



Desempenho diagnóstico do VEGF-D para linfangioleiomiomatose: meta-análise

Min Li^{1,2,3}, Wen-Ye Zhu⁴, Ji Wang^{1,3,5}, Xiao-Dong Yang¹, Wei-Min Li^{1,6}, Gang Wang^{1,3}

1. Department of Respiratory and Critical Care Medicine, Clinical Research Center for Respiratory Disease, West China Hospital, Sichuan University, Chengdu, China.
2. Department of Respiratory and Critical Care Medicine, First Affiliated Hospital of Kunming Medical University, Kunming, China.
3. Laboratory of Pulmonary Immunology and Inflammation, Frontiers Science Center for Disease-Related Molecular Network, Sichuan University, Chengdu, China.
4. Department of Pharmacy, First Affiliated Hospital of Kunming Medical University, Kunming, China.
5. Pulmonology Group, Department of Integrated Traditional Chinese and Western Medicine, West China Hospital, Sichuan University, Chengdu, China.
6. Respiratory Microbiome Laboratory, Frontiers Science Center for Disease-related Molecular Network, Sichuan University, Chengdu, China.

Recebido: 19 agosto 2021.

Aprovado: 16 novembro 2021.

RESUMO

Objetivo: O VEGF-D é um potencial biomarcador para linfangioleiomiomatose (LAM); entretanto, seu desempenho diagnóstico ainda não foi sistematicamente estudado.

Métodos: Foram realizadas buscas nos bancos de dados PubMed, EMBASE, Scopus, Web of Science e Cochrane Library para identificar estudos primários sobre o VEGF-D com relação ao diagnóstico de LAM. A qualidade dos estudos foi avaliada por meio da ferramenta *Quality Assessment of Diagnostic Accuracy Studies-2* (QUADAS-2). As estimativas sumárias de acurácia diagnóstica foram combinadas utilizando um modelo bivariado de efeitos aleatórios. Análises de subgrupo e de sensibilidade foram realizadas para explorar possíveis heterogeneidades. O sistema *Grading of Recommendations Assessment, Development, and Evaluation* (GRADE) foi aplicado para avaliar a qualidade das evidências e indicar a força das recomendações. **Resultados:** Dez estudos envolvendo 945 pacientes eram de alto risco em qualidade, segundo a ferramenta QUADAS-2. Os parâmetros diagnósticos combinados foram indicados da seguinte forma: sensibilidade = 0,82 (IC95%: 0,71-0,90); especificidade = 0,98 (IC95%: 0,94-0,99); e OR diagnóstica = 197 (IC95%: 66-587). A ASC da análise *summary ROC* foi de 0,98. As análises de subgrupo e de sensibilidade revelaram que o desempenho global não foi substancialmente afetado pela composição do grupo controle, valor de corte pré-especificado, país de origem ou diferentes valores de corte ($p > 0,05$ para todos). Uma forte recomendação para a dosagem de VEGF-D sérico para auxiliar no diagnóstico de LAM foi feita de acordo com o sistema GRADE. **Conclusões:** O VEGF-D parece ter grandes implicações potenciais para o diagnóstico de LAM na prática clínica em virtude da excelente especificidade e sensibilidade subótima.

Descritores: Linfangioleiomiomatose/diagnóstico; Fator D de crescimento do endotélio vascular; Biomarcadores; Metanálise.

INTRODUÇÃO

A linfangioleiomiomatose (LAM) é uma doença neoplásica sistêmica rara que afeta principalmente mulheres em idade reprodutiva. Caracteriza-se por destruição cística múltipla e progressiva do pulmão e anormalidades do sistema linfático.^(1,2) Histologicamente, as lesões de LAM caracterizam-se pela proliferação de células neoplásicas semelhantes a músculo liso (células de LAM) em pequenos aglomerados localizados nas bordas de cistos pulmonares e ao longo dos vasos sanguíneos e linfáticos. A infiltração de células de LAM causa obstrução das vias aéreas, espessamento das paredes vasculares, rompimento dos vasos linfáticos, oclusão venosa e hemorragia, levando à deterioração da função pulmonar e, eventualmente, à insuficiência respiratória.⁽³⁻⁶⁾

Considerando o dano mortal da LAM e os avanços em tratamentos específicos para doenças, como o sirolimo, é cada vez mais importante fazer um diagnóstico correto e precoce.^(7,8) No entanto, o diagnóstico de

LAM permanece desafiador. De acordo com a *European Respiratory Society* (ERS), o diagnóstico definitivo de LAM depende das alterações típicas nos achados de TCAR mais biópsia pulmonar na ausência de apresentações clínicas compatíveis, como angiomiolipoma renal, quilotórax ou complexo de esclerose tuberosa.⁽⁹⁾ De fato, vários pacientes não apresentam essas características clínicas típicas, e a biópsia geralmente é necessária para a obtenção do diagnóstico definitivo.⁽¹⁰⁾ No entanto, os métodos de biópsia, seja biópsia pulmonar transbrônquica ou cirurgia torácica videoassistida, são invasivos, tecnicamente difíceis e dependentes do operador. Além disso, procedimentos diagnósticos invasivos são inadequados para o rastreamento de LAM em grupos de alto risco.⁽¹¹⁾ Assim, tornou-se imperativa a busca por métodos não invasivos e eficazes para auxiliar no diagnóstico de LAM, entre os quais biomarcadores específicos da doença teriam grande potencial.

Foi constatado que o VEGF-D, uma glicoproteína produzida pelas células de LAM, promove a linfangiogênese

Endereço para correspondência:

Gang Wang, Department of Respiratory and Critical Care Medicine, Clinical Research Center for Respiratory Disease, West China Hospital, Sichuan University, Chengdu, 610041, China.

Tel.: 86 28 85422376. Fax: 86 28 85422376. E-mail: wcums-respiration@hotmail.com

Apoio financeiro: Este estudo foi financiado pelo Plano-Chave Nacional de Desenvolvimento de Pesquisa em Medicina de Precisão (2017YFC091004), pela Fundação Nacional de Ciências Naturais da China (81670023, 81870027 e 81920108002), pela Fundação de Ciência e Tecnologia da Província de Sichuan (2018SZ0167) e pelo Projeto 1.3.5 para Disciplinas de Excelência-Projeto de Incubação de Pesquisa Clínica, Hospital da China Ocidental, Universidade de Sichuan (2018HXFH016).

via receptor VEGF-3.^(12,13) Em 2006, Seyama et al.⁽¹⁴⁾ encontraram pela primeira vez níveis séricos aumentados de VEGF-D em pacientes com LAM. Consequentemente, um número crescente de estudos tem investigado o valor diagnóstico dos níveis de VEGF-D para LAM; no entanto, os resultados não são inteiramente consistentes. Por exemplo, no estudo de Xu et al.,⁽¹⁵⁾ os níveis de VEGF-D demonstraram um excelente desempenho diagnóstico com uma ASC de 0,99; em outro estudo,⁽¹⁶⁾ os níveis de VEGF-D para o diagnóstico de LAM tiveram uma ASC de 0,75. Embora um limiar de 800 pg/mL seja indicado pela *American Thoracic Society* (ATS),⁽⁶⁾ os níveis séricos de VEGF-D parecem variar em diferentes dosagens ou valores de corte. O melhor valor de corte para os níveis de VEGF-D diferiu entre os estudos, variando de 440 pg/mL a 1.239 pg/mL.⁽¹⁶⁾ A identificação de um limiar apropriado para os níveis de VEGF-D em pacientes com LAM, especialmente aqueles comistos pulmonares característicos na TCAR, mas sem achados clínicos adicionais, é essencial. Até o momento, nenhum estudo combinou sistematicamente esses dados para estimar o desempenho diagnóstico global dos níveis de VEGF-D para LAM ou para explorar ainda mais os potenciais problemas que afetam sua acurácia diagnóstica. Portanto, realizamos uma meta-análise com base nas evidências atualmente disponíveis com o intuito de estabelecer estimativas sumárias de acurácia diagnóstica dos níveis de VEGF-D para LAM, o que teria implicações relevantes na prática clínica.

MÉTODOS

Trata-se de um estudo com desenho de meta-análise cujo protocolo foi registrado no *International Prospective Register of Systematic Reviews* (PROSPERO; Protocolo n. CRD42020164137), um banco de dados internacional de revisões sistemáticas registradas prospectivamente. Nosso estudo foi realizado e os resultados foram relatados segundo os métodos recomendados pelo *Cochrane Diagnostic Test Accuracy Working Group* e pelo *Preferred Reporting Items for Systematic Reviews and Meta-Analyses* (PRISMA; Quadro S1).⁽¹⁷⁻¹⁹⁾ A aprovação ética não foi necessária em razão da natureza retrospectiva da meta-análise.

Estratégia de busca e critérios de seleção

Dois revisores, de forma independente, realizaram buscas nos seguintes bancos de dados: PubMed, EMBASE, Scopus, *Web of Science* e *Cochrane Library* (até 26 de junho de 2021) para identificar estudos relacionados. Foram utilizadas combinações das seguintes *strings* de busca: ("lymphangioleiomyomatosis" or "LAM") AND ("endothelial growth factor-D" or "VEGF-D"). As referências listadas nos artigos ou nos artigos de revisão considerados também foram verificadas manualmente para obtenção de artigos adicionais relevantes. A LAM foi descrita e definida segundo as diretrizes da *ATS/Japanese Respiratory Society* (JRS; Quadro S2)⁽⁶⁾ e as diretrizes da ERS (Quadro S3).⁽⁹⁾ Os estudos foram incluídos na meta-análise se

atendessem a todos os seguintes critérios: 1) estudos originais que examinaram a capacidade diagnóstica do VEGF-D para LAM; 2) dados suficientes para construir uma tabela 2 × 2 de resultados verdadeiro-positivos, verdadeiro-negativos, falso-positivos e falso-negativos; e 3) estudos que incluíram pelo menos 20 indivíduos (estudos com menor tamanho amostral podem ser vulneráveis a viés de seleção). Não houve limitação de idiomas, e a data limite mínima foi restrita ao ano de 1997, quando o VEGF-D foi identificado pela primeira vez.⁽²⁰⁾ Foram incluídos apenas os dados associados ao melhor desempenho diagnóstico provenientes de estudos nos quais vários valores de corte diferentes foram utilizados.

Extração de dados e avaliação de qualidade

Dois revisores, de forma independente, avaliaram os estudos incluídos. Foram extraídas as seguintes informações: primeiro autor, ano de publicação, país de origem, método de teste, valor de corte, padrão diagnóstico e tabelas 2 × 2 para resultados verdadeiro-positivos, falso-positivos, verdadeiro-negativos e falso-negativos. Entramos em contato com os autores responsáveis pela correspondência para identificar informações adicionais, se necessário.

A ferramenta *Quality Assessment of Diagnostic Accuracy Studies-2* (QUADAS-2)⁽²¹⁾ foi aplicada para avaliar a metodologia dos estudos incluídos. Ela abrange quatro domínios principais para avaliar o risco de viés: seleção de pacientes, teste índice, padrão de referência e fluxo e tempo. Preocupações relacionadas à aplicabilidade também foram avaliadas nos domínios seleção de pacientes, teste índice e padrão de referência. A ferramenta fornece questões norteadoras para ajudar a avaliar diferentes estudos nos domínios supracitados. As questões são respondidas com "sim", "não" ou "incerto". Se as respostas a todas as questões norteadoras para um domínio forem "sim", pode-se fazer um julgamento global de "baixo risco de viés". Se qualquer questão norteadora for respondida com "não", existe um potencial viés. Se os dados fornecidos pelo estudo primário forem insuficientes em um ou mais domínios para permitir um julgamento, deve-se utilizar a categoria "incerto". Foram avaliadas ainda as seguintes informações como suplemento para a avaliação da metodologia: registro do estudo, desenho de mundo real, conselho de revisão institucional, cálculo do tamanho amostral, coleta prospectiva de dados, coleta consecutiva de dados, critérios de inclusão/exclusão (rigorosos ou lenientes), estudo multicêntrico, e conflitos de interesse.

O sistema *Grading of Recommendations Assessment, Development, and Evaluation* (GRADE) foi aplicado para avaliar a certeza das evidências. Ele classifica a qualidade das evidências em quatro níveis (alta, moderada, baixa e muito baixa). O sistema GRADE começa com a classificação de "alta qualidade" e então é rebaixado em um nível para cada um dos cinco fatores considerados: risco de viés, *indirectness*, inconsistência, imprecisão e viés de publicação.⁽²²⁾

O risco de viés foi avaliado por meio da ferramenta QUADAS-2. Para avaliar *indirectness*, consideramos nossa avaliação das preocupações relacionadas à aplicabilidade e a *directness* dos resultados do exame de VEGF-D em relação a desfechos importantes para o paciente. A inconsistência foi avaliada de acordo com a heterogeneidade entre estudos (avaliada por meio do teste Q baseado no qui-quadrado e do índice de inconsistência). A amplitude do IC95% foi considerada na avaliação da imprecisão: IC95% < 10% foi considerado “não sério”; 10% ≥ IC95% < 20% foi considerado “sério”; e IC95% ≥ 20% foi considerado “muito sério”.

O sistema GRADE oferece dois graus de recomendações: “forte” e “fraco”.

Quando os efeitos de uma intervenção superam claramente os efeitos indesejáveis, ou claramente não o fazem, ele oferece recomendações fortes. Quando as compensações são menos certas, recomendações fracas tornam-se obrigatórias.⁽²²⁾ Quaisquer discrepâncias na seleção dos estudos, extração dos dados e avaliação da qualidade foram resolvidas por um terceiro revisor.

Análise estatística

A meta-análise foi realizada utilizando um modelo bivariado.^(23,24) Analisamos a acurácia global do teste diagnóstico calculando os seguintes índices: sensibilidade, especificidade, razão de verossimilhança positiva, razão de verossimilhança negativa e OR diagnóstica. Também calculamos a ASC, que sugere o grau de acurácia de uma ferramenta diagnóstica (ruim: 0,50-0,75; boa: 0,75-0,92; muito boa: 0,93-0,96; e excelente: > 0,97).⁽²⁵⁾ Além disso, combinamos o modelo *hierarchical summary* ROC (HSROC) para conduzir a meta-análise, pois esperávamos que os estudos utilizassem diferentes limiares para dicotomizar os resultados dos testes medidos em uma escala contínua. O método pode denotar se um efeito de limiar está presente quando o valor de corte não é pré-especificado. É melhor sumarizar os resultados do que usar uma única estimativa sumária conjunta de sensibilidade e especificidade.^(26,27) Aplicamos ainda o nomograma de Fagan para estimar até que ponto os resultados do teste diagnóstico mudam a possibilidade de um paciente ter uma determinada doença, que se baseia na seguinte fórmula equivalente:

$$\log(\text{odds pós-teste}) = \log(\text{razão de verossimilhança}) + \log(\text{odds pré-teste})$$

Para detectar a heterogeneidade entre estudos, aplicamos o teste Q baseado em qui-quadrado e o índice de inconsistência (I^2). Para sugerir a presença de heterogeneidade além do que seria esperado ao acaso, foi utilizado um valor de $p < 0,10$, e, para quantificar o grau de heterogeneidade, foi utilizado um valor de $I^2 > 50\%$. Realizamos ainda análises de subgrupo e de sensibilidade para explorar possíveis problemas que

podem afetar o desempenho diagnóstico do VEGF-D para LAM. O viés de publicação foi avaliado por meio de gráficos de funil de Deeks.⁽²⁸⁾

Foram utilizados os programas Stata, versão 15.0 (Stata Corp LP, College Station, TX, EUA), Review Manager 5.2 (The Cochrane Collaboration, Copenhagen, Dinamarca) e Meta-Disc (XI Cochrane Colloquium, Barcelona, Espanha), e foi considerado estatisticamente significativo p bilateral $< 0,05$.

RESULTADOS

Características dos estudos

Após a revisão do título e do resumo dos 115 registros identificados na triagem inicial, foram selecionados 33 artigos para revisão do texto completo, dos quais 3 estudos foram excluídos em virtude de dados duplicados. Além disso, 14 desses estudos foram excluídos por não incluírem um grupo controle, e outros 6 estudos foram excluídos por não ter sido possível construir tabelas 2 × 2, embora os autores desses estudos tenham sido contatados. Por fim, 10 estudos foram incluídos na meta-análise (Figura 1).

Os estudos incluídos foram realizados entre 2009 e 2019. O tamanho médio das amostras foi de 95 (variação: 33-173), resultando em uma população total de 945 pacientes. Dois estudos foram realizados na América do Norte,^(16,29) 4 na Europa,⁽³⁰⁻³³⁾ 3 na Ásia^(15,34,35) e 1 na América do Sul.⁽³⁶⁾ A média de idade dos pacientes dos estudos incluídos variou de 30 a 55 anos.

Todos os estudos incluídos foram realizados em hospitais terciários, e utilizou-se soro como amostra para testar os níveis de VEGF-D; o *kit* de ensaio imunossorvente ligado a enzima de VEGF-D humano (R&D System Inc., Minneapolis, MN, EUA) foi utilizado em 8 estudos.^(15,16,29-31,34-36) O padrão diagnóstico de LAM preencheu os critérios diagnósticos de LAM definitiva de acordo com as diretrizes da ERS⁽⁹⁾ em 8 estudos,^(15,16,29-32,35,36) enquanto parte dos pacientes incluídos preencheu o padrão diagnóstico para LAM provável em 2 estudos.^(33,34) Todos os estudos incluídos descreveram a composição do grupo controle em detalhes. Os grupos controle (Tabela 1) foram constituídos por pacientes com outras doenças pulmonares policísticas como histiocitose pulmonar de células de Langerhans, síndrome de Birt-Hogg-Dubé e síndrome de Sjögren em 3 estudos,^(29,31,35) por indivíduos saudáveis em 4 estudos,^(15,16,30,36) e por pacientes com outras doenças pulmonares policísticas e participantes saudáveis em 3 estudos.⁽³²⁻³⁴⁾

Avaliação da qualidade metodológica

Utilizamos a ferramenta QUADAS-2 para avaliar a qualidade metodológica dos estudos incluídos. Pode haver um alto risco nos domínios seleção de pacientes e teste índice. Os estudos apresentaram risco alto ou incerto de viés em virtude do seguinte: 1) 90% dos estudos incluídos não descreveram em detalhes se os

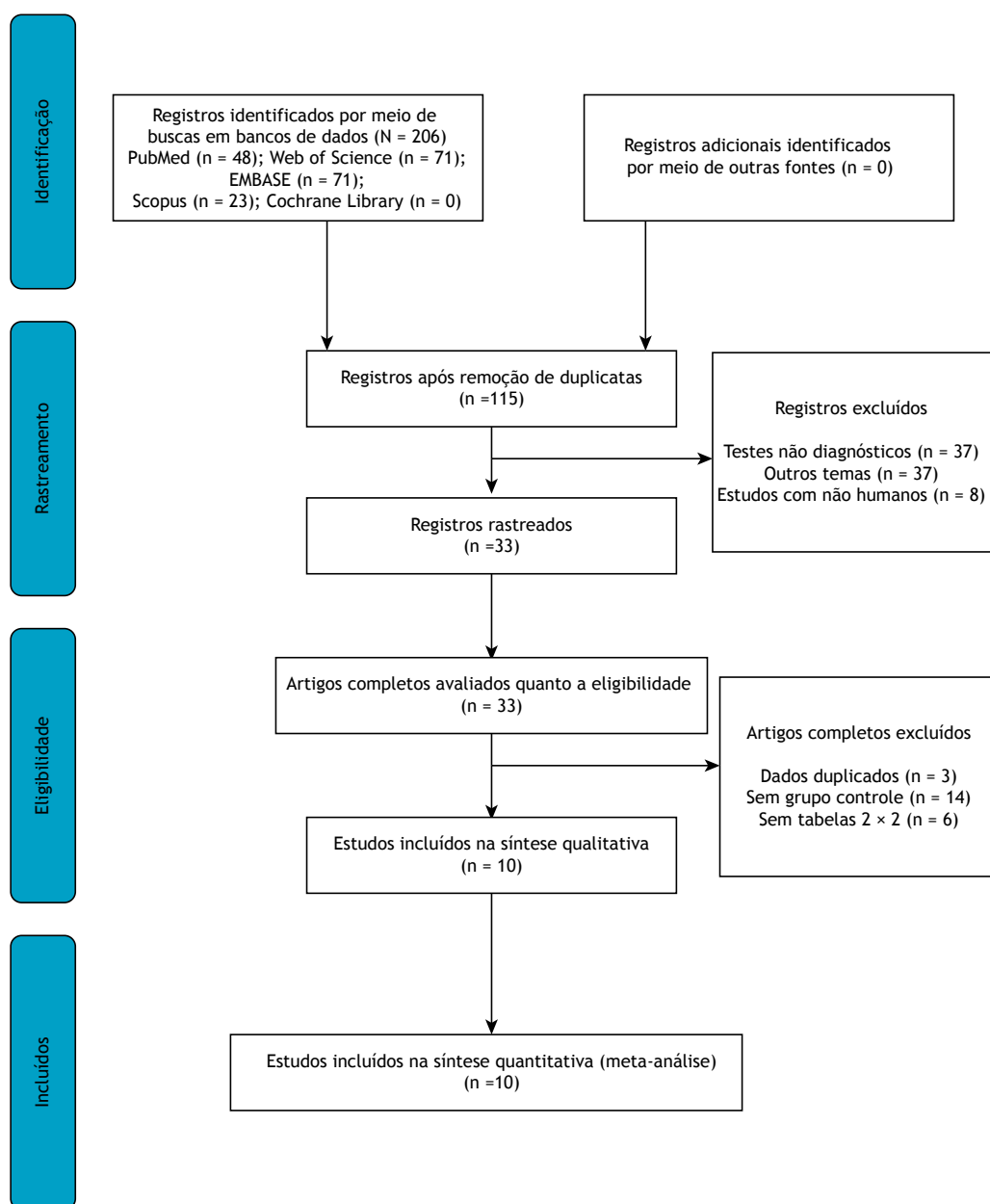


Figura 1. Fluxograma da seleção dos estudos.

participantes foram inscritos consecutivamente ou se foi utilizada amostragem aleatória^(15,16,29-32,34-36); 2) 4 estudos não evitaram o desenho caso-controle^(15,16,30,36); 3) 4 estudos não evitaram exclusão inadequada, nos quais indivíduos saudáveis, em vez de pacientes com doença que precisa ser diferenciada de LAM, foram recrutados como controles^(15,16,30,36); e 4) o valor de corte do VEGF-D no soro era pré-especificado em 33,3% dos estudos incluídos.^(29,32,36) No domínio padrão de referência, todos os estudos incluídos eram de baixo risco. No domínio fluxo e tempo, 1 estudo mostrou alto risco de viés, pois nem todos os pacientes dos estudos foram incluídos na análise.⁽¹⁶⁾ Com relação a preocupações relacionadas à aplicabilidade, 5 estudos eram de alta preocupação em virtude de deficiências

na seleção de pacientes,^(15,16,30,34,36) e 1 estudo era de alta preocupação no teste índice⁽³²⁾ (Figura S1).

Além das informações obtidas com a ferramenta QUADAS-2, também foram extraídas informações adicionais relacionadas à qualidade do desenho dos estudos. Nenhum estudo estava registrado na Plataforma Internacional de Registro de Ensaios Clínicos da OMS, no Registro de Ensaios Clínicos da União Europeia ou no ClinicalTrials.gov. Na amostra total, 7 estudos foram aprovados por um conselho de revisão institucional.^(15,16,29,30,34-36) Dois estudos tinham desenho prospectivo.^(29,32) Os critérios de inclusão/exclusão foram rigorosos em 2 estudos.^(29,36) Podem existir potenciais conflitos de interesse em 2 estudos.^(29,34) Informações mais detalhadas são fornecidas na Tabela S1.

Tabela 1. Características dos estudos incluídos sobre níveis de VEGF-D com relação ao diagnóstico de linfangioleiomiomatose.^a

Autores	País	P/C, n	Média de idade, anos (P/C)	Tabagismo, n (%)	Método de ensaio	Fabricante de VEGF-D	Valor de corte pré-especificado	Valor de corte, pg/mL	LAM + TSC, n (%)	Terapia com inibidor de mTOR, n (%)	Comprometimento linfático, n (%)	Padrão diagnóstico	Controles	VP	FP	FN	VN
Glasgow et al. ⁽¹⁶⁾	EUA	106/40	51,0/49,6	DD	ELISA	R&D Systems	Não	1.239	0 (0)	DD	77 (73)	LAM definitiva	Saudáveis	61	1	45	39
Young et al. ⁽²⁹⁾	EUA	15/18	UN	DD	ELISA	R&D Systems	Sim	800	0 (0)	DD	DD	LAM definitiva	ODPP	12	0	3	18
Cottin et al. ⁽³²⁾	França	45/42	48/?	DD	DD	DD	Sim	800	DI	6 (13%)	DD	LAM definitiva	ODPP+ saudáveis	34	1	11	41
Chang et al. ⁽³⁰⁾	EUA	45/32	46/36	1 (2.2%)	ELISA	R&D Systems	Não	440	2 (4.4)	DD	DD	LAM definitiva	Saudáveis	41	1	4	31
Radzikowska et al. ⁽³¹⁾	Polônia	29/46	41,0/37,6	DD	ELISA	R&D Systems	Não	468	7 (24)	DD	DD	LAM definitiva	ODPP	25	5	4	41
Xu et al. ⁽¹⁵⁾	China	50/40	40,8/41,4	DD	ELISA	R&D Systems	Não	850,7	2 (4%)	7 (14%)	DD	LAM definitiva	Saudáveis	48	0	2	40
Daccord et al. ⁽³³⁾	Suíça	15/21	UN	DD	ELISA	DD	Não	520	DI	3 (20%)	DD	LAM definitivo+ provável	ODPP+ saudáveis	12	1	3	20
Amaral et al. ⁽³⁶⁾	Brasil	104/40	43/43	DD	ELISA	R&D Systems	Sim	800	21 (20)	0 (0%)	20 (24)	LAM definitiva	Saudáveis	52	0	52	40
Hirose et al. ⁽³⁴⁾	Japão	108/65	UN	DD	ELISA	R&D Systems	Não	645	16 (14.8)	35 (32%)	DD	LAM definitivo+ provável	ODPP+ saudáveis	90	2	18	63
Mou et al. ⁽³⁵⁾	China	50/34	39,3/46,9	DD	ELISA	R&D Systems	Não	901	3 (6)	DD	DD	LAM definitiva	ODPP	47	0	3	34

P/C: pacientes/controles; LAM: linfangioleiomiomatose; TSC: *tuberosus sclerosis complex* (complexo da esclerose tuberosa); mTOR: *mammalian target of rapamycin* (alvo da rapamicina em mamíferos); VP: verdadeiro-positivos; FP: falso-positivos; FN: falso-negativos; VN: verdadeiro-negativos; DD: dado desconhecido; ODPP: outras doenças pulmonares polísticas; e DI: dado incerto. ^aTodos os estudos foram realizados em hospitais terciários, e os níveis de VEGF-D foram determinados a partir de amostras de soro.

Acurácia diagnóstica

A Figura 2 mostra que a sensibilidade e especificidade combinadas foram de 0,82 (IC95%: 0,71-0,90) e 0,98 (IC95%: 0,94-0,99), respectivamente. A razão de verossimilhança positiva foi de 35,5 (IC95%: 14,4-87,6), e a razão de verossimilhança negativa foi de 0,18 (IC95%: 0,11-0,30). A OR diagnóstica global foi de 197 (IC95%: 66-587), e a ASC foi de 0,98 (IC95%: 0,97-0,99).

O modelo HSROC mostrou que o valor de β foi igual a $-0,31$ ($p = 0,678$) e o valor de λ foi igual a 5,69, sugerindo que a curva HSROC era simétrica e o valor diagnóstico global, excelente (Figura 3).

Utilizamos o nomograma de Fagan para apoiar a tomada de decisão e avaliar a utilidade clínica (Figura S2). Por exemplo, em uma população de risco médio com probabilidade pré-teste de 20%, o exame de VEGF-D aumentaria a probabilidade de diagnosticar LAM para 90% em caso de resultado positivo e diminuiria essa probabilidade para 4% em caso de resultado negativo.

Análise de subgrupos

Os valores de I^2 para a sensibilidade e especificidade combinadas foram iguais a 92,70% e 57,81%, respectivamente, indicando heterogeneidade significativa entre os estudos incluídos. Primeiramente, procuramos a presença de efeito de limiar, uma fonte

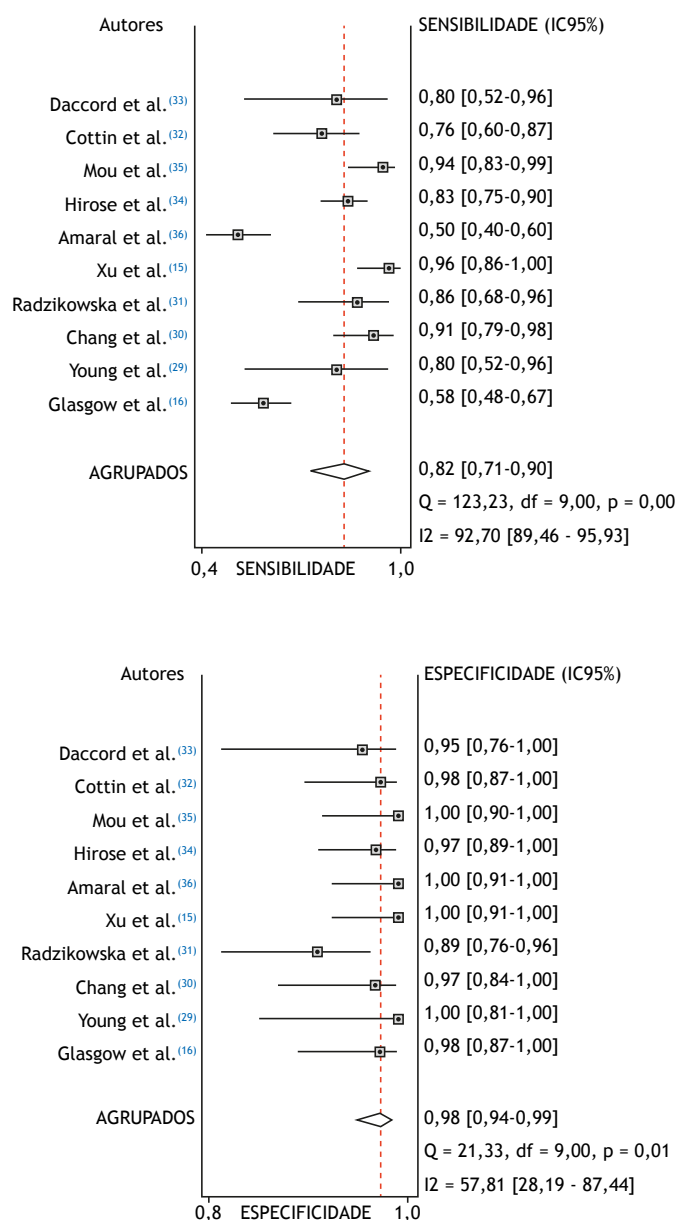


Figura 2. Gráficos de floresta da sensibilidade e especificidade dos níveis de VEGF-D para o diagnóstico de linfangioleiomiomatose. As estimativas pontuais de sensibilidade e especificidade para cada estudo são mostradas como quadrados sólidos. As barras de erro indicam os IC95%. df: *degrees of freedom* (graus de liberdade).

de heterogeneidade referente exclusivamente à meta-análise diagnóstica. Nenhum efeito de limiar foi detectado (coeficiente de correlação de Spearman = 0,097; $p = 0,789$). Investigamos ainda os potenciais fatores que podem afetar os resultados globais da meta-análise. Conforme mostrado na Tabela 2, a ASC não foi significativamente afetada pela composição do grupo controle ($p = 0,115$), valor de corte pré-especificado ($p = 0,839$), país de origem ou diferentes valores de corte; isso indicou que essas covariáveis não afetaram substancialmente a acurácia diagnóstica do VEGF-D para LAM em nosso estudo ($p = 0,889$).

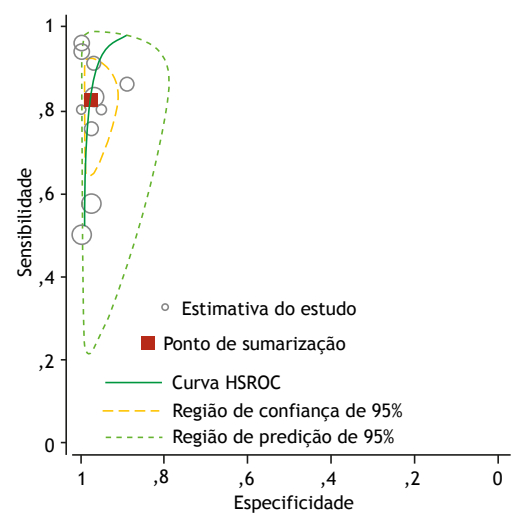


Figura 3. Gráfico da curva *hierarchica summary* ROC (HSROC) de sensibilidade *versus* especificidade dos níveis de VEGF-D para o diagnóstico de linfangioleiomiomatose.

Tabela 2. Análise de subgrupos dos estudos incluídos sobre níveis de VEGF-D com relação ao diagnóstico de linfangioleiomiomatose.

Covariáveis	Estudos, n	ASC	ep (ASC)	Z	p
Composição do grupo controle					
Não saudáveis	6	0,96	0,02	1,576	0,115
Saudáveis	4	0,99	0,01		
Valor de corte pré-especificado					
Sim	3	0,98	0,04	0,203	0,839
Não	7	0,97	0,02		
País de origem					
Países asiáticos	3	0,99	0,00	0,139	0,889
Países não asiáticos	7	0,96	0,02		
Valores de corte					
≤ 600 pg/mL	3	0,94	0,04	1,558 ^a	0,119 ^a
> 800 pg/mL	3	0,99	0,00	0,739 ^a	0,460 ^a
600-800 pg/mL	4	0,98	0,02	Ref.	Ref.

^aEm comparação com os valores de corte (> 600 a ≤ 800 pg/mL).

Tabela 3. Sumário da análise global e da análise de sensibilidade.

Variáveis	Estudos, n	ASC	ep (ASC)	Z	p
Resultados globais da meta-análise	10	0,98	0,005		
Excluindo resumos	8	0,97	0,015	0,632 ^a	0,527 ^a
Excluindo dois estudos com pacientes com LAM provável	8	0,9813	0,01116	0,106 ^a	0,915 ^a

^aEm comparação com os resultados globais da meta-análise.

Análise de sensibilidade

Nem todos os pacientes com LAM atingiram o padrão diagnóstico para LAM definitiva nos 2 estudos primários,^(33,34) e 2 estudos eram resumos.^(32,33) Realizamos uma análise de sensibilidade para explorar os efeitos desses estudos nos resultados globais da meta-análise. Os resultados indicaram que a ASC global não foi materialmente alterada após a exclusão dos 2 estudos primários que incluíram pacientes com LAM provável e dos 2 resumos, sugerindo a robustez do nosso estudo (Tabela 3).

Viés de publicação

O viés de publicação foi explorado por meio do teste para assimetria do gráfico de funil de Deeks. Os resultados apresentaram valor não significativo ($p = 0,92$). A forma dos gráficos de funil não revelou nenhuma evidência de assimetria (Figura S3), indicando baixa probabilidade de viés de publicação.

Qualidade das evidências de acordo com o sistema GRADE

O sistema GRADE foi aplicado para avaliar a qualidade das evidências e indicar a força das recomendações. Em nosso estudo, a qualidade dos resultados combinados foi rebaixada para o risco de viés (com base nos resultados do QUADAS-2), inconsistência ($I^2 > 0,5$ e $p < 0,1$) e imprecisão (amplitude do IC95% > 10%; Tabela S2). Como resultado, a qualidade das evidências foi classificada como baixa com base nos sumários do GRADE.

DISCUSSÃO

Até onde sabemos, este é o primeiro estudo a explorar sistematicamente o desempenho diagnóstico

do VEGF-D para LAM. A análise mostrou que a acurácia diagnóstica global da dosagem de VEGF-D para LAM foi excelente com base nos 10 estudos com alto risco de viés avaliado por meio da ferramenta QUADAS-2; o nível de VEGF-D também exibiu sensibilidade subótima e alta especificidade. As análises de subgrupo e de sensibilidade revelaram que o desempenho global não foi substancialmente afetado pela composição do grupo controle, valor de corte pré-especificado, país de origem ou diferentes valores de corte. Uma forte recomendação para a dosagem de VEGF-D sérico para auxiliar no diagnóstico de LAM na prática clínica foi feita de acordo com o sistema GRADE.

Segundo as diretrizes da ATS/JRS,⁽⁶⁾ o exame de VEGF-D é recomendado antes da biópsia pulmonar diagnóstica para pacientes com suspeita de LAM que apresentam anormalidades císticas na TC, mas sem características clínicas ou radiológicas extrapulmonares confirmatórias. No entanto, as diretrizes citaram apenas 7 estudos. Os estudos foram realizados principalmente nos EUA e na Europa.^(14-16,29-31,37) Os dados do resto do mundo são limitados. Além disso, a sensibilidade ou especificidade global dos níveis de VEGF-D não foi fornecida. Após a publicação das diretrizes, foram relatados mais estudos clínicos, e realizamos a meta-análise incluindo mais estudos realizados em outras partes do mundo, o que reforça o uso do exame de VEGF-D em uma população mais ampla.

Em nosso estudo, a especificidade do VEGF-D foi excelente; mas a sensibilidade foi moderada, o que indica que um resultado positivo pode ser considerado LAM, enquanto um resultado negativo não exclui LAM. De fato, especificidade alta parece ser mais aceitável na prática clínica, pois um diagnóstico incorreto de LAM pode levar a oportunidades perdidas de tratar a doença real, evitando efeitos adversos e gastos desnecessários em virtude de tratamento inadequado.⁽⁶⁾ No entanto, se um paciente com LAM apresenta um resultado falso-negativo, a consequência provável é que o paciente prossiga para o próximo teste diagnóstico para obter o diagnóstico final — biópsia pulmonar, que reduz em grande parte os efeitos indesejáveis do falso-negativo.⁽⁶⁾ Portanto, a dosagem dos níveis de VEGF-D apresenta potencial impacto clínico no algoritmo diagnóstico de LAM. Para pacientes que apresentaram anormalidades císticas pulmonares, mas não apresentam características confirmatórias de LAM, os níveis de VEGF-D devem ser considerados antes de proceder à biópsia pulmonar. Estimou-se que 90% dos casos de LAM provável, segundo as diretrizes da ERS,⁽⁹⁾ poderiam ser elevados para LAM definitiva se o resultado do exame de VEGF-D fosse adicionado aos critérios diagnósticos.⁽³⁸⁾ Portanto, muitos procedimentos diagnósticos invasivos podem ser evitados, o que reduzirá substancialmente o risco para os pacientes e a carga médica.

Em nosso estudo, encontramos 7 estudos que determinaram os melhores valores de corte utilizando o índice de Youden.^(15,16,30-32,34,35) A escolha dos melhores valores de corte baseada em dados para um teste

de uma variável contínua pode levar a estimativas superotimistas de acurácia diagnóstica, especialmente em estudos com amostras pequenas.⁽³⁹⁾ Utilizar um valor de corte pré-especificado é uma abordagem eficiente para reduzir o problema, mas geralmente é difícil conseguir isso. Nas fases iniciais da avaliação de um teste, geralmente há informações limitadas disponíveis sobre o melhor valor de corte provável, especialmente para uma doença rara. Em nosso estudo, constatamos que o desempenho relatado do teste nos estudos com escolha de valores de corte baseada em dados não foi melhor do que o relatado nos estudos com valores de corte pré-especificados.

Embora a ATS recomende um valor de limiar de VEGF-D de 800 pg/mL,⁽⁶⁾ os valores de corte nos estudos incluídos foram variados, indo de 440 pg/mL a 1.239 pg/mL. Isso é razoável porque os níveis de VEGF-D podem ser afetados por muitos fatores. Vários estudos relataram que os níveis de VEGF-D eram mais altos em pacientes com LAM com comprometimento linfático do que naqueles sem comprometimento linfático.⁽²⁹⁾ Além disso, a gravidade da doença,⁽³⁶⁾ o *status* do tratamento com sirolimo⁽⁵⁾ e se há complexo da esclerose tuberosa associado também podem afetar os níveis de VEGF-D.^(40,41) Além disso, diferenças no manuseio da amostra, preparação padrão ou outros aspectos da conduta do ensaio podem levar a discrepâncias nos níveis de VEGF-D entre diferentes estudos.^(31,34) Portanto, é provável que o valor de corte varie com a gravidade da doença, o equipamento e o histórico de tratamento. Por um lado, um valor de corte de 800 pg/mL como limiar diagnóstico específico para LAM pode fornecer um bom valor diagnóstico para pacientes com LAM em geral, e um padrão uniforme para o nível de VEGF-D é necessário se seu uso se tornar a prática padrão para o diagnóstico de LAM. Por outro lado, os investigadores também devem estar abertos à possibilidade de que diferentes valores de corte podem ser necessários para pacientes com diferentes *status* da doença.

Nosso estudo apresenta algumas limitações que precisam ser abordadas. Em primeiro lugar, o número de estudos sobre LAM, por se tratar de uma doença rara, foi relativamente pequeno, embora tenhamos realizado uma busca abrangente na literatura. Em segundo lugar, nem todos os pacientes com LAM desta meta-análise preencheram os critérios diagnósticos de LAM definitiva.^(33,34) Mesmo assim, nossa análise de sensibilidade indicou que os estudos envolvidos não alteraram materialmente os resultados combinados. Por fim, não se deve ignorar o possível viés de seleção, visto que todos os estudos incluídos foram realizados em hospitais terciários. Estudos multicêntricos realizados em hospitais de diferentes níveis são necessários para validar ainda mais nossos achados.

Em conclusão, este estudo demonstrou que a dosagem de VEGF-D é altamente específica e moderadamente sensível para o diagnóstico de LAM. Visto que os efeitos do VEGF-D para LAM superam os efeitos indesejáveis, uma forte recomendação para a dosagem de VEGF-D sérico para auxiliar no diagnóstico de LAM na prática

clínica foi feita de acordo com o sistema GRADE; no entanto, estudos multicêntricos realizados em hospitais de diferentes níveis são necessários para validar ainda mais esses achados.

CONTRIBUIÇÕES DOS AUTORES

ML, WYZ e GW: concepção e desenho. GW: apoio administrativo. ML e WYZ: fornecimento de materiais

de estudo e pacientes. JW: coleta e reunião de dados. ML e JW: análise e interpretação de dados. Todos os autores: redação do manuscrito e aprovação do manuscrito final.

CONFLITO DE INTERESSES

Nenhum declarado.

REFERÊNCIAS

- Elia D, Torre O, Cassandro R, Caminati A, Harari S. Ultra-rare cystic disease. *Eur Respir Rev*. 2020;29(157):190163. <https://doi.org/10.1183/16000617.0163-2019>
- Steagall WK, Pacheco-Rodriguez G, Darling TN, Torre O, Harari S, Moss J. The Lymphangioleiomyomatosis Lung Cell and Its Human Cell Models. *Am J Respir Cell Mol Biol*. 2018;58(6):678-683. <https://doi.org/10.1165/rcmb.2017-0403TR>
- Torre O, Elia D, Caminati A, Harari S. New insights in lymphangioleiomyomatosis and pulmonary Langerhans cell histiocytosis. *Eur Respir Rev*. 2017;26(145):170042. <https://doi.org/10.1183/16000617.0042-2017>
- Avila NA, Kelly JA, Chu SC, Dwyer AJ, Moss J. Lymphangioleiomyomatosis: abdominopelvic CT and US findings. *Radiology*. 2000;216(1):147-153. <https://doi.org/10.1148/radiology.216.1.r00j42147>
- Miller S, Stewart ID, Clements D, Soomro I, Babaei-Jadidi R, Johnson SR. Evolution of lung pathology in lymphangioleiomyomatosis: associations with disease course and treatment response. *J Pathol Clin Res*. 2020;6(3):215-226. <https://doi.org/10.1002/cjp2.162>
- McCormack FX, Gupta N, Finlay GR, Young LR, Taveira-DaSilva AM, Glasgow CG, et al. Official American Thoracic Society/Japanese Respiratory Society Clinical Practice Guidelines: Lymphangioleiomyomatosis Diagnosis and Management. *Am J Respir Crit Care Med*. 2016;194(6):748-761. <https://doi.org/10.1164/rccm.201607-1384ST>
- Davies DM, Johnson SR, Tattersfield AE, Kingswood JC, Cox JA, McCartney DL, et al. Sirolimus therapy in tuberous sclerosis or sporadic lymphangioleiomyomatosis. *N Engl J Med*. 2008;358(2):200-203. <https://doi.org/10.1056/NEJMc072500>
- McCormack FX, Inoue Y, Moss J, Singer LG, Strange C, Nakata K, et al. Efficacy and safety of sirolimus in lymphangioleiomyomatosis. *N Engl J Med*. 2011;364(17):1595-1606. <https://doi.org/10.1056/NEJMoa1100391>
- Johnson SR, Cordier JF, Lazor R, Cottin V, Costabel U, Harari S, et al. European Respiratory Society guidelines for the diagnosis and management of lymphangioleiomyomatosis. *Eur Respir J*. 2010;35(1):14-26. <https://doi.org/10.1183/09031936.00076209>
- Koyama M, Johkoh T, Honda O, Tsubamoto M, Kozuka T, Tomiyama N, et al. Chronic cystic lung disease: diagnostic accuracy of high-resolution CT in 92 patients. *AJR Am J Roentgenol*. 2003;180(3):827-835. <https://doi.org/10.2214/ajr.180.3.1800827>
- Hagaman JT, Schauer DP, McCormack FX, Kinder BW. Screening for lymphangioleiomyomatosis by high-resolution computed tomography in young, nonsmoking women presenting with spontaneous pneumothorax is cost-effective. *Am J Respir Crit Care Med*. 2010;181(12):1376-1382. <https://doi.org/10.1164/rccm.200910-1553OC>
- Stacker SA, Achen MG. Emerging Roles for VEGF-D in Human Disease. *Biomolecules*. 2018;8(1):1. <https://doi.org/10.3390/biom8010001>
- Issaka RB, Oommen S, Gupta SK, Liu G, Myers JL, Ryu JH, et al. Vascular endothelial growth factors C and D induces proliferation of lymphangioleiomyomatosis cells through autocrine crosstalk with endothelium. *Am J Pathol*. 2009;175(4):1410-1420. <https://doi.org/10.2353/ajpath.2009.080830>
- Seyama K, Kumasaka T, Kurihara M, Mitani K, Sato T. Lymphangioleiomyomatosis: a disease involving the lymphatic system. *Lymphat Res Biol*. 2010;8(1):21-31. <https://doi.org/10.1089/lrb.2009.0018>
- Xu KF, Zhang P, Tian X, Ma A, Li X, Zhou J, et al. The role of vascular endothelial growth factor-D in diagnosis of lymphangioleiomyomatosis (LAM). *Respir Med*. 2013;107(2):263-268. <https://doi.org/10.1016/j.rmed.2012.10.006>
- Glasgow CG, Avila NA, Lin JP, Stylianou MP, Moss J. Serum vascular endothelial growth factor-D levels in patients with lymphangioleiomyomatosis reflect lymphatic involvement. *Chest*. 2009;135(5):1293-1300. <https://doi.org/10.1378/chest.08-1160>
- Leeflang MM, Deeks JJ, Gatsonis C, Bossuyt PM; Cochrane Diagnostic Test Accuracy Working Group. Systematic reviews of diagnostic test accuracy. *Ann Intern Med*. 2008;149(12):889-897. <https://doi.org/10.7326/0003-4819-149-12-200812160-00008>
- Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation [published correction appears in *BMJ*. 2016 Jul 21;354:i4086]. *BMJ*. 2015;350:g7647. <https://doi.org/10.1136/bmj.g7647>
- Shrestha BM. Systematic Reviews and Meta-analysis: Principles and Practice. *JNMA J Nepal Med Assoc*. 2019;57(215):1-2. <https://doi.org/10.31729/jnma.3986>
- Yamada Y, Nezu J, Shimane M, Hirata Y. Molecular cloning of a novel vascular endothelial growth factor, VEGF-D. *Genomics*. 1997;42(3):483-488. <https://doi.org/10.1006/geno.1997.4774>
- Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med*. 2011;155(8):529-536. <https://doi.org/10.7326/0003-4819-155-8-201110180-00009>
- Schünemann HJ, Oxman AD, Brozek J, Glasziou P, Jaeschke R, Vist GE, et al. Grading quality of evidence and strength of recommendations for diagnostic tests and strategies. *BMJ*. 2008;336(7653):1106-1110. <https://doi.org/10.1136/bmj.39500.677199.AE>
- Leeflang MM. Systematic reviews and meta-analyses of diagnostic test accuracy. *Clin Microbiol Infect*. 2014;20(2):105-113. <https://doi.org/10.1111/1469-0691.12474>
- Negeri ZF, Shaikh M, Beyene J. Bivariate random-effects meta-analysis models for diagnostic test accuracy studies using arcsine-based transformations. *Biom J*. 2018;60(4):827-844. <https://doi.org/10.1002/bimj.201700101>
- Rosman AS, Korsten MA. Application of summary receiver operating characteristics (sROC) analysis to diagnostic clinical testing. *Adv Med Sci*. 2007;52:76-82.
- Jia CE, Zhang HP, Lv Y, Liang R, Jiang YQ, Powell H, et al. The Asthma Control Test and Asthma Control Questionnaire for assessing asthma control: Systematic review and meta-analysis. *J Allergy Clin Immunol*. 2013;131(3):695-703. <https://doi.org/10.1016/j.jaci.2012.08.023>
- Willis BH, Quigley M. Uptake of newer methodological developments and the deployment of meta-analysis in diagnostic test research: a systematic review. *BMC Med Res Methodol*. 2011;11:27. <https://doi.org/10.1186/1471-2288-11-27>
- Deeks JJ, Macaskill P, Irwig L. The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. *J Clin Epidemiol*. 2005;58(9):882-893. <https://doi.org/10.1016/j.jclinepi.2005.01.016>
- Young LR, Vandyke R, Gulleman PM, Inoue Y, Brown KK, Schmidt LS, et al. Serum vascular endothelial growth factor-D prospectively distinguishes lymphangioleiomyomatosis from other diseases. *Chest*. 2010;138(3):674-681. <https://doi.org/10.1378/chest.10-0573>
- Chang WY, Cane JL, Blakey JD, Kumaran M, Pounton KS, Johnson SR. Clinical utility of diagnostic guidelines and putative biomarkers in lymphangioleiomyomatosis. *Respir Res*. 2012;13(1):34. <https://doi.org/10.1186/1465-9921-13-34>

31. Radzikowska E, Jaguś P, Skoczylas A, Sobiecka M, Chorostowska-Wynimko J, Wiatr E, et al. Role of serum vascular endothelial growth factor D in discrimination of patients with polycystic lung diseases. *Pol Arch Med Wewn.* 2013;123(10):533-538. <https://doi.org/10.20452/pamw.1927>
32. Cottin V, Lacroix J, Denis L, Jaffray P, Khouatra C, Guelminger R, et al. Diagnostic value of serum VEGF-D in consecutive patients with lymphangioleiomyomatosis. *Am J Respir Crit Care Med.* 2011;183:A1636. https://doi.org/10.1164/ajrcm-conference.2011.183.1_MeetingAbstracts.A1636
33. Daccord C, Aubert V, Félix N, Lazor R. Serum VEGF-D level as a diagnostic biomarker of lymphangioleiomyomatosis in clinical practice. *Chest.* 2017;151(Suppl):P235. <https://doi.org/10.1016/j.chest.2017.04.142>
34. Hirose M, Matsumuro A, Arai T, Sugimoto C, Akira M, Kitaichi M, et al. Serum vascular endothelial growth factor-D as a diagnostic and therapeutic biomarker for lymphangioleiomyomatosis. *PLoS One.* 2019;14(2):e0212776. <https://doi.org/10.1371/journal.pone.0212776>
35. Mou Y, Huang J, Ye L, Wang J, Jin M. Role of serum vascular endothelial growth factor-D (VEGF-D) in the differential diagnosis of polycystic lung disease (PLD). *Fudan Univ J Med Sci.* 2019;46(01):47-52. <https://doi.org/10.3969/j.issn.1672-8467.2019.01.008>
36. Amaral AF, de Oliveira MR, Dias OM, Arimura FE, Freitas CSG, Acencio MMP, et al. Concentration of Serum Vascular Endothelial Growth Factor (VEGF-D) and Its Correlation with Functional and Clinical Parameters in Patients with Lymphangioleiomyomatosis from a Brazilian Reference Center. *Lung.* 2019;197(2):139-146. <https://doi.org/10.1007/s00408-018-00191-3>
37. Young LR, Inoue Y, McCormack FX. Diagnostic potential of serum VEGF-D for lymphangioleiomyomatosis. *N Engl J Med.* 2008;358(2):199-200. <https://doi.org/10.1056/NEJMc0707517>
38. Xu KF, Lo BH. Lymphangioleiomyomatosis: differential diagnosis and optimal management. *Ther Clin Risk Manag.* 2014;10:691-700. <https://doi.org/10.2147/TCRM.S50784>
39. Leeftang MM, Moons KG, Reitsma JB, Zwinderman AH. Bias in sensitivity and specificity caused by data-driven selection of optimal cutoff values: mechanisms, magnitude, and solutions. *Clin Chem.* 2008;54(4):729-737. <https://doi.org/10.1373/clinchem.2007.096032>
40. Hu S, Wu X, Xu W, Tian X, Yang Y, Wang ST, et al. Long-term efficacy and safety of sirolimus therapy in patients with lymphangioleiomyomatosis. *Orphanet J Rare Dis.* 2019;14(1):206. <https://doi.org/10.1186/s13023-019-1178-2>
41. Nijmeh J, El-Chemaly S, Henske EP. Emerging biomarkers of lymphangioleiomyomatosis. *Expert Rev Respir Med.* 2018;12(2):95-102. <https://doi.org/10.1080/17476348.2018.1409622>