



Universitat de Lleida

## Precisión diagnóstica de la FDG-PET para diferenciar derrames pleurales malignos de benignos: revisión sistemática y meta-análisis

Paula Judit Hernández Ferrer

<http://hdl.handle.net/10803/385620>

**ADVERTIMENT.** L'accés als continguts d'aquesta tesi doctoral i la seva utilització ha de respectar els drets de la persona autora. Pot ser utilitzada per a consulta o estudi personal, així com en activitats o materials d'investigació i docència en els termes establerts a l'art. 32 del Text Refós de la Llei de Propietat Intel·lectual (RDL 1/1996). Per altres utilitzacions es requereix l'autorització prèvia i expressa de la persona autora. En qualsevol cas, en la utilització dels seus continguts caldrà indicar de forma clara el nom i cognoms de la persona autora i el títol de la tesi doctoral. No s'autoritza la seva reproducció o altres formes d'explotació efectuades amb finalitats de lucre ni la seva comunicació pública des d'un lloc aliè al servei TDX. Tampoc s'autoritza la presentació del seu contingut en una finestra o marc aliè a TDX (framing). Aquesta reserva de drets afecta tant als continguts de la tesi com als seus resums i índexs.

**ADVERTENCIA.** El acceso a los contenidos de esta tesis doctoral y su utilización debe respetar los derechos de la persona autora. Puede ser utilizada para consulta o estudio personal, así como en actividades o materiales de investigación y docencia en los términos establecidos en el art. 32 del Texto Refundido de la Ley de Propiedad Intelectual (RDL 1/1996). Para otros usos se requiere la autorización previa y expresa de la persona autora. En cualquier caso, en la utilización de sus contenidos se deberá indicar de forma clara el nombre y apellidos de la persona autora y el título de la tesis doctoral. No se autoriza su reproducción u otras formas de explotación efectuadas con fines lucrativos ni su comunicación pública desde un sitio ajeno al servicio TDR. Tampoco se autoriza la presentación de su contenido en una ventana o marco ajeno a TDR (framing). Esta reserva de derechos afecta tanto al contenido de la tesis como a sus resúmenes e índices.

**WARNING.** Access to the contents of this doctoral thesis and its use must respect the rights of the author. It can be used for reference or private study, as well as research and learning activities or materials in the terms established by the 32nd article of the Spanish Consolidated Copyright Act (RDL 1/1996). Express and previous authorization of the author is required for any other uses. In any case, when using its content, full name of the author and title of the thesis must be clearly indicated. Reproduction or other forms of for profit use or public communication from outside TDX service is not allowed. Presentation of its content in a window or frame external to TDX (framing) is not authorized either. These rights affect both the content of the thesis and its abstracts and indexes.

# **Precisión diagnóstica de la FDG-PET para diferenciar derrames pleurales malignos de benignos:**

Revisión sistemática y meta-análisis



**PAULA JUDIT HERNÁNDEZ FERRER**

TESIS DOCTORAL, 2015

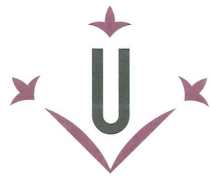
Programa de Doctorado en Recerca Clínica en Medicina

Departamento de Medicina, Universitat de Lleida

**DIRECTOR: JOSÉ MANUEL PORCEL PÉREZ**

Profesor Titular de Medicina. Departamento de Medicina,

Universitat de Lleida



**D. José Manuel Porcel Pérez**, Profesor Titular de Medicina de la Universitat de  
Lleida

### CERTIFICA

Que el estudio que lleva por título **“Precisión diagnóstica de la FDG-PET para diferenciar derrames pleurales malignos de benignos: revisión sistemática y meta-análisis”** y del que es autora la licenciada **Paula Judith Hernández Ferrer**, se ha realizado bajo mi dirección para aspirar al grado de **Doctor en Medicina** y se encuentra en condiciones de ser defendido ante el tribunal correspondiente.

Lo que firmo en Lleida a 30 de septiembre de 2015

Fdo: José Manuel Porcel Pérez

# ÍNDICE

## ÍNDICE

<b>TÍTULO.....</b>	<b>2</b>
<b>ÍNDICE.....</b>	<b>4</b>
<b>AGRADECIMIENTOS.....</b>	<b>7</b>
<b>ÍNDICE DE TABLAS.....</b>	<b>8</b>
<b>ÍNDICE DE FIGURAS.....</b>	<b>9</b>
<b>ABREVIATURAS.....</b>	<b>10</b>
<b>RESUMEN.....</b>	<b>11</b>
<b>INTRODUCCIÓN.....</b>	<b>18</b>
<b>A. Derrame pleural maligno.....</b>	<b>19</b>
<b>B. Tomografía por emisión de positrones.....</b>	<b>25</b>
<b>C. Meta-análisis de precisión diagnóstica.....</b>	<b>27</b>
<b>OBJETIVOS E HIPÓTESIS.....</b>	<b>33</b>
<b>PACIENTES Y MÉTODOS.....</b>	<b>35</b>
<b>A. Estrategia de búsqueda.....</b>	<b>36</b>
<b>B. Criterios de inclusión.....</b>	<b>38</b>
<b>C. Criterios de exclusión.....</b>	<b>39</b>
<b>D. Extracción de datos.....</b>	<b>40</b>
<b>E. Evaluación de la calidad metodológica.....</b>	<b>41</b>
<b>F. Análisis estadístico.....</b>	<b>41</b>

<b>RESULTADOS.....</b>	<b>44</b>
<b>A. Características de los estudios incluidos.....</b>	<b>45</b>
<b>B. Índices de precisión diagnóstica para los DPs malignos</b>	<b>55</b>
1. Método de interpretación de la imagen.....	55
1.1. Análisis cualitativo.....	55
1.2. Análisis semicuantitativo.....	56
2. Uso de la PET vs PET-TC integrado.....	58
2.1. PET.....	58
2.2. PET-TC integrado.....	59
3. Etiología de la patología pleural.....	60
<b>C. Riesgo de sesgo y heterogeneidad.....</b>	<b>62</b>
<b>D. Aplicación clínica.....</b>	<b>65</b>
<b>DISCUSIÓN.....</b>	<b>67</b>
<b>A. Resumen de la evidencia.....</b>	<b>68</b>
<b>B. Puntos fuertes y limitaciones de la revisión.....</b>	<b>70</b>
<b>C. Comparación con investigaciones previas.....</b>	<b>72</b>
<b>D. Estudio posterior.....</b>	<b>76</b>
<b>CONCLUSIONES.....</b>	<b>78</b>
<b>BIBLIOGRAFÍA.....</b>	<b>80</b>
<b>APÉNDICES.....</b>	<b>88</b>
<b>Apéndice 1: QUADAS-2.....</b>	<b>89</b>
<b>Apéndice 2: Artículos excluidos.....</b>	<b>96</b>
<b>Apéndice 3: Artículo relacionado con la tesis .....</b>	<b>101</b>

## AGRADECIMIENTOS

- Me gustaría que estas líneas sirvieran para expresar mi más profundo y sincero agradecimiento a todas aquellas personas que con su ayuda han colaborado en la realización del presente trabajo; en especial, al Dr José Manuel Porcel, director de esta tesis, por la orientación, seguimiento y la supervisión de la misma, pero sobretodo por la motivación y el apoyo recibido a lo largo de estos años. Trabajar a su lado es un estímulo constante para mejorar. Gracias por el ánimo infundido y la confianza en mí depositada.
- A los médicos, residentes y adjuntos, con los que he compartido incontables horas de trabajo estos años y en la actualidad; en especial, a la Dra Silvia Bielsa, trabajadora incansable, amiga y gran persona y a la Dra Silvia Caro por pensar en mi. Ha sido un honor contar con su apoyo, amistad y consejos.
- A Montserrat Martínez, por su fundamental desarrollo del análisis estadístico, sin su ayuda esta tesis no hubiera podido llevarse a cabo.
- A la Dra Sara Muñoz y a Marta Rueda, por su colaboración en las traducciones del alemán, inglés y catalán.
- A mi familia, padres y hermana, por su paciencia y por ser fuente de apoyo constante e incondicional en todos los aspectos de mi vida. Todo lo que soy, es gracias a ellos.
- A José Mari, por su apoyo y amor sincero, por acompañarme y escucharme en los buenos y malos momentos, por estar siempre a mi lado y por ser el mejor compañero de viaje junto a nuestra preciosa hija Candela y el bebé que está en camino.

## ÍNDICE DE TABLAS

<b>Tabla 1:</b> Etiología del derrame pleural (Porcel et al, 2014) .....	20
<b>Tabla 2:</b> Causas de derrame pleural maligno y rentabilidad de la citología del líquido pleural (Porcel et al, 2014) .....	21
<b>Tabla 3:</b> Lista de verificación de la guía PRISMA .....	36
<b>Tabla 4:</b> Estrategia de búsqueda bibliográfica en PUBMED .....	37
<b>Tabla 5:</b> Datos analizados .....	40
<b>Tabla 6:</b> Características de los 14 estudios incluidos en la evaluación de la precisión diagnóstica de la PET en DPs .....	47
<b>Tabla 7:</b> Características de los 13 estudios excluidos en la evaluación de la precisión diagnóstica de la PET en DPs .....	50
<b>Tabla 8:</b> Etiología de los DPs malignos .....	53
<b>Tabla 9:</b> Etiología de los DPs benignos .....	53
<b>Tabla 10:</b> Etiología y método de interpretación de la imagen .....	54
<b>Tabla 11:</b> Resumen de las medidas de precisión diagnóstica de la FDG-PET para identificar DPs malignos .....	61
<b>Tabla 12:</b> Sensibilidad y especificidad de los 27 estudios iniciales según el método de análisis de la imagen de la FDG-PET .....	70
<b>Tabla 13:</b> Características de los 16 artículos inicialmente incluidos (Treglia et al, 2014-A) .....	72
<b>Tabla 14:</b> Aspectos técnicos de los 16 artículos inicialmente incluidos (Treglia et al, 2014-A) .....	73
<b>Tabla 15:</b> Características de los 8 artículos (Treglia G, 2014-B) ....	74
<b>Tabla 16:</b> Aspectos técnicos de los 8 artículos(Treglia G, 2014-B)..	74



## ÍNDICE DE FIGURAS

<b>Figura 1:</b> PET normal (Suárez et al, 2004) .....	26
<b>Figura 2:</b> Forest Plot (Bravata et al, 2007) .....	30
<b>Figura 3:</b> Curva ROC-resumen (Van den Bruel et al, 2011) .....	32
<b>Figura 4:</b> Diagrama de flujo de la estrategia de búsqueda y selección de los estudios .....	46
<b>Figura 5:</b> Forest plot de sensibilidad y especificidad de la FDG-PET con el método visual para DPs malignos .....	56
<b>Figura 6:</b> Curva S-ROC de la metodología visual del FDG-PET en DPs malignos .....	56
<b>Figura 7:</b> Forest plot de sensibilidad y especificidad de la FDG-PET con el método semicuantitativo para DPs malignos .....	57
<b>Figura 8:</b> Curva S-ROC de la metodología semicuantitativa de la FDG-PET para DPs malignos .....	58
<b>Figura 9:</b> Curva S-ROC de la metodología visual y semicuantitativa de la FDG-PET en DPS malignos .....	62
<b>Figura 10:</b> Calidad metodológica mediante QUADAS-2 .....	63
<b>Figura 11:</b> Funnel plot LR positiva para el método cualitativo .....	64
<b>Figura 12:</b> Funnel plot LR negativa para el método cualitativo .....	64
<b>Figura 13:</b> Funnel plot LR negativa para el método SUV .....	65
<b>Figura 14:</b> Forest plot de sensibilidad y especificidad de la FDG-PET o PET-TC (Treglia et al, 2014-A) .....	73
<b>Figura 15:</b> Curva ROC-resumen de la valoración de las anormalidades pleurales (Treglia et al, 2014-B) .....	75
<b>Figura 16-19</b> Presentación de los resultados del QUADAS-2 .....	94-95

## **ABREVIATURAS**

- AUC = area under the curve = área bajo la curva
- DOR = diagnostic odds ratio = razón de probabilidades de diagnóstico
- DP = derrame pleural
- DPM = derrame pleural maligno
- FDG = fluorodeoxiglucosa
- IC 95% = intervalo de confianza del 95%
- LP = líquido pleural
- LR = likelihood ratio = razón de verosimilitud
- PET = tomografía por emisión de positrones
- QUADAS = quality assessment of diagnostic accuracy studies
- SUV = standardized uptake value = valor de captación estandarizada
- TC = tomografía computarizada

# **RESUMEN**

## **ABSTRACT**

### **Background**

The role of fluorodeoxyglucose (FDG)-PET imaging for diagnosing malignant pleural effusions is not well defined. The aim of this study was to summarize the evidence for its use in ruling in or out the malignant origin of a pleural effusion or thickening.

### **Methods**

A meta-analysis was conducted of diagnostic accuracy studies published in the Cochrane Library, PubMed, and Embase (inception to June 2013) without language restrictions. Two investigators selected studies that had evaluated the performance of FDG-PET imaging in patients with pleural effusions or thickening, using pleural cytopathology or histopathology as the reference standard for malignancy. Subgroup analyses were conducted according to FDG-PET imaging interpretation (qualitative or semiquantitative), PET imaging equipment (PET vs integrated PET-CT imaging), and/or target population (known lung cancer or malignant pleural mesothelioma). Study quality was assessed using Quality Assessment of Diagnostic Accuracy Studies-2. We used a bivariate random-effects model for the analysis and pooling of diagnostic performance measures across studies.

## **Results**

Fourteen non-high risk of bias studies, comprising 407 patients with malignant and 232 with benign pleural conditions, met the inclusion criteria. Semiquantitative PET imaging readings had a significantly lower sensitivity for diagnosing malignant effusions than visual assessments (82% vs 91%;  $P=0.026$ ). The pooled test characteristics of integrated PET-CT imaging systems using semiquantitative interpretations for identifying malignant effusions were: sensitivity, 81%; specificity, 74%; positive likelihood ratio (LR), 3.22; negative LR, 0.26; and area under the curve, 0.838. Resultant data were heterogeneous, and spectrum bias should be considered when appraising FDG-PET imaging operating characteristics

## **Conclusions**

The moderate accuracy of PET-CT imaging using semiquantitative readings precludes its routine recommendation for discriminating malignant from benign pleural effusions. (Porcel et al, 2015-A).

## **RESUMEN**

### **Objetivo**

El papel de la FDG-PET para diagnosticar derrames pleurales malignos no está bien definido. El objetivo de este estudio fue el de resumir la evidencia bibliográfica que existe acerca del uso de esta técnica de imagen para diagnosticar o descartar la naturaleza maligna de un derrame o engrosamiento pleural.

### **Métodos**

Este estudio es un meta-análisis de artículos de eficacia diagnóstica extraídos de las bases de datos de la Biblioteca Cochrane, PubMed y Embase, desde su comienzo hasta junio de 2013, sin restricciones de lenguaje. Dos investigadores (JMP y PHF) seleccionaron los estudios que evaluaban la FDG-PET en pacientes con derrame y/o engrosamientos pleurales, usando la citología o histopatología como método de referencia para diagnosticar malignidad. Se llevaron a cabo una serie de análisis por subgrupos de acuerdo al método de interpretación de la imagen de la FDG-PET (cualitativa/visual o semicuantitativa), al tipo de PET (PET o PET-TC integrado) y/o a la población de estudio (pacientes con cáncer de pulmón conocido o mesotelioma). La calidad del meta-análisis se determinó mediante el QUADAS-2. Se usó un modelo bivariado de efectos aleatorios para el análisis y el cálculo de las medias ponderadas de eficacia diagnóstica de los estudios.

## **Resultados**

Catorce estudios que no poseían un alto riesgo de sesgo cumplieron los criterios de inclusión. Incluyeron un total de 407 pacientes con procesos pleurales malignos y 232 con enfermedades pleurales benignas. La lectura semicuantitativa de la PET tuvo significativamente menor sensibilidad para el diagnóstico de derrame maligno que la evaluación visual (82% vs 91%,  $p=0,026$ ). Las características globales del sistema PET-TC integrado usando una interpretación visual para identificar derrames malignos fueron: sensibilidad 81%, especificidad 74%, *likelihood ratio* (LR) positiva 3,22, LR negativa 0,26, y área bajo la curva 0,838. Los datos fueron heterogéneos y se debe considerar la existencia de un sesgo de espectro diagnóstico cuando se analizan las características operativas de la FDG-PET.

## **Conclusiones**

La precisión moderada de la PET-TC, con la interpretación semicuantitativa de las imágenes, hace que no se pueda recomendar de forma rutinaria para la diferenciación de los derrames pleurales malignos de los benignos.

## **RESUM**

### **Objectiu**

El paper de la FDG-PET per a diagnosticar els vessaments pleurals malignes no est3 ben definit. L'objectiu d'aquest estudi va ser resumir la evid3ncia bibliogr3fica que existeix sobre l'ús d'aquesta t3cnica d'imatge per a diagnosticar o descartar la naturalesa maligna d'un vessament i/o engruiximent pleural.

### **M3todes**

Aquest estudi 3s un meta-an3lisi d'articles d'efic3cia diagn3stica extrets de les bases de dades de la Biblioteca Cochrane, PubMed i Embase, des dels seus inicis fins al juny de 2013, i sense restriccions de llenguatge. Dos investigadors (JMP i PHF) van seleccionar els estudis que avaluaven la FDG-PET en pacients amb vessament i/o engruiximent pleural, utilitzant la citologia o histopatologia com a m3tode de refer3ncia per a diagnosticar malignitat. Es van dur a terme un seguit d'an3lisis per subgrups d'acord amb el m3tode d'interpretaci3 de la imatge de la FDG-PET (qualitativa o semiquantitativa), el tipus de PET (PET o PET-TC integrat) i/o la poblaci3 d'estudi (pacients amb c3ncer de pulm3 conegut o mesotelioma). La qualitat del meta-an3lisi es va determinar mitjançant el QUADAS-2. Es va utilitzar un model bivariabile d'efectes aleatoris per a l'an3lisi i el c3lcul de les mitges ponderades d'efic3cia diagn3stica dels estudis.



## **Resultats**

Catorze estudis que no posseïen un alt risc de biaix complien els criteris d'inclusi3. Van incloure un total de 407 pacients amb processos pleurals malignes i 232 amb malalties pleurals benignes. La lectura semiquantitativa de la PET va tenir significativament menys sensibilitat per al diagn3stic de vessaments malignes que l'avaluaci3 visual (82 % vs 91%,  $p=0,026$ ). Les característiques globals del sistema PET-TC integrat usant una interpretaci3 visual per a identificar vessaments malignes van ser: sensibilitat 81%, especificitat 74%, *likekihood ratio* (LR) positiva 3,22, LR negativa 0,26, i l'àrea sota la corba 0,838. Les dades van ser heterogènies i s'ha de considerar l'existència d'un biaix d'espectre diagn3stic quan s'analitzen les característiques operatives de la FDG-PET.

## **Conclusions**

La precisi3 moderada de la PET-TC, amb la interpretaci3 semiquantitativa de la imatge, suggereix que no es pot recomenar de forma rutinària per a diferenciar els vessaments pleurals malignes dels benignes.

# INTRODUCCIÓN

## **A. DERRAME PLEURAL MALIGNO**

El derrame pleural (DP) es un problema clínico frecuente entre los pacientes que se atienden en los servicios de Medicina Interna y Neumología. Sus causas pueden variar según:

- *el área geográfica*: la tuberculosis pleural, dentro del estado español, predomina en regiones con una mayor prevalencia de la enfermedad, como Galicia (Valdés et al, 1996).

- *el ámbito de atención sanitaria*: los pacientes evaluados en un Servicio de Medicina Interna presentan una mayor proporción de insuficiencia cardiaca como etiología del DP, respecto a los Servicios de Neumología (Porcel, 2013-A).

- *la edad del sujeto*: la mitad de los DPs en pacientes jóvenes (menores de 34 años) son tuberculosos y, en los octogenarios, la causa más frecuente de DP es la insuficiencia cardiaca (45%) (Porcel et al, 2014).

- *el periodo de tiempo estudiado*: en esta última década existe un descenso progresivo en el número de casos de tuberculosis (Valdés et al, 2012). Por lo otro lado, se ha observado un aumento de los casos de insuficiencia cardiaca (29% de los DPs en 2011 frente a un 15% antes del 2000, en una serie reciente. (Porcel et al, 2014).

Un reciente estudio epidemiológico retrospectivo evaluó más de 3.000 pacientes consecutivos que se sometieron a una toracocentesis diagnóstica en un Servicio de Medicina Interna, durante un período de 19 años (Porcel et al, 2014). Las tres cuartas partes de los pacientes con DP tenían una de las siguientes cuatro causas: neoplasia (27%), insuficiencia cardiaca (21%), neumonía (19%) o tuberculosis (9%). No se observaron diferencias significativas en la etiología del DP según el sexo (Tabla 1).

**Tabla 1: Etiología del derrame pleural (Porcel et al, 2014)**

<b>Causas</b>	<b>N (%)</b>	<b>Edad, mediana (cuartiles)</b>
Cáncer	840 (27)	70 (59-79)
Insuficiencia cardíaca	640 (21)	80 (74-85)*
Neumonía/Infección	582 (19)	57 (42-73)
<i>DPPNC</i>	202	
<i>DPPC</i>	216	
<i>Empiema</i>	164	
Tuberculosis	275 (9)	32 (25-44)**
<i>Demostrada</i>	96	
<i>Probable</i>	179	
Poscirugía abdominal	110 (4)	69 (53-77)
Enfermedades del pericardio	109 (4)	68 (54-77)
Hidrotórax hepático	97 (3)	67 (55-76)
Idiopático	94 (3)	70 (53-79)
Traumatismo	77 (2,5)	71 (42-79)
Tromboembolia pulmonar	48 (1,6)	70 (48-81)
Poscirugía cardíaca	32 (1)	68 (60-72)
Conectivopatías	27 (0,9)	50 (44-67)
Miscelánea <sup>a</sup>	146 (5)	70 (52-79)
<b>Total</b>	<b>3077</b>	<b>69 (52-79)</b>

DPPC=derrame pleural paraneumónico complicado; DPPNC=derrame pleural paranéumónico no complicado.

Miscelánea<sup>a</sup>: Incluye, entre otros, hipoproteinemia (17), síndrome nefrótico (15), atelectasia (14), pancreatitis (14), pulmón atrapado (13), sobrecarga volumétrica (10), hipertensión pulmonar (8) y uremia (8).

\*: Significativamente superior al resto de los grupos (p<0,01).

\*\* : Significativamente inferior al resto de los grupos (p<0,01).

De los 840 pacientes con DP maligno, en aproximadamente la mitad, los tumores primarios fueron pulmonares (37%) o mamario (16%), como muestra la Tabla 2. La prevalencia de mesotelioma fue baja (3% de todos los DPs malignos). Tanto los tumores de pulmón como los mesoteliomas predominaron en varones (80% y 86% de los casos, respectivamente). Al menos la mitad de los cánceres de pulmón que metastatizaron a pleura fueron adenocarcinomas.

**Tabla 2: Causas de derrame pleural maligno y rentabilidad de la citología del líquido pleural (Porcel et al, 2014)**

<b>Tumor primario</b>	<b>N (%)</b>	<b>Rentabilidad citología del líquido pleural (%)</b>
Pulmón	309 (37)	177/304 (58)
<i>Adenocarcinoma</i>	150	116/148 (78)
<i>Carcinoma escamoso</i>	67	17/67 (25)*
<i>Carcinoma de célula pequeña</i>	37	19/36 (53)
<i>Otros<sup>a</sup></i>	55	25/53 (47)
Mama	137 (16)	92/136 (68)
Origen desconocido	82 (10)	57/82 (70)
Hematológico <sup>b</sup>	82 (10)	49/80 (61)
Gastrointestinal <sup>c</sup>	71 (8)	34/71 (48)
Ovario	59 (7)	41/59 (70)
Mesotelioma	22 (3)	6/22 (27)**
Miscelánea <sup>d</sup>	78 (9)	35/77 (45)
<b>Total</b>	<b>840</b>	<b>491/831 (59)</b>

LP: líquido pleural

<sup>a</sup>: comprende desconocido (34), célula no pequeña sin concretar estirpe específica (18), indiferenciado (2) y sarcoma (1).

<sup>b</sup>: Incluye linfoma no Hodgkin (61), leucemias (9), síndromes mielodisplásicos (5), mieloma múltiple (5) y linfoma Hodgkin (2).

<sup>c</sup>: Incluye colon (25), estómago (22), páncreas (15), esófago (5), vía biliar (3) y tumor del estroma gastrointestinal (1).

<sup>d</sup>: Incluye riñón (21), endometrio (12), melanoma (10), cabeza y cuello (9), próstata (6), sarcomas (6), vejiga (4), timoma (3), tiroides (3), vulva (2), suprarrenal (1) y hemangioendotelioma pleural (1).

\*: Significativamente inferior que el resto de los tipos histológicos de cáncer de pulmón ( $p < 0,001$ ).

\*\* : Significativamente inferior que el resto de los grupos ( $p = 0,002$ ).

En esta serie, una primera citología del líquido pleural fue positiva en 424 (51%) de 840 pacientes con DP maligno. En 214 pacientes con una primera citología pleural negativa una segunda resultó diagnóstica en 55 casos (26%). Finalmente, en 52 pacientes con dos citologías negativas, una tercera muestra fue positiva en 12 casos (23%). En conjunto, teniendo en cuenta todas las citologías que se realizaron, la rentabilidad de esta prueba para identificar malignidad fue del 59% (491/831). Estas cifras concuerdan con estudios previos sobre la rentabilidad de las sucesivas citologías descritos con anterioridad por el mismo grupo (Bielsa et al, 2008). La utilidad del estudio citológico del líquido pleural resultó significativamente inferior en mesoteliomas (27%) y carcinomas escamosos de pulmón (25%). Por lo tanto, la citología pleural es un método sencillo para diagnosticar malignidad pero ofrece resultados falsos negativos en al menos el 40% de los casos.

De este modo, en muchas ocasiones es necesaria una biopsia pleural guiada para conseguir el diagnóstico definitivo de malignidad. Este procedimiento es invasivo y, cuando se realiza mediante toracoscopia, no está disponible en todos los centros. Este es uno de los motivos, por el que sería de gran ayuda a los clínicos:

- Identificar aquellos pacientes con una baja probabilidad de que el DP sea maligno. De este modo se disminuiría el número de procedimientos invasivos innecesarios.
- A la inversa, también es importante reconocer aquellos pacientes con alta probabilidad de DP maligno para justificar la realización urgente de una toracoscopia.

Las pruebas de imagen pueden ayudar a perfilar la probabilidad de que un DP sea o no maligno (Porcel et al, 2013-B):

- *Radiografía de tórax*: existen ciertos signos que nos pueden ayudar al diagnóstico diferencial de un DP.
  - DP bilateral: su causa más frecuente es la insuficiencia cardiaca. Otros signos radiológicos que apoyan este diagnóstico son la redistribución vascular, edema intersticial o alveolar y la cardiomegalia. El diagnóstico diferencial de un DP bilateral con silueta cardiaca normal, debe incluir malignidad y, menos frecuentemente, pleuritis lúpica y pericarditis constrictiva.
  - DP masivo: el 60% de los DPs masivos son secundarios a cáncer, seguidos de los DPs paraneumónicos (20%) y, en menor proporción, tuberculosis e hidrotórax hepáticos.
  - DP loculados: ocurren en aquellas situaciones que provocan una gran inflamación pleural, como el DP paraneumónico, empiemas, hemotórax y tuberculosis. No obstante, también pueden localarse los DPs malignos, tromboembolias pulmonares e incluso la insuficiencia cardiaca. En este último caso, el líquido localizado en una cisura puede simular una masa intraparenquimatosa que desaparece con tratamiento diurético (“tumor fantasma o evanescente”).
- *Ultrasonografía*: la ecografía pleural tiene una mayor sensibilidad que la radiografía de tórax a la hora de detectar un DP. Además, puede distinguir entre líquido pleural, engrosamiento pleural o consolidación pulmonar, evitando una toracocentesis o inserción de un drenaje pleural

innecesarios. Existen ciertos hallazgos que sugieren malignidad: engrosamiento de la pleura parietal >1 cm, nódulos pleurales y engrosamiento diafragmático >7 mm (sensibilidad 42% y especificidad 95% para cada criterio) (Qureshi et al, 2009).

- *Tomografía computarizada (TC)*: para visualizar correctamente el realce de la pleura en la TC, debemos administrar contraste endovenoso y adquirir la imagen a los 60-90 segundos (fase venosa tardía). La presencia de nódulos o engrosamientos pleurales >1 cm, son muy sugestivos de malignidad (Porcel et al, 2015-B). La TC también es útil para mostrar DPs loculados, alteraciones del parénquima pulmonar, adenopatías mediastínicas, afectación pericárdica, tromboembolia pulmonar o enfermedad subdiafragmática. Sin embargo, la presencia de nódulos o engrosamientos pleurales >1 cm es un dato muy específico (>95%), pero no lo suficientemente sensible (43%).
- *Resonancia magnética nuclear (RM)*: su uso está poco extendido en patología pleural y se reserva para casos muy seleccionados o con alergia al contraste yodado. Puede ser útil:
  - Caracterizando un hemotórax, por su capacidad para detectar sangre y determinar el tiempo de sangrado.
  - Para detectar invasión tumoral de la pared torácica o diafragma (superior a la TC), particularmente en sujetos con mesotelioma.

En resumen, estas técnicas de imagen tradicionales tienen limitaciones a la hora de discriminar entre condiciones pleurales malignas y benignas. Los nódulos o engrosamientos pleurales, que son altamente sugestivos de malignidad, están presentes en menos de la mitad de los



casos. Además estas técnicas son dependientes del radiólogo que informa la ecografía o TC.

La tomografía por emisión de positrones (PET), objetivo de este trabajo, ha sido menos estudiada en las enfermedades de la pleura.

## **B. TOMOGRAFÍA POR EMISIÓN DE POSITRONES**

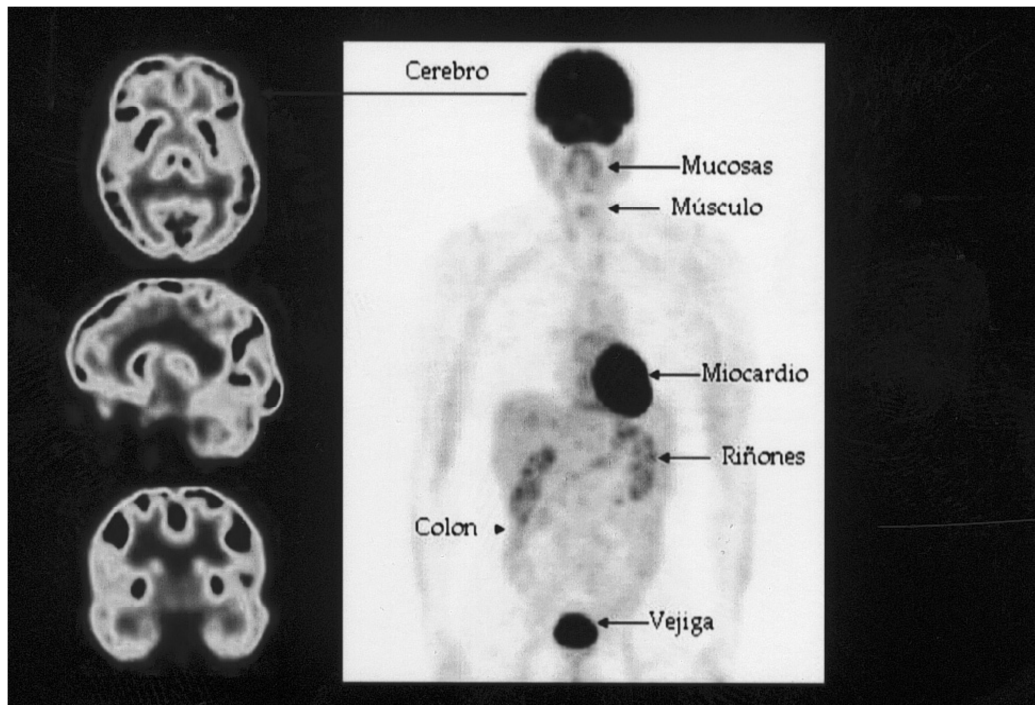
La PET es una técnica de Medicina Nuclear que permite obtener imágenes de la distribución *in vivo* de diferentes moléculas; es lo que se ha denominado “imagen molecular”, un complemento de la información anatómica que proporcionan la TC o la RM. Esto se consigue mediante la administración intravenosa de un radiofármaco, que es la unión de un isótopo radioactivo con una molécula determinada, tras lo cual se realiza la adquisición de las imágenes en una cámara PET.

En la práctica clínica, la molécula que se utiliza es la fluorodeoxiglucosa (FDG) unida al  $^{18}\text{F}$ , isótopo radioactivo emisor de positrones de período de semidesintegración muy corto (110 minutos). La FDG es un análogo de la glucosa y es captada tanto por las células normales como por las cancerígenas, pero no sigue el ciclo bioquímico normal de la glucosa, por lo que queda atrapado en el interior celular.

De forma fisiológica, algunos tejidos como el cerebro, el miocardio o el hígado tienen una alta demanda de glucosa, lo que se va a traducir en una alta retención de FDG (Figura 1). En el caso de las células malignas, su alta tasa metabólica implica igualmente una gran avidéz por la FDG, lo que permite *a priori* diferenciar los tejidos malignos de los benignos (Suárez et al, 2004). La  $^{18}\text{F}$ -FDG es el radiofármaco que ha permitido la introducción de la técnica PET, a pesar de que existen en la actualidad cerca de 3.000 radiofármacos sintetizados.

Por lo tanto, la PET mide la captación de FDG en una lesión, reflejando su consumo glicídico, que se correlaciona con la vascularización tisular, la celularidad y la capacidad proliferativa.

**Figura 1: PET normal (Suárez et al, 2004)**



Los resultados de la PET se expresan:

- de manera *cualitativa*: considera malignas a aquellas lesiones con hipermetabolismo superior al del tejido circundante, utilizando la inspección visual.

- de manera *semicuantitativa*: en forma de valores de captación estandarizados (*Standardized Uptake Value*, [SUV]). Se considera patológico un  $SUV > 3$  en las lecturas de PET; además, por sí mismo, es un parámetro pronóstico independiente del estadio clínico y del tamaño lesional (las lesiones más activas presentan SUV mayores, y valores  $\geq 7$  se asocian a un peor pronóstico) (Cano et al, 2012).

No obstante, la imagen obtenida tiene limitaciones en cuanto a la detección de tumores de muy bajo grado de malignidad, debido

precisamente a su menor avidez por la glucosa. Del mismo modo, también puede existir zonas de elevada actividad metabólica en procesos benignos, inflamatorios o infecciosos, que son causa de falsos positivos.

Además, aunque las cámaras PET poseen una aceptable resolución espacial, cabe señalar que la imagen obtenida no detecta enfermedad microscópica.

La PET surgió en la década de los 70 en Estados Unidos, implantándose en la práctica clínica durante la década de los 90. Se introdujo en España en 1995 (Centro PET Complutense, Madrid). Desde el año 2000, los hospitales públicos de diferentes comunidades autónomas han incorporado esta tecnología a la batería de pruebas diagnósticas existentes.

Las enfermedades oncológicas son la principal indicación clínica de la PET-FDG (> 90%). Dado que refleja el metabolismo de las células cancerígenas, la detección de tumores se presume más precoz que las alteraciones morfológicas detectadas por la TC o la RM. Además, puede evaluar también de forma más precoz la respuesta a la quimioterapia o radioterapia, en comparación con otras técnicas convencionales. La falta de definición anatómica de la PET se puede mitigar con los equipos híbridos PET-TC, que fusionan la información aportada por la PET y la TC (metabólica y anatómica).

### **C. META-ANÁLISIS DE PRECISIÓN DIAGNÓSTICA**

El meta-análisis resume los resultados de varios estudios estimando un promedio o un efecto común que es más preciso que cada estudio de forma independiente (Cochrane Handbook, 2010).

Existen dos tipos de revisiones sistemáticas: de intervención y de eficacia de una prueba diagnóstica. Ésta última, motivo de nuestro estudio, está diseñada para:

1. Evaluar toda la información existente sobre la precisión de una prueba diagnóstica,
2. Estimar lo más exactamente posible dicha precisión e,
3. Investigar las posibles diferencias en la precisión diagnóstica según diferentes subgrupos

Los métodos para la realización de una revisión sistemática para estudios diagnósticos no están tan bien establecidos como aquéllos que se usan en las revisiones sistemáticas de intervenciones. La tendencia actual en las revisiones sistemáticas es comparar las cualidades técnicas entre pruebas diagnósticas, más que intentar proponer una estimación de precisión definitiva para una prueba concreta.

Resumir los resultados de los estudios de precisión diagnóstica es complicado por varias razones (Perera et al, 2011):

- a. Los resultados de las pruebas diagnósticas se obtienen a través de dos valores: *sensibilidad* y *especificidad*, que no son independientes uno del otro.
- b. El punto de corte de positividad del test muchas veces difiere en los diferentes estudios.
- c. La heterogeneidad de los estudios de precisión diagnóstica es frecuente y, generalmente, mayor que en los estudios de intervención. Esta heterogeneidad se suele deber a diferencias en el diseño o en el desarrollo de los estudios. Por este motivo, la valoración precisa de la calidad de cada uno de los estudios incluidos es esencial.

Por lo tanto, las revisiones sistemáticas de los estudios de eficacia diagnóstica comparan toda la evidencia disponible sobre la precisión de una prueba y permiten la investigación de la variabilidad entre subgrupos. Si se cumplen ciertas condiciones, el meta-análisis nos proporcionará una

estimación más precisa de la eficacia de la prueba y diferencias estadísticas del test entre los subgrupos.

Para poder ilustrar los conceptos relativos a la interpretación de un meta-análisis, ponemos como ejemplo la revisión sistemática de los efectos de los pedómetros para aumentar la actividad física (Bravata et al, 2007).

Una buena revisión sistemática incluye:

- una búsqueda sistemática de todos los estudios,
- una evaluación de su calidad metodológica, y
- una selección de los mejores estudios para responder a la pregunta de investigación

La lectura de un meta-análisis puede dividirse en 4 pasos básicos:

### **1. ¿Cuál es la medida usada?**

Se debe realizar una lectura de la escala horizontal del forest plot (Figura 2) para comprobar cual es la medida del efecto que ha sido utilizada y que lado es “bueno” y “malo” para la intervención o prueba diagnóstica. En la revisión de Bravata, et al (2007), el resultado que nos interesa es la diferencia significativa en el número de pasos al día.

### **2. ¿Qué muestra el forest plot?**

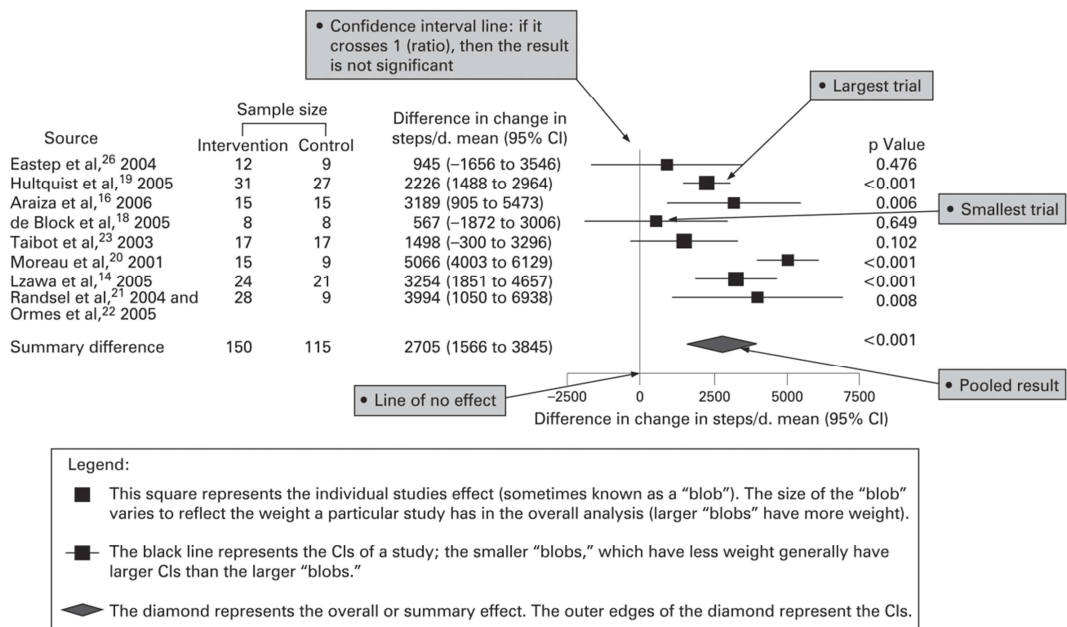
El forest plot o diagrama de bosque es el método de gráfica usado de forma más convencional en los meta-análisis. Representa los resultados de cada estudio por separado y, habitualmente, el efecto promedio de todos los estudios.

La Figura 2 presenta un forest plot pudiendo identificar los componentes básicos. Una línea vertical de “no efecto”, en el centro de la gráfica, representa el resultado si la intervención no ha tenido impacto. Cada estudio incluido se muestra con una línea horizontal y con un cuadrado en el centro. El cuadrado corresponde al efecto del estudio y la línea horizontal a su intervalo de confianza del 95% (IC 95%). Normalmente, los estudios con peor precisión y menor peso tienen IC mayores y, por lo tanto, destacan más en el forest plot a primera vista (Perera et al,

2011). Para intentar remediar esto, el tamaño del cuadrado central es equivalente a la proporción del peso que tiene el estudio en el efecto combinado global del análisis; es decir, los mayores pesos en importancia (cuadrados más grandes) corresponden a aquellos estudios que ofrecen más información (normalmente los de mayor tamaño muestral).

Hay que tener en cuenta que, según el resultado que se esté analizando, un efecto “positivo” puede significar que el cuadrado y el IC 95% caigan a la izquierda o a la derecha de la línea del no efecto; es decir, un efecto positivo no tiene porqué colocarse siempre a un mismo lado. En nuestro ejemplo, un aumento de los pasos al día es considerado como un hallazgo positivo y, por consiguiente, esperamos observar que los estudios positivos se encuentren en el lado derecho, como indica la Figura 2.

**Figura 2: Forest Plot (Bravata et al, 2007)**



### 3. ¿Qué significa el efecto combinado o promedio?

El efecto promedio aparece en forma de rombo o diamante al final del forest plot. El centro del rombo es equivalente al resultado promedio, y sus extremos representan el IC 95% de este mismo efecto.

Este rombo se obtiene de la combinación de las medidas resumen recogidas de todos los estudios incluidos. Con este dato conseguimos un único valor, más preciso, de la pregunta o intervención que nos interesa. Es decir, cada estudio proporciona información real sobre el efecto y, con la combinación de todos ellos, estamos más cerca de “la verdad”. Ya que la información obtenida de cada estudio es diferente, al agrupar la información también se obtiene un promedio del peso que tiene cada estudio.

#### 4. ¿Es válida esta combinación de estudios?

Al obtener un resultado global tras la combinación de varios estudios, surge una pregunta obvia: *¿Es válido combinar estos estudios?* Dos estudios nunca son iguales; todos ellos varían, en menor o mayor manera, en sus “PICO” (población, intervención, comparador y resultado). Analizamos la anterior pregunta en 3 pasos:

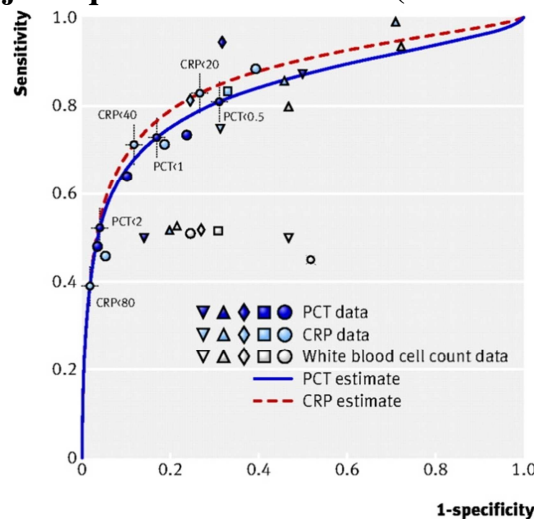
- A. ¿Los estudios son clínicamente similares? Para esto es necesario tener una cierta experiencia en los contenidos de los estudios.
- B. ¿Los estudios son metodológicamente similares? Es necesaria una habilidad para la evaluación crítica de los estudios.
- C. ¿Los resultados estadísticos son similares, es decir, existe homogeneidad en los resultados? Para esto, usamos un test estadístico si la variación es mayor de lo esperado. El chi-cuadrado es la medida más comúnmente utilizada y nos evalúa el grado de heterogeneidad. Como regla general, un chi-cuadrado >50% indica que la combinación de estudios no es válida (aunque este dato depende del número y tamaño de los estudios incluidos). En la revisión de Bravata et al (2007) del pedómetro, la heterogeneidad del resultado principal (aumento en el número de pasos al día) fue elevada: el incremento de pasos en el grupo de intervención fue 2491 (IC 95% 1098 a 3885) y el chi-cuadrado del 91%. Para mejorar estas cifras, los investigadores excluyeron un estudio que aportaba un mayor incremento en la

actividad física respecto a los otros estudios. Con esto se consiguió reducir el resultado a 2004 (IC 95% 878 a 3129).

Además del forest plot, el otro método gráfico que se recomienda a la hora de la realización de un meta-análisis es la curva jerárquica ROC-resumen (S-ROC) que representa la eficacia de la prueba diagnóstica en diferentes puntos de corte (Perera et al, 2011). Esta gráfica, que muestra la relación que existe entre las parejas de sensibilidad y especificidad, es básicamente un diagrama de dispersión de los puntos de estimación de la sensibilidad (eje vertical) y de 1-especificidad (eje horizontal), y, dentro de un estudio, nos presenta los distintos resultados para los diferentes puntos de corte. La Figura 3 es un ejemplo de curva ROC, donde cada estudio está representado por un símbolo (Van den Bruel et al, 2011).

El objetivo de la prueba, en el ámbito clínico, determina la interpretación de los resultados. Por ejemplo, si el fin de la prueba es el cribado de cierta enfermedad, la sensibilidad es más importante que la especificidad y, por lo tanto, la curva s-ROC debería ser lo más alta posible en el espacio ROC. Por lo contrario, cuando la especificidad es más importante porque, por ejemplo, la prueba seleccionará a los pacientes que deberán realizar un tratamiento invasivo, la curva s-ROC deberá estar lo más cerca posible al eje vertical.

**Figura 3: Curva jerárquica ROC-resumen (Van den Bruel et al, 2011)**





## **OBJETIVOS E HIPÓTESIS**

## **OBJETIVOS**

- Realizar una búsqueda sistemática y un resumen de la literatura sobre la eficacia diagnóstica de la FDG-PET para diferenciar malignidad de benignidad en los derrames y/o engrosamientos pleurales con análisis de las variables sensibilidad, especificidad y LR.
- Evaluar cómo la técnica de imagen de la PET puede ayudarnos a disminuir el número de procedimientos invasivos innecesarios (toracoscopia), al descartar malignidad. Asimismo, detectar los pacientes con alta sospecha de malignidad para confirmar este diagnóstico y acelerar la realización de pruebas invasivas para no demorar el diagnóstico definitivo
- Evaluar si existen diferencias en la rentabilidad diagnóstica de la PET según el método de análisis de la imagen (método cualitativo vs semicuantitativo), el tipo de tumor primario del paciente (cáncer de pulmón o mesotelioma) o el uso de máquinas híbridas PET-TC.

## **HIPÓTESIS**

- La FDG-PET puede contribuir al diagnóstico de los DPs de causa incierta para evitar pruebas invasivas.
- Existen diferencias en la precisión diagnóstica de la FDG-PET según la etiología del DP.
- El método de adquisición de la imagen semicuantitativa conlleva una mayor precisión diagnóstica.
- La PET tiene menor rendimiento diagnóstico que el sistema PET-TC integrado.

## **PACIENTES Y MÉTODOS**

## **A. ESTRATEGIA DE BÚSQUEDA**

Se realizó una búsqueda sistemática para identificar todos los artículos originales que hacían referencia al uso de la PET para la diferenciación de DPs o engrosamientos pleurales malignos y benignos, según los métodos y procedimientos de las guías Cochrane y PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) (Tabla 3) (Moher et al, 2009).

**Tabla 3: Lista de verificación de la guía PRISMA**

<b>SECCIÓN</b>	<b>PUNTO</b>	<b>PÁGINA DE LA TESIS</b>
<i>TÍTULO</i>	1	1
<i>RESUMEN</i>	2	9
<i>INTRODUCCIÓN</i>		
Justificación	3	16
Objetivos	4	31
<i>MÉTODOS</i>		
Protocolo y Registro	5	34
Criterios de elegibilidad	6	35
Fuente de información	7	34
Búsqueda	8	34
Selección de estudios	9	34
Proceso de extracción de datos	10	36
Lista de datos	11	35
Riesgo de sesgo en los artículos individuales	12	37
Medidas de resumen	13	37
Método de análisis	14	37
Riesgo de sesgo entre los artículos	15	38

Precisión diagnóstica de la FDG-PET para diferenciar DPs malignos de benignos

Análisis adicionales	16	38
<i>RESULTADOS</i>		
Selección de estudios	17	42
Características de los estudios	18	41
Riesgo de sesgo entre los estudios	19	57
Resultados de los estudios individuales	20	43
Síntesis de los resultados	21	56
Riesgo de sesgo entre los estudios	22	57
Resultados de análisis adicionales	23	51
<i>DISCUSIÓN</i>		
Resumen de la evidencia	24	63
Limitaciones	25	66
Conclusiones	26	73
<i>FINANCIACIÓN</i>	27	No se obtuvo

Se realizó una búsqueda en las bases de datos MEDLINE (PubMed), EMBASE y la Biblioteca Cochrane desde su creación hasta el 17 de junio de 2013, sin restricción de lenguaje, usando los siguientes términos para la misma (Tabla 4):

**Tabla 4: Estrategia de búsqueda bibliográfica en PUBMED**

Estrategia de búsqueda
(pleura or pleural or pleural disease or pleural diseases or pleural effusion or pleural effusions or malignant pleural or malignant pleural effusion or malignant pleural effusions or mesothelioma or malignant mesothelioma or benign pleural or benign pleural effusion or benign pleural effusions) AND (PET or positron emission tomography or fluorodeoxyglucose or 18-fluorodeoxyglucose or FDG PET or 18-FDG PET or 18F-FDG).

FDG=fluorodeoxyglucosa

Asimismo, se realizó una búsqueda manual de las referencias incluidas en las publicaciones originales elegidas.

La búsqueda inicial de los artículos en las bases de datos se realizó por dos investigadores: José Manuel Porcel, internista con más de 25 años de experiencia laboral y científica, y director de esta tesis; y por Paula Hernández, internista con unos 8 años de experiencia. Ambos investigadores seleccionaron de forma independiente los artículos potencialmente aptos y extrajeron los datos de los estudios incluidos. Las discrepancias se resolvieron de forma consensuada.

## **B. CRITERIOS DE INCLUSIÓN**

Los criterios de inclusión que debían cumplir todos los artículos seleccionados fueron:

1. Estudio con un diseño diagnóstico.
2. Inclusión de al menos 10 pacientes adultos para su análisis en total ( $n > 10$ ).
3. Los pacientes debían presentar afectación pleural en forma de DP, engrosamiento o nódulos/masas pleurales.
4. Uso de la FDG-PET como prueba diagnóstica (índice test) para la diferenciación entre lesiones pleurales malignas y benignas.
5. El diagnóstico estándar de referencia en los DPs malignos debía ser la demostración citológica o histológica de células malignas.
6. Los estudios debían proporcionar los datos necesarios para construir tablas de contingencia 2x2 de las que se pudieran calcular las cifras de eficacia diagnóstica.

### **C. CRITERIOS DE EXCLUSIÓN**

Si los estudios no cumplían los criterios de inclusión se descartaron para la elaboración del meta-análisis.

El idioma de los estudios no fue un criterio de exclusión. De los 27 artículos seleccionados inicialmente, 25 estaban escritos en inglés, 1 en alemán y otro en chino (se contactó con el autor de este último artículo para la obtención de datos sobre su estudio) (Liao et al, 2012).

## **D. EXTRACCIÓN DE DATOS**

Los datos que se recogieron de cada artículo seleccionado fueron (Tabla 5):

### **TABLA 5: Datos analizados**

- Título y autores
- Año de publicación
- País de publicación
- Tamaño muestral
- Características demográficas de los pacientes:
  - Edad media
  - Sexo
- Diseño del estudio (prospectivo o retrospectivo)
- Estrategia de reclutamiento (consecutivo o caso-control)
- Tipo de lesiones pleurales por imagen radiológica (DP, engrosamiento y/o masa-nódulo pleural)
- Prevalencia de malignidad pleural en la población a estudio
  - (probabilidad pretest)
- Etiología de los DPs benignos
- Tipo de tumor primario en los DPs malignos
- Estándar de referencia
- Uso de FDG-PET o tecnología combinada PET-TC
- Método de interpretación de la imagen del PET:
  - Cualitativo o visual
  - Semicuantitativo: usando el valor de captación estandarizado (SUV)
  - Dual time point: medición de 2 valores de SUV separados en el tiempo
- Puntuación del QUADAS-2 (Quality Assessment of Diagnostic Accuracy Studies)
- Datos necesarios para la construcción de la tabla 2x2



## **E. EVALUACIÓN DE LA CALIDAD METODOLÓGICA**

Ambos investigadores evaluaron de forma independiente la calidad metodológica de los artículos incluidos mediante el QUADAS-2 (Quality Assessment of Diagnostic Accuracy Studies), herramienta rediseñada del original QUADAS (Whiting et al, 2011), disponible en el página web [www.quadas.org](http://www.quadas.org).

El QUADAS-2 consta de 4 dominios que analizan la posibilidad de de sesgo:

1. Selección de los pacientes (*patient selection*)
2. Prueba diagnóstica a estudio (*index test*)
3. Estándar de referencia (*reference standard*)
4. Flujo de los pacientes a lo largo del estudio y temporalidad de la prueba diagnóstica y de la prueba considerada *gold estándar* (*flow and timing*)

Cada uno de estos dominios analiza el posible sesgo como bajo, alto o poco claro (*low, high and unclear*).

Los tres primeros dominios, además, también son evaluados respecto a su concordancia con la pregunta del estudio (*applicability*) (Apéndice 1).

## **F. ANÁLISIS ESTADÍSTICO**

Las variables continuas se expresaron como medias o medianas, y las variables categóricas como frecuencias o porcentajes. Las sensibilidades, especificidades, razones de verosimilitud o *likelihood ratio* (LR) y la razón de probabilidades (*odds*) diagnósticas (DOR) de los estudios, con sus intervalos de confianza del 95%, se calcularon de las tablas 2x2, usando modelos bivariados de efectos aleatorios. En las tablas en las que existía alguna celda con valor cero, se añadió 0,5 a todas las celdas para poder calcular las LR.

Hemos presentado los datos a través de diagramas de bosque (forest plot) y mediante curvas jerárquicas ROC-resumen o SROC (*summary receiver operating characteristic*).

Las características operativas de la FDG-PET fueron analizadas según:

- El equipo utilizado (PET o PET-TC híbrido)
- La interpretación de la captación del trazador:
  - o Visual o cualitativa: comparación visual con el nivel de actividad del trazador en los tejidos sanos circundantes.
  - o Semicuantitativa: con unidades SUV.
  - o Técnica del “*dual time point*”: donde se comparan las imágenes de captación del trazador en dos momentos diferentes, de forma semicuantitativa.
- Diferentes subgrupos de pacientes: DP asociado a cáncer de pulmón o mesotelioma confirmados.

Los cálculos estadísticos se llevaron a cabo una vez fueron excluidos aquellos artículos con un alto riesgo de sesgo según el QUADAS-2.

Una meta-regresión de efectos aleatorios permitió la comparación de las sensibilidades y especificidades ponderadas entre las diferentes metodologías de lectura de la FDG-PET y los subgrupos de pacientes.

El acuerdo interobservador, en cuanto a la valoración del QUADAS-2, se determinó mediante el coeficiente de Cohen  $k$  no ponderado. El coeficiente kappa de Cohen ( $k$ ) corresponde a la proporción de concordancias observadas sobre el total de observaciones, habiendo excluido las concordancias atribuibles al azar. El coeficiente kappa toma valores entre -1 y +1; cuanto más cercano a +1 se encuentre nuestro valor, mayor es el grado de concordancia interobservador. Por el contrario, un valor de  $k=0$  es la concordancia que se espera a causa exclusivamente del

azar (Cerdeira et al, 2008). Este coeficiente indica un acuerdo “bueno” si el resultado es mayor de 0,6 y un acuerdo “excelente” si es mayor de 0,8.

La heterogeneidad entre los estudios se analizó mediante un modelo de efectos aleatorios con el uso de la función del enlace logit (*logit link function*).

El sesgo de publicación se evaluó mediante un gráfico de embudo (funnel plot) para las LRs positivas y negativas.

La significación estadística se asumió con una  $p < 0,05$ . Los cálculos se realizaron con el software estadístico R-package “mada”. (R-project; <http://cran.r-project.org/web/packages/mada/index.html>).

# **RESULTADOS**

## **A. CARACTERÍSTICAS DE LOS ESTUDIOS INCLUIDOS**

Se realizó la búsqueda sistemática en las distintas bases de datos, identificando un total de 619 artículos. Se realizó una lectura del título y abstract de todos ellos, pudiendo ser descartados 574 artículos por no tener relación con el objetivo a estudio (Figura 4).

Llevamos a cabo una lectura a texto completo de los 45 artículos restantes. Se seleccionaron para la revisión sistemática y meta-análisis 27 artículos publicados entre 1997 y 2012, por cumplir con todos los criterios de inclusión. Los motivos de exclusión de 18 artículos (Apéndice 2) se exponen en la Figura 4.

En los 27 artículos que cumplían los criterios de inclusión, se calculó la calidad metodológica mediante el QUADAS-2. Se excluyeron 12 estudios con alto riesgo de sesgo, así como el único estudio basado en una técnica “dual time” (Mavi et al, 2009). Las Tablas 6 y 7 resumen los datos extraídos del análisis de estos 27 estudios.

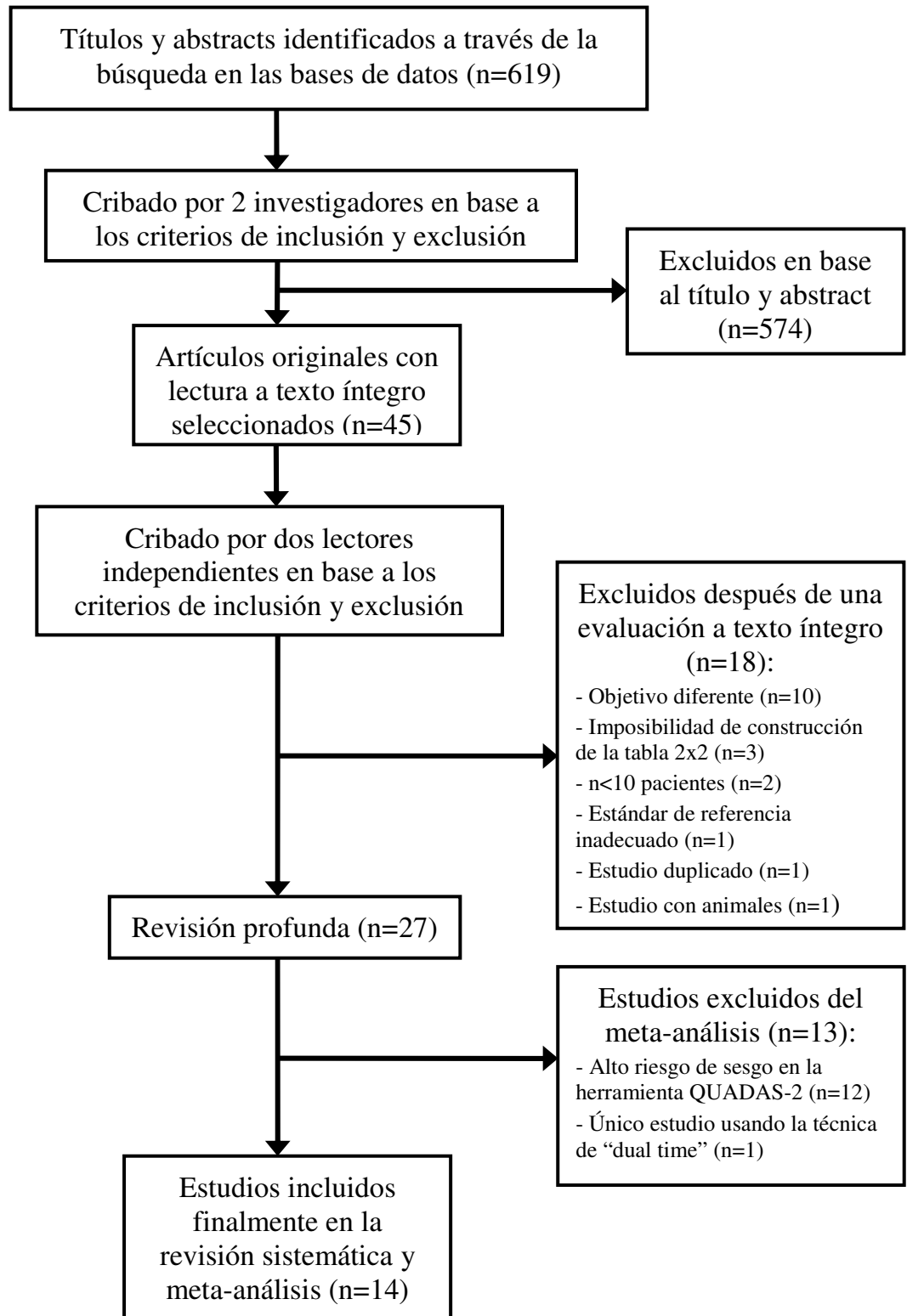
Por lo tanto, esta revisión sistemática y meta-análisis, consta finalmente de 14 estudios con diseño:

- Prospectivo: Duysinx et al, 2004; Duysinx et al, 2006; Kurata et al, 2010; Coolen et al, 2012.
- Retrospectivo: Schaffler et al, 2004; Alkhaldeh et al, 2011; Jung et al, 2012; Liao et al, 2012; Elboga et al, 2012.
- No se aporta este dato en 5 estudios

La suma total de pacientes fue 639, de los cuales el 66% fueron hombres. La edad media ponderada fue de 62 años.

407 pacientes (64%) tenían un DP o engrosamiento pleural maligno, y 232 pacientes, enfermedades pleurales benignas.

**Figura 4: Diagrama de flujo de la estrategia de búsqueda y selección de los estudios**



**Tabla 6: Características de los 14 estudios incluidos en la evaluación de la precisión diagnóstica de la PET en DPs**

AUTOR, AÑO	PAÍS	EDAD MEDIA (rango)	GÉNERO (M/F)	DISEÑO / ESTRATEGIA	NÚMERO DPM / DPB	TIPO DE TUMOR EN DPM	ETIOLOGÍA DE LOS DPB	PET vs PET-TC	ANÁLISIS DE LA IMAGEN	VP	FP	FN	VN
Bury et al, 1997	Bélgica	60 (36-80)	NC	NC / NC	16/9	13 metástasis, 3 MPM	5 paraneumónicos 4 miscelánea (1 TBC, 2 fibromas, 1 asbestosis benigna)	PET	Cualitativo	16	2	0	7
Buchman et al, 1999	Alemania	58 (38-68)	13/3	NC/ Consecutivo	12/4	9 MPM y 3 ADC pulmón	1 fibroma, 1 TBC, 1 fibrosis pleural, 1 empiema	PET	Cualitativo	12	0	0	4
Carretta et al, 2000	Italia	58 (35-73)	11/3	NC / NC	13/1	10 MPM, 2 ADC y 1 liposarcoma	1 fibrosis pleural	PET	Cualitativo	12	0	1	1
Gerbaudo et al, 2002	USA	60 (37-74)	11/4	NC/ Consecutivo	11/4	11 MPM	2 fibrosis pleurales, 1 pleurodesis con talco, 1 asbestosis	PET	Cualitativo	11	1	0	3

Precisión diagnóstica de la FDG-PET para diferenciar DPs malignos de benignos

<b>Schaffler et al, 2004</b>	Austria	58 (43-81)	62/30 (población total)	Retrospectivo/ Consecutivo	30/38	30 CPCNP	16 empiemas, 22 DP inespecíficos	PET	Cualitativo	30	15	0	23
<b>Duysinx et al, 2004</b>	Bélgica	61 (36-83)	67/31	Prospectivo/ Consecutivo	63/35	47 metástasis, 13 MPM, 2 linfomas, 1 sarcoma	12 paraneumónicos, 8 pleuritis crónica, 6 asbestosis benigna, 3 TBC, 6 otros	PET	Cualitativo	61	4	2	31
<b>Duysinx et al, 2006</b>	Bélgica	63 (21-84)	47/32	Prospectivo/ Consecutivo	51/28	25 cánceres de pulmón (21 CPCNP), 8 MPM, 25 metástasis	5 paraneumónicos, 12 pleuritis crónica, 3 TBC, 8 otros	PET	Semi- cuantitativo SUV >2.2	44	7	7	21
<b>Kurata et al, 2010</b>	Japón	65 (50-84)	9/1	Prospectivo/ Consecutivo	6/4	6 MPM	4 Patología pleural relacionada con el asbesto	PET-TC	Semi- cuantitativo SUV >2.2	6	2	0	2
<b>Alkhwald et al, 2011</b>	Alemania	63 (45-88)	39/22	Retrospectivo/ NC	29/32	29 CPCNP	NC	PET-TC	Cualitativo SUV ≥2.4 Dual time (SUV) en un grupo de 31 pacientes	25	9	4	23
										27	9	2	23
										10	1	5	15



Precisión diagnóstica de la FDG-PET para diferenciar DPs malignos de benignos

<b>Jung et al, 2012</b>	Corea	63.5 (37-81)	38/25	Retrospect/ NC	51/12	51 CPCNP	3 empiemas (los otros no consta)	PET-TC	Cualitativo SUV >2,8	47	5	4	7
<b>Elboga et al, 2012</b>	Turquía	58 (24-79)	18/32	Retrospect/ NC	37/13	37 MPM	7 pleuritis crónica, 3 TBC, 3 fibrosis	PET-TC	Cualitativo SUV >2.5	34	5	3	8
<b>Coolen et al, 2012</b>	Bélgica	60 (32-86)	24/7	Prospectivo/ Consecutivo	14/17	12 MPM, 2 ADC	NC (4 pleurodesis con talco, el resto tuvieron pleuritis)	PET-TC	Cualitativo	14	11	0	6
<b>Terada et al, 2012</b>	Japón	67.5	60/16	NC/NC	47/29	47 MPM	6 pleuritis crónica, 3 asbestosis benigna, 1 TBC y 1 pleuritis infecciosa. Resto no consta	PET-TC	SUV ≥ 3,5	28	2	19	27
<b>Liao et al, 2012</b>	China	61	24/9	Retrospect/ Consecutivo	27/6	27 CPCNP	5 paraneumónicos, 1 TBC	PET-TC	Cualitativo SUV>2,5	22	1	5	5

ADC=adenocarcinoma; CPCNP=carcinoma pulmonar de célula no pequeña; DPM/DPB=derrame pleural maligno/derrame pleural benigno; M/F=masculino/femenino; MPM=mesotelioma pleural maligno; NC=no consta; SUV=standardized uptake value; TBC=tuberculosis; VP/FP/FN/VN=verdadero positivo/falso negativo/verdadero negativo.

**Tabla 7: Características de los 13 estudios excluidos en la evaluación de la precisión diagnóstica de la PET en DPs**

AUTOR, AÑO	PAÍS	EDAD MEDIA (rango)	GÉNERO (M/F)	DISEÑO / ESTRATEGIA	NÚMERO DPM/DPB	TIPO DE TUMOR EN DPM	ETIOLOGÍA DE LOS DPB	PET vs PET-TC	ANÁLISIS DE LA IMAGEN	VP	FP	FN	VN
<b>Bénaud et al, 1998</b>	USA	NC	NC	NC/ Consecutivo	24/4	22 MPM y 2 ADC	2 pleuritis inespecífica, 1 asbestosis benigna, 1 angiolipoma	PET	Cualitativo  Semi-cuantitativo SUV>2	22	1	2	3
<b>Erasmus et al, 2000</b>	USA	65 (37-86)	18/7	Retrospect/ NC	22/3	22 CPCNP	3 paramaligno (1 gold standard dudoso)	PET	Cualitativo	21	1	1	2
<b>Gupta et al, 2002</b>	USA	70 (33-84)	25/10	NC/ Consecutivo	18/17	18 cánceres de pulmón	NC	PET	Cualitativo	16	1	2	16

Precisión diagnóstica de la FDG-PET para diferenciar DPs malignos de benignos

<b>Kramer et al, 2004</b>	Holanda	NC	27/5	Prospectivo/ Consecutivo	19/13	10 MPM, 2 sarcomas, 7 otros	6 fibrosis, 3pleuritis reactiva, 2pleuritis infecciosa, 2 otros	PET	Cualitativo	18	1	1	12
<b>Melloni et al, 2004</b>	Francia	61 (43-77)	NC	Retrospect/ Consecutivo	8/3	6 MPM, 1cáncer de pulmón, 1sarcoma	3 fibrosis pleural	PET	Cualitativo	7	3	1	0
<b>Toaff et al, 2005</b>	Israel	64 (29-83)	25/6	Retrospect/ NC	21/10	14 cánceres de pulmón (12 CPCNP), 5 linfoma, 2 otros	NC	PET-TC	Cualitativo	18	1	3	9
<b>Mavi et al, 2009</b>	USA	61.4 (NC)	48/7	NC/ Consecutivo	44/11	44 MPM	11 Pleuritis reactiva o inflamación inespecífica	PET	Dual time (SUV)	43	2	1	9
<b>Yildirim et al, 2009</b>	Turquía	61 (40-82)	20/11	NC/ Consecutivo	17/14	17 MPM	9 asbestosis benigna, 5 engrosamiento pleural difuso	PET-TC	Cualitativo  SUV >2.2	15	1	2	13 14

Precisión diagnóstica de la FDG-PET para diferenciar DPs malignos de benignos

<b>Yamamoto et al, 2009</b>	Japón	64 (45-84)	30/3	Retrospect/ NC	17/16	17 MPM	2 TBC, 1 pleuritis, 1 fibrosis pleural, 12 no consta	PET	Cualitativo	15	2	2	14
-----------------------------	-------	---------------	------	-------------------	-------	--------	---	-----	-------------	----	---	---	----

Precisión diagnóstica de la FDG-PET para diferenciar DPs malignos de benignos

<b>Orki et al, 2009</b>	Turquía	47 (NC)	63/20	NC/ Consecutivo	44/39	25 MPM, 18 metástasis de ADC, 1 CPCP	30 Pleuritis crónica, 9 TBC	PET-TC	Semi- cuantitativo SUV >3	44	2	0	37
<b>Kim et al, 2011</b>	Corea	67 (38-84)	20/13	Retrospect/ Consecutivo	24/9	24 cánceres de pulmón	NC	PET-TC	Cualitativo	21	1	3	8
<b>Abe et al 2012</b>	Japón	67 (47-79)*	27/4**	Retrospect/ NC	31/59	31 MPM	NC	PET-TC	Dual time (SUV)	31	7	0	52
<b>Letovanec et al, 2012</b>	Suiza	60 (NC)	29/18	Retrospect/ NC	17/33	6 CPCNP, 5 linfomas, 2 metástasis de cáncer de mama, 2 MPM, 2 teratomas	NC	PET-TC	Semi- cuantitativo SUV ≥2,2	9	3	8	30

Las etiologías de los DPs malignos fueron las siguientes (Tabla 8):

**Tabla 8: Etiología de los DPs malignos**

DP MALIGNOS	MESOTELIOMA	98 epiteliales		
		20 sarcomatoide o mixto		
		38 no especificados		
	METÁSTASIS	169 pulmón	135 CPCNP	
			7 CPCP	
			27 no especificados	
		23 miscelánea	5 mama, 5 GI, 3 ovario, 3 sarcoma, 3 linfoma, 2 timoma y 2 riñón	
59 no especificados				

CPCNP=cáncer de pulmón de célula no pequeña; CPCP=cáncer de pulmón de célula pequeña; GI=gastrointestinal.

Mientras que las etiologías de los DPs benignos fueron (Tabla 9):

**Tabla 9: Etiología de los DPs benignos**

DP BENIGNOS	76 no especificados		
	49 paraneumónicos o empiemas		
	40 pleuritis inespecíficas		
	23 fibrosis pleurales		
	22 asbestosis benignas		
	13 tuberculosos		
	9 tumores pleurales	8 fibromas	
		1 lipoma	

DP=derrame pleural

Se analizó por separado la utilidad de la PET para discriminar entre DPs y/o engrosamientos pleurales de origen maligno y benigno en pacientes con (Tabla 10):

- DP de etiología indeterminada en 6 artículos: Bury et al, 1997; Buchmann et al, 1999; Carretta et al, 2000; Duysinx et al, 2004; Duysinx et al, 2006; Coolen et al, 2012.

Precisión diagnóstica de la FDG-PET para diferenciar DPs malignos de benignos

- Diagnóstico previo de cáncer de pulmón en 4 artículos: Liao et al, 2012; Schaffler et al, 2004; Alkhaldeh et al, 2011; Jung et al, 2012.
- Sospecha o confirmación de mesotelioma maligno pleural en 4 artículos: Gerbaudo et al, 2002; Kurata et al, 2010; Elboga et al, 2012; Terada et al, 2012.

El método de interpretación de la actividad metabólica de la PET (Tabla 10) fue:

- Cualitativa en 7 estudios: Bury et al, 1997; Buchmann et al, 1999; Carretta et al, 2000; Gerbaudo et al, 2002; Schaffler et al, 2004; Duysinx et al, 2004; Coolen et al, 2012.
- Semicuantitativa en 3 estudios: Duysinx et al, 2006; Kurata et al, 2010; Terada et al, 2012.
- Ambos métodos en 4 estudios: Alkhaldeh et al, 2011; Jung et al, 2012; Liao et al, 2012; Elboga et al, 2012.

En el método semicuantitativo, el punto de corte óptimo de SUV para diferenciar benignidad de malignidad varió de 2,2 a 3,5 según los estudios.

**Tabla 10: Etiología y método de interpretación de la imagen de los estudios**

ETIOLOGÍA DEL DP	NÚMERO DE ESTUDIOS	MÉTODO DE INTERPRETACIÓN DE LA IMAGEN	NÚMERO DE ESTUDIOS
Indeterminada	6	Cualitativa	7
Cáncer de pulmón	4	SUV	3
Mesotelioma	4	Ambos métodos	4

En 7 artículos (los más recientes; de 2010 en adelante) se usó el PET-TC híbrido: Kurata et al, 2010; Alkhaldeh et al, 2011; Liao et al, 2012; Jung et al, 2012; Elboga et al, 2012; Coolen et al, 2012; Terada et al, 2012.

Todos los estudios registraban la captación de FDG en las membranas pleurales, independientemente de si eran morfológicamente anormales. En dos estudios, además, se evaluó la tasa metabólica del líquido pleural (Alkhaldeh et al, 2011; Elboga et al, 2012).

Se debe tener en cuenta que la mayoría de los pacientes evaluados en los estudios presentaban engrosamiento y/o nodularidad pleural en las imágenes de la TC.

## **B. ÍNDICES DE PRECISIÓN DIAGNÓSTICA PARA LOS DPs MALIGNOS**

### **1. MÉTODO DE INTERPRETACIÓN DE LA IMAGEN**

#### **1.1. ANÁLISIS CUALITATIVO**

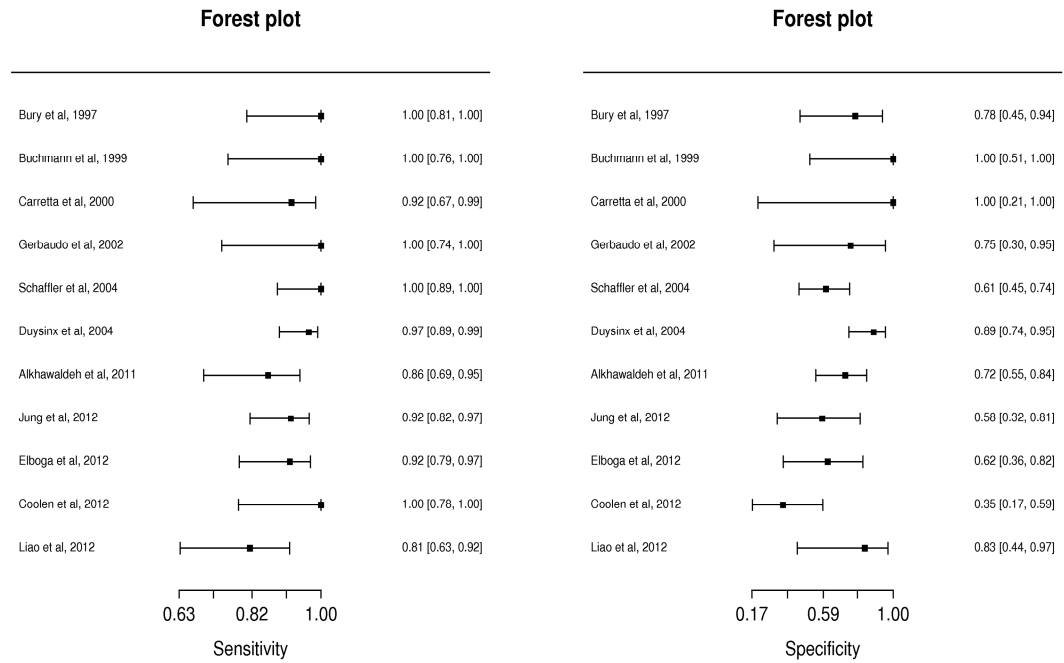
Los 11 estudios que analizaban las imágenes de la PET mediante un método cualitativo o visual, fueron: Bury et al, 1997; Buchmann et al, 1999; Carretta et al, 2000; Gerbaudo et al, 2002; Schaffler et al, 2004; Duysinx et al, 2004; Alkhaldeh et al, 2011; Jung et al, 2012; Elboga et al, 2012; Coolen et al, 2012; Liao et al, 2012. Presentaron los siguientes resultados:

- Sensibilidad del 91%
- Especificidad del 67%
- LR positiva 2,83
- LR negativa 0,14
- DOR 22
- AUC 0,893

La Figura 5 representa la sensibilidad y especificidad de la FDG-PET obtenidas en los artículos incluidos en el meta-análisis con el método de interpretación visual, para el diagnóstico de DP maligno. La Figura 6 representa los valores diagnósticos de los estudios en una curva jerárquica ROC-resumen, con el método de interpretación visual. Cuanto más se acerca la curva al ángulo izquierdo, mejor precisión diagnóstica se observa; mientras que los datos más cercanos a la diagonal implican una peor precisión diagnóstica.

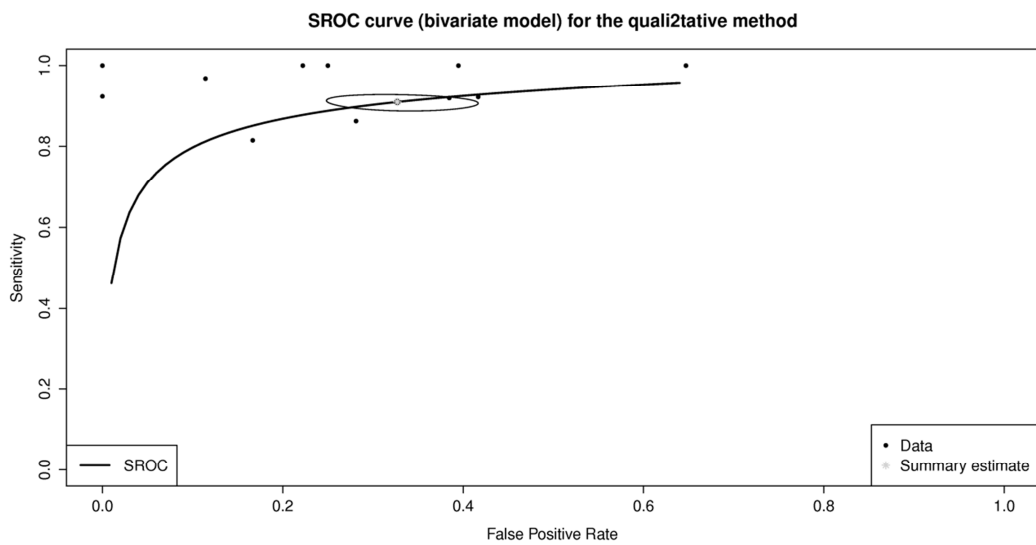


**Figura 5: Forest plot de sensibilidad y especificidad de la FDG-PET con el método de interpretación visual para diagnosticar DPs malignos**



La sensibilidad y especificidad de cada estudio se representa mediante un cuadrado, a través del cual discurre una línea horizontal (IC del 95%).

**Figura 6: Curva S-ROC de la metodología visual del FDG-PET en DPs malignos**



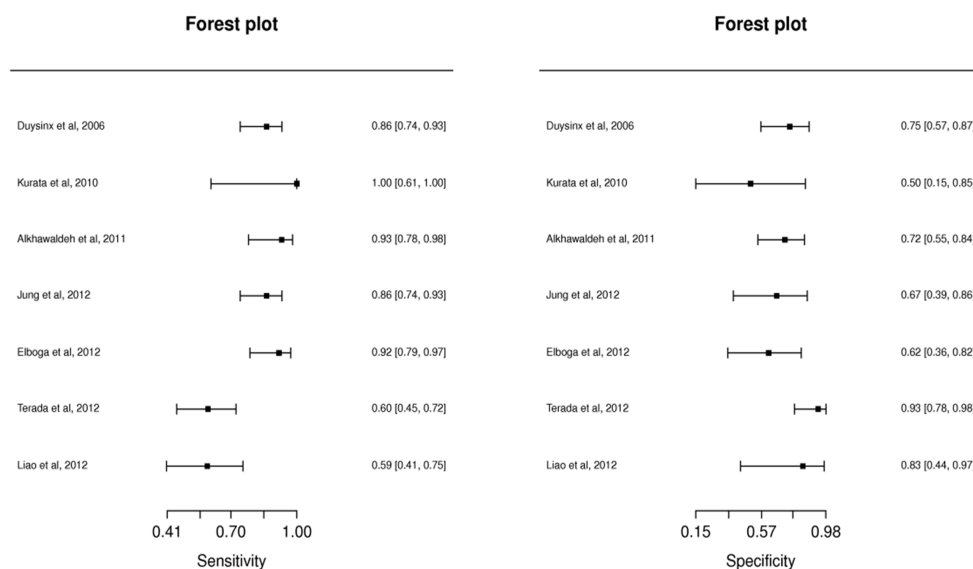
## 1.2. ANÁLISIS SEMICUANTITATIVO

Los 7 estudios que analizaban las imágenes de la PET mediante un método semicuantitativo fueron: Duysinx et al, 2006; Kurata et al, 2010; Alkhalwaleh et al, 2011; Jung et al, 2012; Elboga et al, 2012; Terada et al, 2012; Liao et al, 2012. Presentaron los siguientes resultados:

- Sensibilidad del 82%
- Especificidad del 74%
- LR positiva 3,24
- LR negativa 0.25
- DOR 13,8
- AUC 0,840

La Figura 7 representa la sensibilidad y especificidad de la FDG-PET obtenidas en los artículos incluidos en el meta-análisis con el método de interpretación semicuantitativo, para el diagnóstico de DP maligno y la Figura 8 muestra la curva S-ROC de los mismos artículos.

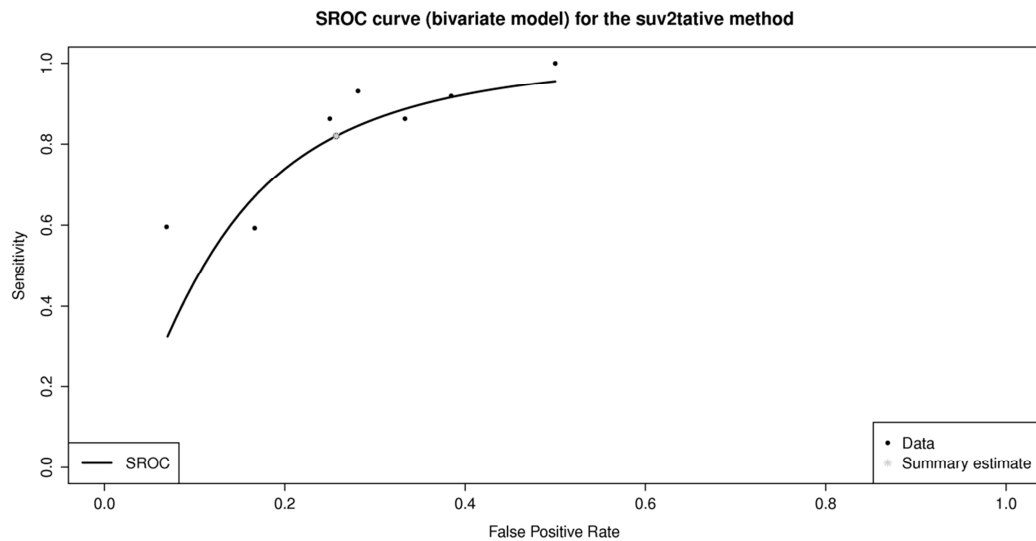
**Figura 7: Forest plot de sensibilidad y especificidad de la FDG-PET con el método de interpretación semicuantitativo para diagnosticar DPs malignos**



La sensibilidad y especificidad de cada estudio se representa mediante un cuadrado, a través del cual discurre una línea horizontal (IC del 95%).

En general, la sensibilidad de la FDG-PET fue significativamente mayor con el método cualitativo que con el semicuantitativo ( $p=0,026$ ). La especificidad no variaba en el análisis de estas dos metodologías.

**Figura 8: Curva S-ROC de la metodología semicuantitativa del FDG-PET en DPs malignos**



## 2. USO DE LA PET vs PET-TC INTEGRADO

### 2.1. PET

Los 6 estudios que usaban el sistema de PET como método de prueba diagnóstica fueron: Bury et al, 1997; Buchmann et al, 1999; Carretta et al, 2000; Gerbaudo et al, 2002; Schaffler et al, 2004; Duysinx et al, 2004. Presentaron los siguientes resultados:

- Sensibilidad del 96%
- Especificidad del 75%
- LR positiva 4,09
- LR negativa 0,06
- DOR 76,7
- AUC 0,950

El método de interpretación de la imagen en los estudios que usaban el sistema PET aislado era el análisis visual o cualitativo.

A pesar de que el artículo de Duysinx et al (2006) debería entrar a formar parte de este análisis, se excluyó ya que es el único estudio realizado con el sistema PET que interpretaba la imagen mediante el método semicuantitativo.

## **2.2. PET-TC**

### **- Visual**

Los 5 estudios que usaban el sistema de PET-TC integrado como método de prueba diagnóstica y un análisis de la imagen visual, fueron: Alkhaldeh et al, 2011; Jung et al, 2012; Elboga et al, 2012; Coolen et al, 2012; Liao et al, 2012. Presentaron los siguientes resultados:

- Sensibilidad del 89%
- Especificidad del 61%
- LR positiva 2,32
- LR negativa 0.19
- DOR 13
- AUC 0,950

### **- Semicuantitativo**

Los 6 estudios que usaban el sistema de PET-TC integrado como método de prueba diagnóstica y un análisis de la imagen semicuantitativo, fueron: Kurata et al, 2010; Alkhaldeh et al, 2011; Jung et al, 2012; Elboga et al, 2012; Terada et al, 2012; Liao et al, 2012. Presentaron los siguientes resultados:

- Sensibilidad del 81%
- Especificidad del 74%
- LR positiva 3,22
- LR negativa 0,26
- DOR 13,3
- AUC 0,838

Los estudios que emplearon la PET sin asociación a TC contribuyeron en mayor grado a la mejor sensibilidad de las lecturas con el método cualitativo (sensibilidad del 96%).

De hecho, cuando se analizaron únicamente los estudios que usaron el sistema PET-TC integrado, la precisión diagnóstica no fue

significativamente diferente entre las lecturas cualitativa y semicuantitativa ( $p=0.208$ ).

En cuanto a la especificidad, no se observaron diferencias estadísticamente significativas entre los grupos analizados (Tabla 11).

### **3. ETIOLOGÍA DE LA PATOLOGÍA PLEURAL**

En cuanto a las etiologías benignas incluidas en este meta-análisis se ha detectado una falta en la proporción de alguna de ellas para considerarse representativas de nuestra población.

Cabe destacar, la existencia de falsos positivos; es decir, pacientes diagnosticados de patología benigna que presentaron una ávida captación en la FDG-PET: 5 (38,5%) de los 13 pacientes con tuberculosis y 21 (43%) de los 49 DPs paraneumónicos o empiemas.

Un análisis por meta-regresión mostró que la capacidad de la FDG-PET para discriminar entre benignidad y malignidad pleural se mantenía sin cambios independientemente de que la población diana fueran pacientes con mesotelioma o con cáncer de pulmón ( $p>0,4$ ).

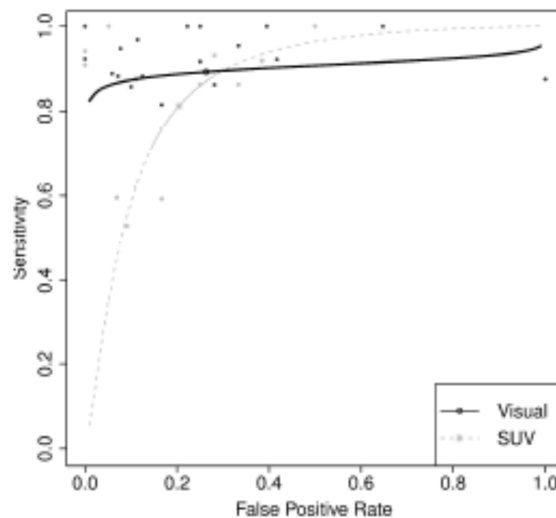
**Tabla 11: Resumen de las medidas de precisión diagnóstica de la FDG-PET para identificar DPs malignos**

Medidas	PET			PET-TC integrado	
	Cualitativo	Semicuantitativo	Cualitativo	Cualitativo	Semicuantitativo
Número de estudios	11	7	6	5	6
Sensibilidad, % (IC 95%)	91 (86-94)	82 (69-90)	96 (90-98)	89 (81-94)	81 (66-91)
Especificidad, % (IC 95%)	67 (56-77)	74 (62-84)	75 (59-87)	61 (44-75)	74 (58-85)
LR positiva (IC 95%)	2,83 (2,04-3,98)	3,24 (2,24-4,70)	4,09 (2,34-7,1)	2,32 (1,61-3,45)	3,22 (2-5)
LR negativa (IC 95%)	0,14 (0,08-0,22)	0,25 (0,14-0,39)	0,06 (0,03-0,13)	0,19 (0,11-0,31)	0,26 (0,14-0,43)
DOR (IC 95%)	22 (10,2-41,7)	13,8 (7,3-23,9)	76,7 (22,1-195)	13 (6-24,8)	13,3 (6,32-24,8)
AUC	0,893	0,840	0,950	0,859	0,838

AUC=área bajo la curva; DOR=odds ratio diagnóstica; IC=intervalo de confianza; LR=razón de verosimilitud

La Figura 9 representa los valores diagnósticos de los estudios en una curva jerárquica ROC-resumen. Cuanto más se acerca la curva al ángulo izquierdo, mejor precisión diagnóstica se observa; mientras que los datos más cercanos a la diagonal implican una peor precisión diagnóstica.

**Figura 9: Curva S-ROC de la metodología visual y semicuantitativa del FDG-PET en DPS malignos**



Cada símbolo representa un estudio del meta-análisis. El círculo indica los estudios de metodología visual (o cualitativa), y el asterisco los de metodología semicuantitativa.

### **C. RIESGO DE SESGO Y HETEROGENEIDAD**

Se llevó a cabo la valoración de la calidad metodológica en los 27 artículos inicialmente incluidos, siendo buena la correlación interobservador ( $k=0,77$ ).

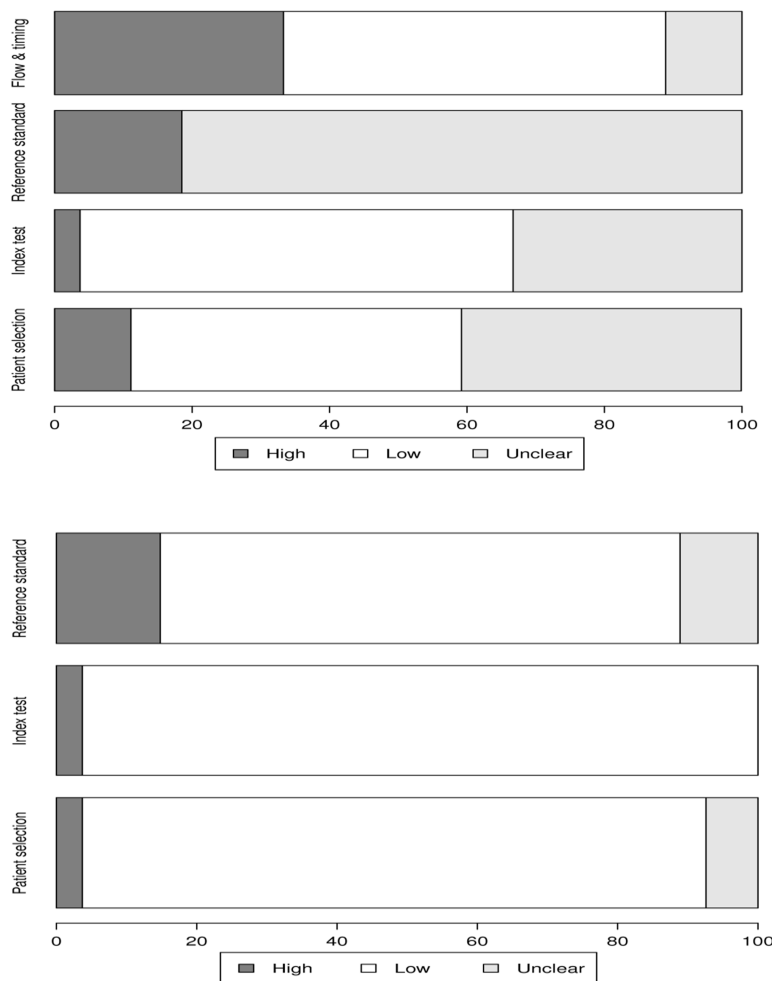
Quince de los 27 estudios incluidos, no presentaron ningún riesgo alto de sesgo al analizar todos sus dominios mediante el QUADAS-2, por lo que se consideraron aptos para la revisión sistemática y meta-análisis. Los 15 estudios incluidos fueron: Bury et al, 1997; Buchmann et al, 1999; Carretta et al, 2000; Gerbaudo et al, 2002; Schaffler et al, 2004; Duysinx et al, 2004; Duysinx et al, 2006; Mavi et al, 2009; Kurata et al, 2010;

Alkhaldeh et al, 2011; Jung et al, 2012; Elboga et al, 2012;Coolen et al, 2012: Terada et al, 2012; Liao et al, 2012.

Sin embargo, la calidad global de todos los estudios fue modesta, ya que ninguno de estos 15 estudios fue clasificado con bajo riesgo de sesgo en los cuatro dominios (alguno de los dominios fue evaluado como “poco claro”). Solamente en cinco de los estudios, se determinó un bajo riesgo de sesgo en tres de los cuatro dominios existentes y éstos fueron: Schaffler et al, 2004; Duysinx et al, 2004; Duysinx et al, 2006; Kurata et al, 2010; Coolen et al, 2012.

La Figura 10 resume la calidad metodológica de los 27 estudios incluidos usando los criterios del QUADAS-2.

**Figura 10: Calidad metodológica de los estudios mediante QUADAS-2**

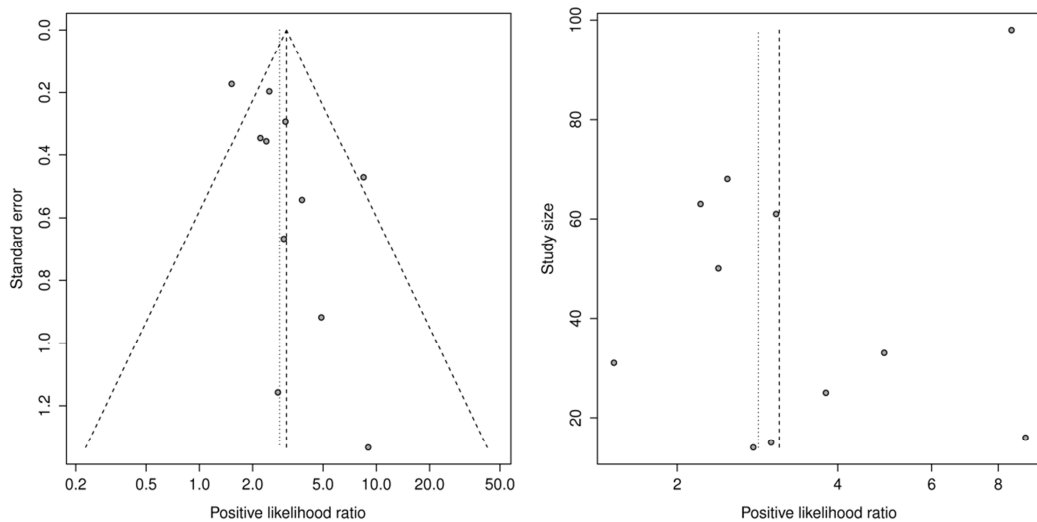


F&T=*flow and timing*=flujo de pacientes y temporalidad de la prueba diagnóstica y gold estándar. I=índice

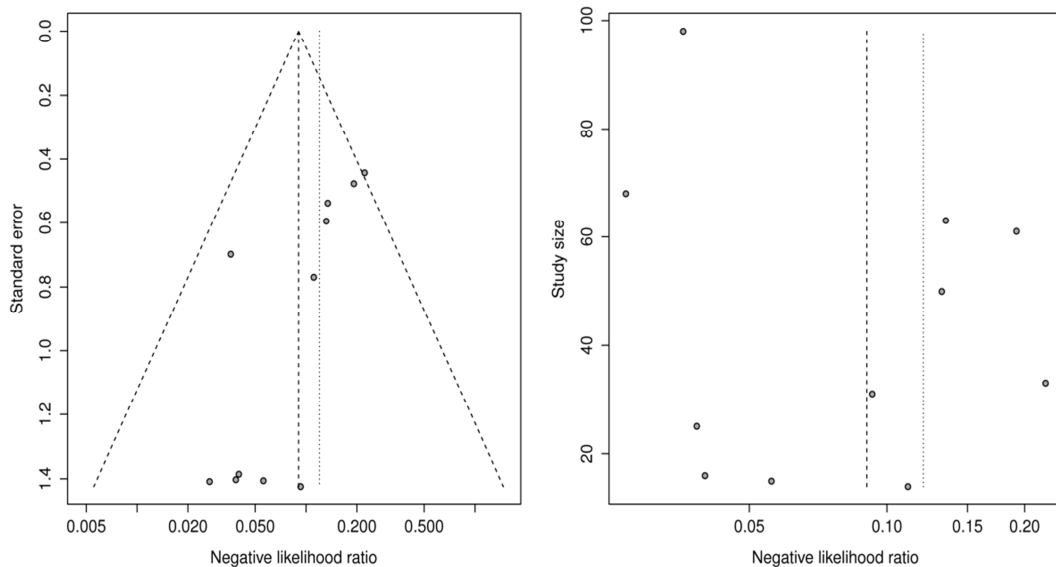


El funnel plot mostró un posible sesgo de publicación para las LR positivas y negativas de los estudios con lecturas de imagen de la FDG-PET cualitativas ( $p < 0,05$ ) tal y como muestran las Figuras 11 y 12. Igual ocurrió con las LR negativas de los estudios que usaban el método semicuantitativo ( $p = 0,03$ ) como demuestra la Figura 13.

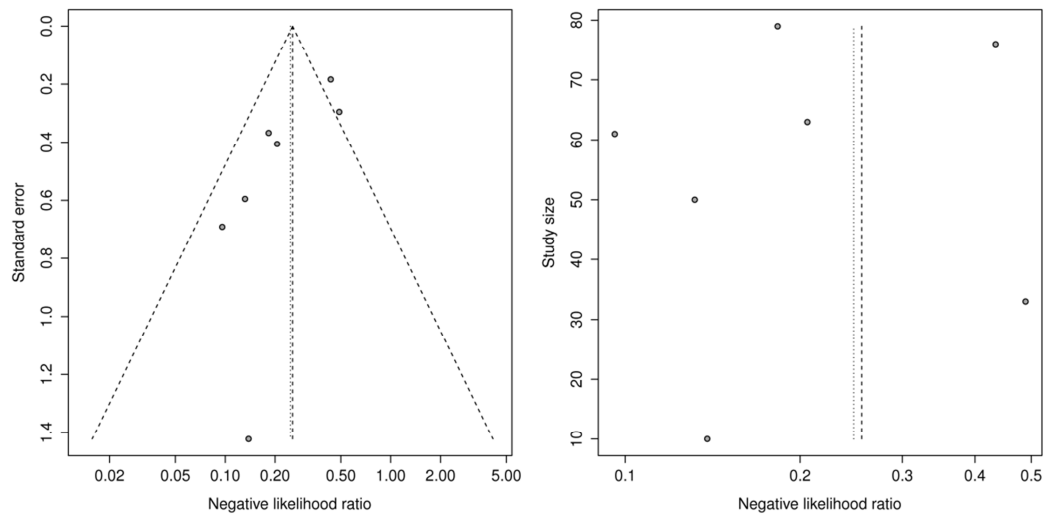
**Figura 11: Funnel plot de la LR positiva para el método cualitativo**



**Figura 12: Funnel plot de la LR negativa para el método cualitativo**



**Figura 13: Funnel plot de la LR negativa para el método semicuantitativo**



De forma contraria, no existió riesgo de sesgo de publicación estadísticamente significativo para las lecturas cualitativa y semicuantitativa cuando se tenía en cuenta el tipo de máquina utilizada, PET aislada o combinada con TC (todos  $p > 0,10$ ).

Finalmente, se observó que existe heterogeneidad estadísticamente significativa para las medidas de sensibilidad y especificidad en todos los subgrupos analizados, excepto en el subgrupo de metodología de interpretación de la imagen cualitativa usando un sistema PET (desviación estándar de sensibilidad=0,062).

#### **D. APLICACIÓN CLÍNICA**

De acuerdo con el método Bayesiano, la estimación de la probabilidad post-test de DP maligno en pacientes a los que se les realiza una FDG-PET, se relaciona con la prevalencia de la enfermedad (probabilidad pre-test).

Por ejemplo, en una hipotética población con una prevalencia de malignidad del 35% entre los pacientes que se presentan con un exudado pleural, el hallazgo de una PET-TC positiva para malignidad usando el método de interpretación de imagen semicuantitativo (LR positiva de 3,22), haría incrementar esta probabilidad hasta un 63,5%, mientras que un resultado negativo (LR negativa de 0,26), la disminuiría al 12,3%. Estos cálculos habrían cambiado al 68,7% (LR positiva=4,09) y al 3% (LR negativa=0,06), respectivamente, si se usa una FDG-PET con interpretación visual de la imagen.

# DISCUSIÓN

## **A. RESUMEN DE LA EVIDENCIA**

Este meta-análisis recoge 14 estudios, sin riesgo de sesgo alto, que analizan la utilidad diagnóstica de la FDG-PET en pacientes con DP (o engrosamiento pleural).

De acuerdo con los datos extraídos, el mejor de los casos es el rendimiento del sistema de PET aislado para identificar la naturaleza maligna de un DP, usando la interpretación de la imagen visual o cualitativa, cuyos resultados son:

- Sensibilidad 96%
- Especificidad 76%
- LR positiva 4,09
- LR negativa 0,06
- AUC 0,95

Pero el sistema PET aislado rara vez se utiliza en la actualidad porque:

- El sistema PET-TC híbrido proporciona información metabólica y anatómica en una misma sesión.
- La valoración visual es propensa a la subjetividad y a la falta de reproducibilidad. Esta limitación se mitiga parcialmente con una técnica algo más consistente como es la medida semicuantitativa (SUV), que actualmente domina el uso clínico del sistema PET-TC integrado o híbrido.

En esta tesitura clínica más real, la eficacia del sistema PET-TC integrado, usando un método de interpretación de la imagen semicuantitativo, para identificar malignidad pleural se redujo a:

- Sensibilidad 81%
- Especificidad 74%
- LR positiva 3,22
- LR negativa 0,26
- AUC 0,95

La razón de porqué las interpretaciones visuales proporcionan mayor sensibilidad que los métodos semicuantitativos puede estar en

relación con los muchos factores biológicos y tecnológicos que influyen en los resultados de estos últimos.

Por ejemplo, el SUV se calcula normalmente sobre una pequeña área de la pleura (ROI=region of interest=región de interés), lo que puede no ser lo más adecuado por la frecuente afectación parcheada o difusa pleural que caracteriza a las lesiones malignas. Por el contrario, la técnica volumétrica (VOI=volumen of interest=volumen de interés) es una herramienta que determina la concentración del volumen del radiofármaco en la totalidad del tumor, incluyendo áreas de diferente SUV y, por lo tanto, puede representar mejor el comportamiento biológico o metabólico de toda la lesión pleural. Desafortunadamente ninguno de los estudios incluidos en el meta-análisis utilizaron el análisis cuantitativo basado en el VOI.

La razón fundamental de que exista una superior rentabilidad del uso de la PET aislada frente al PET-TC híbrido es especulativa. Sin tener en cuenta el sesgo de publicación inherente de los estudios más antiguos, podríamos argüir una mayor experiencia de los radiólogos que interpretaban la PET en un tiempo en el que el uso de estos aparatos estaba altamente restringido a centros de referencia.

En conjunto, la capacidad de los sistemas de PET-TC integrados para diferenciar los procesos pleurales benignos de malignos está plagada de potenciales errores y debe considerarse moderada, como demuestra un AUC de 0,84.

La LR negativa de la PET es mayor que la positiva. Esto supone que la PET sea más útil para descartar DPs malignos cuando la prueba demuestre que no existe captación anormal del trazador. En este sentido, la FDG-PET reduciría teóricamente el número de procedimientos invasivos innecesarios en pacientes con DP de etiología benigna.

La imagen de la FDG-PET podría guiar las biopsias con aguja fina o por toracoscopia, hacia el área de mayor captación del trazador dentro de una pleura engrosada, minimizando los errores de muestreo.

## **B. PUNTOS FUERTES Y LIMITACIONES DE LA REVISIÓN**

### **Puntos fuertes**

Podemos decir que la fortaleza de esta revisión sistemática es que se ha llevado a cabo una estrategia de búsqueda bibliográfica robusta y se ha realizado un análisis estadístico en profundidad.

Nuestros cálculos se restringieron a aquellos estudios que no presentaban ningún riesgo alto de sesgo al aplicar la herramienta del QUADAS-2, lo que ha permitido una mejor evaluación de la literatura disponible.

De hecho, la inclusión de los 27 estudios iniciales en el análisis estadístico hubiera producido unos índices diagnósticos superiores, conllevando de este modo a una conclusión potencialmente engañosa acerca de las propiedades discriminativas de la FDG-PET. Específicamente, la sensibilidad y especificidad de las interpretaciones de la FDG-PET hubieran alcanzado los resultados que se resumen en la Tabla 12.

**Tabla 12: Sensibilidad y especificidad de los 27 estudios iniciales según el método de análisis de la imagen de la FDG-PET**

<b>Método de análisis de la imagen</b>	<b>Sensibilidad</b>	<b>Especificidad</b>
Cualitativa (20 estudios)	92%	81%
Semicuantitativa (11 estudios)	91%	83%
Dual time point (3 estudios)	98%	90%

## **Limitaciones**

La fiabilidad y la generalización de los datos están limitadas por el pequeño número de estudios seleccionados y por las deficiencias metodológicas de los mismos. Además, las fuentes potenciales de sesgo y la heterogeneidad de los estudios se añaden a las limitaciones.

Ciertamente, en la mayoría de los estudios se detectaron sesgos de selección. Por ejemplo, la mayor parte de los pacientes con DP maligno presentaban nódulos o engrosamientos pleurales. En estos pacientes la PET se interpretó como positiva cuando existía un incremento de la captación en las membranas pleurales anormales y no en el derrame. Sin embargo, tales alteraciones morfológicas pleurales caracterizan a menos de la mitad de los DP's malignos (Porcel et al, 2013-B).

De forma similar, las patologías pleurales benignas comunes, como el DP paraneumónico, tuberculoso o secundario a insuficiencia cardiaca, estuvieron claramente infrarrepresentadas o incluso ausentes en las ramas comparadoras de los estudios.

Se demostró un sesgo de publicación en aquellos estudios que interpretaban la PET tanto con el método cualitativo como semicuantitativo. El hecho de que los estudios de lectura de imagen cualitativa sean los más antiguos introduce un factor de confusión para la interpretación de su superior sensibilidad en comparación con los de lectura semicuantitativa. Las posibles fuentes de heterogeneidad incluyeron el diseño del estudio, la población a estudio, la variación de los puntos de corte para la positividad en las evaluaciones semicuantitativas y las diferencias en la adquisición de la imagen de la PET o el equipo utilizado.

Algunos de estos factores pudieron ser controlados en la metarregresión.



## **C. COMPARACIÓN CON INVESTIGACIONES PREVIAS**

Existen dos meta-análisis previos, realizados por el mismo grupo de investigadores, que analizan el rendimiento diagnóstico de la PET para discriminar benignidad y malignidad pleural. (Treglia et al, 2014-A. Treglia et al, 2014-B).

En el primer estudio (Treglia et al, 2014-A), se eligieron inicialmente para el meta-análisis a 16 artículos, cuyas características se describen en las Tablas 13 y 14. Cinco de estos artículos fueron excluidos por no presentar suficientes datos para calcular la sensibilidad y la especificidad (Duysinx et al, 2004; Mavi et al, 2009; Kurata et al, 2010; Abe et al, 2012; Terada et al, 2012). Por lo tanto, finalmente se incluyeron 11 artículos con un total de 440 pacientes.

**Tabla 13: Características de los 16 artículos inicialmente incluidos en (Treglia et al, 2014-A)**

Authors	Year	Country	Study Design	Patients Performing <sup>18</sup> F-FDG PET or PET/CT	Population	Mean Age (y)	Gender (%Male)
Terada et al (38)	2012	Japan	NR	76	Patients with suspicious MPM	67	79
Coolen et al (37)	2012	Belgium	Prospective	31	Patients undergoing evaluation for pleural disease	60	77
Abe et al (36)	2012	Japan	Retrospective	90	Patients with suspicious MPM	NR	NR
Elboga et al (35)	2012	Turkey	Retrospective	50	Patients undergoing evaluation for pleural disease	58	36
Kurata et al (34)	2010	Japan	Prospective	17	Patients with suspicious MPM	69	94
Yamamoto et al (33)	2009	Japan	Retrospective	33	Patients with suspicious MPM	64	91
Mavi et al (32)	2009	USA	NR	55	Patients with suspicious MPM	61	87
Yildirim et al (31)	2009	Turkey	NR	31	Patients with suspicious MPM	61	65
Orki et al (30)	2009	Turkey	Prospective	83	Patients undergoing evaluation for pleural disease	47	76
Duysinx et al (29)	2006	Belgium	Prospective	79	Patients undergoing evaluation for pleural effusions	63	59
Duysinx et al (28)	2004	Belgium	Prospective	98	Patients undergoing evaluation for pleural disease	61	68
Kramer et al (27)	2004	Netherlands	Prospective	32	Patients undergoing evaluation for pleural disease	NR	84
Gerbaudo et al (26)	2002	United States	Retrospective	15	Patients with suspicious MPM	60	73
Carretta et al (25)	2000	Italy	NR	14	Patients undergoing evaluation for pleural disease	58	79
Bénard et al (24)	1998	United States	NR	28	Patients with suspicious MPM	NR	NR
Bury et al (23)	1997	Belgium	NR	25	Patients undergoing evaluation for pleural disease	60	NR

CT=tomografía computarizada; MPM=mesotelioma maligno pleural; NR=no informado

**Tabla 14: Aspectos técnicos de los 16 artículos inicialmente incluidos (Treglia et al, 2014-A)**

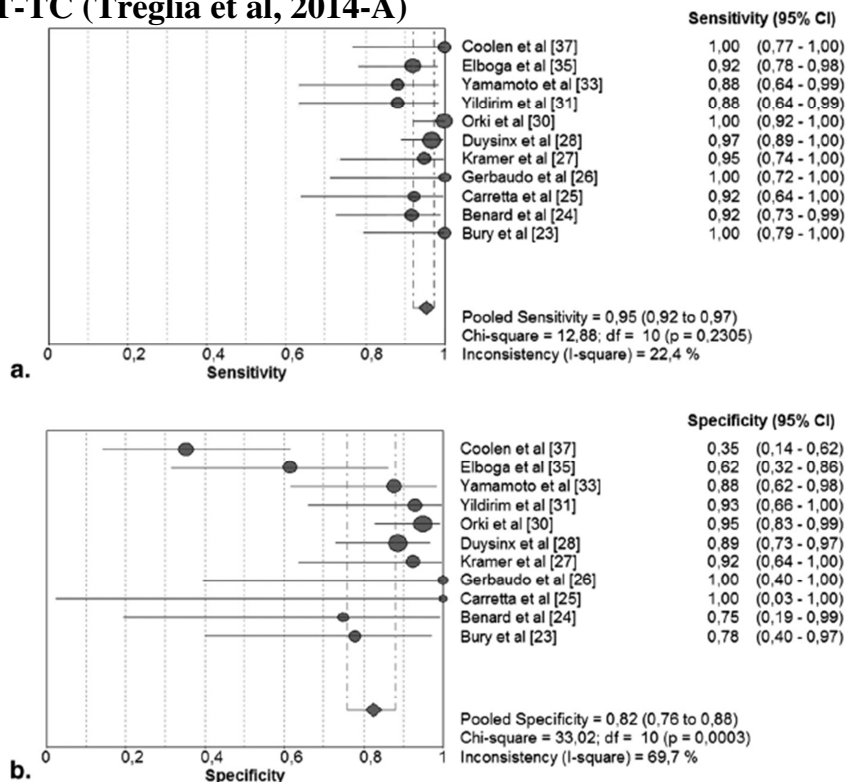
Authors	Device	Mean <sup>18</sup> F-FDG Injected Activity	Mean Time between <sup>18</sup> F-FDG Injection and Acquisition (min)	Image Analysis
Terada et al (38)	PET/CT	NR	60	Semiquantitative
Coolen et al (37)	PET/CT	370 MBq	50	Visual
Abe et al (36)	PET/CT	3.7 MBq/kg	60 and 120	Visual and semiquantitative at early and delayed phase
Elboga et al (35)	PET/CT	296–555 MBq	60 and 120	Visual and semiquantitative at early and delayed phase
Kurata et al (34)	PET/CT	5.2 MBq/kg	60	Visual and semiquantitative
Yamamoto et al (33)	PET	3.5 MBq/kg	60 and 120	Visual and semiquantitative at early and delayed phase
Mavi et al (32)	PET	5.2 MBq/kg	60 and 90	Semiquantitative at early and delayed phase
Yildirim et al (31)	PET/CT	350–400 MBq	NR	Visual and semiquantitative
Orki et al (30)	PET/CT	370–666 MBq	60–120	Visual and semiquantitative
Duysinx et al (29)	PET	2.1–3.7 MBq/kg	73	Semiquantitative
Duysinx et al (28)	PET	NR	NR	Visual
Kramer et al (27)	PET	400–600 MBq	90	Visual and semiquantitative
Gerbaudo et al (26)	PET	375 MBq	90	Visual
Carretta et al (25)	PET	3.7 MBq/kg	60	Visual and semiquantitative
Bénard et al (24)	PET	4.25 MBq/kg	60–90	Visual and semiquantitative
Bury et al (23)	PET	NR	NR	Visual

NR=no informado.

Como indica la Figura 14, la eficacia diagnóstica de la PET y PET-TC integrado juntos para el diagnóstico de DPs benignos y malignos, considerando únicamente la lectura de la imagen visual, fue:

- Sensibilidad 95%
- Especificidad 82%
- LR positiva 5,3
- LR negativa 0,09
- AUC 0,95

**Figura 14: Forest plot de sensibilidad y especificidad de la FDG-PET o PET-TC (Treglia et al, 2014-A)**



En el segundo metaanálisis (Treglia et al, 2014-B) se seleccionaron 8 artículos que incluyeron un total de 360 pacientes con cáncer confirmado (90% de origen pulmonar) que desarrollaron alteraciones pleurales. Sus características se resumen en las Tablas 15 y 16.

**Tabla 15: Características de los 8 artículos incluidos (Treglia et al, 2014-B)**

Characteristics of the included studies.								
Authors	Year	Country	Study design	Patients performing <sup>18</sup> F-FDG-PET or PET/CT	Population	Mean age (years)	Gender (% male)	Type of primary tumor
Liao et al. [20]	2012	China	Retrospective	33	Patients with lung cancer and pleural effusion	61	73	29 NSCLC, 4 SCLC
Letovanec et al. [19]	2012	Switzerland	Retrospective	50	Patients with known tumors and pleural effusion	60	62	24 NSCLC, 7 Ly, 5 BC, 4 GIST, 3 Mes, 2 HNC, 2 MT, 1 CC, 1 EC, 1 Mel
Alkhalwaldeh et al. [18]	2011	Germany	Retrospective	61	Patients with NSCLC and pleural effusion	63	64	61 NSCLC
Kim et al. [17]	2011	Korea	Retrospective	33	Patients with NSCLC and pleural effusion	67	61	33 NSCLC
Toaff et al. [16]	2005	Israel	Retrospective	31	Patients with known tumors and pleural effusion	64	81	17 NSCLC, 3 SCLC, 9 Ly, 1 OC, 1 Mel
Schaffler et al. [15]	2004	Austria	Retrospective	92	Patients with NSCLC and pleural abnormalities	58	67	92 NSCLC
Gupta et al. [14]	2002	USA	Prospective	35	Patients with lung cancer and pleural effusion	70	71	27 NSCLC, 8 SCLC
Erasmus et al. [13]	2000	USA	Retrospective	25	Patients with NSCLC and pleural effusions	65	72	25 NSCLC

BC=cáncer de mama; CC=carcinoma colorrectal; EC=carcinoma esofágico; GIST=tumor estromal gastrointestinal; HNC=cáncer de cuello y cabeza; Ly=linfoma; Mel=melanoma; Mes=mesotelioma; MT=teratoma maligno; NR=no informado; NSCLC=cáncer pulmonar célula no pequeña; OC=cáncer de ovario; SCLC=cáncer pulmonar célula pequeña.

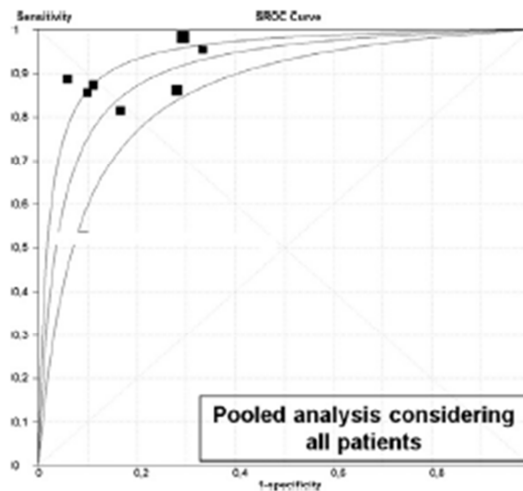
**Tabla 16: Aspectos técnicos de los 8 artículos (Treglia et al, 2014-B)**

Authors	Device	Mean <sup>18</sup> F-FDG injected activity	Mean time between <sup>18</sup> F-FDG injection and acquisition (min)	Image analysis
Liao et al. [20]	PET/CT	7.4 MBq/kg	60	Visual and semiquantitative
Letovanec et al. [19]	PET/CT	5 MBq/kg	60	Visual and semiquantitative
Alkhalwaldeh et al. [18]	PET/CT	2.52 MBq/kg	60 and 90	Visual and semiquantitative
Kim et al. [17]	PET/CT	384 MBq	60	Visual and semiquantitative
Toaff et al. [16]	PET/CT	370-666 MBq	60-120	Visual and semiquantitative
Schaffler et al. [15]	PET	300-375 MBq	60-70	Visual
Gupta et al. [14]	PET	370 MBq	60	Visual
Erasmus et al. [13]	PET	370 MBq	60	Visual

Como se observa en la Figura 15, la eficacia diagnóstica de la PET y PET-TC integrado juntos, considerando la metodología de interpretación de la imagen visual, fue:

- Sensibilidad 86%
- Especificidad 80%
- LR positiva 3,7
- LR negativa 0,18
- AUC 0,90

**Figura 15: Curva S-ROC de la valoración de las anomalías pleurales (Treglia et al, 2014-B)**



En ambos artículos, la conclusión de los autores fue que la FDG-PET es un método preciso para el diagnóstico diferencial de los DPs.

### **Limitaciones de los meta-análisis publicados**

Estos dos meta-análisis previos son, a nuestro parecer, imperfectos por una serie de limitaciones metodológicas. Esto implica que sus conclusiones, tan alentadoras respecto a la precisión diagnóstica de la PET, son cuestionables.

Por ejemplo, respecto al primer y más amplio estudio (Treglia et al, 2014-A):

1. La estrategia de búsqueda fue defectuosa: nosotros fuimos capaces de encontrar y, por lo tanto, incluir 3 estudios adicionales, durante el mismo periodo de tiempo, en nuestra revisión sistemática (Buchmann et al, 1999; Melloni et al, 2004; Jung et al, 2012)
2. Los autores no incluyeron en el meta-análisis 5 estudios por la supuesta imposibilidad de construir las tablas 2x2, cuando, de hecho, estos datos podían extraerse fácilmente (Duysinx et al,

2006; Mavi et al, 2009; Kurata et al, 2010; Abe et al, 2012; Terada et al, 2012)

3. Los autores no excluyeron ningún artículo por razones de calidad metodológica. Nosotros, sin embargo, tuvimos que eliminar de nuestra revisión sistemática 5 de los 11 estudios que los autores incluyeron, por presentar en el QUADAS-2 riesgo alto de sesgo en algún dominio. (Bénard et al, 1998; Kramer et al, 2004; Yildirim et al, 2009; Yamamoto et al, 2009; Orki et al, 2009)
4. La tabla 2x2 de uno de los 11 estudios incluidos en el meta-análisis contenía datos erróneos. (Gerbaudo et al, 2002)
5. Los autores eligieron no analizar los estudios en los que se había llevado a cabo una lectura semicuantitativa de la PET. Aún así, un estudio con metodología semicuantitativa fue interpretado de forma errónea como si fuera de técnica visual. (Orki et al, 2009).

## **D. ESTUDIO POSTERIOR**

Tras la realización de nuestro trabajo, se ha publicado un nuevo estudio aislado que evalúa la eficacia diagnóstica de la FDG-PET-TC en la detección del DP maligno (Nakajima et al, 2015).

Se trata de un estudio retrospectivo que incluyó a 36 pacientes consecutivos con diagnóstico de cáncer histológicamente confirmado (excluido mesotelioma pleural) y a los que se realizó una PET-TC ante la sospecha de DP maligno. Catorce pacientes tenían DP maligno comprobado por citología y los otros 22 pacientes presentaron benignidad pleural (9 pacientes considerados benignos por tener al menos 2 citologías negativas y 13 pacientes con seguimiento clínico mayor de 6 meses).

Se llevaron a cabo varias mediciones de la captación del líquido pleural mediante:

- La técnica de imagen cualitativa: obteniendo una sensibilidad del 57%, especificidad del 91% y una precisión diagnóstica del 78%.
- La técnica de imagen semicuantitativa: análisis del ratio entre los SUVs del líquido pleural y del tejido normal (TNR=target-to-normal tissue ratio). La TNR de la 12<sup>a</sup> vértebra torácica presentó un valor  $>0,95$ , con una sensibilidad del 93%, especificidad del 68% y una precisión diagnóstica del 75%, siendo considerado el mejor parámetro.

Los autores concluyen que la PET-TC puede ser un método útil en el diagnóstico diferencial de los DPs malignos y benignos.

## **CONCLUSIONES**

## **CONCLUSIONES**

Los resultados de este estudio sobre la precisión diagnóstica de la FDG-PET para diferenciar DPs malignos de benignos están basados en 14 estudios originales, que incluyeron a más de 600 pacientes durante los últimos 16 años. Estos resultados sugieren que:

- La FDG-PET no parece cambiar la probabilidad de malignidad pleural lo suficiente como para ser recomendada de rutina en el proceso diagnóstico de los DPs pleurales de causa incierta. (LR positiva 3,22 y LR negativa 0,26 en el análisis del PET-TC con el método de interpretación semicuantitativo).
- Nuevos estudios donde se apliquen las mejorías técnicas recientemente descritas en la FDG-PET podrían modificar las conclusiones sobre su aplicabilidad en el estudio de los derrames pleurales:
  - Mejores técnicas de adquisición de imagen para reducir los artefactos de movimiento.
  - Análisis de imagen mediante VOI para mejorar las medidas de imagen cuantitativas de la PET.
  - La técnica de la FDG-PET necesita una evaluación más extensa usando un adecuado grupo comparador de benignidad pleural, que refleje la epidemiología real de los DPs.
- El papel de la FDG-PET en los escenarios clínicos más problemáticos, como son los DPs malignos sin nodularidad o engrosamiento pleural y/o con citología falsamente negativa del líquido pleural, necesita estudiarse en el futuro.



# **BIBLIOGRAFÍA**

## **BIBLIOGRAFÍA**

1. Abe Y, Tamura K, Sakata I, et al. Clinical implications of 18F-fluorodeoxyglucose positron emission tomography/computed tomography at delayed phase for diagnosis and prognosis of malignant pleural mesothelioma. *Oncol Rep.* 2012;27(2):333-338.
2. Alkhalaf K, Biersack HJ, Henke A, Ezziddin S. Impact of dual-time-point F-18 FDG PET/CT in the assessment of pleural effusion in patients with non-small-cell lung cancer. *Clin Nucl Med.* 2011;36(6):423-428.
3. Bénard F, Sterman D, Smith RJ, Kaiser LR, Albelda SM, Alavi A. Metabolic imaging of malignant pleural mesothelioma with fluorodeoxyglucose positron emission tomography. *Chest.* 1998;114(3):713-722.
4. Bielsa S, Panadés MJ, Egido R, Rue M, Salud A, Matías-Guiu X, Rodríguez-Panadero F, Porcel JM. Rentabilidad del estudio citológico del líquido pleural en el derrame maligno. *An Med Interna.* 2008;25(4):173-177.
5. Bravata DM, Smith-Spangler C, Sundaram V, Gienger AL, Lin N, Lewis R, Stave CD, Olkin I, Sirard JR. Using pedometers to increase physical activity and improve health: a systematic review. *JAMA.* 2007;298(19):2296-304.
6. Bristol University/QUADAS/Quality assessment tool for diagnostic accuracy studies. [www.bris.ac.uk/quadas](http://www.bris.ac.uk/quadas)
7. Buchmann I, Guhlmann CA, Elsner K, et al. F-18-FDG PET for differential diagnosis of pleural processes. *Nucl Med.* 1999;38(8):319-322.
8. Bury T, Paulus P, Dowlati A, Corhay JL, Rigo P, Radermecker MF. Evaluation of pleural diseases with FDG-PET imaging: preliminary report. *Thorax.* 1997;52(2):187-189.

9. Cano Alonso R, Herráiz Hidalgo L, Álvarez Moreno E, Paniagua Correa C, Martínez de Vega V. Papel de las técnicas de imagen en la nueva clasificación TNM del carcinoma broncogénico no microcítico. *Radiología*. 2012;54(4):306-320.
10. Carretta A, Landoni C, Melloni G, et al. 18-FDG positron emission tomography in the evaluation of malignant pleural diseases – a pilot study. *Eur J Cardiothorac Surg*. 2000;17(4):377-383.
11. Cerda J, Villarroel del P.L. Evaluación de la concordancia inter-observador en investigación pediátrica: Coeficiente de Kappa. *Rev Chil Pediatr*. 2008;79(1):54-58.
12. Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 1.0. The Cochrane Collaboration, 2010.
13. Coolen J, De Keyzer F, Naftoux P, et al. Malignant pleural disease: diagnosis by using diffusion-weighted and dynamic contrast-enhanced MR imaging-initial experience. *Radiology*. 2012;263(3):884-892.
14. Duysinx B, Nguyen D, Louis R, et al. Evaluation of pleural disease with 18-fluorodeoxyglucose positron emission tomography imaging. *Chest*. 2004;125(2):489-493.
15. Duysinx BC, Larock MP, Nguyen D, et al. 18F-FDG PET imaging in assessing exudative pleural effusions. *Nucl Med Commun*. 2006;27(12):971-976.
16. Elboga U, Yilmaz M, Uyar M, Zeki Çelen Y, Bakir K, Dikensoy O. The role of FDG PET-CT in differential diagnosis of pleural pathologies. *Rev Esp Med Nucl Imagen Mol*. 2012;31(4):187-191.
17. Erasmus JJ, McAdams HP, Rossi SE, Goodman PC, Coleman RE, Patz EF. FDG PET of pleural effusions in patients with non-small cell lung cancer. *AJR Am J Roentgenol*. 2000;175(1):245-249.
18. Gerbaudo VH, Sugarbaker DJ, Britz-Cunningham S, Di Carli MF, Mauceri C, Treves ST. Assessment of malignant pleural

- mesothelioma with (18)F-FDG dual-head gamma-camera coincidence imaging: comparison with histopathology. *J Nucl Med.* 2002;43(9):1144-1149.
19. Gupta NC, Rogers JS, Graeber GM, et al. Clinical role of F-18 fluorodeoxyglucose positron emission tomography imaging in patients with lung cancer and suspected malignant pleural effusion. *Chest.* 2002;122(6):1918-1924.
20. Jung MY, Chong A, Seon HJ, et al. Indeterminate pleural metastasis on contrast-enhanced chest CT in non-small cell lung cancer: improved differential diagnosis with (18)F-FDG PET/CT. *Ann Nucl Med.* 2012;26(4):327-336.
21. Kim BS, Kim IJ, Kim SJ, Pak K, Kim K. Predictive value of F-18 FDG PET/CT for malignant pleural effusion in non-small cell lung cancer patients. *Onkologie.* 2011;34(6):298-303.
22. Kramer H, Pieterman RM, Slebos DJ, et al. PET for the evaluation of pleural thickening observed on CT. *J Nucl Med.* 2004;45(6):995-998.
23. Kurata S, Ishibashi M, Azuma K, et al. Preliminary study of positron emission tomography/computed tomography and plasma osteopontin levels in patients with asbestos-related pleural disease. *Jpn J Radiol.* 2010;28(6):446-452.
24. Letovanec I, Allenbach G, Mihaescu A, et al. <sup>18</sup>F-fluorodeoxyglucose PET/CT findings in pleural effusions of patients with known cancer. A cytopathological correlation. *Nucl Med.* 2012;51(5):186-193.
25. Liao R, Yang X, Wang S, et al. Clinical role of F-18 FDG PET/CT in differentiating malignant and benign pleural effusion in patients with lung cancer. *Zhongguo Fei Ai Za Zhi.* 2012;15(11):652-655.

26. Mavi A, Basu S, Cermik TF, et al. Potential of dual time point FDG-PET imaging in differentiating malignant from benign pleural disease. *Mol Imaging Biol.* 2009;11(5):369-378.
27. Melloni B, Monteil J, Vincent F, et al. Assessment of 18F-fluorodeoxyglucose dual-head gamma camera in asbestos lung diseases. *Eur Respir J.* 2004;24(5):814-821.
28. Moher D, Liberati A, Tetzlaff J, Altman DG. PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med.* 2009;151(4):264-269.
29. Nakajima R, Abe K, Sakai S. Diagnostic ability of FDG-PET/TC in the detection of malignant pleural effusion. *Medicine.* 2015;94(29):e1010.
30. Orki A, Akin O, Tasci AE, et al. The role of positron emission tomography/computed tomography in the diagnosis of pleural diseases. *Thorac Cardiovasc Surg.* 2009;57(4):217-221.
31. Perera R, Heneghan C. ACP Journal Club. Interpreting meta-analyses in systematic reviews. *Ann Intern Med.* 2009;150(4):JC2-2. JC2-3.
32. Perera R, Van Den Bruel A. Interpreting Diagnostic Meta-analyses. *International Pleural Newsletter.* 2011;9(3): 11-13.
33. Porcel JM. Diagnosis of pleural effusions. *Hosp Med Clin.* 2013;2:337-357(A).
34. Porcel JM, Light RW. Pleural effusions. *Dis Mon.* 2013;59(2):29-57 (B).
35. Porcel JM, et al. Etiología del derrame pleural: análisis de más de 3000 toracocentesis consecutivas. *Arch Bronconeumol.* 2014;50(5):161-165.
36. Porcel JM, Hernández P, Martínez-Alonso M, Bielsa S, Salud A. Accuracy of fluorodeoxyglucose-PET imaging for differentiating

- benign from malignant pleural effusions: a meta-analysis. *Chest*. 2015;147(2):502-512 (A).
37. Porcel JM, Pardina M, Bielsa S, González A, Light RW. Derivation and validation of a CT scan scoring system for discriminating malignant from benign pleural effusions. *Chest*. 2015;147(2):513-519 (B).
38. Qureshi NR, Rahman NM, Gleeson FV. Thoracic ultrasound in the diagnosis of malignant pleural effusion. *Thorax*. 2009;64(2):139-143
39. Schaffler GJ, Wolf G, Schoellnast H, et al. Non-small cell lung cancer: evaluation of pleural abnormalities on CT scans with 18F FDG PET. *Radiology*. 2004;231(3):858-865.
40. Smith H, Pickering RM, Struthers A, Simpson I, Mant D. Biochemical diagnosis of ventricular dysfunction in elderly patients in general practice: observational study. *BMJ*. 2000;320(7239):906-908
41. Suárez Fernández JP, Maldonado Suárez A, Domínguez Grande ML, Serna Macías JA, Kostvinseva O, Ordovás Oromendia A, Castell E, Martín C, Gorospe E, Alfonso Alfonso JM. La Tomografía por Emisión de Positrones (PET) en la práctica clínica oncológica. *Oncología*. 2004;27(8):479-489.
42. Terada T, Tabata C, Tabata R, et al. Clinical utility of 18-fluorodeoxyglucose positron emission tomography/computed tomography in malignant pleural mesothelioma. *Exp Ther Med*. 2012;4(2):197-200.
43. Toaff JS, Metser U, Gottfried M, et al. Differentiation between malignant and benign pleural effusion in patients with extrapleural primary malignancies. Assessment with positron emission tomography-computed tomography. *Invest Radiol*. 2005;40(4):204-209.

44. Treglia G, Sadeghi R, Annunziata S, et al. Diagnostic accuracy of  $^{18}\text{F}$ -FDG-PET and PET/CT in the differential diagnosis between malignant and benign pleural lesions: a systematic review and meta-analysis. *Acad Radiol*. 2014;21(1):11-20 (A).
45. Treglia G, Sadeghi R, Annunziata S, et al. Diagnostic performance of fluorine-18-fluorodeoxyglucose positron emission tomography in the assessment of pleural abnormalities in cancer patients: a systematic review and meta-analysis. *Lung Cancer*. 2014;83(1):1-7(B).
46. Valdés L, Álvarez D, Valle JM, Pose A, San José E. The etiology of pleural effusions in a area with high incidence of tuberculosis. *Chest*. 1996;109(1):158-162.
47. Valdés L, Ferreiro L, Cruz-Ferro E, González-Barcala FJ, Gude F, Ursúa MI, et al. Recent epidemiological trends in tuberculous pleural effusion in Galicia, Spain. *Eur J Intern Med*. 2012;23(8):727-732.
48. Van den Bruel A, Thompson MJ, Haj-Hassan T, Stevens R, et al. Diagnostic value of laboratory tests in identifying serious infections in febrile children: systematic review. *BMJ*. 2011;342:d3082.
49. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, Leeflang MM, Sterne JA, Bossuyt PM. QUADAS-2 Group. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies.. *Ann Intern Med*. 2011;155(8):529-536.
50. Yamamoto Y, Kameyama R, Togami T, et al. Dual time point FDG PET for evaluation of malignant pleural mesothelioma. *Nucl Med Commun*. 2009;30(1):25-29.
51. Yildirim H, Metintas M, Entok E, et al. Clinical value of fluorodeoxyglucose-positron emission tomography/computed tomography in differentiation of malignant mesothelioma from

Precisión diagnóstica de la FDG-PET para diferenciar DPs malignos de benignos

asbestos-related benign pleural disease: an observational pilot study. *J Thorac Oncol.* 2009;4(12):1480-1484.



# APÉNDICES

## **APÉNDICE 1: QUADAS-2**

## **QUADAS-2**

Las conclusiones de las revisiones sistemáticas de los estudios de eficacia diagnóstica en ocasiones se caracterizan por resultados muy heterogéneos debido a diferencias en el diseño o en el desarrollo de los estudios. Por este motivo, la valoración precisa de la calidad de cada uno de los estudios incluidos es esencial.

La calidad metodológica de los artículos incluidos en revisiones sistemáticas se evalúa mediante el QUADAS-2 (Quality Assessment of Diagnostic Accuracy Studies), herramienta rediseñada del original QUADAS que fue desarrollado en el año 2003. (Whiting et al, 2011).

El QUADAS-2 separa la calidad metodológica en “riesgo de sesgo” y “aplicabilidad”. La calidad de los estudios debe cumplir ambas características: el grado de estimación de cómo se evitan los riesgos de sesgo y el grado en el que los estudios pueden ser aplicables a la pregunta de búsqueda del investigador. Un *riesgo de sesgo* tiene lugar cuando existen defectos sistemáticos o limitaciones en el diseño o realización del estudio que distorsionan los resultados. Por otro lado, existe evidencia de que un estudio presenta una *aplicabilidad limitada* para la revisión cuando, comparada con la pregunta de búsqueda del investigador, ha sido llevada a cabo:

- En un grupo de pacientes con diferentes características demográficas o clínicas,
- La prueba diagnóstica ha sido aplicada o interpretada de forma diferente,
- O la definición de la patología objetivo es diferente

A continuación, adjuntamos la plantilla del QUADAS-2 que se puede obtener en la siguiente dirección:

<http://www.bris.ac.uk/quadas/quadas-2/>

## QUADAS-2

Phase 1: State the review question:

<i>Patients (setting, intended use of index test, presentation, prior testing):</i>
<i>Index test(s):</i>
<i>Reference standard and target condition:</i>

Phase 2: Draw a flow diagram for the primary study

Phase 3: Risk of bias and applicability judgments

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

<b>DOMAIN 1: PATIENT SELECTION</b>	
<b>A. Risk of Bias</b>	
Describe methods of patient selection:	
❖ Was a consecutive or random sample of patients enrolled?	Yes/No/Unclear
❖ Was a case-control design avoided?	Yes/No/Unclear
❖ Did the study avoid inappropriate exclusions?	Yes/No/Unclear
Could the selection of patients have introduced bias?	<b>RISK: LOW/HIGH/UNCLEAR</b>
<b>B. Concerns regarding applicability</b>	
Describe included patients (prior testing, presentation, intended use of index test and setting):	
Is there concern that the included patients do not match the review question?	<b>CONCERN: LOW/HIGH/UNCLEAR</b>

<b>DOMAIN 2: INDEX TEST(S)</b>	
If more than one index test was used, please complete for each test.	
<b>A. Risk of Bias</b>	
Describe the index test and how it was conducted and interpreted:	
❖ Were the index test results interpreted without knowledge of the results of the reference standard?	Yes/No/Unclear
❖ If a threshold was used, was it pre-specified?	Yes/No/Unclear
Could the conduct or interpretation of the index test have introduced bias?	RISK: LOW /HIGH/UNCLEAR
<b>B. Concerns regarding applicability</b>	
Is there concern that the index test, its conduct, or interpretation differ from the review question?	CONCERN: LOW /HIGH/UNCLEAR

<b>DOMAIN 3: REFERENCE STANDARD</b>	
<b>A. Risk of Bias</b>	
Describe the reference standard and how it was conducted and interpreted:	
❖ Is the reference standard likely to correctly classify the target condition?	Yes/No/Unclear
❖ Were the reference standard results interpreted without knowledge of the results of the index test?	Yes/No/Unclear
Could the reference standard, its conduct, or its interpretation have introduced bias?	RISK: LOW /HIGH/UNCLEAR
<b>B. Concerns regarding applicability</b>	
Is there concern that the target condition as defined by the reference standard does not match the review question?	CONCERN: LOW /HIGH/UNCLEAR

<b>DOMAIN 4: FLOW AND TIMING</b>	
<b>A. Risk of Bias</b>	
Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2x2 table (refer to flow diagram):	
Describe the time interval and any interventions between index test(s) and reference standard:	
❖ Was there an appropriate interval between index test(s) and reference standard?	Yes/No/Unclear
❖ Did all patients receive a reference standard?	Yes/No/Unclear
❖ Did patients receive the same reference standard?	Yes/No/Unclear
❖ Were all patients included in the analysis?	Yes/No/Unclear
Could the patient flow have introduced bias?	<b>RISK: LOW /HIGH/UNCLEAR</b>

## **INCORPORACIÓN DE LA EVALUACIÓN DEL QUADAS-2 EN LAS REVISIONES DE PRECISIÓN DIAGNÓSTICA**

Si un estudio se clasifica como “bajo” en todos los dominios relacionados con el sesgo o aplicabilidad, entonces es apropiado realizar un juicio global como “bajo riesgo de sesgo” o “bajo problema de aplicabilidad” para ese estudio. Si un estudio es determinado como “alto” o “poco claro” en uno o más dominios, entonces el juicio global será de “alto riesgo de sesgo” o de tener “problemas en cuanto a la aplicabilidad”.

Como mínimo, los revisores deberán presentar un resumen de los resultados de la evaluación del QUADAS-2 de todos los artículos incluidos. Podría incluir un resumen del número de estudios en los que se ha determinado que tienen un bajo, alto o poco claro riesgo de sesgo/problema de aplicabilidad. Los revisores pueden destacar aquellas preguntas de señalización en las que los estudios son evaluados mejor o peor de forma reiterada. Estos resultados pueden representarse con los gráficos que se muestran en las Figuras 16 a 19.

**Figuras 16-19: Presentación de los resultados del QUADAS-2**

**Figura 16:**

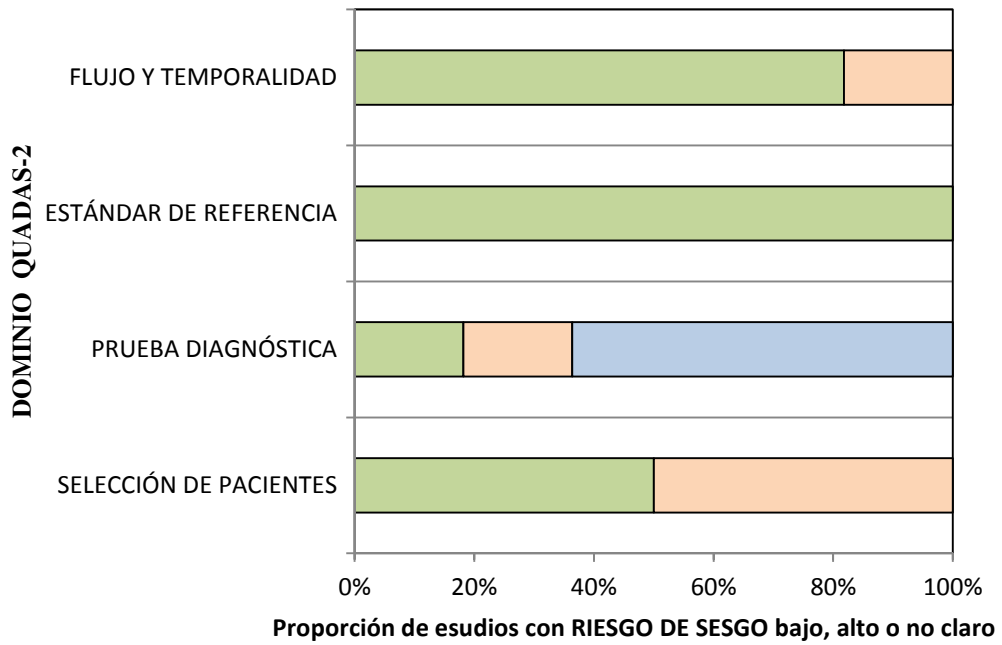
Estudio	RIESGO DE SESGO				APLICABILIDAD		
	SELECCIÓN DE PACIENTES	PRUEBA DIAGNÓSTICA	ESTÁNDAR DE REFERENCIA	FLUJO Y TEMPORALIDAD	SELECCIÓN DE PACIENTES	PRUEBA DIAGNÓSTICA	ESTÁNDAR DE REFERENCIA
Estudio 1	J	J	J	J	L	J	J
Estudio 2	J	J	J	J	L	J	J
Estudio 3	L	L	J	J	L	J	J
Estudio 4	L	L	J	J	L	J	J
Estudio 5	L	?	J	J	L	J	J
Estudio 6	L	?	J	J	L	?	J
Estudio 7	L	?	J	J	L	J	J
Estudio 8	L	?	J	J	L	?	J
Estudio 9	L	?	J	J	L	J	J
Estudio 10	L	?	J	L	L	J	J
Estudio 11	J	?	J	L	J	J	J

J Riesgo bajo    L Riesgo alto    ? Riesgo poco claro

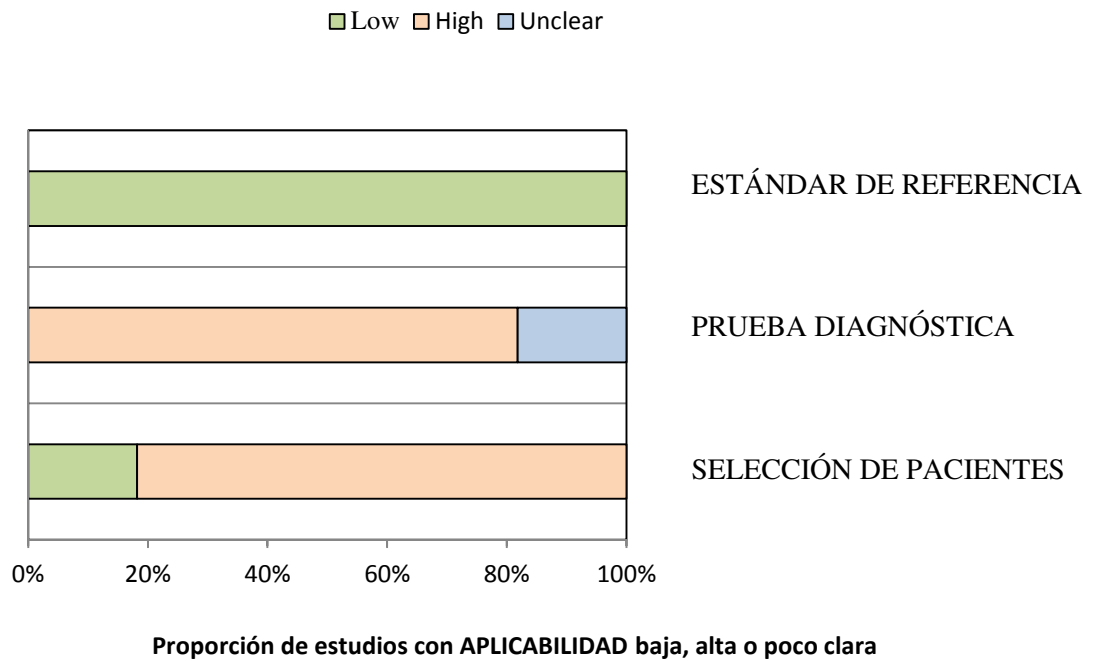
**Figura 17:**

	SELECCIÓN DE PACIENTES	PRUEBA DIAGNÓSTICA	ESTÁNDAR DE REFERENCIA	FLUJO Y TEMPORALIDAD
Estudio 1	Bajo	Bajo	Bajo	Bajo
Estudio 2	Bajo	Bajo	Bajo	Bajo
Estudio 3	Alto	Alto	Bajo	Bajo
Estudio 4	Alto	Alto	Bajo	Bajo
Estudio 5	Alto	Poco claro	Bajo	Poco claro
Estudio 6	Alto	Poco claro	Bajo	Bajo
Estudio 7	Alto	Poco claro	Poco claro	Bajo
Estudio 8	Alto	Poco claro	Bajo	Bajo
Estudio 9	Alto	Poco claro	Bajo	Bajo
Estudio 10	Bajo	Poco claro	Bajo	Alto
Estudio 11	Poco claro	Poco claro	Bajo	Alto

**Figura 18:**



**Figura 19:**





**APÉNDICE 2:**  
**ARTÍCULOS EXCLUIDOS**

**ARTÍCULOS NO INCLUIDOS EN LA REVISIÓN SISTEMÁTICA Y  
META-ANÁLISIS POR CAUSAS DE EXCLUSIÓN**

A. OBJETIVO DIFERENTE AL DE LA PREGUNTA DE REVISIÓN:

1. Al-Sarraf N, Gately K, Lucey J, Aziz R, Doddakula K, Wilson L, McGovern E, Young V. Clinical implication and prognostic significance of standardised uptake value of primary non-small cell lung cancer on positron emission tomography: analysis of 176 cases. *Eur J Cardiothorac Surg.* 2008;34(4):892-897.
2. Bénard F, Serman D, Smith RJ, Kaiser LR, Albelda SM, Alavi A. Prognostic value of FDG PET imaging in malignant pleural mesothelioma. *J Nucl Med.* 1999;40(8):1241-1245.
3. Higashi K, Ueda Y, Ayabe K, Sakurai A, Seki H, Nambu Y, Oguchi M, Shikata H, Taki S, Tonami H, Katsuda S, Yamamoto I. FDG PET in the evaluation of aggressiveness of pulmonary adenocarcinoma: correlation with histopathological features. *Nucl Med Commun.* 2000;21(8):707-714.
4. Kanauchi N, Oizumi H, Honma T, Kato H, Endo M, Suzuki J, Fukaya K, Sadahiro M. Role of diffusion-weighted magnetic resonance imaging for predicting of tumor invasiveness for clinical stage IA non-small cell lung cancer. *Eur J Cardiothorac Surg.* 2009;35(4):706-710.
5. Maeda R, Isowa N, Onuma H, Miura H, Harada T, Touge H, Tokuyasu H, Kawasaki Y. The maximum standardized 18F-fluorodeoxyglucose uptake on positron emission tomography predicts lymph node metastasis and invasiveness in clinical stage IA non-small cell lung cancer. *Interact Cardiovasc Thorac Surg.* 2009;9(1):79-82.

6. Maeda R, Isowa N, Onuma H, Miura H, Harada T, Touge H, Tokuyasu H, Kawasaki Y. The maximum standardized uptake values on positron emission tomography to predict the Noguchi classification and invasiveness in clinical stage IA adenocarcinoma measuring 2 cm or less in size. *Interact Cardiovasc Thorac Surg.* 2009;9(1):70-73.
7. McAdams HP, Erasmus JJ, Patz EF, Goodman PC, Coleman RE. Evaluation of patients with round atelectasis using 2-(18F)-fluoro-2-deoxy-D-glucose PET. *J Comput Assist Tomogr.* 1998;22(4):601-604.
8. Risum S, Høgdall C, Loft A, Berthelsen AK, Høgdall E, Nedergaard L, Lundvall L, Engelholm SA. Prediction of suboptimal cytoreduction in primary ovarian cancer with combined positron emission tomography/computed tomography—a prospective study. *Gynecol Oncol.* 2008;108(2):265-270.
9. Schmid RA, Hautmann H, Poellinger B, Kellner W, Moisseev A, Brinkbaeumer K, Weiss M, Hahn K, Dresel S. Staging of recurrent and advanced lung cancer with 18F-FDG PET in a coincidence technique (hybrid PET). *Nucl Med Commun.* 2003;24(1):37-45.
10. Vesselle H, Pugsley JM, Vallières E, Wood DE. The impact of fluorodeoxyglucose F18 positron-emission tomography on the surgical staging of non-small cell lung cancer. *J Thorac Cardiovasc Surg.* 2002;124(3):511-519.

#### B. IMPOSIBILIDAD DE CONSTRUCCIÓN DE LA TABLA 2x2:

1. Balogova S, Grahek D, Kerrou K, Montravers F, Younsi N, Aide N, Jacob T, Talbot JN. (18F)-FDG imaging in apparently isolated pleural lesions. *Rev Pneumol Clin.* 2003;59(5 Pt 1):275-288.

2. Otsuka H, Terazawa K, Morita N, Otomi Y, Yamashita K, Nishitani H. Is FDG-PET/CT useful for managing malignant pleural mesothelioma? *J Med Invest*. 2009;56(1-2):16-20.
3. Schneider DB, Clary-Macy C, Challa S, Sasse KC, Merrick SH, Hawkins R, Caputo G, Jablons D. Positron emission tomography with f18-fluorodeoxyglucose in the staging and preoperative evaluation of malignant pleural mesothelioma. *J Thorac Cardiovasc Surg*. 2000;120(1):128-133.

C. NÚMERO TOTAL DE PACIENTES INFERIOR A 10:

1. Kutlu CA, Pastorino U, Maisey M, Goldstraw P. Selective use of PET scan in the preoperative staging of NSCLC. *Lung Cancer*. 1998;21(3):177-184.
2. Oh JK, Ahn MI, Kim CH, Cho KD, Cho DG, Kang CU, Yoo J. The value of F-18-FDG-PET/CT in diagnosis of chronic empyema-associated malignancy. *Clin Radiol*. 2008;63(10):1177-1180.

D. CRITERIOS INADECUADOS PARA EL ESTÁNDAR DE REFERENCIA:

1. Fischer BM, Mortensen J, Langer SW, Loft A, Berthelsen AK, Petersen BI, Daugaard G, Lassen U, Hansen HH. A prospective study of PET/CT in initial staging of small-cell lung cancer: comparison with CT, bone scintigraphy and bone marrow analysis. *Ann Oncol*. 2007;18(2):338-345.

E. ESTUDIO DUPLICADO:

1. Gerbaudo VH, Britz-Cunningham S, Sugarbaker DJ, Treves ST. Metabolic significance of the pattern, intensity and kinetics of 18F-FDG uptake in malignant pleural mesothelioma. *Thorax*. 2003;58(12):1077-1082.

F. ESTUDIO CON ANIMALES:

1. Zhuang H, Pourdehnad M, Lambright ES, Yamamoto AJ, Lanuti M, Li P, Mozley PD, Rossman MD, Albelda SM, Alavi A. Dual time point 18F-FDG PET imaging for differentiating malignant from inflammatory processes. *J Nucl Med.* 2001;42(9):1412-1417.

**APÉNDICE 3:**  
**ARTÍCULO RELACIONADO**

# Accuracy of Fluorodeoxyglucose-PET Imaging for Differentiating Benign From Malignant Pleural Effusions

## A Meta-analysis

José M. Porcel, MD, FCCP; Paula Hernández, MD; Montserrat Martínez-Alonso, BSc; Silvia Bielsa, MD; and Antonieta Salud, MD

**BACKGROUND:** The role of fluorodeoxyglucose (FDG)-PET imaging for diagnosing malignant pleural effusions is not well defined. The aim of this study was to summarize the evidence for its use in ruling in or out the malignant origin of a pleural effusion or thickening.

**METHODS:** A meta-analysis was conducted of diagnostic accuracy studies published in the Cochrane Library, PubMed, and Embase (inception to June 2013) without language restrictions. Two investigators selected studies that had evaluated the performance of FDG-PET imaging in patients with pleural effusions or thickening, using pleural cytopathology or histopathology as the reference standard for malignancy. Subgroup analyses were conducted according to FDG-PET imaging interpretation (qualitative or semiquantitative), PET imaging equipment (PET vs integrated PET-CT imaging), and/or target population (known lung cancer or malignant pleural mesothelioma). Study quality was assessed using Quality Assessment of Diagnostic Accuracy Studies-2. We used a bivariate random-effects model for the analysis and pooling of diagnostic performance measures across studies.

**RESULTS:** Fourteen non-high risk of bias studies, comprising 407 patients with malignant and 232 with benign pleural conditions, met the inclusion criteria. Semiquantitative PET imaging readings had a significantly lower sensitivity for diagnosing malignant effusions than visual assessments (82% vs 91%;  $P = .026$ ). The pooled test characteristics of integrated PET-CT imaging systems using semiquantitative interpretations for identifying malignant effusions were: sensitivity, 81%; specificity, 74%; positive likelihood ratio (LR), 3.22; negative LR, 0.26; and area under the curve, 0.838. Resultant data were heterogeneous, and spectrum bias should be considered when appraising FDG-PET imaging operating characteristics.

**CONCLUSIONS:** The moderate accuracy of PET-CT imaging using semiquantitative readings precludes its routine recommendation for discriminating malignant from benign pleural effusions.

CHEST 2015; 147(2):502-512

Manuscript received April 3, 2014; revision accepted August 13, 2014; originally published Online First September 4, 2014.

**ABBREVIATIONS:** AUC = area under the curve; FDG = fluorodeoxyglucose; LR = likelihood ratio; MPM = malignant pleural mesothelioma; QUADAS = Quality Assessment of Diagnostic Accuracy Studies; SUV = standardized uptake value; SUV<sub>max</sub> = maximum standardized uptake value; VOI = volume of interest

**AFFILIATIONS:** From the Pleural Diseases Unit, Departments of Internal Medicine (Drs Porcel, Hernández, and Bielsa), Biostatistics (Ms Martínez-Alonso), and Oncology-Hematology (Dr Salud), Arnau de Vilanova University Hospital, Biomedical Research Institute of Lleida, Lleida, Spain.

**FUNDING/SUPPORT:** The authors have reported to CHEST that no funding was received for this study.

**CORRESPONDENCE TO:** José M. Porcel, MD, FCCP, Department of Internal Medicine, Arnau de Vilanova University Hospital, Avda Alcalde Rovira Roure 80, 25198 Lleida, Spain; e-mail: jporcel@yahoo.es

© 2015 AMERICAN COLLEGE OF CHEST PHYSICIANS. Reproduction of this article is prohibited without written permission from the American College of Chest Physicians. See online for more details.

DOI: 10.1378/chest.14-0820

Cancer is the most frequent cause of pleural effusions subjected to thoracentesis.<sup>1</sup> The diagnosis of malignant effusions can be challenging, since the overall sensitivity of a pleural fluid cytologic examination is, at most, 60%.<sup>1,2</sup> In certain tumor types, such as malignant pleural mesothelioma (MPM), this percentage is much lower.<sup>1</sup> Therefore, a pleural biopsy through a medical thoracoscopy is often necessary to reach a definitive diagnosis of malignancy. However, medical thoracoscopy is an invasive procedure and is not always available. Obviously, the ability to identify patients with a low probability of malignant effusion would reduce the number of unnecessary thoracoscopic procedures. Conversely, clinicians would also appreciate being able to identify those subjects with a sufficiently high likelihood of pleural malignancy, thereby justifying the need for thoracoscopy. Use of a noninvasive screening test to make a presumptive diagnosis of malignant effusion would help to achieve these goals.

Traditional thoracic imaging (ie, CT scanning and ultrasonography) has limitations for discriminating benign

from malignant pleural conditions. Pleural nodules or thickening, being highly suggestive of malignancy, are present in fewer than one-half of cases when detected by operator-dependent ultrasound or CT scan.<sup>3</sup>

Fluorodeoxyglucose (FDG)-labeled metabolic PET imaging has been increasingly used to stage primary malignancies, but its ability to provide information about the benign or malignant nature of an effusion and/or thickening has not been adequately studied.<sup>4</sup> This technique is based on the differential metabolism of normal and abnormal tissues, the uptake of FDG being accelerated in tumor cells. Unfortunately, metabolically active infectious or inflammatory lesions can lead to false-positive results.

The aim of this meta-analysis was to systematically review and summarize the available literature on the diagnostic accuracy of FDG-PET imaging in distinguishing benign from malignant pleural effusions or thickening to evaluate whether this metabolic imaging technique would help reduce the number of unnecessary thorascopies.

## Materials and Methods

### Search Strategy and Study Selection

The systematic review was performed according to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines.<sup>5</sup> This study was registered with the International Prospective Register of Systematic Reviews (<http://www.crd.york.ac.uk/PROSPERO>) as CRD42011001392 ("Accuracy of FDG-PET for diagnosing malignant pleural effusions: a systematic review and meta-analysis"). We searched the Cochrane Library, PubMed/MEDLINE, and EMBASE from inception to June 2013 to identify studies on the diagnostic accuracy of FDG-PET imaging for detecting malignant pleural effusions and/or thickening. The search strategy used "pleural," "positron emission tomography," and related terms, without language restrictions (Table 1). The retrieved articles' accompanying references were also searched.

Two researchers (J. M. P. and P. H.) independently selected eligible publications and extracted data from the included studies, with disagreements resolved by consensus. The inclusion criteria were as follows: diagnostic study design, use of FDG-PET imaging as the index test, use of pleu-

ral cytologic or histologic verification as the reference standard for malignancy, and reporting of sufficient data to construct 2 × 2 contingency tables. A study was excluded if any of the inclusion criteria were not met.

### Data Extraction

The following characteristics were extracted from each selected study: year of publication; study design; country; recruitment strategy; sample size; patient demographic characteristics (age, sex); prevalence of pleural malignancy in the study population (pretest probability); causes of benign effusions; primary tumor types in malignant effusions; reference standard; use of stand-alone FDG-PET imaging or combined PET-CT imaging technologies; methods of FDG-PET image analysis (ie, qualitative or visual, semiquantitative using the standardized uptake value [SUV] or dual time point in which changes in maximum SUV [SUVmax] over time between two time-point scans are recorded); Quality Assessment of Diagnostic Accuracy Studies (QUADAS) 2 score; and data for construction of a 2 × 2 table. One author was contacted to provide data on his study published in Chinese.<sup>6</sup>

### Assessment of Risk of Bias

Two researchers (J. M. P. and P. H.) independently assessed methodologic study quality by using the QUADAS-2 checklist.<sup>7</sup> QUADAS-2 (University of Bristol; [www.quadas.org](http://www.quadas.org)) is an improved, redesigned tool from the original QUADAS, which comprises four domains: (1) patient selection, (2) index test, (3) reference standard, and (4) flow of patients through the study and timing of both the index test and reference standard (flow and timing). QUADAS-2 assesses each domain in terms of risk of bias (low, high, or unclear). The first three domains are also evaluated regarding concerns pertaining to applicability. Tabular displays helped to summarize QUADAS-2 assessments.

### Data Analysis

Continuous variables were expressed as means or medians and categorical variables as frequencies or percentages. Sensitivities, specificities, likelihood ratios (LRs), and diagnostic OR summary estimates

**TABLE 1** Bibliographic Search Strategy: Search Syntax in PUBMED

Search Syntax
(pleura or pleural or pleural disease or pleural diseases or pleural effusion or pleural effusions or malignant pleural or malignant pleural effusion or malignant pleural effusions or mesothelioma or malignant mesothelioma or benign pleural or benign pleural effusion or benign pleural effusions) AND (PET or positron emission tomography or fluorodeoxyglucose or 18-fluorodeoxyglucose or FDG PET or <sup>18</sup> F-FDG PET or <sup>18</sup> F-FDG)

FDG = fluorodeoxyglucose.



of the studies, with their corresponding 95% CIs, were calculated from the  $2 \times 2$  table, using bivariate models with a random-effects approach. For tables in which any cell had a zero value, 0.5 was added to every cell to calculate the LRs. We present the data as forest plots and hierarchical summary receiver operating characteristic curves. The operating characteristics of FDG-PET imaging were examined according to the equipment used (PET or the hybrid PET-CT imaging) and the tracer-uptake interpretation, whether visual (ie, by a qualitative comparison with the level of tracer activity in normal surrounding tissues); semiquantitative, by the SUV; or using a dual time-point technique (where two time-point acquisition images are compared semiquantitatively). In addition, differences in FDG-PET imaging accuracy were explored across different subgroups of patients, namely, those with pleural effusions associated with a confirmed lung cancer or MPM.

## Results

### *Characteristics of the Included Studies*

A total of 619 studies were identified, of which 45 were retrieved for full-text review. Only 27 of those, published between 1997 and 2012, met the inclusion criteria.<sup>6,8-33</sup> Table 2 summarizes the design, participants, and findings of the initially selected studies. Subsequently, 12 studies with a high risk of bias based on the QUADAS-2 tool<sup>9,12,14,17-19,22-24,26,28,31</sup> and another, which was the sole remaining study based on a dual time-point technique<sup>21</sup> were excluded (Fig 1). Consequently, only 14 studies entered the final analysis, of which four were prospective,<sup>16,20,25,32</sup> five were retrospective,<sup>6,15,27,29,30</sup> and five had an unrevealed design.<sup>8,10,11,13,33</sup> These eligible studies comprised 639 patients (66% male patients) with a median age of 62 years, of whom 407 (64%) eventually had malignant effusions (or thickening) and 232 had benign pleural conditions. Specifically, the causes of the malignant pleural effusions included 156 MPMs (98 epithelial, 20 sarcomatoid or mixed, and 38 nonspecified) and 251 pleural metastases, either from lung cancer (135 non-small cell histology, seven small-cell, and 27 nonspecified), nonspecified origin (n = 59), or miscellaneous (n = 23). No data on the precise etiology were reported in 76 cases (33%) of benign effusions or thickening, whereas the remaining ones were attributed to nonspecific pleuritis (n = 40), parapneumonic/empyema (n = 49), benign asbestosis (n = 22), fibrosis (n = 23), TB (n = 13), and benign tumors (n = 9).

The intended use of the index test was to discriminate between benign and malignant pleural effusions (or thickening) either in patients with effusions of uncertain etiology,<sup>8,10,11,16,20,32</sup> a previously diagnosed lung cancer,<sup>6,15,27,29</sup> or a suspected or confirmed MPM.<sup>13,25,30,33</sup> Moreover, FDG-PET imaging metabolic activity was interpreted by a qualitative method in seven studies,<sup>8,10,11,13,15,16,32</sup> a semiquantitative method in three studies,<sup>20,25,33</sup> and by both

Calculations were performed after excluding studies having a high risk of bias. A random effects meta-regression analysis allowed for the comparison of the pooled sensitivities and specificities among different FDG-PET imaging reading methodologies and patient subgroups.

The interobserver agreement in the QUADAS-2 assessment was determined by using the unweighted Cohen  $\kappa$  statistic, in which results suggest "good" agreement if  $> 0.6$  and "excellent" agreement if  $> 0.8$ . Heterogeneity between studies was addressed using a random-effect approach model with a logit link function. Publication bias was assessed using a funnel plot for positive and negative LRs. Statistical significance was assumed at the  $P < .05$  level. Computations were performed with R-package "mada" statistical software (R-project; <http://cran.r-project.org/web/packages/mada/index.html>).

in four studies.<sup>6,27,29,30</sup> For semiquantitative assessments over a region of interest, SUVmax optimal cutoff discriminating values ranged from 2.2 to 3.5. Examinations with hybrid PET-CT imaging machines were performed in seven studies, which were the most recent (from 2010 onward).<sup>6,25,27,29,30,32,33</sup> All studies recorded the FDG uptake in the pleural membranes, regardless of whether they were morphologically abnormal. Yet, in two studies, the tracer uptake of the fluid itself was also assessed.<sup>27,30</sup> It should be noted, however, that the great majority of patients evaluated in the studies had pleural thickening and/or nodularity on CT scans.

### *Diagnostic Accuracy Indexes for Malignant Effusions*

Table 3 presents the operating characteristics of FDG-PET imaging for pleural malignancy according to the PET imaging reading interpretation methodology (qualitative or semiquantitative), through the use of either a PET or an integrated PET-CT imaging system. The 11 studies interpreting FDG-PET imaging visually or qualitatively<sup>6,8,10,11,13,15,16,27,29,30,32</sup> yielded a pooled sensitivity of 91%, a specificity of 67%, a positive LR of 2.83, a negative LR of 0.14, and an AUC of 0.893 (Fig 2). The corresponding diagnostic performance measures for the seven studies using SUV-based readings<sup>6,20,25,27,29,30,33</sup> were as follows: sensitivity, 82%; specificity, 74%; positive LR, 3.24; negative LR, 0.25; and AUC, 0.840 (Fig 3). Overall, the pooled sensitivity of FDG-PET imaging was significantly higher by using qualitative as compared with semiquantitative assessments ( $P = .026$ ). Studies that used a stand-alone PET imaging system contributed the most to the better sensitivity of qualitative readings (96%). In fact, when only integrated PET-CT imaging studies were considered, accuracy was not significantly different between qualitative and semiquantitative readings ( $P = .208$ ). Specificity values did not differ among the groups listed in Table 3. Although etiological data for a significant

**TABLE 2 ] Characteristics of the 14 Included and 13 Excluded Diagnostic Accuracy Studies Evaluating Fluorodeoxyglucose-PET Imaging in Pleural Effusions**

Study/Year	Country	Design	Age, Mean, y	Male Patients, %	Malignant/Benign Conditions, No.	MPM/Lung Cancer, No.	Device	Image Analysis	TP	FP	FN	TN
Included studies												
Bury et al <sup>8</sup> /1997	Belgium	NR	60	NR	16/9	3/NR	PET	Qualitative	16	2	0	7
Buchmann et al <sup>10</sup> /1999	Germany	NR	58	81	12/4	9/3	PET	Qualitative	12	0	0	4
Carretta et al <sup>11</sup> /2000	Italy	NR	58	78	13/1	10/NR	PET	Qualitative	12	0	1	1
Gerbaudo et al <sup>13</sup> /2002	United States	NR	60	73	11/4	11/0	PET	Qualitative	11	1	0	3
Schaffler et al <sup>15</sup> /2004	Austria	Retrospective	58	67	30/38	0/30	PET	Qualitative	30	15	0	23
Duysinx et al <sup>16</sup> /2004	Belgium	Prospective	61	68	63/35	13/4	PET	Qualitative	61	4	2	31
Duysinx et al <sup>20</sup> /2006	Belgium	Prospective	63	59	51/28	8/25	PET	SUV > 2.2	44	7	7	21
Kurata et al <sup>25</sup> /2010	Japan	Prospective	65	90	6/4	6/0	PET-CT	SUV > 2.2	6	2	0	2
Alkhalwaidh et al <sup>27</sup> /2011	Germany	Retrospective	63	64	29/32	0/29	PET-CT	Qualitative	25	9	4	23
								SUV > 2.4	27	9	2	23
								Dual time SUV	10	1	5	15
Jung et al <sup>29</sup> /2012	Korea	Retrospective	63.5	60	51/12	0/51	PET-CT	Qualitative	47	5	4	7
Elboga et al <sup>30</sup> /2012	Turkey	Retrospective	58	36	37/13	37/0	PET-CT	Qualitative	34	5	3	8
								SUV > 2.5	34	5	3	8
Coolen et al <sup>32</sup> /2012	Belgium	Prospective	60	77	14/17	12/NR	PET-CT	Qualitative	14	11	0	6
Terada et al <sup>33</sup> /2012	Japan	NR	67.5	79	47/29	47/0	PET-CT	SUV > 3.5	28	2	19	27
Liao et al <sup>6</sup> /2012	China	Retrospective	61	73	27/6	0/27	PET-CT	Qualitative	22	1	5	5
								SUV > 2.5	16	1	11	5

(Continued)

TABLE 2 ] (continued)

Study/Year	Country	Design	Age, Mean, y	Male Patients, %	Malignant/Benign Conditions, No.	MPM/Lung Cancer, No.	Device	Image Analysis	TP	FP	FN	TN
Excluded studies												
Bénard et al <sup>9</sup> /1998	United States	NR	NR	NR	24/4	22/NR	PET	Qualitative	22	1	2	3
Erasmus et al <sup>12</sup> /2000	United States	Retrospective	65	72	22/3	22/0	PET	SUV > 2 Qualitative	20	0	2	4
Gupta et al <sup>14</sup> /2002	United States	NR	70	71	18/17	0/18	PET	Qualitative	16	1	2	16
Kramer et al <sup>17</sup> /2004	Holland	Prospective	NR	84	19/13	10/NR	PET	Qualitative	18	1	1	12
Melloni et al <sup>18</sup> /2004	France	Retrospective	61	NR	8/3	6/1	PET	Qualitative	7	3	1	0
Toaff et al <sup>19</sup> /2005	Israel	Retrospective	64	81	21/10	0/14	PET-CT	Qualitative	18	1	3	9
Mavi et al <sup>21</sup> /2009	United States	NR	61.4	87	44/11	44/0	PET	Dual-time SUV	43	2	1	9
Yildirim et al <sup>22</sup> /2009	Turkey	NR	61	65	17/14	17/0	PET-CT	Qualitative	15	1	2	13
Yamamoto et al <sup>23</sup> /2009	Japan	Retrospective	64	91	17/16	17/0	PET	SUV > 2.2 Qualitative	16	0	1	14
Orki et al <sup>24</sup> /2009	Turkey	NR	47	76	44/39	25/NR	PET-CT	SUV > 3	44	2	0	37
Kim et al <sup>26</sup> /2011	Korea	Retrospective	67	61	24/9	0/24	PET-CT	Qualitative	21	1	3	8
Abe et al <sup>28</sup> /2012	Japan	Retrospective	67	87	31/59	31/0	PET-CT	Dual-time SUV	31	7	0	52
Letovanec et al <sup>31</sup> /2012	Swiss	Retrospective	60	62	17/33	2/6	PET-CT	SUV > 2.2	9	3	8	30

FN = false negative; FP = false positive; MPM = malignant pleural mesothelioma; NR = not reported; SUV = standardized uptake value; TN = true negative; TP = true positive.

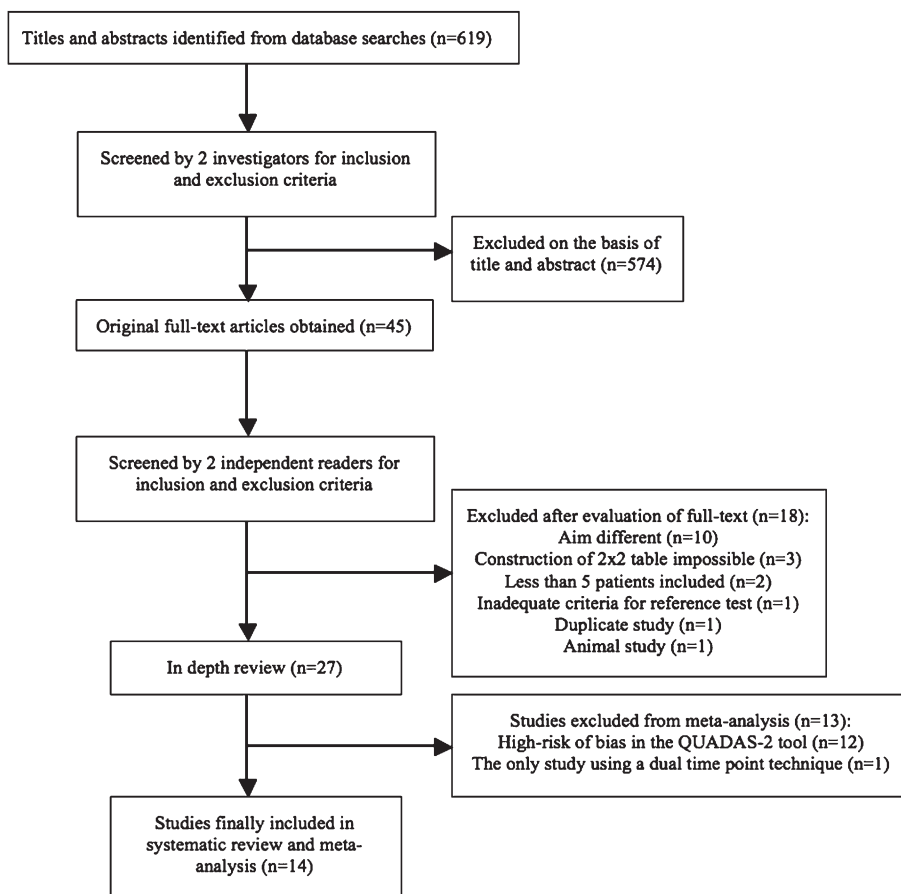


Figure 1 – Flowchart of search strategy and study selection. QUADAS = Quality Assessment of Diagnostic Accuracy Studies.

proportion of benign pleural conditions were lacking, it should be noted that five of 13 (38.5%) tuberculous effusions and 21 of 49 (43%) parapneumonic effusions or empyemas exhibited avid FDG-PET imaging uptake.

Meta-regression analysis showed that the pleural benign-malignant discriminating capacity of FDG-PET imaging remained unchanged regardless of the target population, whether patients with lung cancer or MPM

TABLE 3 ] Summary Measures of Diagnostic Accuracy for FDG-PET Imaging in the Identification of Malignant Pleural Effusions

Measure	Qualitative PET Imaging Readings	Semiquantitative PET Imaging Readings	Qualitative Readings Using PET Imaging Systems	Qualitative Readings Using Integrated PET-CT Imaging Systems	Semiquantitative Readings Using Integrated PET-CT Imaging Systems
Studies, No.	11	7	6	5	6
Sensitivity (95% CI), %	91 (86-94)	82 (69-90)	96 (90-98)	89 (81-94)	81 (66-91)
Specificity (95% CI), %	67 (56-77)	74 (62-84)	75 (59-87)	61 (44-75)	74 (58-85)
Positive LR (95% CI)	2.83 (2.04-3.98)	3.24 (2.24-4.70)	4.09 (2.34-7.1)	2.32 (1.61-3.45)	3.22 (2-5)
Negative LR (95% CI)	0.14 (0.08-0.22)	0.25 (0.14-0.39)	0.06 (0.03-0.13)	0.19 (0.11-0.31)	0.26 (0.14-0.43)
DOR (95% CI)	22 (10.2-41.7)	13.8 (7.3-23.9)	76.7 (22.1-195)	13 (6-24.8)	13.3 (6.32-24.8)
AUC	0.893	0.840	0.950	0.859	0.838

AUC = area under the curve; DOR = diagnostic OR; LR = likelihood ratio. See Table 1 legend for expansion of other abbreviation.

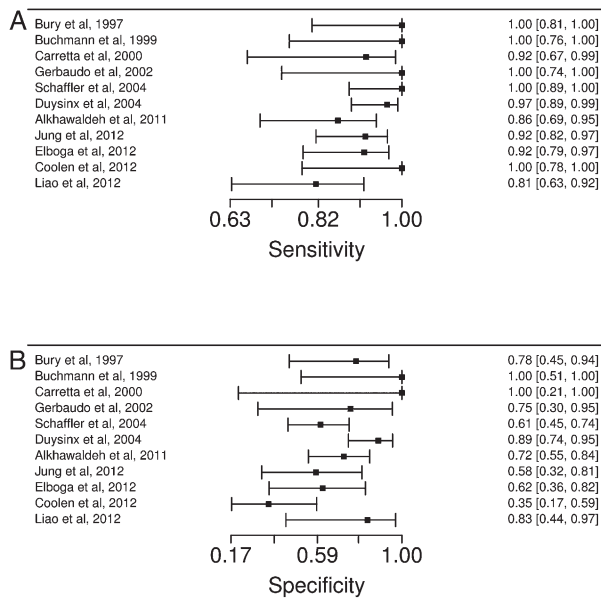


Figure 2 – A, B, Forest plot of the sensitivity (A) and specificity (B) of fluorodeoxyglucose-PET imaging, using a qualitative interpretation, for the diagnosis of malignant pleural effusions. The sensitivity and specificity of individual studies are represented by a square, through which runs a horizontal line (95% CI).

( $P > .4$ ). Figure 4 presents the diagnostic values of the studies in a hierarchical summary receiver operating characteristic graph. The more the curve approaches the left corner, the better the diagnostic accuracy, whereas those closer to the diagonal have poorer diagnostic accuracy.

### Risk of Bias and Heterogeneity

Figure 5 summarizes the QUADAS-2 assessment for the initial 27 studies. The interobserver correlation for this evaluation was good ( $\kappa = 0.77$ ). Fifteen<sup>6,8,10,11,13,15,16,20,21,25,27,29,30,32,33</sup> of the 27 studies lacked a high risk of bias in any of the QUADAS-2 domains. Overall, the quality of the studies was modest because none of the 15 was rated as having a low risk of bias in the four domains, and only five<sup>15,16,20,25,32</sup> met that risk characteristic in three of the domains of the quality tool.

Asymmetric funnel tests for the positive and negative LRs in the studies dealing with qualitative FDG-PET imaging readings ( $P < .05$ ) were consistent with publication bias. The same applied to the negative LR of the studies using semiquantitative readings ( $P = .03$ ). In contrast, there was no significant publication bias for either qualitative or semiquantitative readings when the PET imaging system, whether alone or in combination with CT imaging, was taken into account (all  $P > .10$ ). Finally, there was significant statistical heterogeneity for

sensitivity and specificity measures in all subgroups of studies, with the exception being those that dealt with qualitative readings using a PET imaging system (SD of sensitivity, 0.062).

### Clinical Application

According to the Bayesian method, estimates of the posttest probability of malignant effusion in patients who underwent an FDG-PET imaging examination are a function of disease prevalence (pretest probability). For instance, in a population with a malignancy prevalence of 35% among patients presenting with pleural exudates,<sup>1</sup> the finding of a positive PET-CT scan using a semiquantitative interpretation (LR positive, 3.22) would increase this probability to 63.5%, whereas a negative result (LR negative, 0.26) would decrease the probability of malignant pleuritis to 12.3%. However, these figures would have changed to 68.7% (LR positive, 4.09) and 3% (LR negative, 0.06), respectively, with a PET scan interpreted with qualitative criteria.

### Discussion

#### Summary of Evidence

This meta-analysis summarizes 14 non-high risk of bias studies that assessed the diagnostic usefulness of FDG-PET imaging in patients with pleural effusions (or thickening). According to the data, in the best-case scenario, the yield of stand-alone PET imaging systems for labeling

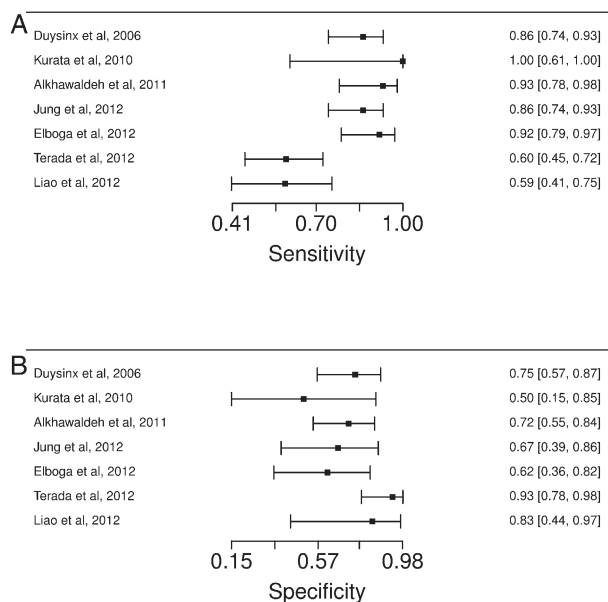


Figure 3 – A, B, Forest plot of the sensitivity (A) and specificity (B) of fluorodeoxyglucose-PET imaging, using a semiquantitative interpretation, for the diagnosis of malignant pleural effusions. The sensitivity and specificity of individual studies are represented by a square, through which runs a horizontal line (95% CI).

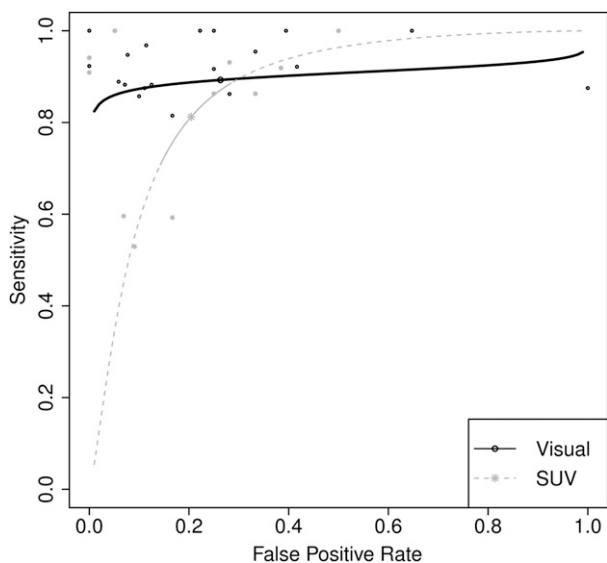


Figure 4 – Hierarchical summary receiver-operating characteristic curve plot of visual and semiquantitative fluorodeoxyglucose-PET imaging studies for malignant pleural effusions. Each symbol represents an individual study in the meta-analysis. ● = visual (or qualitative) interpretation studies; \* = standardized uptake value-based readings.

the malignant nature of pleural fluid accumulation, when qualitative criteria were applied, was sensitivity of 96%; specificity of 76%; LR positive, 4.09; LR negative, 0.06; and AUC, 0.95.

However, the stand-alone PET imaging system is rarely used today because the hybrid PET-CT imaging systems provide metabolic and anatomic information in a single session. Moreover, visual assessments are prone to subjectivity and lack reproducibility. This limitation is partially mitigated by the somewhat more consistent semiquantitative technique of SUV, which currently dominates the clinical use of hybrid imaging. In this more realistic clinical setting, the pooled sensitivity, specificity, LR positive, and LR negative of integrated PET-CT imaging systems, using a semiquantitative method of interpretation, for identifying pleural malignancy dropped to 81%, 74%, 3.22, and 0.26, respectively.

The question of why visual interpretations provide better sensitivity than SUV-based methods may be related to the many biologic and technological factors which hamper the latter.<sup>34</sup> For example, SUV is usually calculated over a small area of the pleura (region of interest), which may not be adequate for the often scattered or diffused pleural involvement that characterizes malignancy. Conversely, the volume of interest (VOI) technique determines the volume concentration of the radiopharmaceutical in the entire tumor, including areas of different SUVmax, and then may better represent the biologic or metabolic behavior of the entire

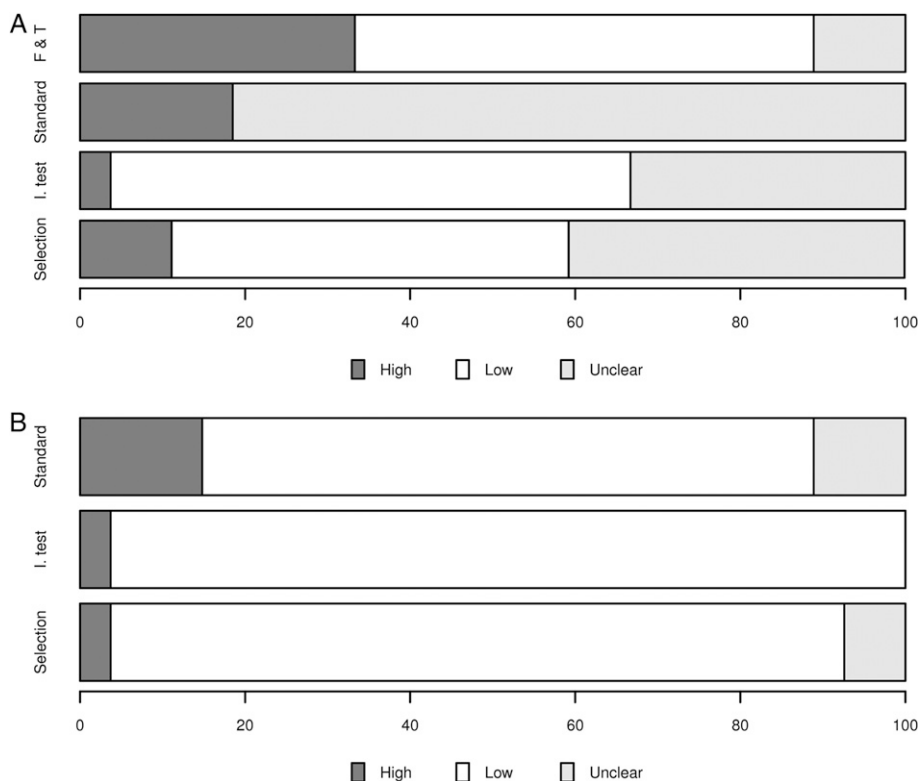


Figure 5 – A, B, Graphic display for the methodologic quality of the included studies using the Quality Assessment of Diagnostic Accuracy Studies-2 criteria. F & T = flow and timing; I. = index.

pleural lesion. Unfortunately, none of the studies included in the meta-analysis used quantitative VOI-based analyses. The rationale behind the outperformance of stand-alone PET imaging over fused PET-CT imaging machines is speculative but, other than the publication bias inherent in the older studies, might be related to the greater experience of radiologists interpreting the former at a time when PET imaging was highly restricted to reference centers.

Overall, the ability of integrated PET-CT imaging systems to separate benign from malignant pleural processes is fraught with potential errors and may be considered moderate, as reflected by an AUC of about 0.84. FDG-PET imaging higher negative than positive LR values makes it slightly more useful for ruling out malignant effusions when the test shows no abnormal tracer uptake. In this sense, FDG-PET imaging would theoretically reduce the number of unnecessary invasive procedures in patients with benign pleural effusions. On the other hand, FDG-PET imaging may guide needle or thoracoscopic biopsy to the highest area of uptake within a thickened pleura, thus minimizing sampling errors.

#### *Strength and Limitations of the Review*

The strength of this review is its robust search strategy and in-depth statistical analysis. Our calculations were restricted to those studies without a high risk of bias, which allowed a better assessment of the available literature. In fact, the inclusion of the initial 27 studies in the analysis would have resulted in superior diagnostic indices, thus pointing to a potentially misleading conclusion about the discriminative properties of FDG-PET imaging. Specifically, the pooled sensitivity and specificity of qualitative (20 studies), semiquantitative (11 studies), and dual-point (three studies) FDG-PET imaging interpretations would have reached 92% and 81%, 91% and 83%, and 98% and 90%, respectively (data not shown).

The reliability and generalization of the data were limited by the small number of eligible studies and the presence of some degree of methodologic weakness in all of them. Potential sources of bias and heterogeneity added to the shortcomings. Certainly, many studies suffered from a spectrum bias. For example, most patients with malignant effusions had pleural nodularity or thickening, with PET imaging being interpreted as positive when there was an increased uptake by the altered pleural membranes and not by the fluid itself. Nevertheless, such morphologic pleural abnormalities characterize fewer than one-half of pleural malignancies.<sup>3</sup>

Similarly, common benign conditions, like parapneumonic effusions, TB, or heart failure, were clearly underrepresented or even absent in the comparator branches.

Publication bias was demonstrated for studies that analyzed functional images by either qualitative or semiquantitative methods. The fact that reports on qualitative readings corresponded with the older studies introduces a confounding factor for interpreting their superior sensitivity as compared with those that are semiquantitative.

Potential sources of heterogeneity included study design, target population, varying cutoff values for test positivity in semiquantitative assessments, and differences in PET image acquisition and/or scanning equipment. Some of them could be controlled in the meta-regression analysis.

#### *Comparison With Previous Research*

Two meta-analyses on the diagnostic performance of pleural PET imaging for benign-malignant discriminations, from the same research group, have been published.<sup>35,36</sup> In the first study,<sup>35</sup> comprising 11 articles<sup>8,9,11,13,16,17,22-24,30,32</sup> and 440 patients, the pooled results for PET and PET-CT imaging, which only considered visual interpretations, were as follows: sensitivity, 95%; specificity, 82%; LR positive, 5.3; LR negative, 0.09; and AUC, 0.95. The second study<sup>36</sup> selected eight articles,<sup>6,12,14,15,19,26,27,31</sup> which included 360 patients with confirmed cancer (90% of lung origin) who also developed pleural abnormalities. Under the same assumptions (ie, pooled results for PET and PET-CT imaging using visual analysis), the test yielded a sensitivity of 86%, specificity of 80%, LR positive of 3.7, LR negative of 0.18, and AUC of 0.90.<sup>36</sup> In both articles, the authors concluded that FDG-PET imaging is an accurate imaging method for the differential diagnosis of pleural effusions.<sup>35,36</sup>

These previous meta-analyses were, we feel, flawed by a number of methodologic limitations, which make their encouraging conclusions questionable. For instance, concerning the largest study<sup>35</sup>: (1) the search strategy was faulty, as we were able to find and include three additional articles for the same time period for our systematic review<sup>10,18,29</sup>; (2) the authors did not meta-analyze five studies<sup>20,21,25,28,33</sup> because of an apparent lack of data to construct  $2 \times 2$  tables when, in fact, these data could have been easily extracted; and (3) they excluded no studies for quality reasons. However, we felt compelled to eliminate five<sup>9,17,22-24</sup> of the 11 articles composing their meta-analysis, due to a high risk of bias. Additionally,

the 2 × 2 table of one<sup>13</sup> of the 11 studies meta-analyzed contained erroneous data, and the authors chose not to analyze studies in which semiquantitative PET imaging readings were used. Even so, one semiquantitative study was misinterpreted and included as a visual.<sup>24</sup>

## Conclusions

The results of the present meta-analysis, based on 14 studies with > 600 patients spanning the last 16 years, suggest that, although of some value, FDG-PET imaging does not seem to change the probability of pleural malignancy sufficiently as to be recommended in the

routine workup of effusions of undetermined cause. The FDG-PET imaging technique warrants broader prospective evaluations using: (1) gating acquisition techniques to reduce motion artifacts, (2) VOI analyses to enhance quantitative PET imaging measurements, and (3) an adequate benign comparator group, which should reflect the local epidemiology of pleural effusions. In addition, its role in the most challenging clinical scenario (ie, malignant effusions without pleural nodularity or thickening and/or a false-negative cytologic examination of the pleural fluid) needs to be determined in future studies.

## Acknowledgments

**Author contributions:** J. M. P. had full access to all of the data in the study, takes responsibility for the integrity of the data and the accuracy of the data analysis, is guarantor for the entire manuscript, contributed to the study concept and design and to drafting the manuscript, and served as principal author. J. M. P. and P. H. contributed to the systematic review. M. M.-A. contributed to the statistical analysis. S. B. and A. S. contributed to revision of the manuscript. J. M. P., P. H., M. M.-A., S. B., and A. S. approved the final version of the manuscript.

**Financial/nonfinancial disclosures:** The authors have reported to *CHEST* that no potential conflicts of interest exist with any companies/organizations whose products or services may be discussed in this article

**Other contributions:** The authors appreciate the assistance of Cristina Gámez, MD, for critical review of the manuscript.

## References

- Porcel JM, Esquerda A, Vives M, Bielsa S. Etiology of pleural effusions: analysis of more than 3,000 consecutive thoracenteses. *Arch Bronconeumol*. 2014;50(5):161-165.
- Porcel JM. Pearls and myths in pleural fluid analysis. *Respirology*. 2011;16(1):44-52.
- Porcel JM, Light RW. Pleural effusions. *Dis Mon*. 2013;59(2):29-57.
- Duysinx B, Corhay JL, Larock MP, et al. Contribution of positron emission tomography in pleural disease. *Rev Mal Respir*. 2010;27(8):e47-e53.
- Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med*. 2009;151(4):264-269, W64.
- Liao R, Yang X, Wang S, et al. Clinical role of F-18 FDG PET/CT in differentiating malignant and benign pleural effusion in patients with lung cancer [in Chinese]. *Zhongguo Fei Ai Za Zhi*. 2012;15(11):652-655.
- Whiting PF, Rutjes AWS, Westwood ME, et al; QUADAS-2 Group. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med*. 2011;155(8):529-536.
- Bury T, Paulus P, Dowlati A, Corhay JL, Rigo P, Radermecker MF. Evaluation of pleural diseases with FDG-PET imaging: preliminary report. *Thorax*. 1997;52(2):187-189.
- Bénard F, Sterman D, Smith RJ, Kaiser LR, Albelda SM, Alavi A. Metabolic imaging of malignant pleural mesothelioma with fluorodeoxyglucose positron emission tomography. *Chest*. 1998;114(3):713-722.
- Buchmann I, Guhlmann CA, Elsner K, et al. F-18-FDG PET for differential diagnosis of pleural processes [in German]. *Nucl Med (Stuttg)*. 1999;38(8):319-322.
- Carretta A, Landoni C, Melloni G, et al. 18-FDG positron emission tomography in the evaluation of malignant pleural diseases - a pilot study. *Eur J Cardiothorac Surg*. 2000;17(4):377-383.
- Erasmus JJ, McAdams HP, Rossi SE, Goodman PC, Coleman RE, Patz EF. FDG PET of pleural effusions in patients with non-small cell lung cancer. *AJR Am J Roentgenol*. 2000;175(1):245-249.
- Gerbaudo VH, Sugarbaker DJ, Britz-Cunningham S, Di Carli MF, Mauceri C, Treves ST. Assessment of malignant pleural mesothelioma with (18)F-FDG dual-head gamma-camera coincidence imaging: comparison with histopathology. *J Nucl Med*. 2002;43(9):1144-1149.
- Gupta NC, Rogers JS, Graeber GM, et al. Clinical role of F-18 fluorodeoxyglucose positron emission tomography imaging in patients with lung cancer and suspected malignant pleural effusion. *Chest*. 2002;122(6):1918-1924.
- Schaffler GJ, Wolf G, Schoellnast H, et al. Non-small cell lung cancer: evaluation of pleural abnormalities on CT scans with 18F FDG PET. *Radiology*. 2004;231(3):858-865.
- Duysinx B, Nguyen D, Louis R, et al. Evaluation of pleural disease with 18-fluorodeoxyglucose positron emission tomography imaging. *Chest*. 2004;125(2):489-493.
- Kramer H, Pieterman RM, Slebos DJ, et al. PET for the evaluation of pleural thickening observed on CT. *J Nucl Med*. 2004;45(6):995-998.
- Melloni B, Monteil J, Vincent F, et al. Assessment of 18F-fluorodeoxyglucose dual-head gamma camera in asbestos lung diseases. *Eur Respir J*. 2004;24(5):814-821.
- Toaff JS, Metser U, Gottfried M, et al. Differentiation between malignant and benign pleural effusion in patients with extra-pleural primary malignancies: assessment with positron emission tomography-computed tomography. *Invest Radiol*. 2005;40(4):204-209.
- Duysinx BC, Larock MP, Nguyen D, et al. 18F-FDG PET imaging in assessing exudative pleural effusions. *Nucl Med Commun*. 2006;27(12):971-976.
- Mavi A, Basu S, Cermik TF, et al. Potential of dual time point FDG-PET imaging in differentiating malignant from benign pleural disease. *Mol Imaging Biol*. 2009;11(5):369-378.
- Yildirim H, Metintas M, Entok E, et al. Clinical value of fluorodeoxyglucose-positron emission tomography/computed tomography in differentiation of malignant mesothelioma from asbestos-related benign pleural disease: an observational pilot study. *J Thorac Oncol*. 2009;4(12):1480-1484.
- Yamamoto Y, Kameyama R, Togami T, et al. Dual time point FDG PET for evaluation of malignant pleural mesothelioma. *Nucl Med Commun*. 2009;30(1):25-29.
- Orki A, Akin O, Tasci AE, et al. The role of positron emission tomography/computed tomography in the diagnosis of pleural diseases. *Thorac Cardiovasc Surg*. 2009;57(4):217-221.
- Kurata S, Ishibashi M, Azuma K, et al. Preliminary study of positron emission tomography/computed tomography and plasma osteopontin levels in patients with asbestos-related pleural disease. *Jpn J Radiol*. 2010;28(6):446-452.
- Kim BS, Kim IJ, Kim SJ, Pak K, Kim K. Predictive value of F-18 FDG PET/CT for malignant pleural effusion in non-small cell lung cancer patients. *Onkologie*. 2011;34(6):298-303.
- Alkhalaf K, Biersack HJ, Henke A, Ezziddin S. Impact of dual-time-point



- F-18 FDG PET/CT in the assessment of pleural effusion in patients with non-small-cell lung cancer. *Clin Nucl Med*. 2011;36(6):423-428.
28. Abe Y, Tamura K, Sakata I, et al. Clinical implications of 18F-fluorodeoxyglucose positron emission tomography/computed tomography at delayed phase for diagnosis and prognosis of malignant pleural mesothelioma. *Oncol Rep*. 2012;27(2):333-338.
  29. Jung MY, Chong A, Seon HJ, et al. Indeterminate pleural metastasis on contrast-enhanced chest CT in non-small cell lung cancer: improved differential diagnosis with (18)F-FDG PET/CT. *Ann Nucl Med*. 2012;26(4):327-336.
  30. Elboga U, Yılmaz M, Uyar M, Zeki Çelen Y, Bakır K, Dikensoy O. The role of FDG PET-CT in differential diagnosis of pleural pathologies. *Rev Esp Med Nucl Imagen Mol*. 2012;31(4):187-191.
  31. Letovanec I, Allenbach G, Mihaescu A, et al. <sup>18</sup>F-fluorodeoxyglucose PET/CT findings in pleural effusions of patients with known cancer. A cytopathological correlation. *Nucl Med (Stuttg)*. 2012; 51(5):186-193.
  32. Coolen J, De Keyzer F, Naftoux P, et al. Malignant pleural disease: diagnosis by using diffusion-weighted and dynamic contrast-enhanced MR imaging—initial experience. *Radiology*. 2012;263(3): 884-892.
  33. Terada T, Tabata C, Tabata R, et al. Clinical utility of 18-fluorodeoxyglucose positron emission tomography/computed tomography in malignant pleural mesothelioma. *Exp Ther Med*. 2012;4(2):197-200.
  34. Adams MC, Turkington TG, Wilson JM, Wong TZ. A systematic review of the factors affecting accuracy of SUV measurements. *AJR Am J Roentgenol*. 2010;195(2):310-320.
  35. Treglia G, Sadeghi R, Annunziata S, et al. Diagnostic accuracy of 18F-FDG-PET and PET/CT in the differential diagnosis between malignant and benign pleural lesions: a systematic review and meta-analysis. *Acad Radiol*. 2014;21(1): 11-20.
  36. Treglia G, Sadeghi R, Annunziata S, et al. Diagnostic performance of fluorine-18-fluorodeoxyglucose positron emission tomography in the assessment of pleural abnormalities in cancer patients: a systematic review and a meta-analysis. *Lung Cancer*. 2014;83(1):1-7.