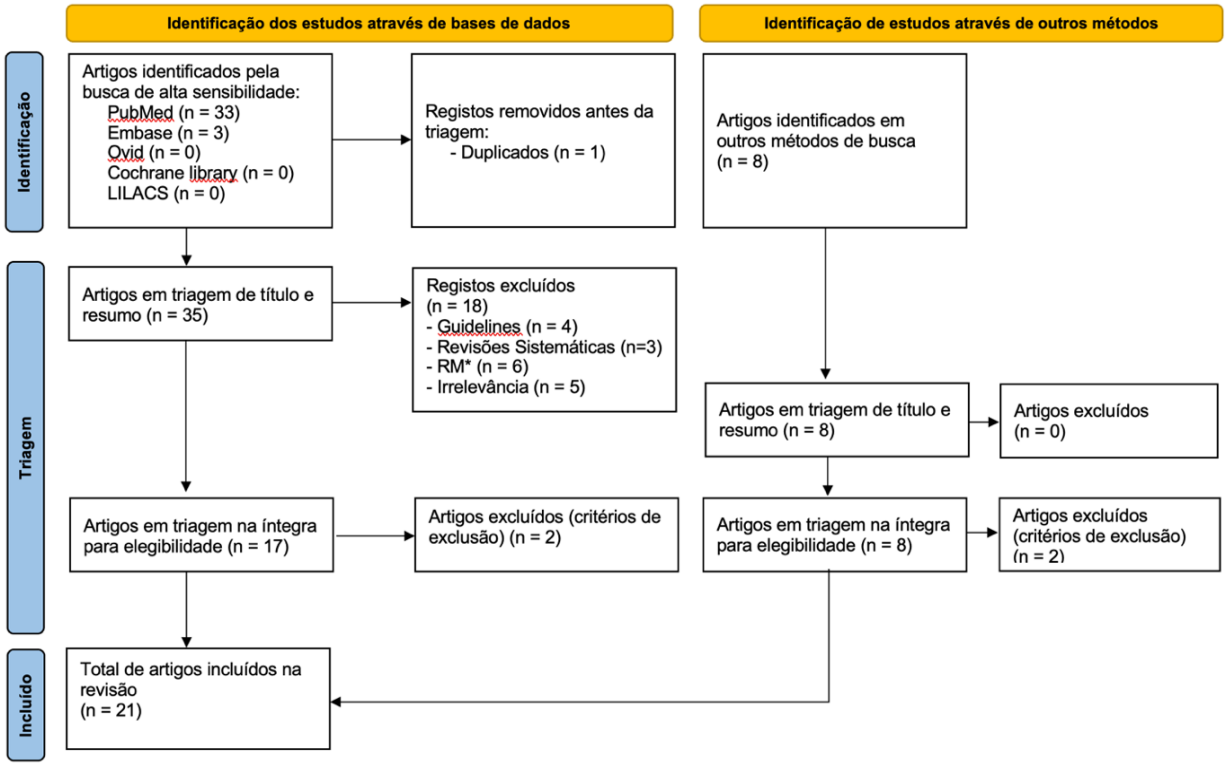
5. RESULTADOS

5.1. Seleção dos estudos

A estratégia de busca sistematizada resultou na identificação de 36 registros. Após a remoção de uma duplicata, 35 artigos foram submetidos à triagem por título e resumo. Nesta fase, 18 artigos foram excluídos por não preencherem os critérios de inclusão, resultando em 17 artigos selecionados para a próxima etapa. Paralelamente, as referências dos artigos selecionados foram manualmente avaliadas para buscar outros estudos potencialmente elegíveis, o que permitiu a identificação, também pelo título e resumo, de mais oito estudos.

Um total de 25 artigos (17 da busca inicial e 8 da busca manual) foram submetidos à análise de texto completo para avaliação de elegibilidade. Dentre estes, quatro foram excluídos com base nos critérios pré-definidos, cujos motivos de exclusão são detalhados no Apêndice 2, resultando em uma amostra final de 21 estudos incluídos na presente revisão sistemática.

Figura 1- Fluxograma de seleção dos estudos pela metodologia PRISMA.



Fonte: Elaboração própria (2025), adaptada de Page *et al.* (2020). **Legenda:** *n* = número; RM = ressonância nuclear magnética; LILACS = Literatura Latino-Americana e do Caribe em Ciências da Saúde.

5.2. Coleta de dados

Os dados coletados foram organizados em duas tabelas: a Tabela 2 detalha as características dos 21 estudos incluídos na revisão sistemática da literatura, enquanto a Tabela 3 compila os estudos elegíveis para a síntese quantitativa (metanálise) comparando os modelos IOTA e O-RADS.

Para a condução da metanálise, cinco dos 21 estudos foram excluídos; inicialmente, dois foram considerados inelegíveis por não compararem diretamente os modelos ultrassonográficos O-RADS e IOTA: Wang *et al.* (2023) comparou o O-RADS ecográfico com o ADNEX na ressonância nuclear magnética, enquanto Li *et al.* (2024) validou o modelo O-RADS isoladamente, sem realizar comparação com o IOTA. Três estudos foram excluídos posteriormente em razão da ausência de dados completos disponíveis para a análise estatística: Pelayo *et al*. (2023a), Pelayo *et al*. (2023b) e Filiz; Kahyaoglu e Atalay (2024). Todos os autores correspondentes foram contatados por correio eletrônico, porém não houve retorno até o encerramento do período de coleta.

Tabela 2 – Características dos estudos incluídos na Revisão Sistemática

Autor (ano)	País	Design do estudo	Período do estudo	Equipamento	Avaliadores (anos experiência)	Cego	Modelos avaliados	Padrão de referência
Basha, *et al*. (2021)	Egito	Coorte Restrospectivo Multicêntrico (3 centros)	05/2016 a 12/2019	GE LOGIC 9 Philips ClearVue 650 GE Voluson S8 BT18 SonoScape S40	5 (>15)	Sim	O-RADS GI-RADS IOTA SR	(1) histologia pós-operatório (2) follow-up 2 anos (3/3meses)
Timmerman S, *et al*. (2022)	EUA	Coorte Restrospectivo Multicêntrico (36 centros-14 países)	01/2012 a 03/2015	N/A	N/A Experientes (curso IOTA)	Sim	O-RADS IOTA 2 etapas (SR ou ADNEX)	(1) histologia pós-operatória (2) Follow-up 12 meses
Hiett, *et al* (2022)	EUA	Coorte Restrospectivo (centro único oncológico)	03/2018 a 02/2021	GE Voluson E8 Philips EPIC 7 Siemens S2000	2 (25)	N/A	O-RADS ADNEX IOTA SR IOTA SRR	(1) histologia pós-operatória (180 dias após US)
Hack, *et al*. (2022)	N/A Canadá	Coorte Retrospectivo (centro único oncológico)	08/2015 a 04/2017	Toshiba Aplio XG Philips iU22	2 (9 e 30)	Sim	O-RADS ADNEX	(1) histologia pós-operatória
Xie, *et al.* (2022)	China	Coorte Restrospectivo (Centro único)	01/2017 a 09/2020	Toshiba Aplio XG GE LOGIC E9	2 (>5)	Sim	IOTA SR O-RADS CA125	(1) histologia pós-operatória

Lai, *et* al. (2022)	China	Coorte Restrospectivo (Centro único)	01/2017 a 11/2020	GE Voluson E8 GE Voluson E10 GE Voluson 730 Expert Toshiba Aplio 770/790 ALOKA SSD 5500	2 (5)	Sim	ADNEX O-RADS GI-RADS	(1) histologia pós-operatória
Chen, *et al*. (2022)	Taiwan	Coorte Restrospectivo (Centro único)	01/2020 a 10/2020	GE Voluson S8 GE P6 Philips EPIC Philips Affiniti-50 Siemens Acuson x300PE	2 (> 5)	Sim	O-RADS ADNEX	(1) histologia pós-operatória
Yoeli-Bik, *et al.* (2023)	EUA	Coorte Restrospectivo (Centro único)	01/2017 a 10/2022	GE Voluson E8 GE Voluson E10 Samsung Elite WS80	2 (>40)	N/A	IOTA SR ADNEX O-RADS	(1) histologia pós-operatória (180 dias pós US) (2) Follow up
Poonyakanok; Tanmahasamut; Jaishuen, (2023)	Tailândia	Coorte Prospectiva (Centro único)	05/2018 a 05/2019	GE Voluson E6 BT13.5	1 dos 6 (5 a 18)	Sim	O-RADS ADNEX RMI	(1) histologia pós-operatória (120 dias pós US)
Shang, *et al*. (2023)	China	Coorte Restrospectivo (Centro único)	01/2003 a 12/2021	N/A (variedade)	1 (N/A) Experiente	Sim	IOTA SR ADNEX O-RADS	(1) histologia pós-operatória Brenner (14 dias pós US)
Lu, *et al.* (2023)	China	Coorte prospectivo (2 centros)	01/2018 a 12/2021	GE E10 Phillips EPIQ 7 Siemens ACUSON OXANA Contraste: SonoVue (3 min após)	2 (N/A) experiente	Sim	IOTA O-RADS CEUS	(1) histologia pós-operatória
Wang, *et al.* (2023)	China	Coorte Restrospectivo (Centro único)	05/2017 a 07/2022	Philips EPIQ7 e Acuson Sequoia Contraste: SonoVue	N/A (2 a 12)	N/A	O-RADS O-RADS + CEUS ADNEX (RM)	(1) histologia pós-operatória (2) follow up (6 a 12 meses).
Li, *et al*. (2024)	China	Coorte prospectiva (Centro único)	05/2020 a 08/2021	GE LOGIQ 9 GE Voluson E8, Mindray Resona 8EXP	1 (20)	Não	O-RADS	(1) histologia pós operatória (120 dias pós US)
Lu, *et al.* (2024)	China	Coorte restrospectivo (2 centros)	01/2018 a 02/2023	GE E10 Phillips EPIQ 7 Siemens ACUSON OXANA Contraste: SonoVue (3 min após)	2 (>20)	Sim	IOTA SR O-RADS CEUS	(1) histologia pós-operatória (três meses do exame CEUS.)

Liu, *et al*. (2024)	China	Coorte Restrospectivo (Centro único)	01/2018 a 12/2023	GE LOGIQ S8 GE Esaote MyLab70	2 revisaram (+10 e +5)	Sim	IOTA SR O-RADS (v2022) O-RADS (V2020)	(1) histologia pós-operatória (120 dias pós US)
Guo, *et al.* (2022)	China	Coorte Restrospectivo (Centro único)	01/2017 a 10/2020	N/A (Transdutores intracavitários de 9-15 MHz)	2 (>10) 2 (1)	Sim	O-RADS RMI4 IOTA LR2 IOTA SR	(1) histologia pós-operatória
Pelayo, *et al.* (2023a)	Espanha	Coorte Restrospectivo (Centro único)	01/2021 a 12/2022	GE Voluson E8 Canon Aplio A	2 (>15)	Sim	Subjetiva IOTA SR IOTA SRR ADNEX O-RADS	(1) histologia pós-operatória (180 dias pós US)
Pelayo, *et al.* (2023b)	Espanha	Coorte Restrospectivo (Centro único)	01/2021 a 12/2022	GE Voluson E8 Canon Aplio A Canon Xario 100	2 (N/A)	Sim	Subjetiva IOTA SR ADNEX O-RADS	(1) histologia pós-operatória (180 dias pós US)
Yang; Ju; Huang, (2023)	China	Coorte Restrospectivo (Centro único)	09/2021 a 02/2022	Philips iU22 Philips EPIQ 5 Toshiba Aplio400 Toshiba Aplio500 GE LOGIQ E9 SuperSonic Imagine Aixplorer	1 (N/A) 1 (10)	Sim	CA125 HE4 ROMA IOTA SR O-RADS	(1) histologia pós-operatória (120 dias pós US)
Filiz, *et al.* (2023)	Turquia	Coorte Restrospectivo (Centro único)	09/2019 a 09/2021	GE Voluson S10	N/A	N/A	ADNEX O-RADS	(1) histologia pós-operatória
Spagnol, *et al. (*2023)	Itália	Coorte Restrospectivo (Centro único)	01/2018 a 12/2021.	GE Voluson E8 GE Voluson S10	2 (curso IOTA – N/A) 1 (sob supervisão)	Sim	IOTA SR IOTA SRR ADNEX O-RADS	(1) histologia pós-operatória (3 meses após US)

Fonte: Elaboração própria (2025). **Legenda:** ADNEX = *Assessment of Different Neoplasias in the adnexa*; CA 125 = *Cancer Antigen 125*; CEUS = *Contrast-Enhanced Ultrasound*; EUA = Estados Unidos da América; GE = *General Electric Health Care*; GI-RADS = *Gynecology Imaging Reporting and Data System*; HE4 = *Human Epididymis Protein 4*; IOTA = *International Ovarian Tumor Analysis*; min = minuto; N/A = *not available*; O-RADS = *Ovarian-Adnexal Reporting and Data System*; RM = ressonância nuclear magnética; RMI = *Risk of Malignancy Index*; ROMA = *Risk of Ovarian Malignancy Algorithm*; SR = *Simple Rules*; SSR = *Simple Rules Risk*; US = ultrassom; v2020 = Versão 2020 do sistema O-RADS; v2022 = Versão 2022 do sistema O-RADS.

Tabela 3 – Estudos elegíveis para a síntese quantitativa (metanálise)

Autor (ano)	Total MA	n malignas	n beninas	Modelo avaliado	Cutoff	TP	FP	TN	FN
Basha, *et al.* (2021)	647	178	469	O-RADS IOTA SR	O-RADS >3 IOTA SR M ou I	172 164	34 32	435 437	6 14
Timmerman S, *et al.* (2022) (O-RADS 9 MA não classificadas)	4905	1038	3867	O-RADS IOTA 2 etapas	O-RADS >3 IOTA >10%	978 962	865 655	2995 3212	58 76
Hiett, *et al* (2022)	150 100 (O-RADS>3)	40 14	110 86	O-RADS IOTA SR IOTA SRR ADNEX	O-RADS ≥10% (>3) IOTA SR M ou I IOTA SSR ≥10% ADNEX ≥10%	40 14 14 13	59 18 47 33	51 68 39 53	0 0 0 1
Hack, *et al.* (2022)	262	75	187	O-RADS ADNEX	O-RADS >3 ADNEX	74 N/A	56 N/A	131 N/A	1 N/A
Xie, *et al.* (2022)	453	269	184	O-RADS IOTA SR	O-RADS >3 IOTA SR M ou I	254 254	62 52	122 132	15 15
Lai, *et al.* (2022)	734	170	564	O-RADS ADNEX	O-RADS >3 ADNEX ≥10%	159 146	69 30	495 534	11 24
Chen, *et al.* (2022)	322	264	58	O-RADS ADNEX	O-RADS >3 ADNEX ≥10%	56 N/A	51 N/A	213 N/A	2 N/A
Yoeli-Bik, *et al.* (2023)	511	81	430	O-RADS IOTA SR ADNEX	O-RADS >3 IOTA SR M ou I ADNEX ≥10%	80 76 74	110 51 59	320 379 371	1 5 7
Poonyakanok; Tanmahasamut; Jaishuen, (2023)	357	61	296	O-RADS ADNEX	O-RADS >3 ADNEX ≥10%	61 60	30 38	266 258	0 1
Shang, *et al.* (2023)	23	8	15	O-RADS IOTA SRR ADNEX ADNEX +CA125	O-RADS >3 IOTA SRR>10% ADNEX >10% ADNEX >10%	8 8 8 7	9 9 7 4	6 6 8 11	0 0 0 1
Lu, *et al.* (2023)	35	19	18	O-RADS IOTA SR	O-RADS >3 IOTA SR M ou I	19 19	13 8	3 8	0 0
Lu, *et al.* (2024)	94	48	46	O-RADS IOTA SR	O-RADS >3 IOTA SR M ou I	48 45	33 26	13 20	0 3
Liu, *et al.* (2024)	613	454	159	O-RADS v2022 O-RADS v2020 IOTA SR	O-RADS >3 O-RADS >3 IOTA SR M ou I	450 450 437	49 52 41	110 107 118	4 4 17
Guo, *et al.* (2022)	592	145	447	O-RADS IOTA SR	O-RADS >3 IOTA SR M ou I	132 121	81 54	366 393	13 24
Pelayo, *et al.* (2023[a])*	122	41	81	O-RADS IOTA SR IOTA SRR ADNEX ADNEX +CA125	O-RADS >3 IOTA SRR >10% ADNEX ≥10% ADNEX ≥10%	N/A	N/A	N/A	N/A
Pelayo, *et al.* (2023b)*	187	53	134	O-RADS IOTA SR ADNEX ADNEX +CA125	O-RADS >3 ADNEX ≥10% ADNEX ≥10%	N/A	N/A	N/A	N/A
Yang; Ju; Huang, (2023)	1179	213	966	O-RADS IOTA SR	O-RADS >3 IOTA SR M ou I	204 186	174 112	792 854	9 27
Filiz, *et al.* (2023)*	413	118	295	O-RADS ADNEX	O-RADS >50% ADNEX >50%	N/A	N/A	N/A	N/A

Spagnol, *et al.* (2023)	514	114	400	O-RADS	O-RADS >3	106	43	357	8
				IOTA SR	IOTA SR M ou I	107	62	338	7
				IOTA SRR	IOTA SSR >10%	100	65	335	14
				ADNEX	ADNEX ≥10%	105	47	353	9
				IOTA 2 etapas (SR/SRR)	SR M/SSR >10%	100	33	367	14
				IOTA 2 etapas (SR/ADNEX)	SEM/ADNEX≥10%	103	27	373	11

Fonte: Elaboração própria (2025). **Legenda:** ADNEX = *Assessment of Different Neoplasias in the adnexa*; CA 125 = *Cancer Antigen 125*; FN = *false negative*; FP = *false positive*; I = indeterminada; IOTA = *International Ovarian Tumor Analysis*; M = maligna; MA = massas anexiais; n = número; N/A = *not available*; O-RADS = *Ovarian-Adnexal Reporting and Data System*; SR = *Simple Rules*; SSR = *Simple Rules Risk*; TN = *true negative*; TP = *true positive*; v2020 = Versão 2020 do sistema O-RADS; v2022 = Versão 2022 do sistema O-RADS. **Nota:** (*) Dentre os 19 estudos elegíveis listados, três não foram incluídos na síntese quantitativa (metanálise) devido à ausência de dados brutos, mesmo após tentativa de contato com os autores.

5.3. Metanálise

A metanálise incluiu dados de 16 estudos: 16 com dados completos do O-RADS, dez com dados completos do IOTA-SR e seis com dados completos do IOTA *ADNEX*. Os demais modelos (IOTA SSR e IOTA 2 etapas) apresentaram dados em apenas 3 estudos cada, não sendo metodologicamente possível a realização de análises individuais ou comparativas. O desempenho diagnóstico e a diferença relativa entre os testes (comparação) foram avaliados por meio do modelo bivariado, que ajustou as estimativas de efeito para compensar a diferença amostral entre os grupos, conforme detalhado na seção Metodologia.

A Tabela 4 apresenta a proporção de casos malignos entre todas as massas anexiais incluídas nos estudos e a taxa de detecção correta desses casos (sensibilidade bruta por estudo), de acordo com cada modelo avaliado. Observa-se que a proporção de malignidade variou entre os métodos, sendo mais baixa para o modelo *ADNEX* (média = 21,2%) e mais alta para o IOTA SR (média = 36,1%) e para o *O*-RADS (média = 36,3%). Quanto à detecção de malignos, os resultados foram elevados em todos os métodos, com médias acima de 90%. O melhor desempenho foi observado no IOTA SRR (95,9%), seguido pelo IOTA SR (93,5%), O-RADS (92,3%), *ADNEX* (91,3%) e IOTA 2 etapas (SR/ADNEX e SR/SSR) (90,2%).

Tabela 4 – Proporção e taxa de detecção de massas malignas segundo os diferentes modelos avaliados

Modelo avaliado	n artigos	Proporção de Malignos				Detecção de Malignos			
		Média	D.P.	Mínimo	Máximo	Média	D.P.	Mínimo	Máximo
ADNEX	6	0.212	0.076	0.140	0.348	0.913	0.044	0.859	0.984
IOTA 2 etapas	3	0.218	0.006	0.212	0.222	0.902	0.025	0.877	0.927
IOTA SR	10	0.361	0.215	0.140	0.741	0.935	0.051	0.834	1.000
IOTA SRR	3	0.237	0.105	0.140	0.348	0.959	0.071	0.877	1.000
O-RADS	16	0.363	0.211	0.159	0.820	0.923	0.192	0.212	1.000

Fonte: Elaboração própria (2025). **Legenda:** ADNEX = *Assessment of Different Neoplasias in the adnexa*; D.P. = desvio padrão; IOTA = *International Ovarian Tumor Analysis*; n = número; O-RADS = *Ovarian-Adnexal Reporting and Data System*; SR = *Simple Rules*; SSR = *Simple Rules Risk*;

5.3.1. O-RADS

A Figura 2 apresenta os gráficos *forest plots* de sensibilidade e especificidade do O-RADS. Nota-se que a sensibilidade é consistentemente alta em todos os estudos, geralmente acima de 0,90, o que demonstra a boa capacidade do teste em identificar corretamente os casos positivos. Já a especificidade mostrou mais variabilidade entre os estudos, com valores que oscilaram desde níveis baixos (cerca de 0,20–0,40) até elevados (>0,90). Do ponto de vista estatístico, não foi observada heterogeneidade significativa entre as sensibilidades (p = 0,212), mas a heterogeneidade entre as especificidades foi significativa (p < 0,001), refletindo essa maior instabilidade.

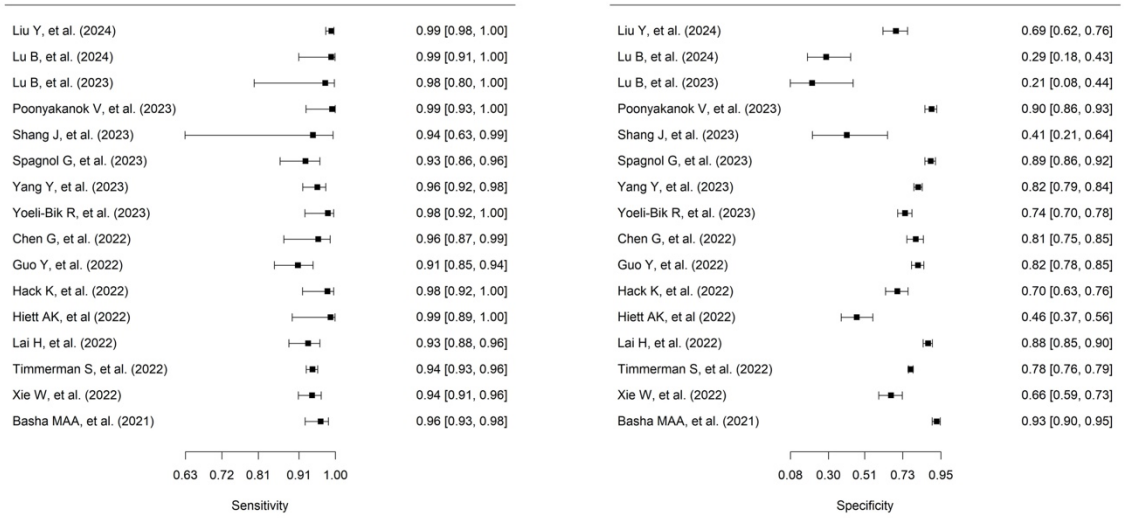
Na análise univariada, o desempenho diagnóstico global, avaliado pela razão de chances diagnóstica (DOR), foi elevado (80,1; IC95%: 54,2–118,3), correspondendo a um ln(DOR) de 4,4 (IC95%: 3,99–4,77). A heterogeneidade do ln(DOR) foi baixa (I^2 = 10,4%), o que também foi confirmado pelo teste de Cochran (p = 0,336). Esses resultados estão ilustrados na Figura 3. A área sob a curva ROC (AUC) estimada pelo método univariado foi de 0,976 (IC95%: 0,968–0,984), reforçando o alto desempenho do O-RADS.

Por sua vez, a análise bivariada, que é considerada mais robusta para metanálises de testes diagnósticos por modelar conjuntamente sensibilidade e especificidade e incorporar a correlação entre elas, apresentou resultados muito próximos. A sensibilidade média estimada foi de 0,961 (IC95%: 0,944–0,974). A taxa de falsos positivos foi de 0,269 (IC95%: 0,180–0,381), o que corresponde a uma especificidade média de aproximadamente 0,731. A AUC bivariada foi de 0,958,

levemente inferior à univariada, mas ainda indicativa de excelente acurácia global. A heterogeneidade também foi baixa (I^2 de Zhou e Dendukuri = 13,5%). A curva ROC resumo (SROC) derivada do modelo bivariado, apresentada na Figura 4, confirma esses achados: observa-se uma curva bastante alta (AUC elevada) e os pontos individuais dos estudos relativamente próximos, sinalizando consistência dos resultados.

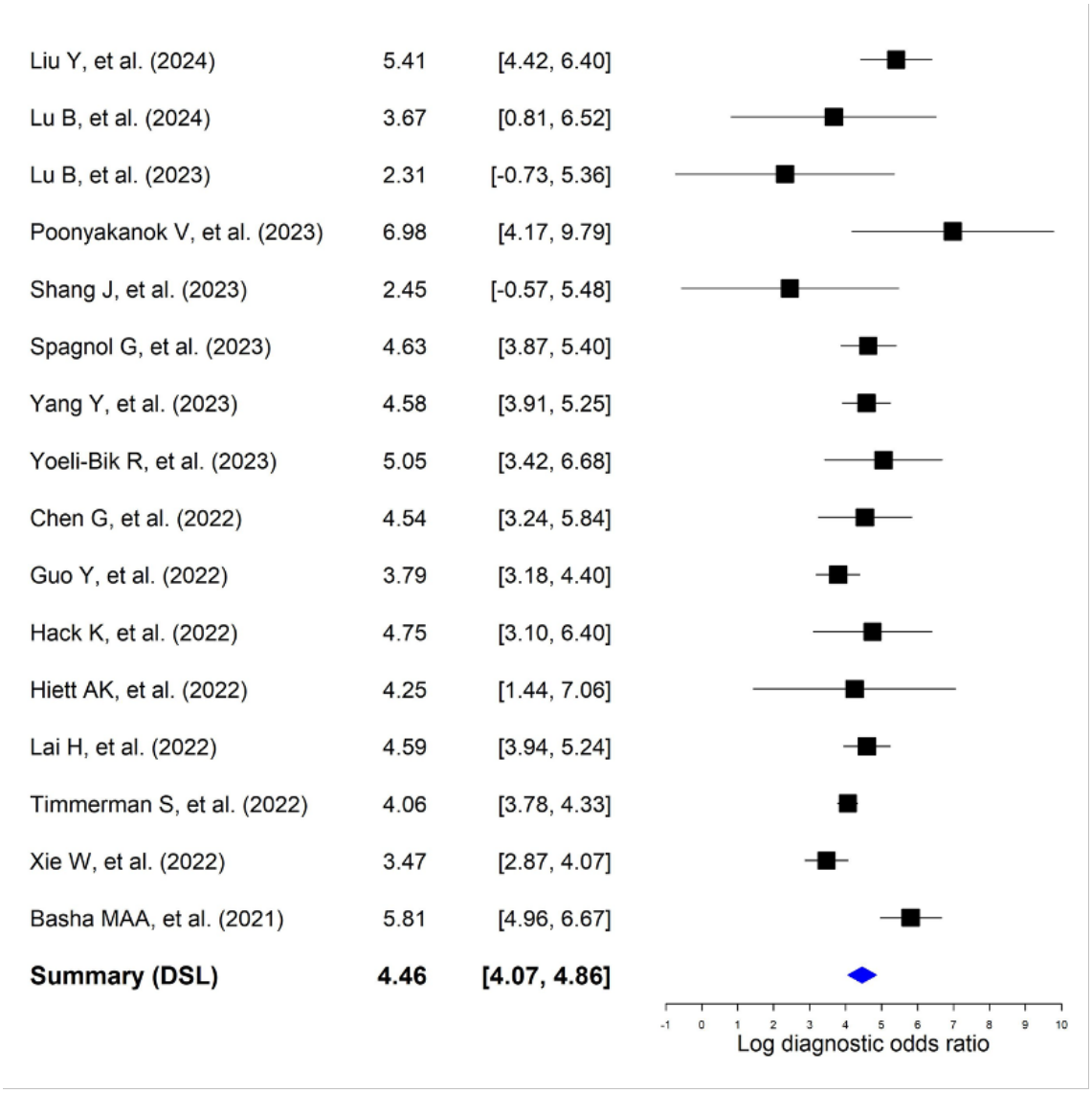
Em resumo, tanto a análise univariada quanto a bivariada apontam para uma performance diagnóstica robusta do O-RADS, com sensibilidade consistentemente elevada e especificidade mais variável.

Figura 2 - *Forest plots* de sensibilidade e especificidade do O-RADS



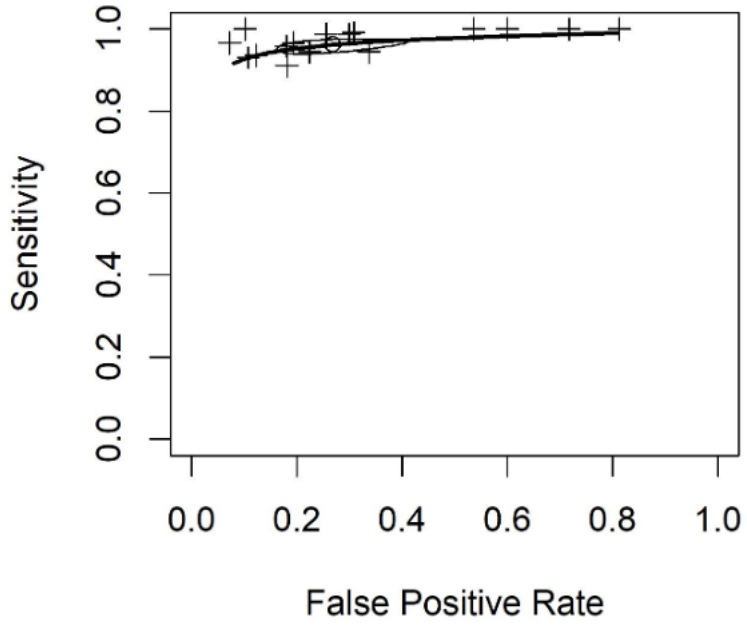
Fonte: Elaboração própria (2025). **Legenda:** Os gráficos exibem as estimativas individuais dos estudos incluídos na metanálise. Os quadrados (■) representam a estimativa pontual da sensibilidade (à esquerda) e da especificidade (à direita), enquanto as linhas horizontais (—) indicam o Intervalo de Confiança de 95% (IC 95%). Os valores numéricos na coluna à direita expressam a proporção estimada seguida pelos limites do intervalo de confiança entre colchetes.

Figura 3 - *Forest plot* de razão de chances diagnóstica (DOR) do O-RADS.



Fonte: Elaboração própria (2025). **Legenda:** O gráfico apresenta as estimativas individuais da Razão de Chances Diagnóstica (*Diagnostic Odds Ratio*) em escala logarítmica. Os quadrados (■) indicam a estimativa pontual de cada estudo e as linhas horizontais (—) o Intervalo de Confiança de 95% (IC 95%). O losango (♦) representa a medida sumarizada (*Summary*) combinada pelo método DerSimonian-Laird. A escala inferior indica o logaritmo da razão de chances e os valores à direita correspondem à estimativa pontual e ao intervalo de confiança entre colchetes.

Figura 4 - Curva SROC do O-RADS obtida a partir do modelo bivariado de Reitsma



Fonte: Elaboração própria (2025). **Legenda:** O gráfico apresenta a Sensibilidade (eixo Y) em função da Taxa de Falso Positivo (eixo X), calculada como *1 – Especificidade*. As cruzes (+) indicam as estimativas de sensibilidade e taxa de falso positivo de cada estudo individual; a linha sólida (—) representa a curva de regressão SROC ajustada pelo modelo bivariado; e o círculo (o) indica o ponto de resumo (*summary point*) das médias conjuntas.

5.3.2. IOTA-SR

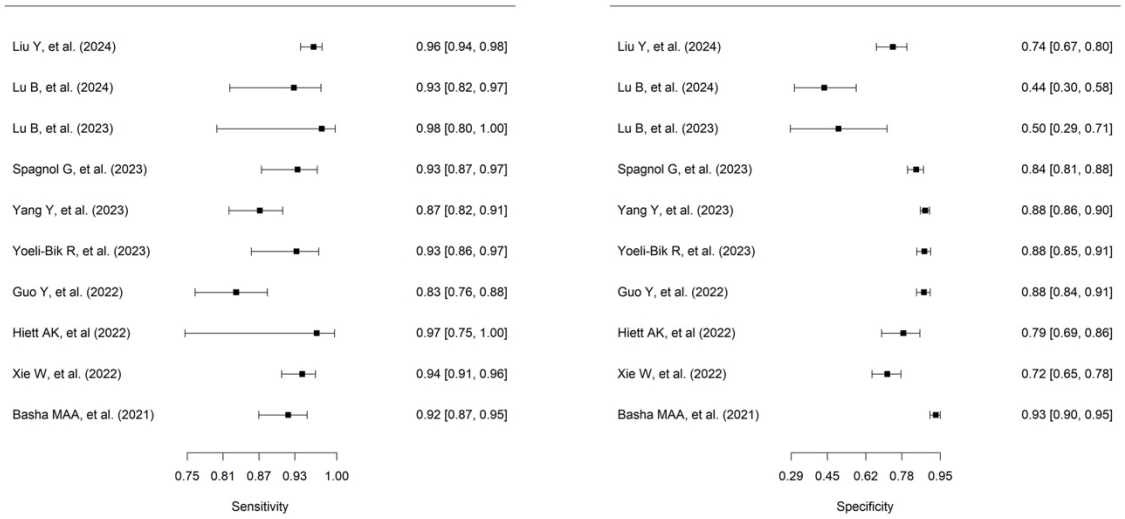
A Figura 5 apresenta os gráficos de *forest plots* de sensibilidade e especificidade do IOTA SR. Observa-se que a sensibilidade foi consistentemente elevada em todos os estudos, em geral acima de 0,90, demonstrando boa capacidade do teste em identificar corretamente os casos positivos. A especificidade, por sua vez, apresentou maior variabilidade, oscilando entre 0,50 e 0,93. Do ponto de vista estatístico, verificou-se heterogeneidade significativa tanto para as sensibilidades (p < 0,001) quanto para as especificidades (p < 0,001).

Na análise univariada, o desempenho diagnóstico global, avaliado pela razão de chances diagnóstica (DOR), foi estimado em 57,4 (IC95%: 38,3–86,0), correspondente a um ln(DOR) de 4,0 (IC95%: 3,65–4,45). A heterogeneidade do ln(DOR) foi baixa (I^2 = 12,3%), achado também confirmado pelo teste de Cochran (p

= 0,330). Esses resultados estão ilustrados na Figura 6. A AUC estimada pelo modelo univariado foi de 0,959 (IC95%: 0,944–0,974), indicando excelente acurácia global.

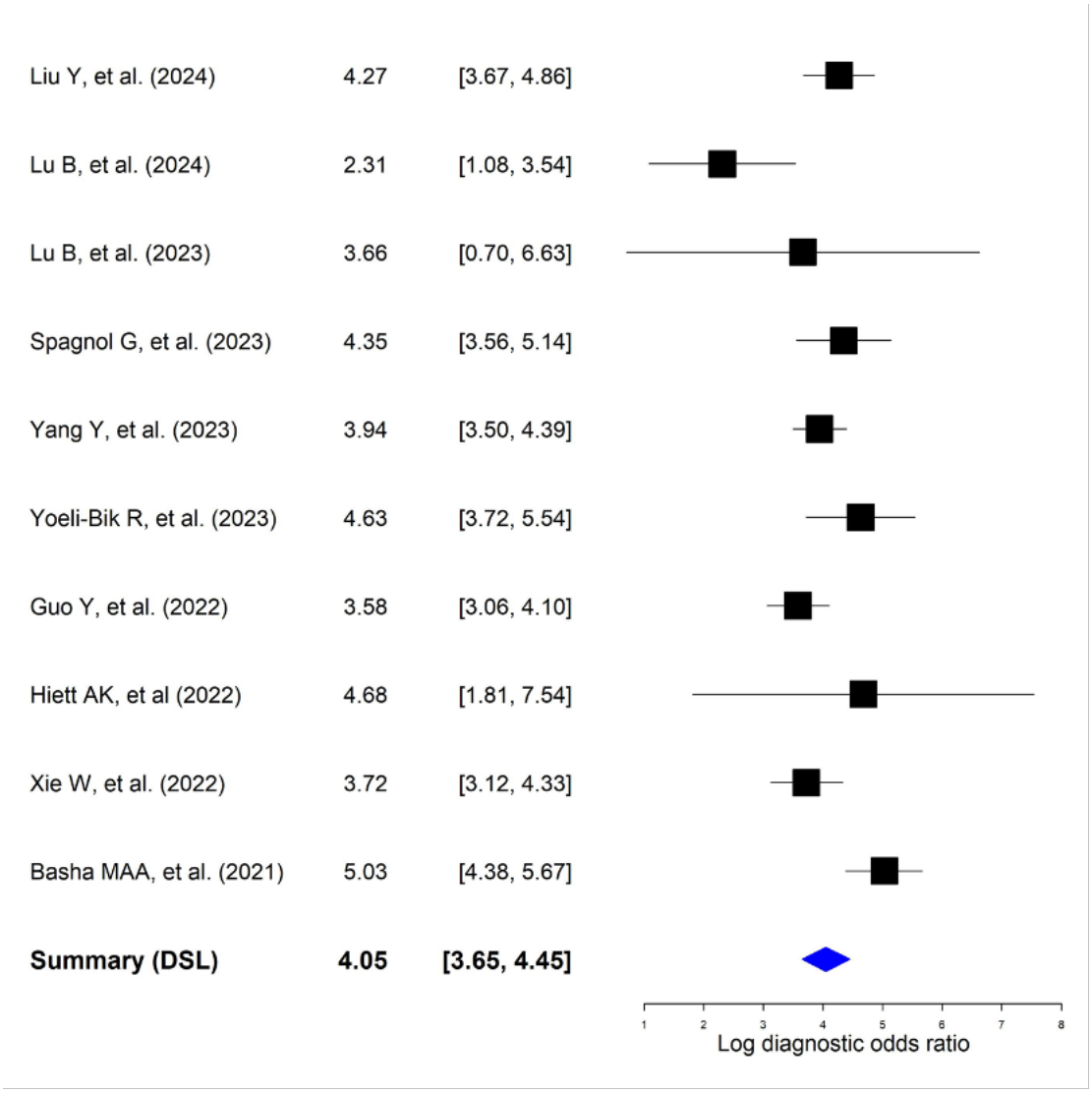
Na análise bivariada, a sensibilidade média foi estimada em 0,928 (IC95%: 0,896–0,951). A taxa de falsos positivos foi de 0,202 (IC95%: 0,128–0,303), o que corresponde a uma especificidade aproximada de 0,798. A AUC bivariada foi de 0,941, levemente inferior à univariada, mas ainda dentro da faixa de excelente desempenho. A heterogeneidade entre os estudos também foi baixa (I^2 de Zhou e Dendukuri = 21,4%). A curva SROC derivada do modelo bivariado, apresentada na Figura 7, confirma esses achados: observa-se uma curva alta (AUC elevada) e os pontos dos estudos próximos entre si.

Figura 5 - *Forest plots* de sensibilidade e especificidade do IOTA-SR



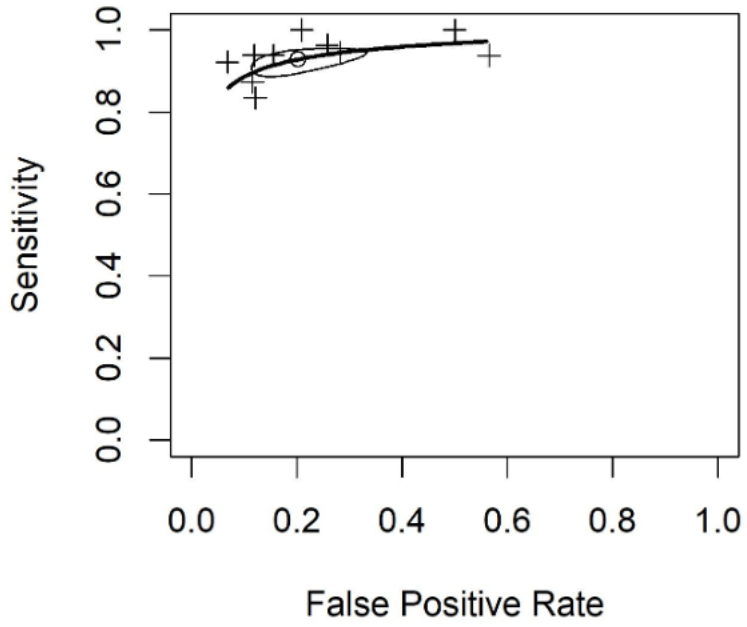
Fonte: Elaboração própria (2025). **Legenda:** Os gráficos exibem as estimativas individuais dos estudos incluídos na metanálise. Os quadrados (■) representam a estimativa pontual da sensibilidade (à esquerda) e da especificidade (à direita), enquanto as linhas horizontais (—) indicam o Intervalo de Confiança de 95% (IC 95%). Os valores numéricos na coluna à direita expressam a proporção estimada seguida pelos limites do intervalo de confiança entre colchetes.

Figura 6 - *Forest plot* de razão de chances diagnóstica (DOR) do IOTA-SR



Fonte: Elaboração própria (2025). **Legenda:** O gráfico apresenta as estimativas individuais da Razão de Chances Diagnóstica (*Diagnostic Odds Ratio*) em escala logarítmica para o modelo IOTA SR. Os quadrados (■) indicam a estimativa pontual de cada estudo e as linhas horizontais (—) o Intervalo de Confiança de 95% (IC 95%). O losango (♦) representa a medida sumarizada (*Summary*) combinada pelo método DerSimonian-Laird. A escala inferior indica o logaritmo da razão de chances e os valores à direita correspondem à estimativa pontual e ao intervalo de confiança entre colchetes.

Figura 7 - Curva SROC do IOTA-SR obtida a partir do modelo bivariado de Reitsma



Fonte: Elaboração própria (2025). **Legenda:** O gráfico apresenta a Sensibilidade (eixo Y) em função da Taxa de Falso Positivo (eixo X), calculada como *1 – Especificidade*. As cruzes (+) indicam as estimativas de sensibilidade e taxa de falso positivo de cada estudo individual; a linha sólida (—) representa a curva de regressão SROC ajustada pelo modelo bivariado; e o círculo (o) indica o ponto de resumo (*summary point*) das médias conjuntas.

5.3.3. IOTA *ADNEX*

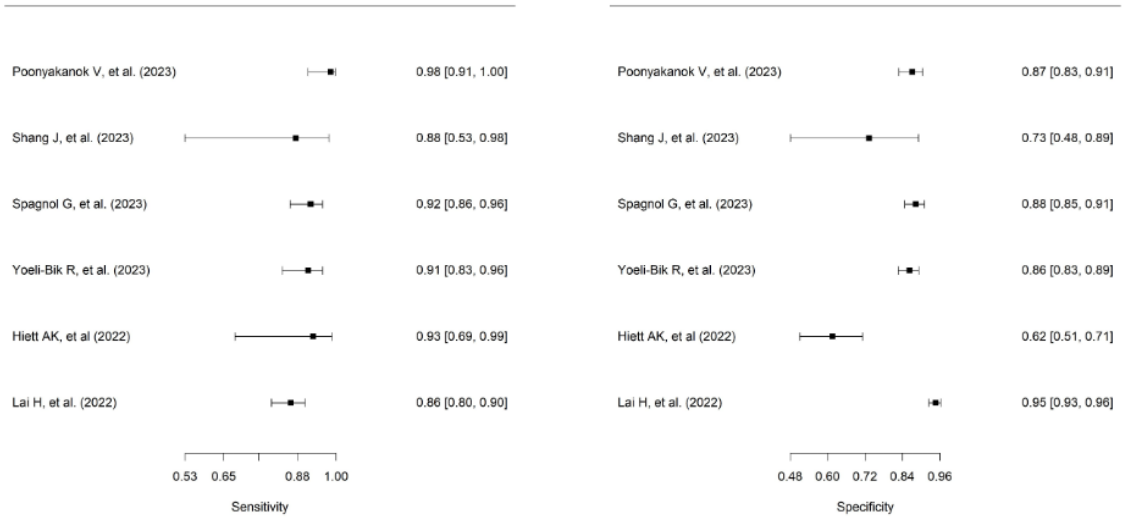
A Figura 8 apresenta os *forest plots* de sensibilidade e especificidade do IOTA *ADNEX*. Observa-se que a sensibilidade foi elevada em todos os estudos, em geral acima de 0,90. A especificidade, por sua vez, apresentou maior variabilidade, oscilando entre 0,62 e 0,95. Do ponto de vista estatístico, não foi observada heterogeneidade significativa entre as sensibilidades (p = 0,105), mas a heterogeneidade entre as especificidades foi significativa (p < 0,001).

Na análise univariada, o desempenho diagnóstico global, avaliado pela razão de chances diagnóstica (DOR), foi estimado em 84,5 (IC95%: 52,1–137,0), correspondente a um ln(DOR) de 4,4 (IC95%: 3,95–4,92). A heterogeneidade do ln(DOR) foi baixa (I^2 = 13,5%), achado também confirmado pelo teste de Cochran (p

= 0,328). Esses resultados estão ilustrados na Figura 9. A AUC estimada pelo modelo univariado foi de 0,966 (IC95%: 0,933–0,999), indicando excelente acurácia global.

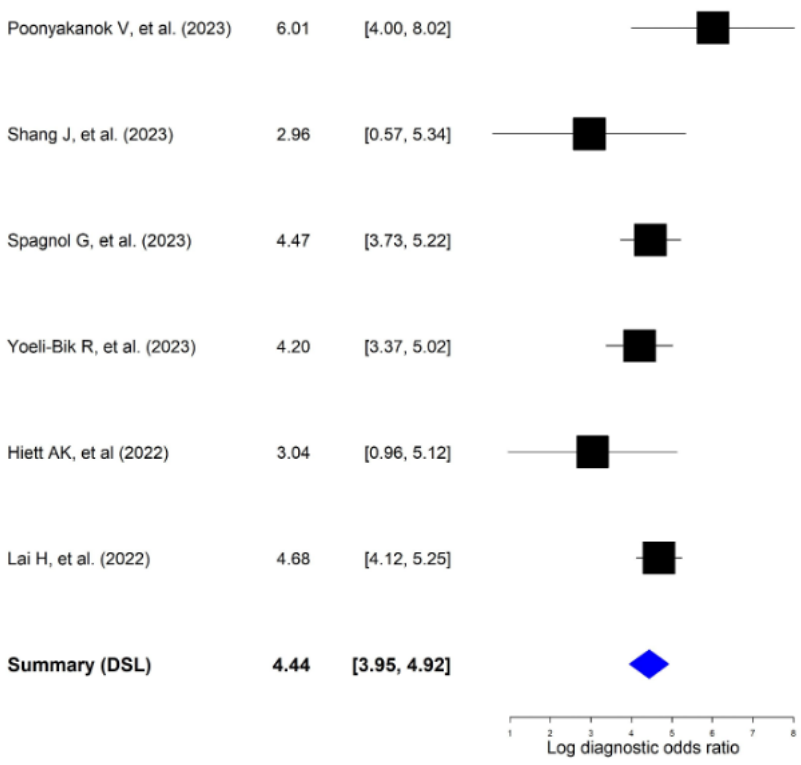
Não foi realizada análise bivariada devido à baixa quantidade de estudos.

Figura 8 - *Forest plots* de sensibilidade e especificidade do IOTA *ADNEX*



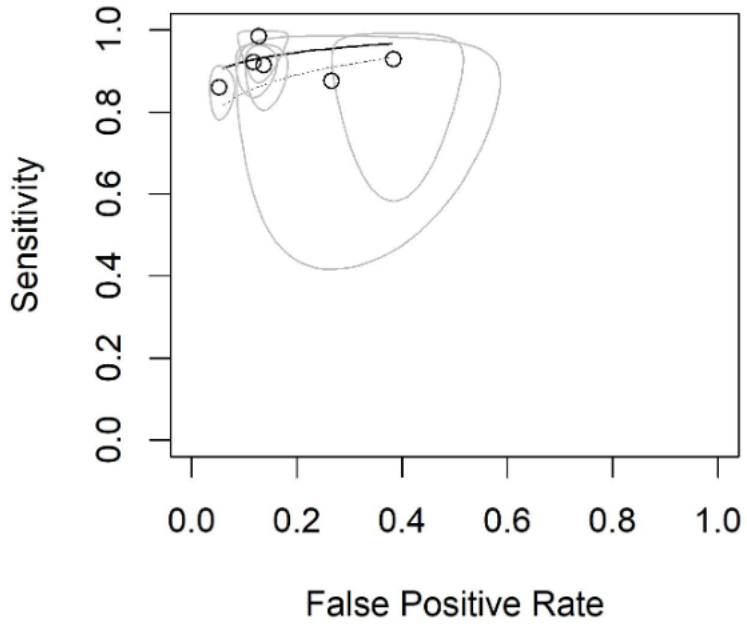
Fonte: Elaboração própria (2025). **Legenda:** Os gráficos exibem as estimativas individuais dos estudos incluídos na metanálise. Os quadrados (■) representam a estimativa pontual da sensibilidade (à esquerda) e da especificidade (à direita), enquanto as linhas horizontais (—) indicam o Intervalo de Confiança de 95% (IC 95%). Os valores numéricos na coluna à direita expressam a proporção estimada seguida pelos limites do intervalo de confiança entre colchetes.

Figura 9 - *Forest plot* de razão de chances diagnóstica (DOR) do IOTA *ADNEX*



Fonte: Elaboração própria (2025). **Legenda:** O gráfico apresenta as estimativas individuais da Razão de Chances Diagnóstica (*Diagnostic Odds Ratio*) em escala logarítmica para o modelo IOTA ADNEX. Os quadrados (■) indicam a estimativa pontual de cada estudo e as linhas horizontais (—) o Intervalo de Confiança de 95% (IC 95%). O losango (♦) representa a medida sumarizada (*Summary*) combinada pelo método DerSimonian-Laird. A escala inferior indica o logaritmo da razão de chances e os valores à direita correspondem à estimativa pontual e ao intervalo de confiança entre colchetes.

Figura 10 - Curva SROC do IOTA *ADNEX* obtida a partir do modelo univariado

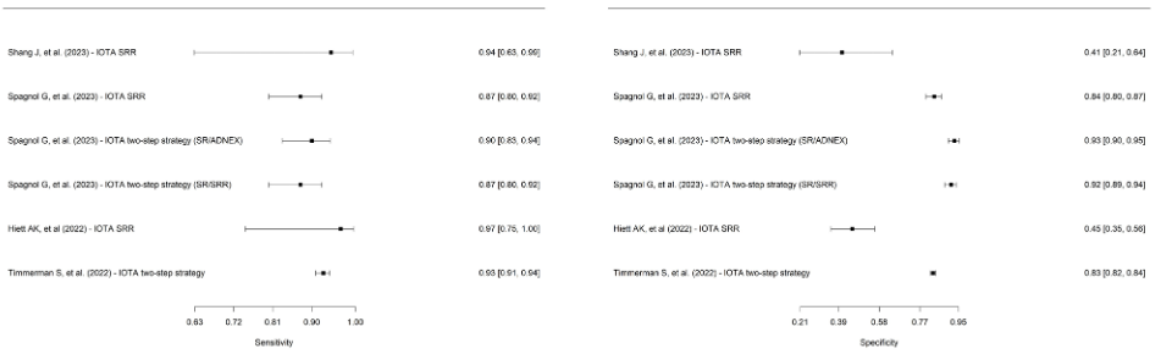


Fonte: Elaboração própria (2025). **Legenda:** O gráfico apresenta a Sensibilidade (eixo Y) em função da Taxa de Falso Positivo (eixo X), calculada como *1 – Especificidade*. As cruzes (+) indicam as estimativas de sensibilidade e taxa de falso positivo de cada estudo individual; a linha sólida (—) representa a curva de regressão SROC ajustada pelo modelo univariado; e o círculo (o) indica o ponto de resumo (*summary point*) das médias conjuntas.

5.3.4. IOTA 2 etapas + IOTA SRR

Os métodos IOTA 2 etapas (SR/*ADNEX* e SR/SRR) e IOTA SRR foram avaliados em apenas três estudos cada, o que impossibilitou a realização de metanálises individuais. Assim, a Figura 11 apresenta os *forest plots* de sensibilidade e especificidade de forma conjunta, como análise descritiva. A sensibilidade do IOTA 2 etapas variou de 0,87 a 0,93, enquanto no IOTA SRR os valores oscilaram entre 0,87 e 0,97. Já a especificidade variou entre 0,83 e 0,93 para o IOTA 2 etapas e entre 0,41 e 0,84 para o IOTA SRR.

Figura 11 - *Forest plots* de sensibilidade e especificidade do IOTA 2 etapas + IOTA SRR

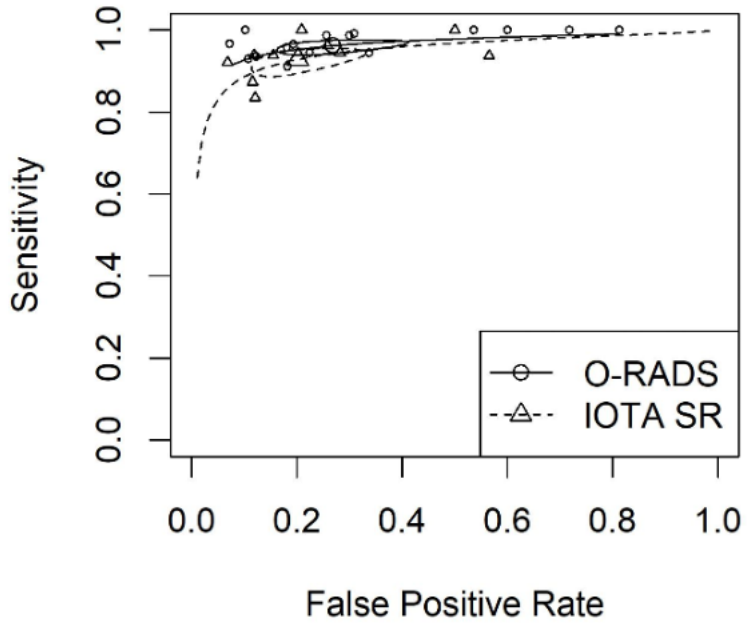


Fonte: Elaboração própria (2025). **Legenda:** Os gráficos exibem as estimativas individuais dos estudos incluídos na metanálise. Os quadrados (■) representam a estimativa pontual da sensibilidade (à esquerda) e da especificidade (à direita), enquanto as linhas horizontais (—) indicam o Intervalo de Confiança de 95% (IC 95%). Os valores numéricos na coluna à direita expressam a proporção estimada seguida pelos limites do intervalo de confiança entre colchetes.

5.3.5. Comparação entre O-RADS e IOTA SR

Os resultados da análise bivariada comparando O-RADS com o IOTA-SR mostram que o O-RADS apresentou sensibilidade significativamente maior (β = 0,652; IC95%: 0,088 - 1,216; *p* = 0,024) em relação ao IOTA-SR. Na prática, isso corresponde a um aumento estimado de cerca de 93% para 97% na sensibilidade média, evidenciando a maior capacidade do O-RADS em identificar corretamente os casos positivos. Por outro lado, não houve diferença estatisticamente significativa na taxa de falsos positivos entre os métodos (β = 0,359; IC95%: -0,419 - +1,137; *p* = 0,366), sugerindo desempenho semelhante em termos de especificidade. A heterogeneidade global foi baixa a moderada (I^2 = 22,4% pelo método de Zhou e Dendukuri), o que demonstra consistência razoável entre os estudos.

Figura 12 - SAUC combinando O-RADS e IOTA SR



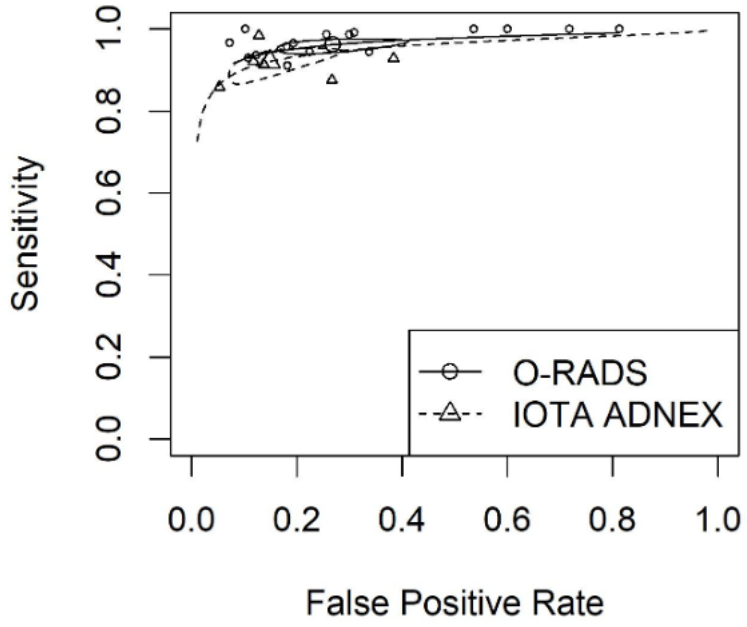
Fonte: Elaboração própria (2025). **Legenda:** O gráfico apresenta as curvas SROC comparativas, com a Sensibilidade no eixo vertical (Y) em função da Taxa de Falso Positivo no eixo horizontal (X; calculada como *1 – Especificidade*). Para o modelo O-RADS, os círculos (o) indicam as estimativas individuais de cada estudo e a linha sólida (—) representa a curva de regressão ajustada. Para o modelo IOTA SR, os triângulos (Δ) indicam as estimativas individuais e a linha tracejada (---) representa a curva de regressão ajustada.

5.3.6. Comparação entre O-RADS e IOTA *ADNEX*

Os resultados da análise bivariada comparando O-RADS com o IOTA *ADNEX* mostram que o O-RADS apresentou sensibilidade significativamente maior (β = 0,839 IC95%: 0,197 - 1,481; *p* = 0,010) em relação ao IOTA *ADNEX*. Na prática, isso corresponde a um aumento estimado de cerca de 91% para 97% na sensibilidade média, evidenciando a maior capacidade do O-RADS em identificar corretamente os casos positivos. Por outro lado, não houve diferença estatisticamente significativa na taxa de falsos positivos entre os métodos (β = 0,708; IC95%: -0,230 - +1,647; *p* = 0,139), sugerindo desempenho semelhante em termos de especificidade. A

heterogeneidade global foi baixa a moderada (I^2 = 28% pelo método de Zhou e Dendukuri), o que demonstra consistência razoável entre os estudos.

Figura 13 - SAUC combinando O-RADS e IOTA *ADNEX*



Fonte: Elaboração própria (2025). **Legenda:** O gráfico apresenta as curvas SROC comparativas, com a Sensibilidade no eixo vertical (Y) em função da Taxa de Falso Positivo no eixo horizontal (X; calculada como *1 – Especificidade*). Para o modelo O-RADS, os círculos (o) indicam as estimativas individuais de cada estudo e a linha sólida (—) representa a curva de regressão ajustada. Para o modelo IOTA *ADNEX*, os triângulos (Δ) indicam as estimativas individuais e a linha tracejada (---) representa a curva de regressão ajustada.

5.4. Risco de Viés

A avaliação do risco de viés foi conduzida segundo o instrumento QUADAS-2 (Whiting *et al.,* 2011) com extensão para o QUADAS-C (Yang *et al.,* 2021), aplicado aos 16 estudos com dados brutos disponíveis, incluídos na síntese quantitativa (metanálise) (Tabela 3).

A maioria dos estudos incluídos apresentou baixo risco de viés em todos os domínios avaliados. O domínio seleção de participantes (D1) mostrou poucas preocupações metodológicas, presentes em apenas dois trabalhos: Xie *et al.* (2022)

e Guo *et al.* (2023). O teste índice (D2) foi consistentemente avaliado como baixo risco em todos os estudos. Em relação ao padrão de referência (D3), houve apenas duas classificações com "*some concerns*": Hiett *et al.* (2022) e Yoeli-Bik *et al*. (2023), indicando pequenas incertezas por falta de dados sobre o cegamento na interpretação do teste padrão ouro. No domínio fluxo e tempo (D4), algumas preocupações foram identificadas em seis estudos: Basha *et al*. (2021); Timmerman S. *et al*. (2022); Hack *et al*. (2022); Chen *et al*, (2022); Guo *et al.* (2022) e Lu *et al.* (2023), relacionadas principalmente à falta de dados de intervalos entre a realização do teste índice e do padrão ouro. No conjunto, os resultados evidenciam que predomina o baixo risco de viés, com preocupações pontuais e não sistemáticas, reforçando a confiabilidade metodológica global da amostra de estudos.

Figura 14 – Risco de viés e aplicabilidade: revisão do julgamento dos autores

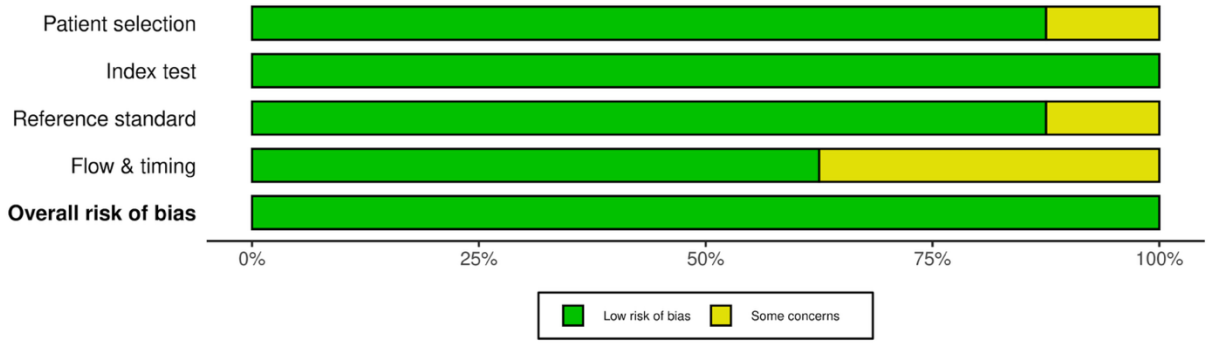


Fonte: Elaboração própria (2025), com base no QUADAS-2 (Whiting *et al.,* 2011)

A avaliação da qualidade com o QUADAS-2 (Whiting *et al.,* 2011) com extensão QUADAS-C (Yang *et al*., 2021) demonstrou um predomínio de baixo risco nos estudos incluídos, conferindo elevada confiança às evidências e indicando alta robustez metodológica. A maioria dos estudos apresentou baixo risco na seleção de pacientes (87,5%), no teste índice (100%) e no padrão de referência (87,5%). O domínio de fluxo e o tempo concentrou as maiores limitações relativas, com 81% classificados como

de baixo risco. Em todos os domínios, nenhum estudo foi classificado como de alto risco e o risco geral de viés foi baixo em 94%, indicando boa qualidade metodológica dos estudos, ainda que recomendações de cautela sejam pertinentes no domínio de fluxo e tempo.

Figura 15 - Gráfico de barras agrupados da avaliação de risco de viés para os estudos incluídos na metanálise



Fonte: Elaboração própria (2025), com base no QUADAS-2 (Whiting *et al.,* 2011)

5.5. Nível de Evidência

A avaliação da certeza da evidência para os desfechos de acurácia diagnóstica foi conduzida pela metodologia GRADE (*Grading of Recommendations Assessment, Development and Evaluation*) (Schünemann *et al*., 2013). Embora a metanálise bivariada comparativa entre O-RADS e IOTA *ADNEX* tenha sido possível, a avaliação da certeza da evidência pelo sistema GRADE foi restrita à comparação entre O-RADS e IOTA-SR. Esta decisão metodológica se baseou no baixo número de estudos disponíveis com dados do IOTA *ADNEX* (n=6). A inclusão dessa análise no GRADE resultaria em uma inevitável redução de nível para 'Baixa' ou 'Muito Baixa' certeza devido à severa imprecisão, o que não agregaria valor à força da evidência principal do estudo.

A metanálise bivariada comparativa do O-RADS e IOTA-SR, base da avaliação GRADE, incluiu um total de 15 estudos (Figura 16) (Tabela 6). É importante ressaltar que, embora 16 estudos tivessem relatado dados completos de acurácia para o O-RADS, o modelo bivariado excluiu automaticamente um estudo. Tal distinção entre o número de estudos qualificados e o número de estudos incluídos na metanálise está totalmente alinhada às diretrizes do Manual Cochrane para Revisões de Acurácia de

Testes Diagnósticos (Deeks *et al.,* 2023), que reconhece que o modelo estatístico pode excluir casos com e sensibilidade ou especificidade perfeitas ou variância não estimável, impedindo a convergência. É fundamental notar que esta exclusão metodológica não alterou as estimativas de efeito nem a direção das conclusões.

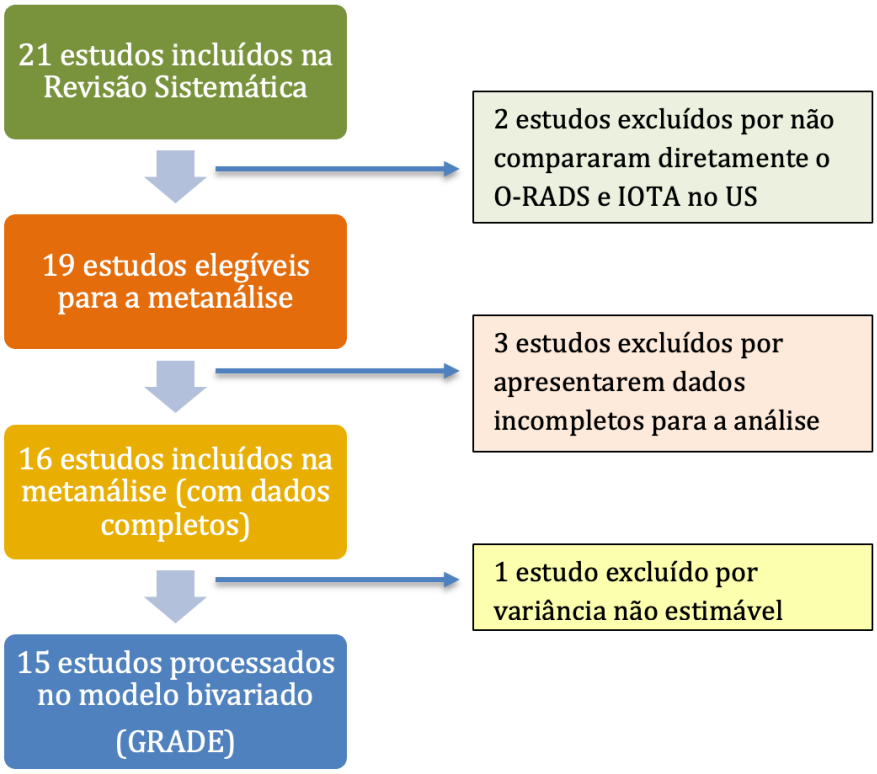
A comparação entre os modelos IOTA-SR e O-RADS revelou um desempenho diagnóstico globalmente elevado para ambos, com sensibilidade de 0,93 (IC95%: 0,88–0,96) e 0,97 (IC95%: 0,94–0,99), respectivamente. Embora a especificidade tenha se mantido similar entre os modelos (0,85 vs. 0,84), o O-RADS demonstrou uma sensibilidade significativamente superior (β = 0,652; IC95%: 0,088–1,216; p = 0,024). Este achado corresponde a um benefício clínico estimado em aproximadamente oito casos malignos adicionais corretamente identificados por 1.000 pacientes avaliadas. A heterogeneidade global observada foi classificada como baixa a moderada (I^2 = 22,4%). A certeza da evidência, conforme o sistema GRADE, foi classificada como moderada para sensibilidade e alta para especificidade, com imprecisão leve associada à amplitude dos intervalos de confiança para a sensibilidade.

Tabela 5 - Nível de evidência pelo GRADE

Desfecho	Nº dos estudos (Nº de pacientes)	Delineamento do estudo	Factors that may decrease certainty of evidence					Effect per 1.000 patients tested		Test accuracy CoE
								Probabilidade pré-teste de 20%		
			Risco de viés	Evidência indireta	Inconsistência	Imprecisão	Viés de publicação	IOTA	O-RADS	
Verdadeiros-positivos (VP) (pacientes corretamente classificados como positivos no US)	15 estudos 1000 pacientes	estudo de acurácia do tipo coorte	não grave	não grave	não grave	grave[a]	nenhum	186 (176 para 192)	194 (188 para 198)	⊕⊕⊕◯ Moderada[a]
								8 menos VP no IOTA		
Falsos-negativos (FN) (pacientes incorretamente classificados como negativos no US)								14 (8 para 24)	6 (2 para 12)	
								8 mais FN no IOTA		
Verdadeiros-negativos (VN) (pacientes corretamente classificados como negativos no US)	15 estudos 800 pacientes	estudo de acurácia do tipo coorte	não grave	não grave	não grave	grave[a]	nenhum	680 (624 para 720)	672 (616 para 712)	⊕⊕⊕◯ Moderada[a]
								8 mais VN no IOTA		
Falsos-positivos (FP) (pacientes incorretamente classificados como positivos no US)								120 (80 para 176)	128 (88 para 184)	
								8 menos FP no IOTA		

Fonte: Elaboração própria (2025), com base no GRADEproGDT (McMaster University; Evidence Prime, 2015). **Legenda:** FN = falsos negativos; FP = falsos positivos; IOTA = International Ovarian Tumor Analysis; O-RADS = Ovarian-Adnexal Reporting and Data System; US = ultrassonografia; VN = verdadeiros negativos; VP = verdadeiros positivos. **Nota:** ([a]) β = 0.652 (0.088-1.216); β = 0.359 (−0.419–1.137).

Figura 16 - Fluxo de inclusão e análise dos estudos na revisão sistemática e na metanálise.



Fonte: Elaboração própria (2025).

Tabela 6 - Interpretação resumida das etapas de seleção, número de estudos e situação final

Etapa	Número de estudos	Situação
Selecionados após triagem	21	Revisão sistemática completa
Elegíveis para a metanálise	19	Base para meta-análise
Utilizados na meta-análise	16	Dados completos (VP, VN, FP, FN)
Efetivamente processados no modelo bivariado (GRADE)	15	Estudos com dados válidos para os modelos

Fonte: Elaboração própria (2025).

6. DISCUSSÃO

O principal resultado desta metanálise foi a demonstração de uma sensibilidade estatisticamente superior do O-RADS quando comparado aos modelos IOTA *Simple Rules* (IOTA-SR) e IOTA *ADNEX* na identificação de massas anexiais malignas. Essa performance foi confirmada na análise bivariada, método robusto para a síntese de estudos de acurácia diagnóstica, que apontou significância estatística tanto na comparação com o IOTA *Simple Rules* (p = 0,024) quanto com o IOTA *ADNEX* (p = 0,010).

Em termos quantitativos, a comparação direta revelou uma sensibilidade média de 97% para o O-RADS frente a 93% do IOTA-SR. Já em relação ao IOTA *ADNEX*, observou-se um comportamento similar, com sensibilidades médias de aproximadamente 97% e 91%, respectivamente. Embora o O-RADS tenha apresentado essa vantagem estatística na discriminação de casos positivos, é fundamental ressaltar que todos os métodos analisados demonstraram altos índices de sensibilidade.

Contudo, no que diz respeito às taxas de falsos positivos, as análises comparativas não evidenciaram distinção estatística significativa do O-RADS em relação ao IOTA-SR (p = 0,366) nem ao IOTA *ADNEX* (p = 0,139), indicando que a especificidade dos modelos permanece comparável.

A validade interna da presente metanálise foi avaliada pela ferramenta QUADAS-2 (Whiting *et al.,* 2011), que indicou que a maior parte da evidência sintetizada (94%) apresenta baixo risco de viés. A principal preocupação se concentrou no domínio de fluxo e tempo, relacionada à falta de informações entre o tempo do teste índice e do teste de referência. No entanto, a aplicação do QUADAS-C (*comparative*) (Yang *et al.,* 2021) assegurou que o risco de viés na comparação direta entre o O-RADS e os modelos IOTA fosse considerado baixo na maioria dos casos, o que fortalece a confiabilidade dos achados de superioridade estatística e sugere que a diferença observada não é majoritariamente atribuível a falhas na condução comparativa dos estudos.

Na prática clínica ginecológica, a sensibilidade ligeiramente superior do O-RADS se traduz em uma menor probabilidade de que uma massa anexial maligna seja erroneamente classificada como benigna.

A interpretação desses achados, no entanto, deve ser feita em consonância com a literatura recente. Duas metanálises publicadas em 2024 e 2025 por Han, Wen e Hu (Han; Wen; Hu, 2024) e Almeida, Bort e Alcázar (Almeida; Bort; Alcázar, 2025), respectivamente, abordaram a mesma questão comparativa. Em ambas, a conclusão foi de que o O-RADS possui um desempenho diagnóstico comparável tanto ao IOTA-SR quanto ao modelo *ADNEX*, sem diferenças estatisticamente significativas.

Almeida, Bort e Alcázar (2025) reportaram uma sensibilidade agrupada de 94% para o O-RADS e 91% para o IOTA-SR (p = 0,2204), e Han, Wen e Hu (2024) encontraram 95% e 91% (p = 0,21), respectivamente. Esses resultados contrastam com o achado do presente estudo, que identificou uma superioridade estatisticamente significativa na sensibilidade do O-RADS (p = 0,024). O mesmo padrão se repete na comparação com o modelo *ADNEX*: o estudo de Almeida, Bort e Alcázar (2025) apontou sensibilidades de 95% para o O-RADS e 91% para o ADNEX (p = 0,2307), e Han, Wen e Hu (2024) reportaram 96% e 94% (p = 0,57), respectivamente, o que também se opõe diretamente ao nosso achado de maior sensibilidade do O-RADS, com significância estatística (p = 0,010).

Uma possível explicação para a divergência entre os achados reside na composição dos estudos incluídos, no tamanho amostral e na sensibilidade inerente ao teste estatístico. As metanálises anteriores, que não encontraram diferença significativa, revelaram p-valores que sugerem uma pequena magnitude de efeito (p ≈ 0,21–0,23). Em cenários de efeito reduzido, a inclusão ou exclusão de estudos específicos pode alterar as estimativas ponderadas o suficiente para cruzar o limiar de significância estatística.

Uma informação relevante nesse aspecto é a exclusão dos estudos de Pelayo et al. (2023a, 2023b) da presente análise — incluídos nos trabalhos de Han, Wen e Hu (2024) e Almeida, Bort e Alcázar (2025) —, devido à indisponibilidade de dados brutos, mesmo após tentativa de contato com os autores. No entanto, esta revisão incorporou um conjunto de trabalhos ligeiramente mais amplo e recente (n=10 vs. n=7 e n=6 para IOTA SR, por exemplo), o que pode ter conferido maior poder estatístico para detectar uma diferença real entre os métodos.

É plausível que a significância estatística aqui encontrada possa não representar uma magnitude de efeito clinicamente substancial do O-RADS, sendo em parte influenciada pela amostra específica. Contudo, é igualmente possível que o

resultado represente uma atualização do panorama, sugerindo que a tendência de melhor desempenho diagnóstico do O-RADS se consolida à medida que o corpo de evidências se expande.

Ainda que as metanálises anteriores não tenham encontrado significância estatística, é notável que os valores de sensibilidade agrupada para o O-RADS foram numericamente superiores em ambas, alinhando-se com a direção do efeito observado em nosso estudo.

A alta acurácia global de todos os modelos avaliados, evidenciada pelas elevadas áreas sob a curva ROC em nossa análise e nas demais, consolida o papel central desses sistemas na prática clínica.

No que tange à especificidade, os achados deste estudo estão em total concordância com a literatura. A ausência de uma diferença estatisticamente significativa na taxa de falsos positivos entre o O-RADS e o IOTA-SR é um resultado consistente com as conclusões de Almeida, Bort e Alcázar (2025) (p = 0,2204) e Han, Wen e Hu (2024) (p = 0,12). Adicionalmente, ao analisarmos o modelo *ADNEX* (p = 0,139), observamos que também não foi encontrada diferença estatisticamente significativa nos outros estudos (Almeida, Bort e Alcázar (2025) (p = 0,2307) e Han, Wen e Hu (2024) (p = 0,87)). Essa consistência, contudo, reforça a limitação inerente à especificidade, que, embora aceitável, pode levar a uma taxa não desprezível de resultados falso-positivos.

Essa questão é particularmente proeminente em lesões de risco intermediário, como as classificadas como O-RADS 4, onde o risco de malignidade pode variar amplamente (10% a 50%). Na prática clínica, essa incerteza diagnóstica frequentemente leva à indicação de procedimentos invasivos. No entanto, essa incerteza difere substancialmente do resultado inconclusivo do IOTA *Simple Rules*. O O-RADS 4, apesar de sua amplitude, delimita o risco em uma faixa conhecida, permitindo ao especialista um ponto de partida para a conduta e integrando a paciente ao algoritmo de gestão. Já o inconclusivo do IOTA-SR não fornece nenhum dado numérico, exigindo uma reavaliação completa para definir o risco inicial.

Uma abordagem promissora para superar essa limitação da especificidade é a utilização da ultrassonografia com contraste (CEUS) (Lu *et al*., 2023; Lu *et al*., 2024; Wang *et al*, 2023). Estudos recentes têm demonstrado consistentemente que a análise direta e detalhada da perfusão vascular em tempo real nos componentes

sólidos das massas anexiais pode aprimorar significativamente a acurácia diagnóstica. Ainda que o *Doppler* colorido, já utilizado pelos modelos, ofereça dados importantes sobre a vascularização da lesão, a avaliação é mais limitada e indireta do fluxo em vasos menores. Para a oncologia, onde a análise da neoangiogênese é fundamental, o CEUS contribui fornecendo informações mais ricas e precisas.

Em um estudo focado em massas ovarianas com aparência sólida, Lu et al. (2023) observaram que o CEUS melhorou a acurácia diagnóstica das lesões classificadas como O-RADS 4 de 47,4% para 87.5%. De forma ainda mais robusta, Wang et al. (2023) propuseram um sistema que combina O-RADS com CEUS, o qual elevou a especificidade de 81,3% para 92,3% e a acurácia geral de 84,9% para 92,8%. Corroborando esses achados, Lu et al. (2024) concluíram que a combinação aumentou a Área sob a Curva (AUC) de 0,641 para 0,871.

As implicações clínicas desses achados são diretas: o ultrassom com contraste (CEUS) emerge como uma ferramenta valiosa no algoritmo de avaliação de massas anexiais, capaz de reclassificar uma parcela significativa de lesões benignas, evitando, assim, cirurgias desnecessárias. Wang et al. (2023) sugerem que a eficácia diagnóstica dessa combinação é comparável à de sistemas baseados em ressonância nuclear magnética (RM). Contudo, é essencial reconhecer que, na prática clínica diária, a implementação do CEUS enfrenta barreiras como o custo do contraste, o tempo adicional de sala e a necessidade de treinamento específico. Tais fatores, somados ao fato de que o sistema de estratificação de risco O-RADS US já recomenda explicitamente a avaliação por RM em categorias de risco intermediário a alto (O-RADS US 4 e 5), ainda levam à priorização da Ressonância Magnética (RM) como o método de eleição na segunda linha da avaliação.

A RM, que possui protocolos bem estabelecidos (O-RADS por Ressonância Magnética - O-RADS MRI) (Reinhold, C*. et al*, 2021) (Sadowski, E. A. et al, 2022) e oferece uma visão panorâmica e tecidual superior, demonstra um desempenho robusto. A metanálise sobre o O-RADS na RM por Rizzo et al. (2023) comprovou sua alta capacidade discriminatória para lesões indeterminadas, reportando uma sensibilidade sumária de 92% (IC 95%: 88-95) e, crucialmente, uma especificidade sumária de 91% (IC 95%: 89-93) para o diagnóstico de malignidade (corte ≥ O-RADS MRI 4). Essa elevada especificidade é o diferencial necessário para resolver a incerteza da categoria O-RADS US 4.

Apesar da superioridade discriminatória da RM em relação ao ultrassom (US) devido à sua especificidade mais alta (Zhang; Dai; Li, 2023), o US é mantido como o método de primeira linha por ser amplamente acessível, de baixo custo e por utilizar apenas ondas sonoras, sem o uso de radiação ionizante (RGOG, 2011) (Carvalho; Moretti-Marques; Silva Filho, 2020; Pelayo *et al*., 2023a).

No entanto, diante de uma possível ambiguidade diagnóstica ultrassonográfica, a RM atua como um filtro poderoso, sendo o método de segunda linha preferencial para a avaliação das classificações indeterminadas (O-RADS US 4) e de alto risco do ultrassom, permitindo ao especialista uma decisão mais segura e otimizada sobre o manejo cirúrgico.

Para além da acurácia na estratificação de risco, uma vantagem pragmática e distintiva do sistema O-RADS é a sua capacidade de fornecer recomendações de conduta claras e padronizadas para cada categoria de risco (Andreotti *et al.,* 2018, 2020). Enquanto os modelos do grupo IOTA, como o *Simple Rules* (Timmerman, D. *et al.,* 2008) e o *ADNEX*, são ferramentas excepcionais para a estimativa da probabilidade de malignidade, deixando a decisão sobre o manejo subsequente inteiramente a critério da equipe clínica, o O-RADS integra ao laudo ultrassonográfico um algoritmo de gestão bem definido, sugerindo desde o acompanhamento expectante para lesões de baixo risco (O-RADS 3) até o encaminhamento mandatório para um especialista em ginecologia oncológica para as lesões de alto risco (O-RADS 5) (Andreotti *et al.,* 2018, 2020; Strachowski *et al*., 2023).

Esse componente de gestão possui um valor na prática clínica diária: ele promove a padronização do cuidado entre diferentes instituições e níveis de experiência profissional, oferecendo um suporte de decisão crucial, sobretudo para o ginecologista geral ou o ultrassonografista menos familiarizado com a complexidade das massas anexiais. Ao reduzir a ambiguidade e a variabilidade na conduta pós-diagnóstico, o O-RADS transcende seu papel como uma ferramenta de classificação para se tornar um guia clínico mais completo, fornecendo uma proposta de laudo bem estruturado e recomendação de conduta, o que representa uma evolução significativa na abordagem e no manejo das massas anexiais.

Embora a introdução de sistemas de estratificação como o IOTA e O-RADS represente um marco na ultrassonografia ginecológica, a implementação universal na prática clínica diária ainda enfrenta barreiras significativas (Pelayo *et al*., 2023a).

Diferentemente de outros métodos de imagem, como a RM, em que a aquisição da imagem é padronizada por técnicos e a interpretação é realizada sobre imagens fixas e registradas, a ultrassonografia é um exame dinâmico, em que a aquisição e a interpretação ocorrem simultaneamente pelo médico executante, em tempo real.

A avaliação subjetiva por um ultrassonografista experiente é frequentemente citada como o método mais preciso para a discriminação de massas anexiais (Andreotti *et al*., 2020; Liu *et al.*, Sokalska *et al.*, 2009). Contudo, essa expertise não está universalmente disponível, e foi precisamente para mitigar essa lacuna que os sistemas padronizados foram desenvolvidos como ferramentas educativas para encurtar a curva de aprendizado. O paradoxo reside no fato de que esses mesmos sistemas são, por vezes, percebidos como complexos. Essa percepção é alimentada pela necessidade de interpretar características ultrassonográficas sofisticadas de forma consistente e precisa, o que pode ser desafiador para examinadores menos experientes.

Apesar de o O-RADS possuir uma estrutura de classificação mais complexa se comparada à simplicidade e praticidade de aplicação do IOTA *Simple Rules*, essa diferença representa uma escolha de prioridade diagnóstica. O sistema opta por sacrificar parte da facilidade de uso em prol da maior sensibilidade, com uma clara ênfase na segurança oncológica ao minimizar os falsos negativos.

Diante do desafio envolvendo essa percepção de complexidade, a solução, no entanto, não reside na simplificação excessiva dos métodos, o que poderia comprometer sua acurácia, mas sim no investimento em programas de capacitação e treinamento, que podem ser objetivamente mensurados pela concordância interobservador, avaliada, em alguns estudos, pelo coeficiente Kappa (κ) (Guo *et al*., 2022; Lai *et al.,* 2022; Liu *et al.,* 2024; Yang *et al.,* 2023).

Valores de Kappa elevados demonstram que, após a devida instrução, profissionais diferentes chegam à mesma conclusão ao avaliar a mesma lesão. A literatura recente fornece evidências contundentes nesse sentido: Yang et al. (2023) encontraram uma concordância 'muito boa' entre um ultrassonografista menos experiente e um especialista na aplicação do O-RADS e IOTA SR, com valores de Kappa de 0,926 e 0,903, respectivamente. Da mesma forma, Liu et al. (2024) reportaram concordância substancial para O-RADS v2022 (κ = 0,933) e IOTA SR (κ = 0,923). Lai et al. (2022) demonstraram uma 'boa' concordância (κ > 0.83) para o O-

RADS entre sonografistas júniores que receberam treinamento específico para o estudo, e Guo et al. (2022), ao comparar diretamente médicos seniores e júniores, concluíram que os sistemas de diagnóstico puderam compensar a inexperiência, obtendo uma concordância 'moderada' a 'alta' (κ de 0,71 para O-RADS e 0,77 para IOTA SR).

Coletivamente, esses dados formam um argumento de que a complexidade percebida dos sistemas é uma barreira superável e que, até mesmo avaliadores inexperientes podem alcançar uma boa confiabilidade após a devida instrução. Isso sugere que, embora exista uma necessidade de familiarização, os sistemas são projetados para serem aprendidos e aplicados de forma reprodutível. Logo, o sucesso da implementação dos modelos depende criticamente do desenvolvimento e da disseminação de programas de educação estruturados, capazes de garantir a aplicação consistente e confiável dessas valiosas ferramentas diagnósticas.

No entanto, é fundamental reconhecer que a fronteira da pesquisa está avançando para além dos sistemas baseados em regras ou critérios definidos pelo examinador. Nesta era de rápida transformação digital, onde a Inteligência Artificial (IA) vem sendo utilizada como uma ferramenta facilitadora e estratégica em diversas áreas do conhecimento, o próprio grupo IOTA está na vanguarda dessa evolução no campo da ginecologia, por meio de iniciativas de IA e *Deep Learning* (Aprendizado Profundo).

Com a hipótese central de que a IA e novos modelos translacionais apoiarão o diagnóstico mais precoce de malignidade e o tratamento ideal em centros oncológicos especializados, pesquisas, formalizadas em projetos multicêntricos (KU Leuven, 2024–2029), incluem o módulo "IOTA-AI", focado especificamente na contribuição da IA para a avaliação ultrassonográfica de tumores ovarianos. Esta abordagem representa um salto metodológico, pois o *Deep Learning* é capaz de aprender diretamente com os dados, analisando imagens ultrassonográficas completas, e extrair características preditivas de forma autônoma, eliminando a necessidade de o examinador definir a priori as regras ou critérios manuais.

Modelos baseados em IA/DL prometem mitigar a variabilidade interobservador e, ao integrar a precisão da máquina, potencialmente elevar ainda mais a acurácia diagnóstica do ultrassom, sugerindo que o futuro da área reside na fusão dos sistemas

padronizados atuais com a objetividade e o poder de processamento da inteligência artificial.

Contudo, o sucesso da implementação tanto dos modelos padronizados atuais quanto das futuras ferramentas depende diretamente da validação e da força da evidência científica que suportam suas conclusões. Neste sentido, a pesquisa com IA, embora promissora, ainda é um projeto relativamente de longo prazo (com conclusão prevista para 2029), o que significa que as decisões clínicas atuais dependem da melhor evidência que pode ser extraída dos dados já publicados, e é neste contexto que essa revisão sistemática com metanálise busca contribuir.

Desta forma, para determinar o nível de certeza atual dessa evidência, a força dos achados foi avaliada pela metodologia GRADE, revelando certeza alta para a especificidade, corroborando a estabilidade desse achado e a concordância com a literatura, e moderada para a superioridade na sensibilidade.

É crucial notar que o rebaixamento pontual na sensibilidade se deve exclusivamente à imprecisão dos intervalos de confiança. Essa imprecisão reflete uma heterogeneidade residual classificada como baixa a moderada ($I^2 = 22,4\%$), atribuível a diferenças inevitáveis no delineamento dos estudos primários e à heterogeneidade clínica das populações.

Contudo, tal variabilidade não refuta a significância estatística encontrada ($p = 0,024$); ela indica que, embora a superioridade do O-RADS seja uma evidência confiável, a magnitude exata desse benefício deverá ser refinada por estudos multicêntricos internacionais prospectivos, a despeito do que já foi realizado pelo grupo IOTA, confirmando sua validação externa.

Em suma, a evidência atual respalda a preferência pelo O-RADS quando o objetivo é a otimização da sensibilidade. Ressalva-se, contudo, que este sistema ainda carece da mesma maturidade de validação prospectiva multicêntrica já consolidada pelo IOTA-SR (Timmerman, D. *et al*, 2016). Assim, esta metanálise documenta a vantagem estatística em sensibilidade do O-RADS no nível terciário, ao passo que delineia a necessidade de grandes coortes futuras para confirmar se essa acurácia se sustenta de forma homogênea em diferentes cenários clínicos.

Apesar do rigor metodológico adotado nesta revisão sistemática, algumas limitações devem ser consideradas. O viés de publicação não pode ser descartado, visto que resultados negativos tendem a ser subnotificados, o que pode superestimar

levemente a acurácia. Além disso, a predominância de estudos retrospectivos (Tabela 2) é um fator restritivo: a análise baseada em imagens estáticas ou laudos prévios impede a avaliação dinâmica em tempo real, essencial na ultrassonografia, podendo introduzir vieses de informação.

Há também o viés de verificação diferencial; a inclusão do seguimento clínico como referência para casos não operados, embora evite excluir benignos óbvios, gera assimetria: casos suspeitos têm confirmação histopatológica, enquanto os aparentemente benignos dependem apenas da imagem. Isso pode superestimar a sensibilidade caso malignidades indolentes não sejam detectadas durante o acompanhamento.

Adicionalmente, a indisponibilidade de dados brutos resultou na exclusão de três estudos elegíveis para a metanálise, devido à ausência de resposta dos autores. Somado a isso, o reduzido número de estudos avaliando o IOTA *ADNEX* (n=6), o IOTA SRR (n=3) e a estratégia de 2 etapas do IOTA (n=3) restringiu as análises e inviabilizou a modelagem bivariada para estas comparações e a aplicação do sistema GRADE, limitando a robustez das conclusões sobre tais estratégias.

Em última análise, as restrições metodológicas apontadas não ofuscam a principal mensagem deste trabalho: independentemente da escolha entre o O-RADS ou os modelos IOTA-SR e IOTA ADNEX, a presente metanálise corrobora que a avaliação padronizada de massas anexiais é, hoje, uma prática baseada em evidências de alta qualidade. A vantagem estatística observada no O-RADS, portanto, posiciona-o não apenas como um competidor à altura dos modelos IOTA, mas como uma estratégia refinada para minimizar falsos negativos. O estudo contribui, assim, para a sedimentação do O-RADS na prática clínica, oferecendo suporte quantitativo para sua utilização segura em centros de referência.

7. CONCLUSÃO

A presente revisão sistemática com metanálise demonstrou que tanto o sistema O-RADS quanto os modelos IOTA (*Simple Rules* e *ADNEX*) são ferramentas diagnósticas robustas para a avaliação de massas anexiais, apresentando elevado desempenho global evidenciado por altas Áreas Sob a Curva (AUC) e Razões de Chances Diagnósticas (DOR).

Entretanto, na análise comparativa, o O-RADS exibiu sensibilidade estatisticamente superior aos modelos IOTA analisados (*Simple Rules* e *ADNEX*), resultando em uma menor taxa de falsos negativos na detecção de malignidade, com especificidade estatisticamente equivalente entre os métodos.

Conclui-se, portanto, que o O-RADS permite otimizar a detecção de casos malignos, mantendo uma acurácia global comparável à dos modelos IOTA.

8. REFERÊNCIAS

ALMEIDA, G.; BORT, M.; ALCÁZAR, J. L. Comparison of the diagnostic performance of Ovarian Adnexal Reporting Data System (O-RADS) with IOTA Simple Rules and ADNEX model for classifying adnexal masses: a head-to-head meta-analysis. **Journal of Clinical Ultrasound**, v. 53, n. 7, p. 1574-1583, set. 2025. DOI: 10.1002/jcu.24048. Disponível em: https://doi.org/10.1002/jcu.24048.

AMERICAN COLLEGE OF OBSTETRICIANS AND GYNECOLOGISTS. Committee on Practice Bulletins—Gynecology. Practice Bulletin No. 174: evaluation and management of adnexal masses. **Obstetrics & Gynecology**, v. 128, n. 5, p. e210-e226, nov. 2016. DOI: 10.1097/AOG.0000000000001768. Disponível em: https://doi.org/10.1097/AOG.0000000000001768. Acesso em: 06 junho. 2025.

AMERICAN COLLEGE OF RADIOLOGY. **O-RADS™ Ultrasound.** [S. l.], 2022. Disponível em: https://www.acr.org/Clinical-Resources/Clinical-Tools-and-Reference/Reporting-and-Data-Systems/O-RADS/Ultrasound. Acesso em: 09 dez. 2025.

ANDREOTTI, R. F. *et al*. Ovarian-Adnexal Reporting Lexicon for Ultrasound: a white paper of the ACR Ovarian-Adnexal Reporting and Data System Committee. **Journal of the American College of Radiology**, v. 15, n. 10, p. 1415-1429, out. 2018. DOI: 10.1016/j.jacr.2018.07.004. Acesso em: 7 dez. 2025.

ANDREOTTI, R. F. *et al.* O-RADS US Risk Stratification and Management System: a consensus guideline from the ACR Ovarian-Adnexal Reporting and Data System Committee. **Radiology**, v. 294, n. 1, p. 168-185, jan. 2020. DOI: 10.1148/radiol.2019191150. Acesso em: 7 dez. 2025.

ARMSTRONG, D. K. *et al*. Ovarian Cancer, Version 2.2020, NCCN Clinical Practice Guidelines in Oncology. **Journal of the National Comprehensive Cancer Network**, v. 19, n. 2, p. 191-226, fev. 2021. DOI: 10.6004/jnccn.2021.0007. Acesso em: 7 dez. 2025.

BARRA, D. A.; JORGE, N. G.; CONDÉ, E. F. Massas anexiais: descrição e interpretação ultrassonográfica por IOTA. **Femina**, v. 49, n. 1, p. 6-11, 2021.

BASHA, M. A. A. *et al.* Comparison of O-RADS, GI-RADS, and IOTA simple rules regarding malignancy rate, validity, and reliability for diagnosis of adnexal masses. **European Radiology**, v. 31, n. 2, p. 674-684, fev. 2021. DOI: 10.1007/s00330-020-07143-7. Acesso em: 7 dez. 2025.

CARVALHO, J. P.; MORETTI-MARQUES, R.; SILVA FILHO, A. L. da. Adnexal mass: diagnosis and management. **Revista Brasileira de Ginecologia e Obstetrícia**, v. 42, n. 7, p. 438-443, jul. 2020. DOI: 10.1055/s-0040-1715547. Disponível em: https://doi.org/10.1055/s-0040-1715547.

CHEN, G. Y. *et al*. Comparison of the O-RADS and ADNEX models regarding malignancy rate and validity in evaluating adnexal lesions. **European Radiology**, v. 32, n. 11, p. 7854-7864, nov. 2022. DOI: 10.1007/s00330-022-08803-6. Acesso em: 7 dez. 2025.

DEEKS, J. J. et al. (ed.). **Cochrane handbook for systematic reviews of diagnostic test accuracy.** Version 2.0. London: Cochrane, 2023. Disponível em: https://training.cochrane.org/handbook-diagnostic-test-accuracy.

DOEBLER, P. **mada: meta-analysis of diagnostic accuracy**. R package. [S. l.]: CRAN, 2017. Disponível em: https://CRAN.R-project.org/package=mada.

FILIZ, A. A.; KAHYAOGLU, S.; ATALAY, C. R. Comparison of International Ovarian Tumor Analysis ADNEX model and Ovarian-Adnexal Reporting and Data System with final histological diagnosis in adnexal masses: a retrospective study. **Obstetrics & Gynecology Science**, v. 67, n. 1, p. 86-93, jan. 2024. DOI: 10.5468/ogs.23061. Disponível em: https://doi.org/10.5468/ogs.23061.

GUO, Y. et al. A comparison of the diagnostic performance of the O-RADS, RMI4, IOTA LR2, and IOTA SR systems by senior and junior doctors. **Ultrasonography**, v. 41, n. 3, p. 511-518, jul. 2022. DOI: 10.14366/usg.21237. Disponível em: https://doi.org/10.14366/usg.21237.

HACK, K. et al. External validation of O-RADS US risk stratification and management system. **Radiology**, v. 304, n. 1, p. 114-120, jul. 2022. DOI: 10.1148/radiol.211868. Disponível em: https://doi.org/10.1148/radiol.211868.

HAN, J.; WEN, J.; HU, W. Comparison of O-RADS with the ADNEX model and IOTA SR for risk stratification of adnexal lesions: a systematic review and meta-analysis. **Frontiers in Oncology**, v. 14, art. 1354837, 2 maio 2024. DOI: 10.3389/fonc.2024.1354837. Disponível em: https://doi.org/10.3389/fonc.2024.1354837.

HIETT, A. K. et al. Performance of IOTA Simple Rules, Simple Rules risk assessment, ADNEX model and O-RADS in differentiating between benign and malignant adnexal lesions in North American women. **Ultrasound in Obstetrics & Gynecology**, v. 59, n. 5, p. 668-676, maio 2022. DOI: 10.1002/uog.24777. Disponível em: https://doi.org/10.1002/uog.24777.

INSTITUTO NACIONAL DE CÂNCER (Brasil). Câncer de ovário. Rio de Janeiro: INCA, 2025. Disponível em: https://www.gov.br/inca/pt-br/assuntos/cancer/tipos/ovario. Acesso em: 06 dez. 2025.

JIN, C. et al. The predictive value of nomogram for adnexal cystic-solid masses based on O-RADS US, clinical and laboratory indicators. **BMC Medical Imaging**, v. 24, n. 1, art. 315, 18 nov. 2024. DOI: 10.1186/s12880-024-01497-w. Disponível em: https://doi.org/10.1186/s12880-024-01497-w.

KU LEUVEN. **Optimising gynaecological tumour diagnosis and management based on AI and translational studies: the IOTA (International Ovarian Tumour Analysis) and other multicentre projects**. Leuven, 2024–2029. Disponível em: https://research.kuleuven.be/portal/en/project/3M240331. Acesso em: 08 dez.,2025.

LAI, Hong wei *et al.* Comparison of O-RADS, GI-RADS, and ADNEX for Diagnosis of Adnexal Masses: An External Validation Study Conducted by Junior Sonologists. **Journal of ultrasound in medicine: official journal of the American Institute of Ultrasound in Medicine**, v. 41, n. 6, p. 1497–1507, 1 jun. 2022.

LANDOLFO, C. et al. Benign descriptors and ADNEX in two-step strategy to estimate risk of malignancy in ovarian tumors: retrospective validation in IOTA5 multicenter cohort. **Ultrasound in Obstetrics & Gynecology**, v. 61, n. 2, p. 231-242, fev. 2023. DOI: 10.1002/uog.26080. Disponível em: https://doi.org/10.1002/uog.26080.

LEE, S. et al. O-RADS US: a systematic review and meta-analysis of category-specific malignancy rates. **Radiology**, v. 308, n. 2, art. e223269, ago. 2023. DOI: 10.1148/radiol.223269. Disponível em: https://doi.org/10.1148/radiol.223269.

LI, Y. et al. Evaluation of American College of Radiology Ovarian-Adnexal Reporting and Data System ultrasound to predict malignancy risk in adnexal lesions. **Journal of Obstetrics and Gynaecology Research**, v. 50, n. 2, p. 225-232, fev. 2024. DOI: 10.1111/jog.15831. Disponível em: https://doi.org/10.1111/jog.15831.

LIU, Y. et al. Diagnostic accuracy of ultrasound classifications – O-RADS US v2022, O-RADS US v2020, and IOTA SR – in distinguishing benign and malignant adnexal masses: enhanced by combining O-RADS US v2022 with tumor marker HE4. **European Journal of Radiology**, v. 181, art. 111824, dez. 2024. DOI: 10.1016/j.ejrad.2024.111824. Disponível em: https://doi.org/10.1016/j.ejrad.2024.111824.

LU, B. et al. Comparison of contrast-enhanced ultrasound, IOTA simple rules and O-RADS for assessing the malignant risk of sonographically appearing solid ovarian masses. Journal of Gynecology Obstetrics and Human Reproduction, v. 52, n. 4, p. 102564, abr. 2023. DOI: 10.1016/j.jogoh.2023.102564. Acesso em: 14 dez. 2025.

LU, B. et al. Differentiating benign from malignant ovarian masses with solid components: diagnostic performance of CEUS combined with IOTA Simple Rules and O-RADS. **Ultrasound in Medicine & Biology**, v. 50, n. 9, p. 1449-1458, set. 2024. DOI: 10.1016/j.ultrasmedbio.2024.05.024. Disponível em: https://doi.org/10.1016/j.ultrasmedbio.2024.05.024

MCINNES, M. D. F. et al. Preferred reporting items for a systematic review and meta-analysis of diagnostic test accuracy studies: the PRISMA-DTA statement. **JAMA**, v. 319, n. 4, p. 388-396, 23 jan. 2018. DOI: 10.1001/jama.2017.19163. Disponível em: https://doi.org/10.1001/jama.2017.19163.

MCMASTER UNIVERSITY; EVIDENCE PRIME. **GRADEpro GDT**: GRADEpro guideline development tool. [S. l.]: McMaster University; Evidence Prime, 2015. Disponível em: https://www.gradepro.org.

NATIONAL CANCER INSTITUTE. Cancer stat facts: ovarian cancer. Bethesda: National Cancer Institute. Disponível em: https://seer.cancer.gov/statfacts/html/ovary.html. Acesso em: 06 dez. 2025.

PAGE, M. J. et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. **BMJ**, v. 372, art. n71, 29 mar. 2021. DOI: 10.1136/bmj.n71. Disponível em: https://doi.org/10.1136/bmj.n71.

PELAYO, M. et al. Comparison of ultrasound scores in differentiating between benign and malignant adnexal masses. **Diagnostics**, v. 13, n. 7, art. 1307, 30 mar. 2023. DOI: 10.3390/diagnostics13071307. Disponível em: https://doi.org/10.3390/diagnostics13071307.

PELAYO, M. et al. Ultrasound features and ultrasound scores in the differentiation between benign and malignant adnexal masses. **Diagnostics**, v. 13, n. 13, art. 2152, 23 jun. 2023. DOI: 10.3390/diagnostics13132152. Disponível em: https://doi.org/10.3390/diagnostics13132152.

PELAYO, M. et al. Application of ultrasound scores (Subjective Assessment, Simple Rules Risk Assessment, ADNEX model, O-RADS) to adnexal masses of difficult classification. **Diagnostics**, v. 13, n. 17, art. 2785, 28 ago. 2023. DOI: 10.3390/diagnostics13172785. Disponível em: https://doi.org/10.3390/diagnostics13172785.

POONYAKANOK, V.; TANMAHASAMUT, P.; JAISHUEN, A. Prospective comparative trial comparing O-RADS, IOTA ADNEX model, and RMI score for preoperative evaluation of adnexal masses for prediction of ovarian cancer. **Journal of Obstetrics and Gynaecology Research**, v. 49, n. 5, p. 1412-1417, maio 2023. DOI: 10.1111/jog.15624. Disponível em: https://doi.org/10.1111/jog.15624.

R CORE TEAM. **R: a language and environment for statistical computing**. Vienna: R Foundation for Statistical Computing, 2024. Disponível em: https://www.R-project.org/.

REINHOLD, C. et al. Ovarian-Adnexal Reporting Lexicon for MRI: a white paper of the ACR Ovarian-Adnexal Reporting and Data Systems MRI Committee. **Journal of the American College of Radiology**, [S. l.], v. 18, n. 5, p. 713-729, maio 2021. DOI 10.1016/j.jacr.2020.12.022. Disponível em: https://doi.org/10.1016/j.jacr.2020.12.022. Acesso em: 09 dez. 2025.

REITSMA, J. B. et al. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. **Journal of Clinical Epidemiology**, v. 58, n. 10, p. 982-990, out. 2005. DOI: 10.1016/j.jclinepi.2005.02.022

RIZZO, S.; COZZI, A.; DOLCIAMI, M.; DEL GRANDE, F.; SCARANO, A. L.; PAPADIA, A.; GUI, B.; GANDOLFO, N.; CATALANO, C.; MANGANARO, L. O-RADS MRI: a systematic review and meta-analysis of diagnostic performance and category-wise malignancy rates. **Radiology**, v. 307, n. 1, e220795, abr. 2023. DOI: 10.1148/radiol.220795.

ROYAL COLLEGE OF OBSTETRICIANS AND GYNAECOLOGISTS. Management of suspected ovarian masses in premenopausal women. Green-top Guideline No. 62. London: RCOG, 2011. Disponível em: https://www.rcog.org.uk/guidance/browse-all-guidance/green-top-guidelines/management-of-suspected-ovarian-masses-in-premenopausal-women-green-top-guideline-no-62/. Acesso em: 06 jun. 2025.

SADOWSKI, E. A. et al. O-RADS MRI Risk Stratification System: guide for assessing adnexal lesions from the ACR O-RADS Committee. **Radiology**, [S. l.], v. 303, n. 1, p. 35-47, abr. 2022. DOI 10.1148/radiol.204371. Disponível em: https://doi.org/10.1148/radiol.204371. Acesso em: 09 dez. 2025.

SAYASNEH, A. et al. Evaluating the risk of ovarian cancer before surgery using the ADNEX model: a multicentre external validation study. **British Journal of Cancer**, v. 115, n. 5, p. 542-548, 23 ago. 2016. DOI: 10.1038/bjc.2016.227. Disponível em: https://doi.org/10.1038/bjc.2016.227. Acesso em: 06 junho. 2025.

SCHÜNEMANN, H. *et al.* (ed.). **GRADE handbook for grading quality of evidence and strength of recommendations**. Updated October 2013. [S. l.]: The GRADE Working Group, 2013. Disponível em: https://gdt.gradepro.org/app/handbook/handbook.html. Acesso em: 10 out. 2025.

SHANG, J. et al. Comparison of performance between O-RADS, IOTA simple rules risk assessment and ADNEX model in the discrimination of ovarian Brenner tumors. **Archives of Gynecology and Obstetrics**, v. 308, n. 3, p. 961-970, set. 2023. DOI: 10.1007/s00404-022-06903-8. Disponível em: https://doi.org/10.1007/s00404-022-06903-8.

SOKALSKA, A. et al. Diagnostic accuracy of transvaginal ultrasound examination for assigning a specific diagnosis to adnexal masses. **Ultrasound in Obstetrics & Gynecology,** v. 34, n. 4, p. 462-470, out. 2009. DOI: 10.1002/uog.6444. Disponível em: https://doi.org/10.1002/uog.6444.

SPAGNOL, G. et al. Simple rules, O-RADS, ADNEX and SRR model: single oncologic center validation of diagnostic predictive models alone and combined (two-step strategy) to estimate the risk of malignancy in adnexal masses and ovarian tumors. **Gynecologic Oncology,** v. 177, p. 109-116, out. 2023. DOI: 10.1016/j.ygyno.2023.08.012. Disponível em: https://doi.org/10.1016/j.ygyno.2023.08.012.

STRACHOWSKI, L. M. et al. O-RADS US v2022: an update from the American College of Radiology's Ovarian-Adnexal Reporting and Data System US Committee. **Radiology**, v. 308, n. 3, art. e230685, set. 2023. DOI: 10.1148/radiol.230685. Disponível em: https://doi.org/10.1148/radiol.230685.

TIMMERMAN, D. et al. Terms, definitions and measurements to describe the sonographic features of adnexal tumors: a consensus opinion from the International Ovarian Tumor Analysis (IOTA) group. **Ultrasound in Obstetrics & Gynecology**, v. 16, n. 4, p. 500-505, out. 2000. DOI: 10.1046/j.1469-0705.2000.00287.x.

TIMMERMAN, D. et al. Logistic regression model to distinguish between the benign and malignant adnexal mass before surgery: a multicenter study by the International Ovarian Tumor Analysis Group. **Journal of Clinical Oncology**, v. 23, n. 34, p. 8794-8801, 1 dez. 2005. DOI: 10.1200/JCO.2005.01.7632. Disponível em: https://doi.org/10.1200/JCO.2005.01.7632. Acesso em: 06 jun. 2025.

TIMMERMAN, D. et al. Inclusion of CA-125 does not improve mathematical models developed to distinguish between benign and malignant adnexal tumors**. Journal of Clinical Oncology**, v. 25, n. 27, p. 4194-4200, 20 set. 2007. DOI: 10.1200/JCO.2006.09.5943. Disponível em: https://doi.org/10.1200/JCO.2006.09.5943. Acesso em: 06 jun. 2025.

TIMMERMAN, D. et al. Simple ultrasound-based rules for the diagnosis of ovarian cancer. **Ultrasound in Obstetrics & Gynecology,** v. 31, n. 6, p. 681-690, jun. 2008. DOI: 10.1002/uog.5365. Disponível em: https://doi.org/10.1002/uog.5365. Acesso em: 06 jun. 2025.

TIMMERMAN, D. et al. Ovarian cancer prediction in adnexal masses using ultrasound-based logistic regression models: a temporal and external validation study by the IOTA group. **Ultrasound in Obstetrics & Gynecology**, v. 36, n. 2, p. 226-234, ago. 2010. DOI: 10.1002/uog.7636. Disponível em: https://doi.org/10.1002/uog.7636. Acesso em: 06 jun. 2025.

TIMMERMAN, D. et al. Predicting the risk of malignancy in adnexal masses based on the Simple Rules from the International Ovarian Tumor Analysis group. **American Journal of Obstetrics and Gynecology**, v. 214, n. 4, p. 424-437, abr. 2016. DOI: 10.1016/j.ajog.2016.01.007. Disponível em: https://doi.org/10.1016/j.ajog.2016.01.007. Acesso em: 06 jun. 2025.

TIMMERMAN, S. et al. External validation of the Ovarian-Adnexal Reporting and Data System (O-RADS) lexicon and the International Ovarian Tumor Analysis 2-step strategy to stratify ovarian tumors into O-RADS risk groups. **JAMA Oncology**, v. 9, n. 2, p. 225-233, 1 fev. 2023. DOI: 10.1001/jamaoncol.2022.5969. Disponível em: https://doi.org/10.1001/jamaoncol.2022.5969.

US PREVENTIVE SERVICES TASK FORCE. **Screening for ovarian cancer**: US Preventive Services Task Force recommendation statement. Rockville: USPSTF, 2018. Disponível em: https://www.uspreventiveservicestaskforce.org/uspstf/recommendation/ovarian-cancer-screening. Acesso em: 25 jun. 2025.

VAN CALSTER, B. et al. Evaluating the risk of ovarian cancer before surgery using the ADNEX model to differentiate between benign, borderline, early and advanced stage invasive, and secondary metastatic tumours: prospective multicentre diagnostic study. **BMJ**, v. 349, art. g5920, 15 out. 2014. DOI: 10.1136/bmj.g5920. Disponível em: https://doi.org/10.1136/bmj.g5920. Acesso em: 06 jun. 2025.

VAN CALSTER, B. et al. Practical guidance for applying the ADNEX model from the IOTA group to discriminate between different subtypes of adnexal tumors. **Facts, Views & Vision in ObGyn**, v. 7, n. 1, p. 32-41, 2015. Disponível em: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4402441. Acesso em: 06 jun. 2025.

VARA, J. et al. O-RADS classification for ultrasound assessment of adnexal masses: agreement between IOTA lexicon and ADNEX model for assigning risk group. **Diagnostics**, v. 13, n. 4, art. 673, 10 fev. 2023. DOI: 10.3390/diagnostics13040673. Disponível em: https://doi.org/10.3390/diagnostics13040673.

WANG, T. et al. Comparative study of the efficacy of the Ovarian-Adnexa Reporting and Data System Ultrasound combined with contrast-enhanced ultrasound and the ADNEX MR scoring system in the diagnosis of adnexal masses. **Ultrasound in Medicine & Biology**, v. 49, n. 9, p. 2072-2080, set. 2023. DOI: 10.1016/j.ultrasmedbio.2023.05.012. Disponível em: https://doi.org/10.1016/j.ultrasmedbio.2023.05.012.

WHITING, P. F. et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. **Annals of Internal Medicine**, v. 155, n. 8, p. 529-536, 18 out. 2011. DOI: 10.7326/0003-4819-155-8-201110180-00009. Disponível em: https://doi.org/10.7326/0003-4819-155-8-201110180-00009.

XIE, W. T. et al. Efficacy of IOTA simple rules, O-RADS, and CA125 to distinguish benign and malignant adnexal masses. **Journal of Ovarian Research**, v. 15, n. 1, art. 15, 23 jan. 2022. DOI: 10.1186/s13048-022-00947-9. Disponível em: https://doi.org/10.1186/s13048-022-00947-9.

XIE, W. et al. Developing a deep learning model for predicting ovarian cancer in Ovarian-Adnexal Reporting and Data System Ultrasound (O-RADS US) Category 4 lesions: a multicenter study. **Journal of Cancer Research and Clinical Oncology**, v. 150, n. 7, art. 346, 9 jul. 2024. DOI: 10.1007/s00432-024-05872-6. Disponível em: https://doi.org/10.1007/s00432-024-05872-6.

YANG, B. et al. QUADAS-C: a tool for assessing risk of bias in comparative diagnostic accuracy studies. **Annals of Internal Medicine**, v. 174, n. 11, p. 1592-1599, nov. 2021. DOI: 10.7326/M21-2234. Disponível em: https://doi.org/10.7326/M21-2234.

YANG, Y.; JU, H.; HUANG, Y. Diagnostic performance of IOTA SR and O-RADS combined with CA125, HE4, and risk of malignancy algorithm to distinguish benign and malignant adnexal masses. **European Journal of Radiology**, v. 165, art. 110926, ago. 2023. DOI: 10.1016/j.ejrad.2023.110926. Disponível em: https://doi.org/10.1016/j.ejrad.2023.110926.

YOELI-BIK, R. et al. Diagnostic performance of ultrasonography-based risk models in differentiating between benign and malignant ovarian tumors in a US cohort. **JAMA Network Open**, v. 6, n. 7, art. e2323289, 3 jul. 2023. DOI: 10.1001/jamanetworkopen.2023.23289. Disponível em: https://doi.org/10.1001/jamanetworkopen.2023.23289.

ZHANG, Q.; DAI, X.; LI, W. Systematic review and meta-analysis of O-RADS ultrasound and O-RADS MRI for risk assessment of ovarian and adnexal lesions. **American Journal of Roentgenology**, [S. l.], v. 221, n. 1, p. 21-33, jul. 2023. DOI 10.2214/AJR.22.28396. Disponível em: https://doi.org/10.2214/AJR.22.28396. Acesso em: 09 dez. 2025.

ZHOU, X. H.; DENDUKURI, N. Confidence intervals for the summary ROC curve. **Biostatistics**, v. 11, n. 2, p. 365–377, 2010.

9. APÊNDICES

Apêndice 1 - *Check List* PRISMA

Section and Topic	Item #	Checklist item	Location where item is reported
TITLE			
Title	1	Identify the report as a systematic review.	Page 1
ABSTRACT			
Abstract	2	See the PRISMA 2020 for Abstracts checklist.	Page 8
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of existing knowledge.	Page 16
Objectives	4	Provide an explicit statement of the objective(s) or question(s) the review addresses.	Page 27
METHODS			
Eligibility criteria	5	Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses.	Page 29
Information sources	6	Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted.	Page 30
Search strategy	7	Present the full search strategies for all databases, registers and websites, including any filters and limits used.	Page 31
Selection process	8	Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process.	Page 31
Data collection process	9	Specify the methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study investigators, and if applicable, details of automation tools used in the process.	Page 32
Data items	10a	List and define all outcomes for which data were sought. Specify whether all results that were compatible with each outcome domain in each study were sought (e.g. for all measures, time points, analyses), and if not, the methods used to decide which results to collect.	Page 29
	10b	List and define all other variables for which data were sought (e.g. participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information.	Page 29
Study risk of bias assessment	11	Specify the methods used to assess risk of bias in the included studies, including details of the tool(s) used, how many reviewers assessed each study and whether they worked independently, and if applicable, details of automation tools used in the process.	Page 32
Effect measures	12	Specify for each outcome the effect measure(s) (e.g. risk ratio, mean difference) used in the synthesis or presentation of results.	Page 29
Synthesis methods	13a	Describe the processes used to decide which studies were eligible for each synthesis (e.g. tabulating the study intervention characteristics and comparing against the planned groups for each synthesis (item #5)).	Page 31
	13b	Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions.	Page 32
	13c	Describe any methods used to tabulate or visually display results of individual studies and syntheses.	Page 32
	13d	Describe any methods used to synthesize results and provide a rationale for the choice(s). If meta-analysis was performed, describe the model(s),	Page 33

Section and Topic	Item #	Checklist item	Location where item is reported
		method(s) to identify the presence and extent of statistical heterogeneity, and software package(s) used.	
	13e	Describe any methods used to explore possible causes of heterogeneity among study results (e.g. subgroup analysis, meta-regression).	Page 33
	13f	Describe any sensitivity analyses conducted to assess robustness of the synthesized results.	Page 33
Reporting bias assessment	14	Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases).	Page 32
Certainty assessment	15	Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome.	Page 33
RESULTS			
Study selection	16a	Describe the results of the search and selection process, from the number of records identified in the search to the number of studies included in the review, ideally using a flow diagram.	Page 34
	16b	Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded.	Page 34
Study characteristics	17	Cite each included study and present its characteristics.	Page 35
Risk of bias in studies	18	Present assessments of risk of bias for each included study.	Page 34
Results of individual studies	19	For all outcomes, present, for each study: (a) summary statistics for each group (where appropriate) and (b) an effect estimate and its precision (e.g. confidence/credible interval), ideally using structured tables or plots.	Page 39
Results of syntheses	20a	For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies.	Page 39
	20b	Present results of all statistical syntheses conducted. If meta-analysis was done, present for each the summary estimate and its precision (e.g. confidence/credible interval) and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect.	Page 39
	20c	Present results of all investigations of possible causes of heterogeneity among study results.	Page 39
	20d	Present results of all sensitivity analyses conducted to assess the robustness of the synthesized results.	Page 39
Reporting biases	21	Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed.	Page 34
Certainty of evidence	22	Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed.	Page 55
DISCUSSION			
Discussion	23a	Provide a general interpretation of the results in the context of other evidence.	Page 59
	23b	Discuss any limitations of the evidence included in the review.	Page 67
	23c	Discuss any limitations of the review processes used.	Page 67
	23d	Discuss implications of the results for practice, policy, and future research.	Page 66-67
OTHER INFORMATION			
Registration and protocol	24a	Provide registration information for the review, including register name and registration number, or state that the review was not registered.	Page 29
	24b	Indicate where the review protocol can be accessed, or state that a protocol was not prepared.	Page 29
	24c	Describe and explain any amendments to information provided at registration or in the protocol.	Page 29

Section and Topic	Item #	Checklist item	Location where item is reported
Support	25	Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review.	Page 6
Competing interests	26	Declare any competing interests of review authors.	
Availability of data, code and other materials	27	Report which of the following are publicly available and where they can be found: template data collection forms; data extracted from included studies; data used for all analyses; analytic code; any other materials used in the review.	

Fonte: Elaboração própria (2025), adaptado de Page *et al.* (2020).

Apêndice 2 – Artigos excluídos pelos critérios de elegibilidade e o motivo

	Autor (ano)	Motivo da exclusão
01.	Xie *et al.* (2024)	Modelo de aprendizado
02.	Jin *et al.* (2024)	Modelo de monograma
03.	Vara *et al.* (2023)	Compara O-RADS com léxico IOTA
04.	Pelayo *et al*. (2023c)	Considera a amostra de estudo anterior.

Fonte: Elaboração própria (2025)

10. ANEXOS

Anexo 1 – Tabela de lesões benignas clássicas (Classic benign lesions)

O-RADS™ US v2022 — Classic Benign Lesions

Release Date: November 2022

Lesion	Descriptors and Definitions For any atypical features on initial or follow-up exam, use other lexicon descriptors (eg, unilocular, multilocular, solid, etc.)	Management If sonographic features are only suggestive, and overall assessment is uncertain, consider follow-up US within 3 months
Typical Hemorrhagic Cyst	Unilocular cyst, **no internal vascularity***, and at least one of the following: • Reticular pattern (fine, thin intersecting lines representing fibrin strands) • Retractile clot (intracystic component with straight, concave, or angular margins)	Imaging: o Premenopausal: • ≤5 cm: None • >5 cm but <10 cm: Follow-up US in 2–3 months o Early postmenopausal (<5 years): • <10 cm, options to confirm include: ▪ Follow-up US in 2–3 months *or* ▪ US specialist (if available) *or* ▪ MRI (with O–RADS MRI score) o Late postmenopausal (≥5 years): • Should not occur; recategorize using other lexicon descriptors. Clinical: Gynecologist**
Typical Dermoid Cyst	Cystic lesion with ≤3 locules, **no internal vascularity***, and at least one of the following: • Hyperechoic component(s) (diffuse or regional) with shadowing • Hyperechoic lines and dots • Floating echogenic spherical structures	Imaging: o ≤3 cm: May consider follow-up US in 12 months*** o >3 cm but <10 cm: If not surgically excised, follow-up US in 12 months*** Clinical: Gynecologist**
Typical Endometrioma	Cystic lesion with ≤3 locules, **no internal vascularity***, homogeneous low–level/ground glass echoes, and smooth inner walls/septation(s) • ± Peripheral punctate echogenic foci in wall	Imaging: o Premenopausal: • <10 cm: If not surgically excised, follow-up US in 12 months*** o Postmenopausal: • <10 cm and initial exam, options to confirm include ▪ Follow–up US in 2–3 months *or* ▪ US specialist (if available) *or* ▪ MRI (with O-RADS MRI score) Then, if not surgically excised, recommend follow-up US in 12 months*** Clinical: Gynecologist**
Typical Paraovarian Cyst	Simple cyst separate from the ovary	Imaging: None Clinical: Gynecologist**
Typical Peritoneal Inclusion Cyst	Fluid collection with ovary at margin or suspended within that conforms to adjacent pelvic organs • ± Septations (representing adhesions)	Imaging: None Clinical: Gynecologist**
Typical Hydrosalpinx	Anechoic, fluid–filled tubular structure • ± Incomplete septation(s) (representing folds) • ± Endosalpingeal folds (short, round projections around inner walls)	

*Excludes vascularity in walls or intervening septation(s)

**As needed for management of clinical issues

***There is a paucity of evidence for defining the need, optimal duration or interval of timing for surveillance. If stable, consider US follow-up at 24 months from initial exam, then as clinically indicated. Specifically, evidence does support **an increased risk of malignancy in endometriomas following menopause and those present greater than 10 years.**

© 2023 American College of Radiology® | All rights reserved

Fonte: American College of Radiology (2022)

Anexo 2 – Tabela de Categorias de avaliação (Assessment Categories)

O-RADS™ US v2022 — Assessment Categories

Release Date: November 2022

O-RADS Score	Risk Category [IOTA Model]	Lexicon Descriptors		Management	
				Pre-menopausal	Post-menopausal
0	Incomplete Evaluation [N/A]	Lesion features relevant for risk stratification cannot be accurately characterized due to technical factors		Repeat US study or MRI	
1	Normal Ovary [N/A]	No ovarian lesion		None	
		Physiologic cyst: follicle (≤3 cm) or corpus luteum (typically ≤3 cm)			
2	Almost Certainly Benign [<1%]	Simple cyst	≤3 cm	N/A (see follicle)	None
			>3 cm to 5 cm	None	Follow-up US in 12 months*
			>5 cm but <10 cm	Follow-up US in 12 months*	
		Unilocular, smooth, non–simple cyst (internal echoes and/or incomplete septations) – – – – – – – – – – – – – – – – Bilocular, smooth cyst	≤3 cm	None	Follow-up US in 12 months*
			>3 cm but <10 cm	Follow-up US in 6 months*	
		Typical benign ovarian lesion (see "Classic Benign Lesions" table)	<10 cm	See "Classic Benign Lesions" table for descriptors and management	
		Typical benign extraovarian lesion (see "Classic Benign Lesions" table)	Any size		
3	Low Risk [1 – <10%]	Typical benign ovarian lesion (see "Classic Benign Lesions" table), ≥10 cm		Imaging: • If not surgically excised, consider follow-up US within 6 months** • If solid, may consider US specialist (if available) *or* MRI (with O-RADS MRI score)*** Clinical: Gynecologist	
		Uni- or bilocular cyst, smooth, ≥10 cm			
		Unilocular cyst, irregular, any size			
		Multilocular cyst, smooth, <10 cm, CS <4			
		Solid lesion, ± shadowing, smooth, any size, CS = 1			
		Solid lesion, shadowing, smooth, any size, CS 2–3			
4	Intermediate Risk [10 – <50%]	Bilocular cyst without solid component(s)	Irregular, any size, any CS	Imaging: Options include: • US specialist (if available) *or* • MRI (with O–RADS MRI score)*** *or* • Per gyn–oncologist protocol Clinical: Gynecologist with gyn–oncologist consultation *or* solely by gyn–oncologist	
		Multilocular cyst without solid component(s)	Smooth, ≥10 cm, CS <4		
			Smooth, any size, CS 4		
			Irregular, any size, any CS		
		Unilocular cyst with solid component(s)	<4 pps or solid component(s) not considered a pp; any size		
		Bi- or multilocular cyst with solid component(s)	Any size, CS 1–2		
		Solid lesion, non-shadowing	Smooth, any size, CS 2–3		
5	High Risk [≥50%]	Unilocular cyst, ≥4 pps, any size, any CS		Imaging: Per gyn-oncologist protocol Clinical: Gyn-oncologist	
		Bi- or multilocular cyst with solid component(s), any size, CS 3–4			
		Solid lesion, ± shadowing, smooth, any size, CS 4			
		Solid lesion, irregular, any size, any CS			
		Ascites and/or peritoneal nodules****			

GLOSSARY

Smooth and irregular: refer to **inner** walls/septation(s) for **cystic** lesions, and **outer** contour for **solid** lesions; irregular inner wall for cysts = <3 mm in height	Solid: excludes blood products and dermoid contents; solid lesion = ≥80% solid; solid component = protrudes ≥3 mm (height) into cyst lumen off wall or septation
Shadowing: must be diffuse or broad to qualify; excludes refractive artifact	pp = papillary projection; subtype of solid component surrounded by fluid on 3 sides
CS = color score; degree of intralesional vascularity; 1 = none, 2 = minimal flow, 3 = moderate flow, 4 = very strong flow	Bilocular = 2 locules; multilocular = ≥3 locules; bilocular smooth cysts have a lower risk of malignancy, regardless of size or CS

*Shorter imaging follow-up may be considered in some scenarios (eg, clinical factors). If smaller (≥10–15% decrease in average linear dimension), no further surveillance. If stable, follow-up US at 24 months from initial exam. If enlarging (≥10–15% increase in average linear dimension), consider follow-up US at 12 and 24 months from initial exam, then management per gynecology. For changing morphology, reassess using lexicon descriptors. **Clinical management with gynecology as needed.**

**There is a paucity of evidence for defining the optimal duration or interval for imaging surveillance. Shorter follow-up may be considered in some scenarios (eg, clinical factors). If stable, follow-up at 12 and 24 months from initial exam, then as clinically indicated. For changing morphology, reassess using lexicon descriptors.

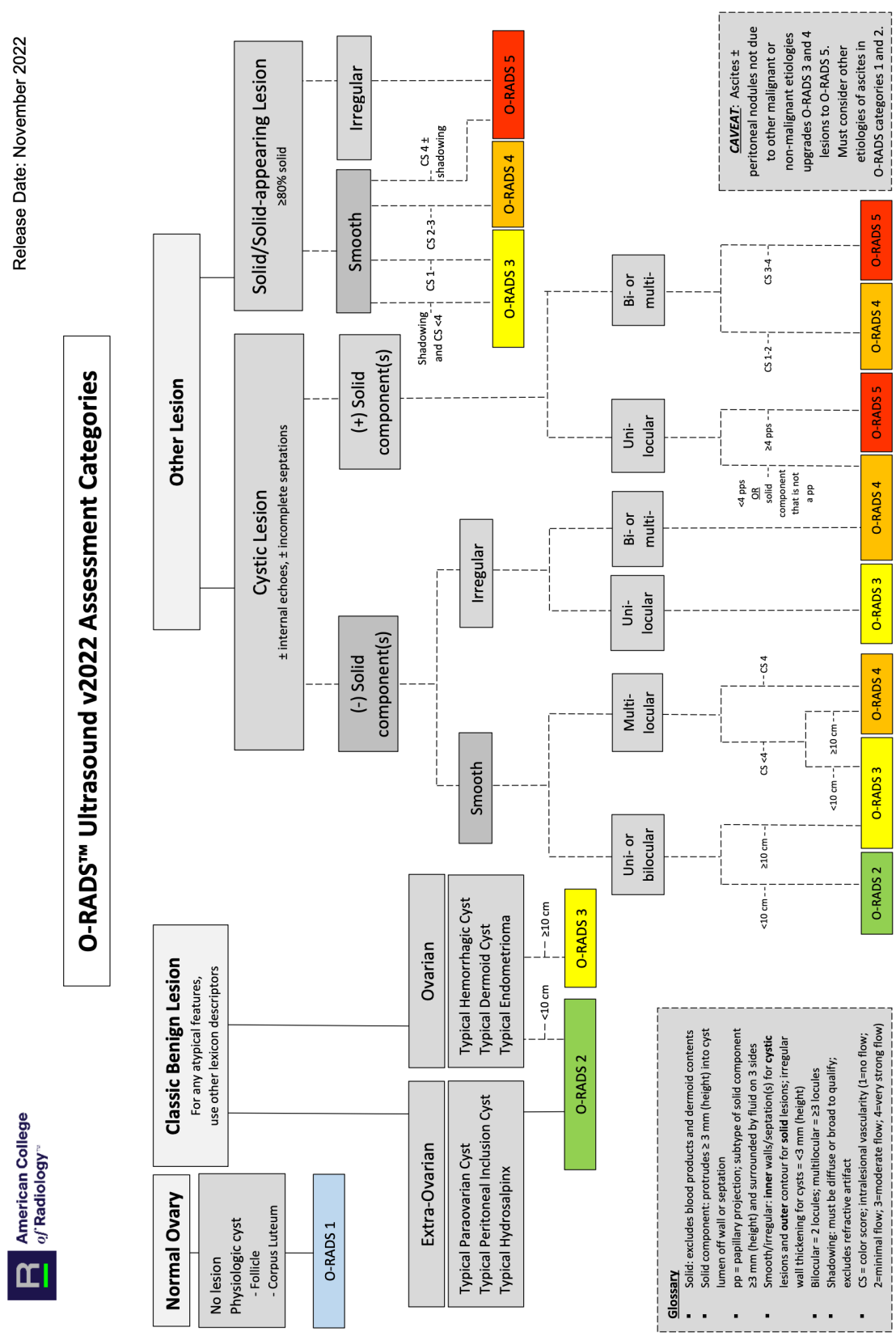
***MRI with contrast has higher specificity for solid lesions, and cystic lesions with solid component(s).

****Not due to other malignant or non-malignant etiologies; specifically, must consider other etiologies of ascites in categories 1–2.

© 2023 American College of Radiology® | All rights reserved

Fonte: American College of Radiology (2022)

Anexo 3 – Algoritmo de Categorias de avaliação (Assessment Categories – Algorithm)



Fonte: American College of Radiology (2022)