



“APLICACIONES DE LA ECOGRAFÍA PULMONAR Y DIAFRAGMÁTICA EN EL PACIENTE CRÍTICO”

Tesis doctoral por compendio de publicaciones.

2019

Autora: Ana María Llamas Álvarez.

Director: Dr. D. Jaime Latour Pérez.

Codirector: Dr. D. Javier Fernández Sánchez.

La presente tesis doctoral por compendio de publicaciones está compuesta por los siguientes estudios:

- Llamas-Álvarez AM, Tenza-Lozano EM, Latour-Pérez J. Accuracy of lung ultrasound in the diagnosis of pneumonia in adults: systematic review and meta-analysis. Chest. 2017 Feb;151(2):374-382. doi: 10.1016/j.chest.2016.10.039.
- Latour-Perez J, Llamas-Álvarez A-M, Tenza-Lozano EM. Lung ultrasound for the diagnosis of pneumonia. Eur J Emerg Med. 2019 Apr;26(2):147-148. doi: 10.1097/MEJ.0000000000000567.
- Llamas-Álvarez AM, Tenza-Lozano EM, Latour-Pérez J. Diaphragm and lung ultrasound to predict weaning outcome: systematic review and meta-analysis. Chest. 2017 Aug 29. pii: S0012-3692(17)31482-4. doi: 10.1016/j.chest.2017.08.028
- Tenza-Lozano E, Llamas Álvarez A, Jaimez Navarro E, Fernández Sanchez J. Lung and diaphragmatic ultrasound as predictors of success in weaning. Crit Ultrasound J. 2018; 10: 12. doi: 10.1186/s13089-018-0094-3



PROGRAMA DE DOCTORADO EN SALUD PÚBLICA, CIENCIAS MÉDICAS Y QUIRÚRGICAS

Don Jaime Latour Pérez, director y Don Francisco Javier Fernández Sánchez, co-director de la tesis doctoral titulada “APLICACIONES DE LA ECOGRAFÍA PULMONAR Y DIAFRAGMÁTICA EN EL PACIENTE CRÍTICO”

CERTIFICAN:

Que Doña Ana María Llamas Álvarez ha realizado bajo nuestra supervisión el trabajo titulado “APLICACIONES DE LA ECOGRAFÍA PULMONAR Y DIAFRAGMÁTICA EN EL PACIENTE CRÍTICO” conforme a los términos y condiciones definidos en su Plan de Investigación y de acuerdo al Código de Buenas Prácticas de la Universidad Miguel Hernández de Elche, cumpliendo los objetivos previstos de forma satisfactoria para su defensa pública como tesis doctoral.

Lo que firmamos en Sant Joan d’Alacant, a treinta de septiembre de dos mil diecinueve

Fdo. Don Jaime Latour Pérez

Director de tesis

Fdo. Don Francisco Javier Fernández Sánchez

Co-director de tesis





PROGRAMA DE DOCTORADO EN SALUD PÚBLICA, CIENCIAS MÉDICAS Y QUIRÚRGICAS

D. Félix Gutiérrez Rodero, Coordinador del Programa de Doctorado en Salud Pública, Ciencias Médicas y Quirúrgicas de la Universidad Miguel Hernández de Elche,

AUTORIZA:

La presentación y defensa como tesis doctoral del trabajo “APLICACIONES DE LA ECOGRAFÍA PULMONAR Y DIAFRAGMÁTICA EN EL PACIENTE CRÍTICO” realizado por Dña. Ana María Llamas Álvarez bajo la dirección del Dr. D. Jaime Latour Pérez y la co-dirección del Dr. D. Francisco Javier Fernández Sánchez.

De acuerdo a la información recibida sobre las evaluaciones previas realizadas en cumplimiento de la normativa general vigente y la propia de la Universidad Miguel Hernández y según lo certificado por las personas que han realizado la tutoría y dirección, la tesis cumple los requisitos para proceder a su defensa pública.

En Sant Joan d’Alacant, a treinta de septiembre de 2019

Firmado: Prof. Félix Gutiérrez Rodero

Coordinador del Programa de Doctorado en Salud Pública, Ciencias Médicas y Quirúrgicas



Agradecimientos.

En primer lugar, me gustaría dar las gracias al Dr Jaime Latour, gran médico y mejor persona, al que estimo a nivel personal y admiro en lo profesional. Has sido un gran apoyo durante estos años y he aprendido mucho de ti. Muchísimas gracias maestro.

Gracias al Dr Javier Fernández, co-director de esta tesis, por tu disponibilidad y amabilidad. Ha sido un placer trabajar contigo.

Me gustaría reconocer el esfuerzo y dedicación de la Dra Eva Tenza, compañera y amiga, por formar parte de este proyecto. Pese a las dificultades ambas lo hemos conseguido!

También a la Dra Eva de Miguel, por su apoyo y sus sabios consejos.

A Marisun, siempre amable y eficiente, por su ayuda y asesoramiento con los trámites y burocracia del programa de doctorado.

A mi familia: mis padres Juani y Antonio, mi hermana Fini y mi sobrino Toni, sois los mejores.

Y sobre todo a Sergio, tú mejor que nadie comprendes el esfuerzo y sacrificio que ha supuesto esta tesis doctoral. Gracias por tu paciencia, tu optimismo y por animarme siempre a avanzar y mejorar. Tengo mucha suerte.

Gracias a todos!



*“El buen médico trata la enfermedad;
el gran médico trata al paciente que tiene la enfermedad”*

William Osler



ÍNDICE

• LISTADO DE ABREVIATURAS.....	5
• LISTADO DE FIGURAS.....	9
• LISTADO DE TABLAS.....	12
• RESUMEN.....	13
1. INTRODUCCIÓN.....	19
1.1. Conceptos básicos de ecografía.....	21
1.2. La ecografía en la Unidad de Cuidados Intensivos (UCI).....	23
1.2.1. Características generales.....	23
1.2.2. Indicaciones de ecografía torácica no cardíaca en el paciente crítico.....	24
1.2.3. Indicaciones no estandarizadas y líneas de investigación actuales.....	28
1.3. Investigación de síntesis	29
1.3.1. El problema de la investigación inútil. Necesidad de revisiones sistemáticas y meta-análisis.....	29
1.3.2. Revisiones sistemáticas y meta-análisis de exactitud diagnóstica.....	30
1.3.3. Meta-análisis sobre reglas de predicción clínica.....	31
1.4. Línea de investigación y justificación de la tesis.....	32
2. OBJETIVOS.....	33
3. HIPÓTESIS E IMPLICACIONES PRÁCTICAS.....	37
4. RESUMEN DE LAS PUBLICACIONES PRESENTADAS.....	41
4.1. ECOGRAFÍA PULMONAR PARA EL DIAGNÓSTICO DE NEUMONÍA EN ADULTOS: REVISIÓN SISTEMÁTICA Y META-ANÁLISIS.....	43
4.1.1. Resumen.....	43
4.1.2. Introducción y objetivos.....	44
4.1.3. Material y métodos.....	45
4.1.3.1. Selección de los estudios.....	45

4.1.3.2. Fuentes de datos y estrategias de búsqueda.....	45
4.1.3.3. Extracción de los datos.....	46
4.1.3.4. Valoración de la calidad de los estudios.....	46
4.1.3.5. Análisis de los datos.....	46
4.1.4. Resultados.....	47
4.1.4.1. Selección de estudios.....	47
4.1.4.2. Características de los estudios incluidos.....	48
4.1.4.3. Riesgo de sesgo y aplicabilidad.....	51
4.1.4.4. Meta-análisis.....	53
4.1.4.5. Exploración de las causas de heterogeneidad.....	54
4.1.4.6. Apéndices.....	56
4.2. ECOGRAFÍA DIAFRAGMÁTICA Y PULMONAR PARA PREDECIR EL RESULTADO DEL DESTETE: REVISIÓN SISTEMÁTICA Y META-ANÁLISIS.....	60
4.2.1. Resumen.....	60
4.2.2. Introducción y objetivos.....	61
4.2.3 Material y métodos.....	62
4.2.3.1. Estrategia de búsqueda y selección de los estudios.....	62
4.2.3.2. Extracción de los datos.....	63
4.2.3.3. Valoración del riesgo de sesgo.....	63
4.2.3.4. Análisis de los datos.....	63
4.2.4. Resultados.....	64
4.2.4.1. Selección de estudios.....	64
4.2.4.2. Características de los estudios incluidos.....	64
4.2.4.3. Riesgo de sesgo y aplicabilidad.....	68
4.2.4.4. Meta-análisis.....	70
4.2.4.5. Exploración de las posibles causas de heterogeneidad.....	74
4.3. ECOGRAFÍA PULMONAR Y DIAFRAGMÁTICA COMO PREDICTORES DE ÉXITO EN LA RETIRADA DE LA VENTILACIÓN MECÁNICA.....	78
4.3.1. Resumen.....	78

4.3.2. Introducción y objetivos.....	79
4.3.3. Material y métodos.....	79
4.3.3.1. Diseño.....	79
4.3.3.2. Población.....	80
4.3.3.3. Mediciones.....	80
4.3.3.4. Procedimiento.....	81
4.3.3.5. Análisis estadístico.....	83
4.3.3.6. Aspectos éticos.....	83
4.3.4 Resultados.....	84
4.3.4.1. Exactitud para predecir el éxito en el weaning.....	84
4.3.4.2. Concordancia interobservador.....	87
5. DISCUSIÓN.....	89
5.1. Resumen de Resultados.....	91
5.1.1. Ecografía pulmonar para el diagnóstico de neumonía en adultos: revisión sistemática y meta-análisis.....	91
5.1.2. Ecografía diafragmática y pulmonar para predecir el resultado del destete: revisión sistemática y meta-análisis.....	91
5.1.3. Ecografía pulmonar y diafragmática como predictores de éxito en la retirada de la ventilación mecánica.....	92
5.2. Validez.....	92
5.3. Aplicabilidad.....	94
6. CONCLUSIONES.....	97
7. REFERENCIAS BIBLIOGRÁFICAS.....	101
8. ANEXOS.....	115
8.1. ANEXO 1. Accuracy of lung ultrasound in the diagnosis of pneumonia in adults: systematic review and meta-analysis.....	117
8.2. ANEXO 2. Lung ultrasound for the diagnosis of pneumonia.....	145
8.3. ANEXO 3. Diaphragm and lung ultrasound to predict weaning outcome: systematic review and meta-analysis.....	149

8.4. ANEXO 4. Lung and diaphragmatic ultrasound as predictors of success in weaning.....187



LISTADO DE ABREVIATURAS.

ACEP: American College of Emergency Physicians.

Ant: anterior.

APACHE-II: Acute Physiology and Chronic Health Evaluation.

AUC: *area under the curve* (área bajo la curva).

BA: broncograma aéreo.

BLUE: Bedside Lung Ultrasound in Emergency.

CB: signo del código de barras.

CCI: coeficiente de correlación intraclase.

CI: consentimiento informado.

Cm H₂O: centímetros de agua.

CNAF: cánula nasal de alto flujo.

CROP: *dynamic Compliance, Respiratory rate, Oxygenation, maximal inspiratory Pressure.*

CSP: consolidación subpleural.

DDIVM: disfunción diafragmática inducida por la ventilación mecánica.

DP: derrame pleural.

DXF: diagnóstico final.

E: especificidad.

EC: ecocardiografía.

ED: excursión diafragmática.

EI: espacio intercostal.

END: experiencia en ecografía pulmonar no descrita.

EPOC: enfermedad pulmonar obstructiva crónica.

Eunetha: *European network for Health Technology Assessment.*

FALLS: *Fluid Administration Limited by Lung Sonography.*

FAST: *Focused Assessment with Sonography for Trauma.*

FC: frecuencia cardíaca.

FEDI: fracción de engrosamiento del diafragma.

FEEL: *Focused Echocardiography Evaluation in Life support.*

FiO₂: fracción inspirada de oxígeno.

FN: *false negatives* (falsos negativos).

FP: *false positives* (falsos positivos).

FR: frecuencia respiratoria.

FVI: fallo ventricular izquierdo.

GNAF: gafas nasales de alto flujo.

HSROC: *hierarchical summary receiver operating characteristic*.

IC: intervalo de confianza.

IRA: insuficiencia respiratoria aguda.

L: litros.

LAA: línea axilar anterior.

Lab: datos de laboratorio.

LAP: línea axilar posterior.

Lat: lateral.

LMA: línea medioaxilar.

LMC: línea medioclavicular.

LP: línea pleural.

LUS score: *lung ultrasound score*.

MHz: megahercios.

Mic: datos microbiológicos.

MIN: Medicina Intena.

Min: minuto.

MmHg: milímetros de mercurio.

Mmol: milimoles.

N: número de pacientes incluidos.

NAC: neumonía adquirida en la comunidad.

NAVM: neumonía asociada a ventilación mecánica.

NE: no especificado.

ORD: odds ratio diagnóstica.

P 0.1: presión de oclusión de la vía aérea.

P25, P50,P75: percentil 25, 50 y 75 respectivamente.

PAI: patrón alveolointersticial.

PaO₂: presión parcial de oxígeno.

PAOP: *pulmonary artery occlusion pressure* (presión de oclusión de la arteria pulmonar).

PCO₂: presión parcial de dióxido de carbono.

Pdi: presión transdiafragmática.

PEEP: *positive end-expiratory pressure* (presión positiva al final de la espiración).

PIMax: presión inspiratoria máxima.

Post: posterior.

PRE: prueba de respiración espontánea.

PRISMA: *Preferred Reporting Items for Systematic Reviews and Meta-Analyses*.

PS: presión de soporte.

PSI: posición semiincorporada.

QUADAS-2: *Quality Assessment of Diagnostic Accuracy Studies*.

ROC: *Receiver Operating Characteristic*.

RSBI: *Rapid Shallow Breathing Index* (índice de respiración superficial rápida).

RxT: radiografía de tórax.

S: sensibilidad.

SAPS-3: *Simplified Acute Physiologic Score 3*.

SatO₂: saturación de oxígeno en sangre periférica.

SDRA: síndrome de distrés respiratorio del adulto.

SP: signo de la playa.

SROC: *Summary Receiver Operating Characteristic*.

TAS: tensión arterial sistólica.

TC: tomografía computerizada.

TDX: tesis doctorales en xarxa.

TEP: tromboembolismo pulmonar.

TN (VN): *true negatives* (verdaderos negativos).

TP (VP): *true positives* (verdaderos positivos).

Tpfr: tasa de falsos positivos.

Tsens: sensibilidad transformada.

UCI: Unidad de Cuidados Intensivos.

UCIP: unidad de cuidados intensivos polivalente.

Urg: Urgencias.

VE: volumen espiratorio.

VMI: ventilación mecánica invasiva.

VMNI: ventilación mecánica no invasiva.

V tidal: volumen tidal o volumen corriente.



LISTADO DE FIGURAS.

- **Figura 1.** Imagen ecográfica en modo B en ventana intercostal derecha, plano oblicuo. (Dra A. Llamas. Servicio de Radiodiagnóstico. Hospital General Universitario de Alicante. 2019).
- **Figura 2.** Imagen ecográfica en modo M en ventana intercostal derecha, plano oblicuo. (Adaptado de: Kim K, Jang D-M, Park J-Y, Yoo H, Kim HS, Choi W-J. Changes of diaphragmatic excursion and lung compliance during major laparoscopic pelvic surgery: A prospective observational study. PLoS ONE. 2018; 13(11): e0207841. <https://doi.org/10.1371/journal.pone.0207841>)
- **Figura 3.** Imagen ecográfica de derrame pleural derecho en ventana intercostal derecha. Disponible en: https://www.ecografiaclinica.es/003_manual_011_008/
- **Figura 4.** Ecografías pulmonares en modo M. A) En condiciones normales se observa la línea pleural (LP) y signo de la playa (SP). B) En presencia de neumotórax se aprecian exclusivamente líneas horizontales (signo de la estratosfera del código de barras). No se ve la línea pleural y ni el signo de la playa. Imágenes extraídas de: Carrillo Esper R et al. Patrones ultrasonográficos pulmonares en el enfermo grave. Rev Asoc Mex Med Crit y Ter Int 2011;25(1):24-32.
- **Figura 5.** Ecografía pulmonar en modo B donde se aprecian las líneas B. Imágenes extraídas de: Ardila-Castellanos, R. et al. Ecografía pulmonar para la valoración del agua pulmonar extravascular en el seguimiento de pacientes con edema pulmonar en ventilación mecánica: estudio piloto. Acta Colomb Cuid Intensivos. 2015. <http://dx.doi.org/10.1016/>
- **Figura 6.** Signo del punto pulmonar: en el modo M se aprecia alternancia del signo de la playa (SP) durante la inspiración y el signo del código de barras (CB) durante la espiración, ver recuadro. Las flechas señalan la línea pleural. Imagen extraída de: Carrillo Esper R et al. Patrones ultrasonográficos pulmonares en el enfermo grave. Rev Asoc Mex Med Crit y Ter Int 2011;25(1):24-32.
- **Figura 7.** Ecografía pulmonar en modo B en el que se aprecia un patrón de aireación normal, con predominio de líneas A (*). Imagen extraída de: Miller, A. Practical approach to lung ultrasound. 2016. BJA Education. Volume 16; 2 (39–45). <https://doi.org/10.1093/bjaceaccp/mkv012>.
- **Figura 8.** Ecografía pulmonar en modo B en el que se aprecia un patrón alveolointerstitial: abundantes líneas B (B) que se originan de la línea pleural (LP). Imagen extraída de: Carrillo Esper R et al. Patrones ultrasonográficos pulmonares en el enfermo grave. Rev Asoc Mex Med Crit y Ter Int 2011;25(1):24-32.
- **Figura 9.** Imagen de consolidación pulmonar. C: consolidación; DP: derrame pleural. Imagen extraída de: Colmenero M, García-Delgado M, Navarrete I, López-Milena G.

Utilidad de la ecografía pulmonar en la Unidad de Medicina Intensiva. *Med Intensiva*. 2010;34(9):620–628.

- **Figura 10.** Resultados de búsqueda en Medline con los descriptores (lung OR diaphragm) AND (ultrasound OR ultrasonography) en el título entre los años 1972 y 2019. Disponible en: <https://www.ncbi.nlm.nih.gov/pubmed> (datos consultados con fecha 1 de agosto de 2019).
- **Figura 11.** Causas de investigación inútil. Traducida de Chalmers I, Glasziou P. Avoidable waste in the production and reporting of research evidence. *Lancet* 2009; 374: 86–89.
- **Figura 12.** Diagrama de flujo de los estudios cribados según PRISMA y motivos de exclusión.
- **Figura 13.** QUADAS-2. En los diferentes dominios se ilustran el número de estudios con riesgo de sesgo alto, bajo o dudoso, así como su porcentaje respecto al total. Igual para la aplicabilidad.
- **Figura 14.** Forest plot de sensibilidad y especificidad de los distintos estudios. Abreviaturas: TP: verdaderos positivos; FP: falsos positivos; FN: falsos negativos; TN: verdaderos negativos.
- **Figura 15.** Curva SROC de la exactitud global de la ecografía pulmonar para el diagnóstico de neumonía y elipse de confianza alrededor del valor promedio. Los distintos estudios se representan con triángulos.
- **Figura 16.** Comparación de curvas ROC y elipse de confianza entre los estudios considerados de alta calidad y el resto de estudios (calidad considerada dudosa o baja). Tsens: sensibilidad transformada; tpfr: tasa de falsos positivos; IC 95%: intervalos de confianza al 95%.
- **Figura 17.** Diagrama de flujo de los estudios cribados según las recomendaciones PRISMA y motivos de exclusión.
- **Figura 18.** Aplicación de QUADAS-2 a los estudios que miden FEDI.
- **Figura 19.** Aplicación de QUADAS-2 a los estudios que miden ED.
- **Figura 20.** Aplicación de QUADAS-2 a los estudios que miden LUS score.
- **Figura 21.** Forest plot de sensibilidad y especificidad de los estudios que miden FEDI. Abreviaturas: TP: verdaderos positivos; FP: falsos positivos; FN: falsos negativos; TN: verdaderos negativos.
- **Figura 22.** Curva SROC de la FEDI para predecir el éxito en el destete y elipse de confianza alrededor del valor promedio. Los distintos estudios se representan con triángulos.

- **Figura 23.** Forest plot de sensibilidad y especificidad de los estudios que miden ED. Abreviaturas: TP: verdaderos positivos; FP: falsos positivos; FN: falsos negativos; TN: verdaderos negativos.
- **Figura 24.** Curva SROC de la ED para predecir el éxito en el destete y elipse de confianza alrededor del valor promedio. Los distintos estudios se representan con triángulos.
- **Figura 25.** Forest plot de sensibilidad y especificidad de los estudios que miden LUS score. Abreviaturas: TP: verdaderos positivos; FP: falsos positivos; FN: falsos negativos; TN: verdaderos negativos.
- **Figura 26.** Análisis de subgrupos de calidad en estudios que miden FEDI (A) y ED (B). Comparación de curvas SROC entre los estudios considerados de alta calidad (círculos) y menor calidad (triángulos). Abreviaturas: tsens: sensibilidad transformada; tfpr: tasa de falsos positivos; IC 95%: intervalo de confianza al 95%.
- **Figura 27.** Análisis de subgrupos de aplicabilidad en estudios que miden FEDI. Comparación de curvas SROC y elipses de confianza entre los estudios considerados aplicables (triángulos) y aquellos que probablemente impliquen problemas de aplicabilidad (círculos). Abreviaturas: tsens: sensibilidad transformada; tfpr: tasa de falsos positivos; IC 95%: intervalo de confianza al 95%.
- **Figura 28.** Análisis de subgrupos de aplicabilidad en estudios que miden ED. Comparación de curvas SROC y elipses de confianza entre los estudios considerados aplicables (triángulos) y aquellos que probablemente impliquen problemas de aplicabilidad (círculos). Abreviaturas: tsens: sensibilidad transformada; tfpr: tasa de falsos positivos; IC 95%: intervalo de confianza al 95%.
- **Figura 29.** Esquema del proceso de evaluación de los pacientes incluidos en el estudio. Abreviaturas: CI: consentimiento informado; PS: presión de soporte; Peep: Presión positiva al final de la espiración (positive end expiratory pressure); FEDI: fracción de engrosamiento diafragmático; LUSm: Lung ultrasound score modificado; PRE: prueba de respiración espontánea; VMI: ventilación mecánica invsiva; VMNI: ventilación mecánica no invasiva; GNAF: gafas nasales de alto flujo.
- **Figura 30.** Resultado del destete y flujo de pacientes. Abreviaturas: PRE: prueba de respiración espontánea; VMNI: ventilación mecánica no invasiva; CNAF: cánula nasal de alto flujo.
- **Figura 31.** Curvas ROC de exactitud de FEDI (rojo) y LUSm (azul) para predecir el éxito del destete. Se resaltan los puntos de corte óptimos para cada índice.
- **Figuras del Apéndice 2.** Análisis de subgrupos, según el riesgo de sesgo alto VS dudoso o bajo en cuanto a la elección del gold standard (2.A), el flujo de pacientes dentro del estudio (2.B) y la gravedad de los pacientes incluidos en cada trabajo (2.C).

LISTADO DE TABLAS.

- **Tabla 1.** Valores de densidad y velocidad de propagación del sonido en los distintos tejidos. Extraído de: García de Casasola G, Torres Macho J. Manual de Ecografía Clínica. Principios físicos de los ultrasonidos. Sociedad Española de Medicina Interna. Madrid, 2014.
- **Tabla 2.** Características de los estudios incluidos. Abreviaturas: DXF: diagnóstico final, entre paréntesis se indica en qué datos se basa éste; RxT: Radiografía de tórax; TC: Tomografía Computerizada de tórax; Cl: datos clínicos, exploración física y/o evolución; Lab: datos de laboratorio; Mic: datos microbiológicos; EC: ecocardiografía.
- **Tabla 3.** Características de los pacientes incluidos en los distintos estudios. Abreviaturas: N: número de pacientes incluidos; NAC: neumonía adquirida en la comunidad; NAVM: neumonía asociada a ventilación mecánica; IRA: insuficiencia respiratoria aguda; VMI: ventilación mecánica invasiva; FVI: fallo ventricular izquierdo. La edad se expresa según los datos de cada estudio como media \pm SD, mediana (P25-P75) o porcentaje dentro de un subgrupo.
- **Tabla 4.** Características de los estudios incluidos. Las medidas y puntos de corte se refieren a la predicción del éxito del destete. Abreviaturas: FEDI: fracción de engrosamiento diafragmático; ED: excursión diafragmática; VM: ventilación mecánica; VMNI: ventilación mecánica no invasiva; LUS score: lung ultrasound score; PRE: prueba de respiración espontánea.
- **Tabla 5.** Características de los pacientes incluidos en los distintos estudios. Abreviaturas: N: número de pacientes incluidos; UCI: unidad de cuidados intensivos (no se especifica el tipo); UCIP: unidad de cuidados intensivos polivalente; VM: ventilación mecánica; PRE: prueba de respiración espontánea; EPOC: enfermedad pulmonar obstructiva crónica. * La edad se expresa como media \pm desviación estándar o mediana (rango intercuartílico), según los datos aportados en cada estudio.
- **Tabla 6.** Criterios de fracaso de la prueba de respiración espontánea. Abreviaturas: PaO₂: presión arterial de oxígeno; mmHg: milímetros de mercurio; PaCo₂: presión arterial de dióxido de carbono; FR: frecuencia respiratoria; FC: frecuencia cardíaca; TAS: tensión arterial sistólica.
- **Tabla 7.** Características basales de los pacientes incluidos en el estudio (n=69).
- **Tabla 8.** Comparativa de los grupos de éxito y fracaso en el destete. Las variables en las que las diferencias alcanzan la significación estadística ($p < 0.05$) se señalan en amarillo. *Las categorías de comorbilidad no son excluyentes.
- **Tabla apéndice 1.** Características de ejecución de la ecografía pulmonar en los distintos estudios.





RESUMEN.

Antecedentes:

La ecografía es una técnica diagnóstica que aporta información morfológica y funcional, que además cuenta con diversas fortalezas como son su portabilidad, rapidez, no invasividad, ausencia de exposición a radiaciones ionizantes y que ofrece información a tiempo real. Estas ventajas la convierten en un método diagnóstico especialmente atractivo en el paciente crítico debido a su complejidad y al elevado riesgo de eventos adversos asociados al traslado, movilización e implementación de técnicas invasivas en estos pacientes. En los últimos años se ha desarrollado la aplicación de los ultrasonidos a la patología pulmonar y diafragmática y diversos estudios sugieren una excelente exactitud diagnóstica y pronóstica de los mismos en dos de los problemas más frecuentes en las unidades de críticos, como son el diagnóstico y manejo de la neumonía y la decisión del momento óptimo para la retirada de la ventilación mecánica.

Objetivos y métodos:

El propósito de esta investigación es avanzar, mediante una investigación de síntesis, el estado del conocimiento sobre la exactitud y aplicabilidad de la ecografía pulmonar y diafragmática en el ámbito de los cuidados intensivos y más concretamente en el diagnóstico de neumonía y la predicción del éxito o fracaso del destete de la ventilación mecánica, así como valorar su reproducibilidad. Para ello los objetivos planteados son:

- Estimar, mediante una revisión sistemática y meta-análisis, la exactitud de la ecografía pulmonar para el diagnóstico de neumonía en pacientes críticos generales
- Estimar mediante una revisión sistemática y meta-análisis, la exactitud de la ecografía pulmonar y diafragmática para predecir el resultado del destete de la ventilación mecánica en pacientes críticos generales sometidos a ventilación mecánica. Los predictores sonográficos propuestos a nivel diafragmático son dos: por un lado la excursión diafragmática (ED), que mide la distancia que es capaz de desplazarse el diafragma durante el ciclo respiratorio, y por otro la fracción de engrosamiento diafragmático (FEDI) que refleja la variación en el grosor del diafragma durante el esfuerzo respiratorio y se calcula mediante la fórmula $FEDI = (\text{grosor al final de la inspiración} - \text{grosor al final de la espiración}) / \text{grosor al final de la espiración}$. A nivel pulmonar podemos detectar la disminución de la aireación del parénquima por ecografía, ya sea por causa cardíaca, respiratoria o diafragmática y una forma de cuantificarlo es mediante el denominado Lung ultrasound score (LUS score), una escala validada cuyo rango de valores oscila entre 0 y 36 puntos obtenidos del sumatorio de puntuaciones asignadas a los distintos patrones de aireación observados en cada área de exploración pulmonar.
- Estimar la exactitud de la ecografía pulmonar y diafragmática como predictores de éxito en la retirada de la ventilación mecánica en nuestro medio a través de un estudio de cohortes prospectivo, así como evaluar la concordancia interobservador mediante un estudio transversal de concordancia.

Resultados:

- Evaluación de la exactitud diagnóstica de la ecografía pulmonar para el diagnóstico de neumonía en pacientes adultos a través de una revisión sistemática y meta-análisis: se incluyeron 16 estudios, con un total de 2359 participantes. Se objetivó heterogeneidad significativa en cuanto a sensibilidad y especificidad de acuerdo con el test Q, sin clara evidencia de efecto umbral. El área bajo la curva de SROC fue 0.93, con una Odds Ratio Diagnóstica (ORD) de 50 en el punto de corte óptimo (IC 95%: 21, 120). Se detectó una tendencia hacia una mayor área bajo la curva SROC en estudios de alta calidad, sin embargo estas diferencias no fueron significativas al aplicar la meta-regresión bivariente.
- Evaluación de la exactitud de la ecografía pulmonar y diafragmática (en concreto de los índices ED, FEDI y LUS score), para predecir el resultado del destete de la ventilación mecánica en pacientes adultos, mediante una revisión sistemática de la literatura y meta-análisis: Se incluyeron 19 estudios con un total de 1071 pacientes. Para la FEDI, el área bajo la curva ROC fue de 0,87, y la Odds Ratio Diagnóstica (ORD) fue de 21 (IC 95%: 11, 40). En cuanto a la ED, la sensibilidad ponderada fue del 75% (IC del 95%: 65, 85), la especificidad ponderada, 75% (IC 95%: 60, 85) y la ORD=10 (IC 95%: 4, 24). Para el LUS score, se obtuvo un área bajo la curva ROC de 0,77, y ORD= 38 (IC95%: 7, 198). Al aplicar la meta-regresión bivariente se detectó una especificidad significativamente mayor para FEDI y mayor sensibilidad para ED en el subgrupo de estudios con posibles problemas de aplicabilidad.
- Evaluación de la exactitud y reproducibilidad de la ecografía pulmonar, en concreto de LUS score modificado y FEDI, para predecir el éxito del destete de la ventilación mecánica en pacientes críticos adultos, mediante un estudio de cohortes prospectivo y un estudio de concordancia interobservador, ambos realizados en la Unidad de Cuidados Intensivos del Hospital General de Elche: Para LUSm se obtuvo un área bajo la curva ROC de 0.80 para predecir el éxito del destete, sensibilidad y especificidad en el punto de corte óptimo 0,76 y 0,73 respectivamente. Para FEDI, el AUC calculada fue de 0.71, sensibilidad de 0.93 y especificidad de 0.48. Se encontró una concordancia interobservador adecuada tanto para LUSm (kappa ponderado 0.95) como para FEDI (ICC 0.78, diferencia según el análisis de Bland-Altman +/- 12.5%).

Conclusiones:

- La ecografía pulmonar presenta una excelente exactitud para el diagnóstico de neumonía.
- FEDI y ED son predictores modestos para estimar el resultado del destete en pacientes críticos generales ya que, en contraposición con los excelentes resultados obtenidos en el global de los estudios, en el análisis de subgrupos la exactitud pronóstica es inferior en pacientes críticos generales que en pacientes con mayor probabilidad pre-prueba de fracaso en el destete. Esto sugiere que los aspectos que afectan a la aplicabilidad de los estudios condicionan una sobreestimación de la exactitud de FEDI y ED.
- LUS score parece ser un buen predictor del resultado del destete, pero de momento existe elevada incertidumbre al respecto y son necesarios más estudios que lo evalúen.
- En nuestro medio, la concordancia interobservador fue excelente para LUSm y moderada-buena para FEDI.

ABSTRACT.

Background:

Ultrasound provides both morphological and functional information and it has several strengths as it is portable and fast, does not use radiation and allows real-time scanning. These advantages make it an especially attractive diagnostic method in the critical patient due to their complexity and the high risk of adverse events associated with the transfer, mobilization and implementation of invasive techniques in these patients. The application of ultrasound to pulmonary and diaphragmatic pathology has been recently developed, and several studies suggest that it has an excellent diagnostic and prognostic accuracy in two of the most frequent problems in critical care units, such as diagnosis and management of pneumonia and the optimal moment for the withdrawal of mechanical ventilation.

Aims and methods:

The aim of this research is to advance, through a synthesis investigation, the state of knowledge about the accuracy and applicability of pulmonary and diaphragmatic ultrasound in the field of Intensive Care and more specifically in the diagnosis of pneumonia and the prediction of mechanical ventilation weaning outcomes (success or failure). The goals set are:

- To estimate the accuracy of pulmonary ultrasound for the diagnosis of pneumonia in critically ill adults through a systematic review and meta-analysis,
- To assess the accuracy of the lung and diaphragm ultrasound for predicting mechanical ventilation weaning outcomes in critically ill adults through a systematic review and meta-analysis. There are two proposed diaphragm sonographic predictors: the diaphragmatic excursion (DE), which measures the distance that the diaphragm is able to move during the respiratory cycle, and the diaphragm thickening fraction (DTF), which reflects variation in the thickness of the diaphragm during respiratory effort and is calculated as $(\text{thickness at end-inspiration} - \text{thickness at the end-expiration}) / \text{thickness at the end of the expiration}$. Ultrasound can detect the decrease in the aeration of the lung parenchyma due to cardiac, respiratory, or diaphragmatic origin. This decrease can be quantified through the so-called lung ultrasound (LUS) score, a validated scale whose values range from 0 to 36 points, obtained from the sum of the grades assigned to different ventilation patterns observed in every area of the lung scan.
- To assess the accuracy of lung and diaphragm ultrasound as predictors of success in the withdrawal of mechanical ventilation in our environment through a prospective cohort study, as well as assessing interobserver concordance through a cross-sectional concordance study.

Results:

- Evaluation of the diagnostic accuracy of lung ultrasound for the diagnosis of pneumonia in adult patients through a systematic review and meta-analysis: 16 studies were included, with a total of 2359 participants. Significant heterogeneity was observed in terms of sensitivity and specificity according to the Q test, without clear evidence of threshold effect. The area under the SROC curve was 0.93, with a Diagnostic Odds Ratio (DOR) of 50 at the optimal cut-off point

(95% CI: 21, 120). A tendency towards a higher area under the SROC curve in high quality studies was detected, however these differences were not significant after applying the bivariate meta-regression.

- Evaluation of the accuracy of lung and diaphragmatic ultrasound (particularly of DE, DTF and LUS score) to predict weaning outcomes in critically ill adults through a systematic review and meta-analysis: Nineteen studies involving 1,071 people were included in the study. For diaphragm thickening fraction, the area under the summary receiver-operating characteristic curve was 0.87, and DOR was 21 (95% CI, 11-40). Regarding diaphragmatic excursion, pooled sensitivity was 75% (95% CI, 65-85); pooled specificity, 75% (95% CI, 60-85); and DOR, 10 (95% CI, 4-24). For lung ultrasound, the area under the summary receiver-operating characteristic curve was 0.77, and DOR was 38 (95% CI, 7-198). Based on bivariate metaregression analysis, a significantly higher specificity for diaphragm thickening fraction and higher sensitivity for diaphragmatic excursion was detected in studies with applicability concerns.

- Evaluation of the accuracy and reproducibility of lung and diaphragm ultrasound (DTF and LUSm) to predict the success of weaning of mechanical ventilation in critically ill adults through a prospective cohort study and an interobserver concordance study, both performed in the Intensive Care Unit of Hospital General de Elche: For LUSm, an area under the ROC curve of 0.80 was obtained to predict the success of weaning, sensitivity and specificity at the optimum cut-off point 0.76 and 0.73 respectively. For DTF, AUC was 0.71, sensitivity was 0.93 and specificity was 0.48. An adequate interobserver agreement was found for both LUSm (weighted kappa 0.95) and DTF (ICC 0.78, difference according to the Bland-Altman analysis +/- 12.5%).

Conclusions:

-Lung ultrasound shows an excellent accuracy for the diagnosis of pneumonia.

- DTF and DT are modest predictors of weaning outcome in the general population of critically ill patients. It contrasts to the excellent results obtained in the current literature, suggesting that accuracy of DTF and DE may be overestimated due to aspects that affect the applicability of the studies, as a large number of them were performed in subpopulations with a higher pretest probability of weaning failure.

- LUS score seems to be an accurate predictor, but more studies are needed to reduce uncertainty.

- In our environment, interobserver concordance was excellent for LUSm and moderate-good for DTF.



Biblioteca
UNIVERSITAS Miguel Hernández

1. INTRODUCCIÓN



1. INTRODUCCIÓN.

1.1. Conceptos básicos de ecografía.

La ecografía es un método diagnóstico que utiliza la energía mecánica de las ondas ultrasónicas y aprovecha las propiedades acústicas de la materia para generar imágenes corporales en escala de grises y representar parámetros de flujo¹. Al realizar un estudio ecográfico aplicamos un haz de ultrasonidos sobre un área anatómica del paciente, produciéndose ecos a medida que profundiza y atraviesa los diferentes tejidos, los cuales son recogidos de nuevo por el transductor y transformados en imágenes. Es posible acceder a infinitos planos de exploración, dado que podemos girar el transductor y realizar barridos en los tres ejes del espacio, además la generación de múltiples imágenes por segundo permite la visualización en tiempo real. Los efectos físicos que sufren los ultrasonidos cuando atraviesan el organismo son complejos y exceden la intención del presente documento, pero conviene mencionar una serie de conceptos básicos²:

- Reflexión: la transmisión del sonido es diferente según la impedancia acústica del tejido por el que se desplaza. La reflexión o eco se produce cuando el sonido pasa de un tejido a otro con diferente impedancia o resistencia. Si el sonido discurre por un medio homogéneo, por ejemplo líquido, al no haber diferencia de impedancia no se producen ecos y la imagen será anecoica o anecogénica (negro). Por el contrario, las estructuras cálcicas (hueso) tienen una impedancia muy distinta al resto de tejidos blandos circundantes y se producirá gran intensidad de ecos que dan lugar a una imagen muy refringente o hipercogénica (blanca). Los distintos órganos están compuestos por diferentes tejidos con impedancias ligeramente diferentes y se producirán por tanto ecos con intensidades variadas que darán lugar a imágenes en matices de grises. El aire tiene un comportamiento particular, ya que conduce mal los ultrasonidos (tabla 1)³ y produce distorsión de la onda, generando un área de reverberación difusa.

Tejido	Velocidad (m/s)	Densidad (g/cm ²)
Grasa	1470	0,97
Músculo	1568	1,04
Hígado	1540	1,05
Cerebro	1530	1,02
Hueso	3600	1,7
Agua	1492	0,99
Aire	332	0,001

Tabla 1. Valores de densidad y velocidad de propagación del sonido en los distintos tejidos.

- Artefactos: son imágenes que no se correlacionan con estructuras anatómicas reales. Las causas son múltiples, la mayoría relacionadas con la no uniformidad del haz de ultrasonido y los efectos físicos del mismo en el organismo.

- Refracción: cambio de dirección del haz de sonido. Puede ser causa de artefactos al reflejar imágenes provenientes de localización y profundidad diferente en la zona que se representa en la imagen.
- Absorción: conversión de parte de la energía cinética del ultrasonido en calor.
- Atenuación: pérdida de energía de las ondas de ultrasonidos al atravesar los tejidos por una combinación de fenómenos de absorción, refracción y reflexión. Tiene relación directa con la frecuencia de los ultrasonidos y de ambas depende la resolución o calidad de la imagen de forma que, si utilizamos frecuencias bajas, entre 2 y 7 MHz, la atenuación es menor y, por lo tanto, el sonido llega más profundo. Sin embargo, la resolución de la imagen es menor. Por el contrario, cuando usamos frecuencias altas, por encima de 7 MHz, la resolución o calidad de la imagen será mayor, pero la atenuación se incrementará y sólo obtendremos imágenes superficiales.

Las señales de ultrasonidos pueden representarse de distintas formas. En la representación en **modo B** (figura 1) la imagen bidimensional en tiempo real se construye a partir de los pulsos de ultrasonidos enviados en forma de sucesivas líneas de barrido, de modo que cada línea de corte es sumada a la imagen. Las distintas intensidades de los ecos se convierten mediante el procesamiento electrónico de densidad variable (escala de grises). Otra forma sencilla de representar la imagen ultrasonográfica es el **modo M** (figura 2)⁴, en el cual se recoge la amplitud de los ecos y el movimiento de los reflectores acústicos en movimiento. Los ultrasonidos en modo M se interpretan evaluando patrones de movimiento de reflectores específicos (por ejemplo el diafragma) y estableciendo relaciones anatómicas a partir de patrones característicos de movimiento⁵.



Figura 1. Imagen ecográfica en modo B en ventana intercostal derecha, plano oblícuo.

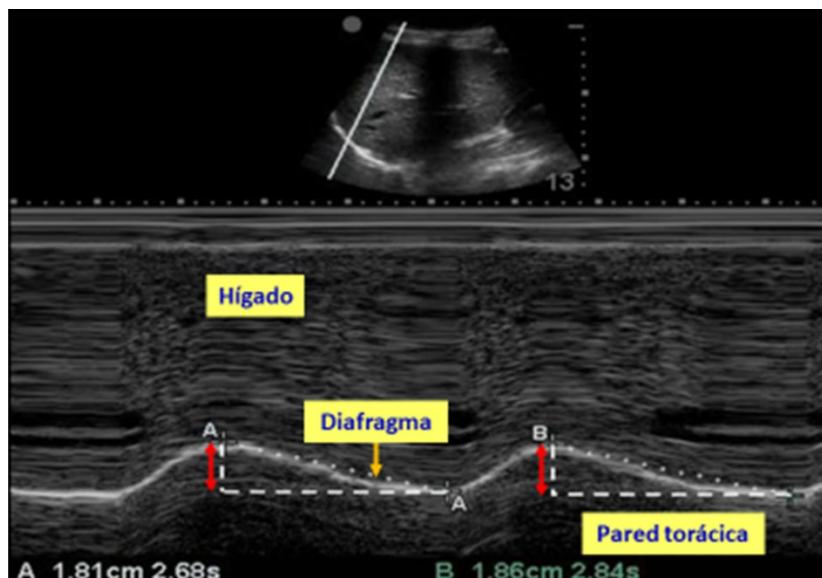


Figura 2. Imagen ecográfica en modo M en ventana intercostal derecha, plano oblicuo.

1.2. La ecografía en la Unidad de Cuidados Intensivos (UCI).

1.2.1. Características generales.

La ecografía es una herramienta ampliamente utilizada en diversas disciplinas médicas. Centrándonos en las UCI podríamos clasificar su uso en términos generales en ecografía cardíaca, torácica, vascular, doppler transcraneal y abdominopélvica, además de servir de guía en procedimientos invasivos diversos. En este contexto también se han desarrollado algoritmos diagnósticos y de toma de decisiones en situaciones agudas basados en los hallazgos ecográficos, como el paciente politraumatizado (FAST: Focused Assessment with Sonography for Trauma)⁶, parada cardiorrespiratoria (FEEL: Focused Echocardiography Evaluation in Life support)⁷, diagnóstico diferencial de la insuficiencia respiratoria aguda (BLUE: Bedside Lung Ultrasound in Emergency)⁸, manejo del shock (FALLS: Fluid Administration Limited by Lung Sonography)⁹, etc. Las características que otorgan a la ecografía un papel tan relevante y despiertan un interés creciente en el ámbito del paciente grave son fundamentalmente¹⁰:

- Portabilidad: los estudios se realizan a pie de cama, evitando movilizar y trasladar al paciente, disminuyendo el riesgo de eventos adversos.
- No invasividad: es inocua y bien tolerada por el paciente.
- Ausencia de exposición a radiaciones ionizantes, tanto de los enfermos como del personal sanitario.
- Disponibilidad y posibilidad de repetir la exploración siempre que sea necesario, sin depender de equipamiento externo ni de otros especialistas.
- Rapidez, pues no precisa de ninguna preparación especial.
- Obtención de imágenes a tiempo real, lo cual permite monitorizar la evolución y respuesta a maniobras terapéuticas.
- Bajo coste.

De entre las múltiples aplicaciones de los ultrasonidos en el paciente crítico, la torácica (excluyendo la ecocardiografía y la ecografía pleural) ha sido la última en desarrollarse, pues clásicamente se ha considerado que el pulmón por su contenido aéreo no era una estructura subsidiaria de ser evaluada por esta técnica. Sin embargo en los años 90 Lichtenstein fue capaz de interpretar los artefactos observados en la ecografía pulmonar y “traducirlos” en una serie de patrones con correlación clínica^{11,12}.

1.2.2. Indicaciones de ecografía torácica no cardíaca en el paciente crítico.

- **Evaluación del derrame pleural:** se visualiza como un espacio libre de ecos (anecoico) entre la pleura parietal y la visceral¹³. Además de permitir identificar y cuantificar el derrame pleural, la ecografía sirve como guía para la realización de toracocentesis, mejorando la seguridad de la misma, sobre todo en pacientes en ventilación mecánica¹⁴.

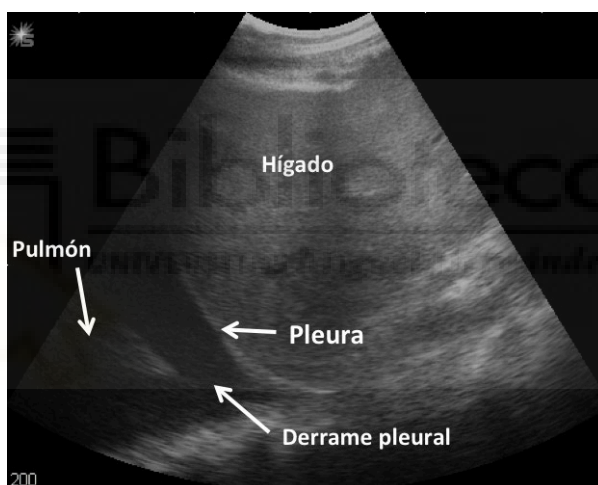


Figura 3. Imagen ecográfica de derrame pleural derecho en ventana intercostal derecha.

- **Diagnóstico de neumotórax:** diversos estudios han demostrado que, para diagnosticar el neumotórax, la ecografía pulmonar en la cabecera del paciente es tan exacta o más que la radiografía convencional realizada con el paciente en decúbito supino¹⁵.

Los signos ecográficos a tener en cuenta son:

- Ausencia de deslizamiento pulmonar (*lung sliding*). En condiciones normales el límite entre la pleura visceral y la superficie pulmonar es visible en la ecografía como una línea ecogénica, la línea pleural. El desplazamiento del pulmón en la cavidad torácica durante la respiración produce una alteración en la línea pleural que denominaremos deslizamiento pulmonar o *lung sliding*, ya que la pleura visceral pierde el contacto con la pleura parietal. La presencia de deslizamiento pulmonar excluye el neumotórax con un valor predictivo negativo y una

sensibilidad del 100%¹¹, sin embargo su ausencia no es específica de neumotórax, apareciendo en otras situaciones como las atelectasias pulmonares o las adherencias pleurales, entre otras.

- Signo de la estratosfera o código de barras. En condiciones normales en modo M se distinguen dos zonas bien diferenciadas dando la apariencia de “orilla de playa”: la parte superior, que corresponde a la pared torácica, formada por líneas horizontales paralelas (agua), y la parte inferior, desde la pleura, de aspecto granuloso (arena) (figura 4A). En presencia de neumotórax se pierde esta apariencia y sólo se visualizan líneas horizontales paralelas (aspecto en código de barras) (Figura 4B).

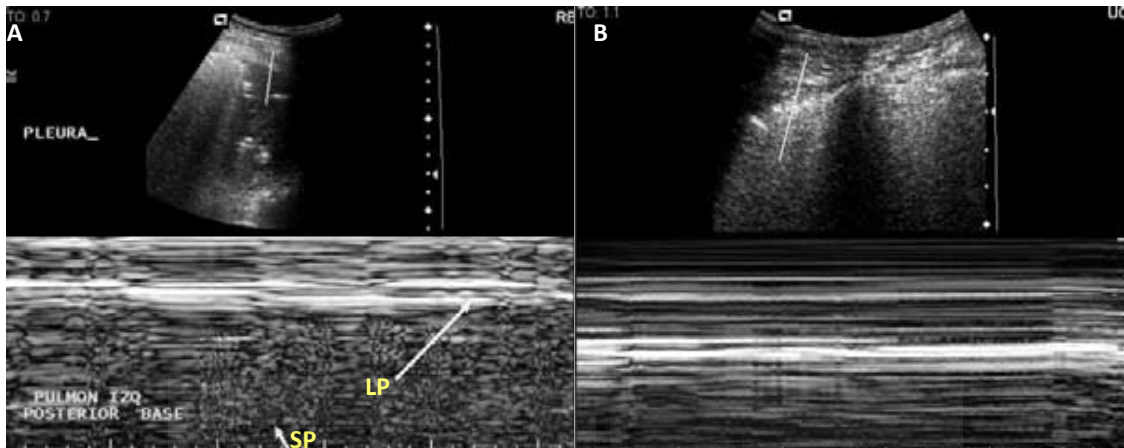


Figura 4. Ecografías pulmonares en modo M. A) En condiciones normales se observa la línea pleural (LP) y signo de la playa (SP). B) En presencia de neumotórax se aprecian exclusivamente líneas horizontales (signo de la estratosfera del código de barras). No se ve la línea pleural y ni el signo de la playa.

- Las líneas B son un artefacto en forma de cola de cometa que se produce en la línea pleural, en la zona de contacto entre la pleura visceral y el pulmón normalmente aireado (Figura 5). Su presencia excluye por lo tanto en neumotórax.

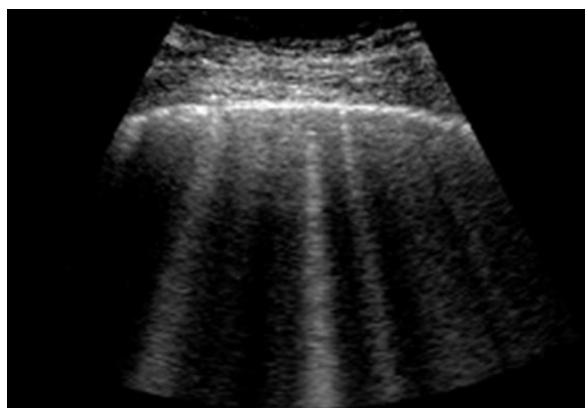


Figura 5. Ecografía pulmonar en modo B donde se aprecian las líneas B.

- Punto pulmonar o *lung point*. Es el punto donde se separan las hojas pleurales visceral y parietal. Es una imagen dinámica, que se identifica como el punto intermedio donde el deslizamiento pulmonar se visualiza de forma intermitente, debido a que con el movimiento respiratorio, el pulmón colapsado se desplaza intermitentemente hacia el área de neumotórax (Figura 6). Su presencia confirma el neumotórax y da una aproximación de su magnitud¹⁶, ya que cuanto más lateral e inferior está el punto pulmonar en la pared torácica, mayor es su extensión.

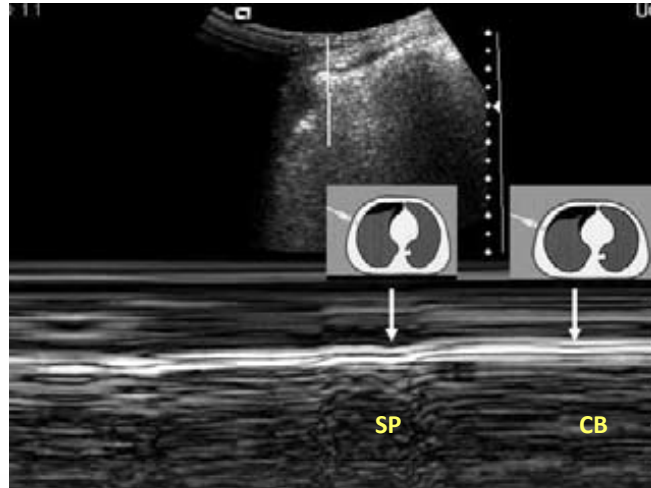


Figura 6. Signo del punto pulmonar: en el modo M se aprecia alternancia del signo de la playa (SP) durante la inspiración y el signo del código de barras (CB) durante la espiración, ver recuadro. Las flechas señalan la línea pleural.

- Exploración del parénquima pulmonar⁸.

-Patrón de aireación pulmonar normal: se caracteriza por el predominio de líneas A (líneas hiperecogénicas horizontales y paralelas que se sitúan a una distancia múltiplo de la que existe entre el transductor y la línea pleural) con deslizamiento pulmonar (Figura 7).

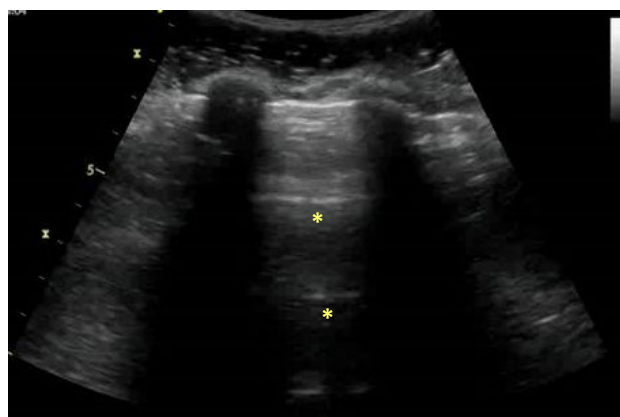


Figura 7. Ecografía pulmonar en modo B en el que se aprecia un patrón de aireación normal, con predominio de líneas A (*).

- **Patrón alveolointersticial:** se caracteriza por un aumento de líneas B en la superficie pleural, de modo que cuanto mayor es la pérdida de aireación pulmonar, mayor es el número de líneas B, y su intensidad se incrementa con los movimientos respiratorios (Figura 8). Este patrón de líneas B representa a un conjunto heterogéneo de entidades, ya que los ultrasonidos no son capaces de diferenciar la naturaleza del fluido que se acumula en los alveolos (líquido, pus) ni del tejido que prolifera. Además del contexto clínico y los factores predisponentes y precipitantes, la separación entre líneas B puede orientar el diagnóstico. Así, líneas B separadas entre sí 7 mm o más, sugieren un engrosamiento de septos interlobulillares por edema o fibrosis intersticial, mientras que una separación menor o igual a 3 mm sugiere áreas en vidrio deslustrado por edema alveolar o síndrome de distrés respiratorio del adulto (SDRA).

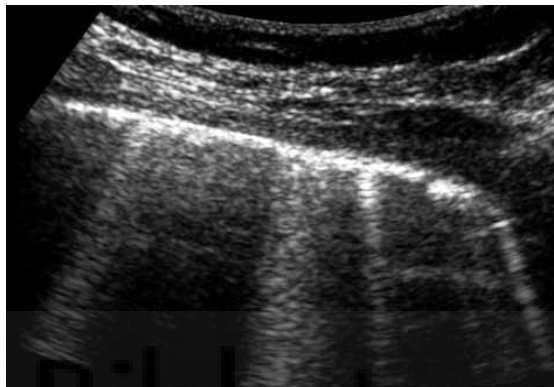


Figura 8. Ecografía pulmonar en modo B en el que se aprecia un patrón alveolointersticial: abundantes líneas B (B) que se originan de la línea pleural (LP).

- **Consolidación alveolar:** los espacios aéreos están rellenos de líquido o células inflamatorias, por lo que el parénquima afectado se convierte en un área densa y con buena transmisión ultrasónica. El pulmón consolidado es hipoecogénico en comparación con el pulmón aireado y también, generalmente, con respecto al hígado y al bazo por su alto contenido acuoso. No se observan artefactos (líneas A o B), pues éstos dependen de la interacción de los ultrasonidos con la interfase aire/fluido¹⁷ (Figura 9). Este patrón comprende entidades diversas, como la neumonía, atelectasia, contusión, infarto o patología tumoral.



Figura 9. Imagen de consolidación pulmonar. C: consolidación; DP: derrame pleural.

1.2.3. Indicaciones no estandarizadas y líneas de investigación actuales.

El importante desarrollo de la ecografía torácica se ve reflejado en el número creciente de artículos publicados en los últimos años. A modo de ilustración, si realizamos una búsqueda sencilla en Medline con la ecuación de búsqueda ((lung[Title] OR diaphragm[Title])) AND (ultrasound[Title] OR ultrasonography[Title]), podemos observar que de 36 estudios publicados y recogidos en esta base de datos en el año 2007, pasamos a 225 estudios en el año 2018 (Figura 10).



Figura 10. Resultados de búsqueda en Medline con los descriptores (lung OR diaphragm) AND (ultrasound OR ultrasonography) en el título entre los años 1972 y 2019.

Estudios recientes sugieren nuevas posibles aplicaciones de la ecografía torácica, entre las que destacan:

- Ajuste óptimo de la presión al final de la espiración (PEEP) en pacientes con SDRA, basándose en la visualización de cambios en el patrón de aireación pulmonar con distintos niveles de PEEP¹⁸.
- Diagnóstico de la presencia y resolución de neumonía en pacientes en ventilación mecánica, basándose en el mismo precepto expuesto en el punto anterior^{19,20}.
- Diagnóstico de tromboembolismo pulmonar²¹: Los hallazgos ecográficos sugestivos de TEP periférico son lesiones hipoecogénicas múltiples subpleurales, de morfología habitualmente de base triangular.
- Evaluación de la presión de oclusión de la arteria pulmonar (PAOP). El análisis de las líneas A y B se correlaciona con la PAOP y puede distinguir a los pacientes con edema pulmonar cardiogénico (PAOP elevado) de aquellos con daño pulmonar agudo (PAOP normal)²².
- Predicción de desarrollo de SDRA tras un traumatismo cerrado de tórax mediante la valoración de la extensión de la contusión pulmonar²³.
- Predicción del fracaso en la retirada de la ventilación mecánica mediante la valoración ecográfica de la función diafragmática²⁴ y el patrón de aireación pulmonar²⁵.

1.3. Investigación de síntesis.

1.3.1. El problema de la investigación inútil. Necesidad de revisiones sistemáticas.

En el año 2009 Chalmers y Glasziou estimaron que el 85% de los fondos destinados a investigación biomédica fueron malgastados de forma evitable en proyectos que no aportaban información relevante²⁶. Se exploraron las posibles causas (Figura 11) y entre ellas destaca que en muchas ocasiones se plantean estudios basados en preguntas de investigación que podrían ser contestadas con la evidencia científica disponible previamente. De hecho, se ha estimado que el 50% de los estudios se diseñan sin hacer referencia a revisiones sistemáticas previas²⁷, con el consiguiente riesgo de generar publicaciones redundantes y exponer a los pacientes a riesgos innecesarios.



Figura 11. Causas de investigación inútil. Traducida de Chalmers, I; Glasziou, P. Avoidable waste in the production and reporting of research evidence. Lancet 2009; 374: 86–89

Se hace imprescindible la conexión entre la evidencia científica disponible y la investigación futura, y en este sentido las revisiones sistemáticas y meta-análisis tienen un papel fundamental, pues permiten sintetizar la información disponible, analizar la calidad de los estudios existentes, proponer nuevas hipótesis e identificar áreas de incertidumbre para guiar nuevas líneas de investigación²⁸. De hecho, el gran volumen de investigación inútil que se genera ha motivado que diversas instituciones académicas y entidades financiadoras exijan la realización de revisiones sistemáticas previamente al inicio de un nuevo ensayo clínico^{29,30}.

1.3.2. Revisiones sistemáticas y meta-análisis de exactitud diagnóstica.

La metodología a seguir a la hora de realizar una revisión sistemática y meta-análisis de estudios de tratamiento o intervención está bien definida y estandarizada, y es en general conocida por la comunidad científica, sin embargo los referentes a estudios de diagnóstico y/o pronóstico implican una mayor complejidad metodológica que a día de hoy se encuentra en desarrollo y no está tan estandarizada. Esto es así atendiendo a ciertas particularidades que los diferencian de los meta-análisis de tratamiento, que se exponen a continuación de forma general:

1. Heterogeneidad. Se asume la presencia de elevada heterogeneidad, por lo que se recomienda por defecto aplicar un modelo de análisis de efectos aleatorios, el cual considera que los efectos de la exposición/intervención en la población son diversos y que los estudios incluidos en la revisión son sólo una muestra aleatoria de todos los posibles efectos³¹.

2. Presencia de efecto umbral: el efecto umbral se produce cuando los estudios incluidos emplean distinto criterio de positividad y suele ser una causa importante de heterogeneidad. Existen distintas formas de explorar la presencia de efecto umbral, por ejemplo:

- Cálculo del coeficiente de correlación de Spearman entre sensibilidad y especificidad. En presencia de efecto umbral existirá una correlación inversa entre ambas³².
- Representación gráfica: en caso de existir efecto umbral, en el diagrama de bosque se apreciaría una imagen especular entre sensibilidad y especificidad. También es útil representar en una gráfica las parejas de sensibilidad y especificidad de cada estudio en el plano ROC (Receiver Operating Characteristic). En este plano, la zona más cercana a la esquina superior izquierda supone un buen rendimiento diagnóstico mientras que la zona central, la diagonal en la que sensibilidad y especificidad son iguales, representa una nula capacidad diagnóstica. Si existiera efecto umbral, los puntos en el plano mostrarían un patrón curvilíneo, cuya elipse de confianza se proyectaría sobre la curva ROC. Cambiando el umbral de positividad de una prueba se obtendría una mayor (o menor) sensibilidad con el consiguiente efecto contrario sobre la especificidad³³.

3. Es fundamental la adecuada elección de la medida resumen. En los estudios de exactitud diagnóstica el resultado no es resumido por un índice, como ocurre en los estudios de evaluación de tratamientos, sino por una pareja de índices que describen la validez de la prueba. Habitualmente, estos dos índices son sensibilidad y especificidad, o bien los cocientes de probabilidad positivo y negativo. El cálculo de la sensibilidad y especificidad ponderadas de los distintos estudios no siempre es adecuada, pues pueden existir diferencias importantes entre los estudios generalmente debido al efecto umbral o a la presencia de cierta subjetividad por parte del evaluador, como ocurre en el caso de la interpretación de pruebas de imagen. Debe huirse, en la medida de lo posible, de utilizar los valores predictivos pues, como es bien sabido, dependen de la prevalencia de la condición clínica que se diagnostica y ésta puede ser muy variable entre estudios³⁴. La odds ratio diagnóstica (ORD) indica en un solo parámetro la probabilidad de un test positivo en enfermos comparada con la probabilidad de un test positivo en sanos. Pese a ser una medida independiente de la prevalencia de enfermedad y que se ve

poco afectada por el efecto umbral, no se recomienda su uso como única medida resumen, ya que al ser tan global implica una dificultad añadida en su interpretación, pues no distingue entre la capacidad del test para detectar enfermos (sensibilidad) y detectar sanos (especificidad) y diferentes pares de sensibilidad y especificidad pueden tener como resultado la misma ORD³⁵. La curva ROC resumen o summary ROC (SROC) es la resultante de combinar en el plano ROC los resultados de los distintos estudios primarios es el método más adecuado a usar en caso de elevada heterogeneidad entre estudios. A nivel gráfico aporta información valiosa sobre el comportamiento del test diagnóstico para distintos puntos corte, permite la estimación del punto de corte óptimo (aquel con el que se consigue la máxima sensibilidad y especificidad de forma conjunta) y da una idea sobre la heterogeneidad entre estudios, posibles indicios de efecto umbral y permite la identificación de estudios con resultados muy dispares del resto (outliers). A partir de la curva ROC es posible calcular el área bajo dicha curva (area under the curve, AUC), considerado un índice de la exactitud global de la prueba: la exactitud máxima correspondería a un valor de AUC de 1 y la mínima a uno de 0.5 (si fuera menor de 0.5 debería invertirse el criterio de positividad de la prueba)³⁶.

4. Uso de modelos estadísticos avanzados. Las revisiones sistemáticas y meta-análisis de pruebas diagnósticas precisan de métodos estadísticos de una complejidad superior a los de intervención, siendo en la actualidad un campo en desarrollo. El método más usado hasta hace unos años era el de Moses-Littenberg, el cual presenta ciertos inconvenientes, pues no tiene en cuenta la distinta precisión con la que se estimaron la sensibilidad y la especificidad en cada estudio ni tampoco la heterogeneidad entre estudios. Para superar estas limitaciones, se han propuesto recientemente modelos de regresión más complejos para estimar la curva sROC. El primero de ellos es un modelo de efectos aleatorios, bivalente, que contempla la eventual correlación entre sensibilidad y especificidad, tiene en cuenta la distinta precisión con la que han sido estimadas e incorpora una fuente de heterogeneidad adicional debida a la varianza entre estudios³⁷. La segunda propuesta se refiere al modelo conocido como HSROC o modelo jerárquico. Es similar al modelo anterior, salvo que hace explícita la relación existente entre sensibilidad y especificidad a través del umbral. Como el anterior, también tiene en cuenta la heterogeneidad entre estudios³⁸.

Ambos modelos, bivalente y jerárquico, permiten obtener estimaciones promedio de sensibilidad y especificidad con sus correspondientes regiones de confianza y predicción. Las diferencias entre ambos modelos son pequeñas y recientemente se ha demostrado que, en ausencia de covariables, ambos abordajes son distintas parametrizaciones del mismo modelo^{33,39}.

1.3.3. Meta-análisis sobre reglas de predicción clínica.

Un objetivo común de los estudios primarios de diagnóstico y/o pronóstico es el desarrollo de los llamados modelos o reglas de predicción clínica. Estos modelos estiman el riesgo de que una determinada condición ocurra en un individuo, para lo cual combinan información de múltiples variables predictoras. A partir de ellos es posible estratificar a los pacientes según su riesgo individual y facilitar la toma de decisiones clínicas⁴⁰.

Desafortunadamente, a menudo la evidencia es contradictoria en cuanto a la exactitud de los modelos de predicción desarrollados. Por esta razón, existe una creciente demanda de

investigación de síntesis que evalúe el desempeño de cada modelo distintos subgrupos de población (aplicabilidad)⁴¹.

1.4. Justificación de la tesis y línea de investigación.

Las técnicas de imagen más habitualmente implementadas en UCI para el diagnóstico de neumonía son la radiografía simple (RxT) y la Tomografía Computerizada (TC). Ambas exploraciones presentan inconvenientes significativos entre los que destacan la exposición del paciente y personal sanitario a radiaciones ionizantes^{42,43}, elevada variabilidad interobservador en el caso de la RxT⁴⁴ y un elevado riesgo de eventos adversos durante el traslado a la sala de Radiología en el caso de la TC. En este contexto y pese a que por el momento no se contempla en las guías de práctica clínica actuales^{45,46}, la ecografía pulmonar podría ser una técnica de imagen adyuvante o incluso alternativa gracias a su portabilidad, rapidez, ausencia de radiaciones ionizantes y ofrecer una exploración a tiempo real, habiendo obtenido resultados prometedores en investigaciones recientes⁴⁷⁻⁵⁰.

A esto se une el hecho de que además de servir para el diagnóstico por imagen, la ecografía nos proporciona información para un diagnóstico funcional, lo cual podría resultar útil para afrontar uno de los problemas habituales en toda unidad de críticos, como es la decisión del momento de retirar la ventilación mecánica invasiva en pacientes intubados, en la que la función diafragmática y la adecuada aireación del parénquima pulmonar juegan un papel fundamental⁵¹. En contraposición a los métodos clásicos para predecir el resultado del destete de la ventilación mecánica, que presentan una baja exactitud pronóstica⁵², en los últimos años se ha desarrollado la aplicación de los ultrasonidos para tal fin con resultados alentadores^{53,54}, no existiendo hasta el momento una evidencia clara acerca de la exactitud real de los índices pronósticos propuestos.

Pese a las cualidades descritas, de momento no se ha generalizado el uso de la ecografía pulmonar y diafragmática, probablemente debido a que se trata de una técnica incipiente, la ausencia de una evidencia sólida que lo justifique (diversidad en los datos primarios) y también de una formación estandarizada en la adquisición de destrezas en ecografía pulmonar y diafragmática entre los profesionales. Además, merece la pena señalar que pese a ser una técnica de imagen que implica cierto componente subjetivo, la concordancia interobservador de la ecografía pulmonar y diafragmática⁵⁵ han sido poco estudiadas hasta el momento.

Según lo expuesto y dado que la ecografía pulmonar y diafragmática se perfila como una herramienta prometedora en el ámbito del paciente crítico, la presente tesis tiene como objeto valorar su exactitud y aplicabilidad en dos de los retos más habituales para el intensivista, como son el diagnóstico de neumonía y la decisión del momento óptimo para la retirada de la ventilación mecánica.



2. OBJETIVOS



2. OBJETIVOS.

El objetivo del presente trabajo es avanzar, mediante una investigación de síntesis, el estado del conocimiento sobre la exactitud y aplicabilidad de la ecografía pulmonar y diafragmática en el ámbito de los cuidados intensivos y más concretamente en el diagnóstico de neumonía y la predicción del éxito o fracaso del destete de la ventilación mecánica.

Objetivos primarios:

- Estimar, mediante una revisión sistemática y meta-análisis, la exactitud de la ecografía pulmonar para el diagnóstico de neumonía en pacientes críticos generales (ver artículos en los anexos 1 y 2: “Accuracy of lung ultrasound in the diagnosis of pneumonia in adults: systematic review and meta-analysis” y “Lung ultrasound for the diagnosis of pneumonia”).
- Estimar mediante una revisión sistemática y meta-análisis, la exactitud de la ecografía pulmonar (LUS score) y diafragmática (fracción de engrosamiento diafragmático y desplazamiento diafragmático) para predecir el resultado del destete de la ventilación mecánica en pacientes críticos generales sometidos a ventilación mecánica (ver artículo en el anexo 3: “Diaphragm and lung ultrasound to predict weaning outcome: systematic review and meta-analysis”).

Objetivos secundarios:

- Explorar la exactitud y concordancia interobservador de la ecografía pulmonar y diafragmática en nuestro medio mediante un estudio de cohortes prospectivo y un análisis de concordancia respectivamente (ver artículo en anexo 4: “Lung and diaphragmatic ultrasound as predictors of success in weaning”).



3. HIPÓTESIS E IMPLICACIONES PRÁCTICAS





3. HIPÓTESIS E IMPLICACIONES PRÁCTICAS.

La ecografía pulmonar y diafragmática presenta una elevada exactitud para el diagnóstico de neumonía y para predecir el éxito o fracaso del destete de la ventilación mecánica en pacientes críticos generales, siendo además una técnica reproducible.

La implantación de la ecografía pulmonar y diafragmática podría modificar significativamente las indicaciones actuales de las distintas pruebas de imagen en diferentes contextos clínicos y dinamizar el proceso diagnóstico y la toma de decisiones terapéuticas, integrando en un mismo especialista (el intensivista) la interpretación de la imagen, monitorización, indicación de tratamiento y valoración de la respuesta al mismo.





4. RESUMEN DE LAS PUBLICACIONES





4. RESUMEN DE LAS PUBLICACIONES PRESENTADAS.

4.1. ECOGRAFÍA PULMONAR PARA EL DIAGNÓSTICO DE NEUMONÍA EN ADULTOS: REVISIÓN SISTEMÁTICA Y META-ANÁLISIS.

4.1.1. RESUMEN.

Antecedentes: Diversos estudios sugieren que la ecografía pulmonar podría resultar útil para evaluar la presencia de neumonía, a lo que se suma un perfil más favorable en términos de seguridad y coste económico que las técnicas de imagen empleadas habitualmente: la radiografía y tomografía axial computerizada de tórax (TC).

Objetivos: Valorar la exactitud de la ecografía pulmonar para el diagnóstico de neumonía en pacientes adultos con sospecha clínica o neumonía confirmada, comparada con otras estrategias diagnósticas que incluyan como técnicas de imagen la radiografía y/o TC, mediante una revisión sistemática de la literatura y meta-análisis.

Métodos: Dos investigadores independientes seleccionaron aquellos estudios que cumplían los criterios de inclusión. La calidad de los estudios se evaluó mediante los criterios QUADAS-2 (Quality Assessment of Diagnostic Accuracy Studies). Se calculó el área bajo la curva SROC (Summary Receiver Operating Characteristic) y una estimación ponderada de la odds ratio diagnóstica mediante un modelo de efectos aleatorios y un análisis bivalente. Se realizó un análisis de subgrupos según calidad global de los estudios, gravedad y riesgo de sesgo por gold standard y flujo de los pacientes dentro de cada estudio, así como una metarregresión bivalente.

Resultados: Se incluyeron dieciséis estudios y se estimó un área bajo la curva SROC de 0.93 y una odds ratio diagnóstica ponderada de 50 con un intervalo de confianza (IC) del 95% de (21, 120). Las diferencias encontradas en el análisis de subgrupos no resultaron significativas al aplicar la metarregresión bivalente.

Conclusiones: La ecografía pulmonar es capaz de contribuir al diagnóstico de neumonía con una elevada exactitud y puede ser un recurso prometedor como adyuvante con las técnicas habituales.

Palabras clave: Ecografía pulmonar, Neumonía, Metaanálisis.

4.1.2. INTRODUCCIÓN Y OBJETIVOS.

La neumonía es una entidad que se asocia a una elevada morbimortalidad, de hecho se sitúa en el tercer puesto de causa de muerte a nivel mundial, precedida únicamente por la cardiopatía isquémica y las enfermedades cerebrovasculares⁵⁶. A nivel europeo su incidencia se estima en torno a 1.07–1.7 casos por 1000 personas-año⁵⁷, siendo dichos datos consistentes con los obtenidos a nivel local en el área del Baix Vinalopó⁵⁸.

En pacientes con sospecha clínica de neumonía es fundamental un correcto diagnóstico diferencial para la adecuada orientación del tratamiento. Un diagnóstico y tratamiento fallido o tardío puede suponer un importante incremento en la mortalidad, mientras que el uso inapropiado de antibióticos para tratar síntomas respiratorios no secundarios a infección bacteriana contribuye al desarrollo de resistencias a los mismos⁵⁹.

El diagnóstico de neumonía se sustenta clásicamente sobre tres pilares: la clínica y datos de laboratorio, las técnicas de imagen y los estudios microbiológicos⁵⁹. Centrándonos en el segundo punto, la radiografía simple de tórax (RxT) a día de hoy es la exploración de imagen inicial ante la sospecha de neumonía y la presencia de una nueva condensación es el hallazgo radiológico característico⁶⁰. No obstante, diversos estudios apuntan hacia una baja sensibilidad de la misma para el diagnóstico de neumonía⁶¹⁻⁶³. Self et al⁶⁴ llevaron a cabo un estudio transversal multicéntrico en el que se incluyeron 3423 pacientes valorados en Urgencias por síntomas respiratorios a los que se realizó RxT y tomografía computerizada (TC) usando la segunda como prueba de referencia para el diagnóstico de consolidación pulmonar. Obtuvieron una sensibilidad del 43.5% (IC 95%: 36.4% – 50.8%) y una especificidad del 93.0% (IC 95%: 92.1%– 93.9%). Al problema de la baja sensibilidad de la RxT en este contexto se suman otras limitaciones, como son la dificultad para ejecutar estudios de calidad en la Unidad de Cuidados Intensivos (UCI) por posición en decúbito, equipos portátiles, etc, la necesidad de movilización del paciente, exposición a radiación y elevada variabilidad interobservador⁴⁴.

La TC de tórax tiene una sensibilidad mayor que la RxT para el diagnóstico de neumonía⁶² y se considera el “gold standard”, pero su uso se reserva para aquellas situaciones en que se produce un fracaso terapéutico o casos de mayor complejidad. Sus principales inconvenientes son la necesidad de traslado del paciente a la sala de Radiodiagnóstico, la mayor exposición a radiación^{42,43} y su elevado coste.

En este contexto y pese a que por el momento no se contempla en las guías de práctica clínica actuales, la ecografía pulmonar podría ser una técnica de imagen adyuvante o incluso alternativa gracias a su portabilidad, rapidez, ausencia de radiaciones ionizantes, ser fácilmente reproducible y ofrecer una exploración a tiempo real⁶⁵. Estas cualidades han impulsado en los últimos años el interés por la aplicación de los ultrasonidos al diagnóstico de la patología pleuropulmonar, habiendo obtenido datos que sugieren un alto rendimiento diagnóstico^{49,66,67}.

La presente revisión sistemática de la literatura y meta-análisis tiene como objetivo valorar la exactitud de la ecografía pulmonar para el diagnóstico de neumonía en pacientes adultos.

4.1.3. MATERIAL Y MÉTODOS.

4.1.3.1. Selección de los estudios.

Se incluyeron aquellos estudios llevados a cabo sobre pacientes mayores de 18 años con sospecha clínica de neumonía o neumonía confirmada en los que se compara la ecografía pulmonar con otras estrategias diagnósticas que incluyan como técnica de imagen la RxT y/o TC para el diagnóstico de neumonía. Se consideró la neumonía alveolar e intersticial, tanto adquirida en la comunidad como nosocomial, incluyendo en este apartado la neumonía asociada a ventilación mecánica (NAVIM). Se excluyeron los estudios referentes a neumonía eosinófila dadas sus distintas características y etiopatogenia.

En la fase de síntesis cualitativa se aplicaron los siguientes criterios de exclusión: estudios no primarios, estudios en los que no es posible definir el diagnóstico de neumonía, el gold standard de comparación no incluye RxT ni TC, imposibilidad para extraer los datos necesarios para el cálculo de una tabla 2x2 de sensibilidad y especificidad, imposibilidad para obtener el trabajo a texto completo y/o su traducción. Igualmente no se tuvieron en cuenta aquellos estudios con menos de 20 pacientes. También se excluyeron aquellos trabajos en los que la unidad de análisis no es el propio paciente, por ejemplo hemitórax, áreas pulmonares o lesiones, asumiendo que dichas unidades no pueden ser independientes unas de otras. En caso de duda se contactó personalmente con los autores.

4.1.3.2. Fuentes de datos y estrategias de búsqueda.

Dos investigadores llevaron a cabo una búsqueda bibliográfica independiente para identificar los artículos potencialmente relevantes en Medline, Scopus, Cochrane library, Web of Science, DARE, HTA Database, Google Scholar, LILACS y ClinicalTrials.gov. Se incluyó la búsqueda de tesis doctorales en Teseo y Tesis Doctorals en Xarxa (TDX), ponencias en congresos en el apartado Conference Proceedings de Web of Science, así como de literatura gris en Opengrey (Sigle) y se revisó la bibliografía de los artículos incluidos.

Para la estrategia de búsqueda se emplearon los términos “pneumonia”, “ultrasound”, “ultrasonography”, “ecograph*” y “sonograph*” y no se usó restricción temporal, incluyendo los estudios publicados sin limitar fecha de inicio hasta abril de 2016. Para proporcionar mayor exhaustividad a la revisión y pese a que los estudios potencialmente elegibles deben referirse necesariamente a pacientes adultos, no se usaron filtros de edad, habiéndose revisado manualmente aquellos trabajos que incluían niños para poder identificar posibles subgrupos de pacientes adultos. Tampoco se limitó la búsqueda por tipo de diseño ni por idioma. En los casos de estudios escritos en lenguas distintas al Inglés, Francés y Español se consultaron servicios de traducción online⁶⁸ y/o traductores externos. Las discrepancias entre ambos investigadores fueron consensuadas con un tercer investigador experto. Los resultados de la búsqueda y flujo de estudios se presentaron según las recomendaciones PRISMA⁶⁹.

4.1.3.3. Extracción de los datos.

De cada uno de los estudios incluidos en el análisis se extrajeron manualmente los datos referentes al número de verdaderos positivos, falsos positivos, verdaderos negativos y falsos negativos. En aquellos casos en los que los datos no se reflejaban literalmente, se construyeron tablas 2x2 para el cálculo de los datos necesarios.

En dos de los artículos^{70,71} se da el caso de que existen exploraciones ecográficas que son clasificadas como “no concluyentes”. En estos casos y con la finalidad de analizar todos los pacientes incluidos, hemos considerado estas exploraciones como “positivas”, siendo conscientes de que esto repercutiría en un aumento de la sensibilidad a costa de disminuir la especificidad. No obstante, en una ocasión una de esas exploraciones no concluyentes para la ecografía también lo fue para el gold standard⁷⁰, por lo que fue preciso en este caso excluir a ese paciente del análisis.

4.1.3.4. Valoración de la calidad de los estudios.

La valoración de la calidad metodológica de los artículos incluidos se realizó según los criterios QUADAS-2⁷² por dos investigadores de forma independiente, siendo consensuadas las discrepancias con un tercer investigador experto. El instrumento QUADAS-2 es un método estandarizado que permite evaluar la calidad de los estudios incluidos en meta-análisis de diagnóstico y clasificar el riesgo de sesgo y factores que puedan afectar a la aplicabilidad en “alto”, “incierto” o “bajo”.

4.1.3.5. Análisis de los datos.

Tras extraer manualmente los datos de cada uno de los estudios incluidos, han sido presentados en el plano ROC y analizados de acuerdo con las directrices de la Asociación Europea de Agencias de Evaluación de Tecnologías³⁵.

En el análisis univariante se construyó un forest plot para la sensibilidad y la especificidad y la odds ratio diagnóstica (ORD). La correlación entre la sensibilidad y la tasa de falsos positivos se exploró gráficamente mediante los correspondientes gráficos de bosque y estadísticamente mediante el examen de la correlación de Spearman y de sus intervalos de confianza al 95%. La heterogeneidad entre los estudios se analizó gráficamente y mediante el test Q.

De acuerdo con la distribución de los estudios en el plano ROC se favoreció el cálculo del área bajo la curva (variabilidad atribuible a efecto umbral) o una estimación ponderada de la sensibilidad y la especificidad en caso contrario. Para ello se utilizó el modelo bivariante de Reitsma⁷³ mediante la aplicación Mada incluida en el paquete estadístico R⁷⁴, habiéndose mostrado este método equivalente al HSROC de Rutter y Gatsonis⁷⁵.

Las posibles causas de heterogeneidad entre los estudios se examinaron mediante el análisis de subgrupos predefinidos: entorno del estudio (UCI frente a otros), calidad global del estudio (bajo frente a alto o dudoso riesgo de sesgo) y las distintas dimensiones de calidad de los estudios incluidos en QUADAS-2. La existencia de interacción entre dichas variables y la exactitud diagnóstica (sensibilidad y especificidad) se examinó mediante una metarregresión bivariante. Las figuras incluidas responden al uso de Mada y RevMan⁷⁶.

4.1.4. RESULTADOS.

4.1.4.1. Selección de estudios.

La búsqueda bibliográfica inicial aportó un total de 2812 estudios, de los cuales 70 fueron seleccionados para una revisión a texto completo al ajustarse a los criterios de inclusión descritos en el apartado anterior. De ellos finalmente se seleccionaron e incluyeron 16 en la fase de síntesis cuali y cuantitativa (Figura 12).

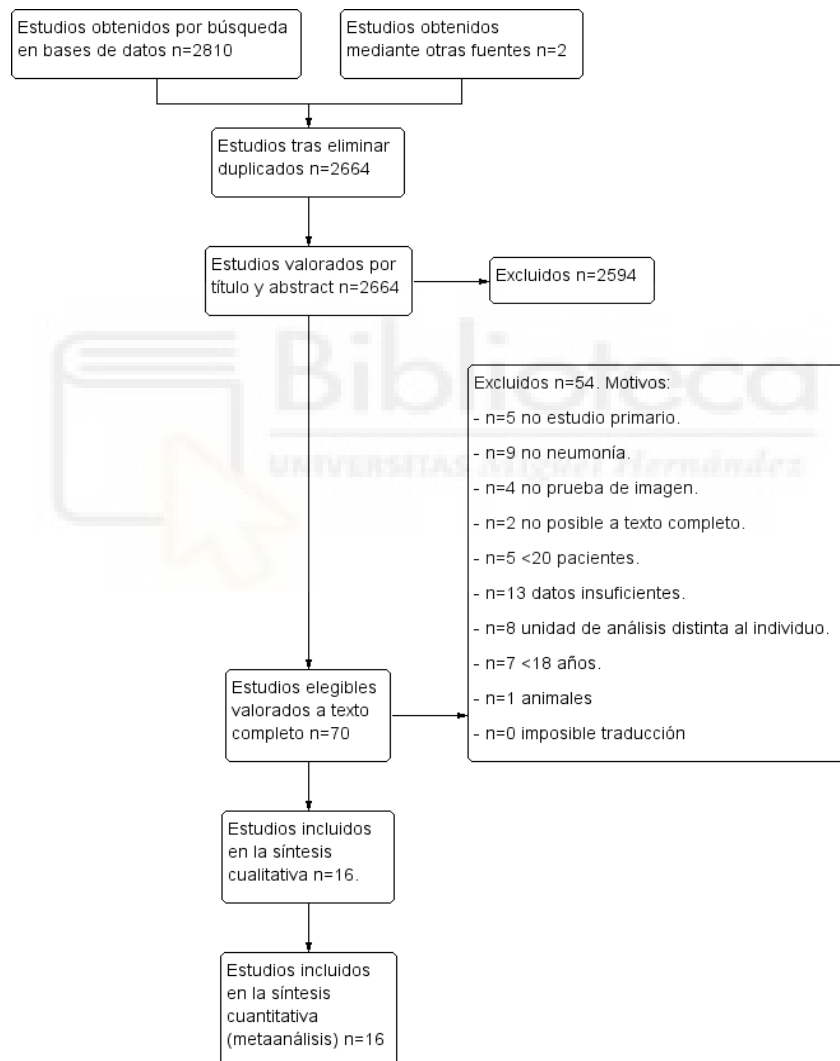


Figura 12. Diagrama de flujo de los estudios cribados según PRISMA y motivos de exclusión.

4.1.4.2. Características de los estudios incluidos.

De los dieciséis estudios incluidos, doce son de tipo cohortes, tres transversales y uno de casos y controles (Tabla 2). Todos se han llevado a cabo entre los años 2008 y 2015, la mayor parte de ellos son europeos y juntos incluyen un total de 2359 pacientes. Pese a que la TC es considerada clásicamente como el test de referencia para el diagnóstico de neumonía, en general no es aplicado de forma sistemática por implicaciones éticas. El gold standard más comúnmente propuesto es el diagnóstico final^{8,66,77-81} teniendo en cuenta datos clínicos, de laboratorio, microbiológicos, RxT y en algunos casos también ecocardiografía⁷⁷ y/o TC^{66,77,80,82}, el cual se realiza únicamente en caso de necesidad clínica y no por motivaciones relacionadas con el estudio. En cuatro estudios la prueba de comparación es la RxT y se reserva la TC para casos dudosos o de discordancia entre los hallazgos de la ecografía y la RxT^{70,83-85} y en otras cuatro referencias se usa la TC como gold standard^{50,86-88}.

En cuanto al ciego del profesional que realiza la ecografía respecto al gold standard y del que interpreta la prueba de referencia en relación a los hallazgos ultrasonográficos, se concreta su presencia en la mayoría de estudios. En ningún caso se especifica la ausencia de cegamiento, pero su rigurosidad no queda claramente expresada en algunas referencias^{77,80,81,84,86,88}. En el apéndice 1 se detallan los aspectos relacionados con la ejecución de la ecografía pulmonar.

En la mitad de los estudios la muestra procedía de Urgencias, cuatro estudios se realizaron sobre pacientes críticos y otros cuatro en pacientes hospitalizados, uno de ellos en Unidad de ictus (Tabla 3). En la mayor parte de los estudios los pacientes incluidos se seleccionaron si presentaban síntomas sugestivos de neumonía^{50,66,70,71,80,84,86,88,89}, aunque otros trabajos se realizaron sobre pacientes con insuficiencia respiratoria^{8,77}, disnea^{79,85} o síntomas respiratorios sin una explicación inicial a los que se les solicita una TC⁸⁷. Uno de los estudios incluye pacientes con ventilación mecánica por causa distinta a patología pulmonar⁹⁰ y en otro se seleccionaron casos confirmados de neumonía asociada a ventilación mecánica⁸¹.

Estudio y año	País	Diseño	Gold standard	Ciego en Ecografía	Ciego en Gold standard
Bataille 2014 ⁷⁷	Francia	Cohortes	DXF (Cl, Lab, Mic, RxT. ± EC ± TC)	Dudoso	Sí
Berlet 2015 ⁹⁰	Suiza	Cohortes	DXF (Cl, Lab, Mic. ± RxT)	Sí	Sí
Bourcier 2014 ⁸⁹	Francia	Cohortes	DXF (Cl, Lab, RxT. ± TC)	Sí	Sí
Busti 2014 ⁷⁰	Italia	Cohortes	RxT. ±TC	Sí	Sí
Corradi 2012 ⁸⁴	Italia	Cohortes	RxT.± TC	Dudoso	Dudoso
Cortellaro 2012 ⁶⁶	Italia	Cohortes	DXF (Cl, RxT, Lab, Mic. ± TC)	Sí	Sí
Fares 2015 ⁸⁶	Egipto	Transversal	TC	Sí	Dudoso
Gallard 2015 ⁷⁹	Francia	Cohortes	DXF (Cl, Lab, RxT)	Sí	Sí
Lichtenstein 2008 ⁸	Francia	Cohortes	DXF (Cl, Lab, RxT, Mic)	Sí	Sí
Liu 2015 ⁵⁰	China	Cohortes	TC	Sí	Sí
Nafae 2013 ⁸⁸	Egipto	Transvesal	TC	Sí	Dudoso
Nazemi 2014 ⁸⁰	Irán	Transversal	DXF (Cl, RxT, Mic. ± TC)	Sí	Dudoso
Nazerian 2015 ⁸⁷	Italia	Cohortes	TC	Sí	Sí
Reissig 2012 ⁷¹	Alemania	Cohortes	RxT ± TC	Sí	Sí
Unluer 2013 ⁸⁵	Turquía	Cohortes	RxT ±TC	Sí	Sí
Zagli 2014 ⁸¹	Italia	Casos y controles	DXF (Cl, RxT y Mic)	Dudoso	Dudoso

Tabla 2. Características de los estudios incluidos. Abreviaturas: DXF: diagnóstico final, entre paréntesis se indica en qué datos se basa éste; RxT: Radiografía de tórax; TC: Tomografía Computerizada de tórax; Cl: datos clínicos, exploración física y/o evolución; Lab: datos de laboratorio; Mic: datos microbiológicos; EC: ecocardiografía

Estudio	N	Procedencia	Edad* (años)	Inclusión	Tipo de neumonía
Bataille ⁷⁷	136	UCI	68±15	IRA	No especificado
Berlet ⁹⁰	57	UCI	61.3 (47.9-71.3)	VMI por causa no pulmonar	NAVVM
Bourcier ⁸⁹	144	Urgencias	77.6±15.2	Sospecha de neumonía	NAC
Busti ⁷⁰	69	Unidad de ictus	77.6±9.3	Sospecha de neumonía	Nosocomial
Corradi ⁸⁴	35	Urgencias	67.09±20.84	Sospecha de neumonía	NAC
Cortellaro ⁶⁶	120	Urgencias	69±18	Sospecha de neumonía	NAC
Fares ⁸⁶	38	UCI	61.02±11.24	Sospecha de neumonía	No especificado
Gallard ⁷⁹	130	Urgencias	79.0±11.1 no FVI 81.9±10.2 FVI	Disnea aguda	NAC
Lichtenstein ⁸	260	UCI	68±16	IRA	NAC y nosocomial
Liu ⁵⁰	179	Urgencias	71.5 (36-88)	Sospecha de neumonía	NAC
Nafae ⁸⁸	100	UCI	≥50 en el 76.25%	Sospecha de neumonía	No especificado
Nazemi ⁸⁰	151	Hospitalizados	61.44±17.40	Sospecha de neumonía	NAC
Nazerian ⁸⁷	285	Urgencias	71±14	Síntomas respiratorios que precisan TC	NAC
Reissig ⁷¹	362	Hospitalizados y Urgencias	63.8 (19-95)	Sospecha de neumonía	NAC
Unluer ⁸⁵	72	Urgencias	Mujeres 68.4±11 Hombres 64.2±12.4	Disnea	NAC
Zagli ⁸¹	221	UCI	56±20.9	Casos confirmados de NAVVM y controles sin NAVVM	NAVVM

Tabla 3. Características de los pacientes incluidos en los distintos estudios.

Abreviaturas: N: número de pacientes incluidos; NAC: neumonía adquirida en la comunidad; NAVVM: neumonía asociada a ventilación mecánica; IRA: insuficiencia respiratoria aguda; VMI: ventilación mecánica invasiva; FVI: fallo ventricular izquierdo.

La edad se expresa según los datos de cada estudio como media ±SD, mediana (P25-P75) o porcentaje dentro de un subgrupo.

4.1.4.3. Riesgo de sesgo y aplicabilidad.

La calidad metodológica de los estudios seleccionados fue analizada mediante el instrumento QUADAS-2 (Figura 13).

En cuanto a la selección de pacientes, se consideró posibilidad de alto riesgo de sesgo en aquellos casos en los que se produjeron exclusiones inadecuadas de pacientes, como pacientes hospitalizados durante menos de 24 horas o aquellos en una situación inestable. Se encontraron problemas de aplicabilidad al considerar que en algunos casos las muestras estudiadas no eran similares al tipo de paciente planteado en esta revisión. Por lo general no se evidenciaron importantes conflictos que pudieran comprometer la aplicabilidad o un alto riesgo de sesgo en lo referente a la prueba índice.

En lo que respecta al gold standard, se considera de calidad alta si se refiere a un diagnóstico final realizado por expertos en base a una síntesis integrada de datos clínicos, radiológicos, de laboratorio y/o microbiológicos, así como aquellos diagnósticos realizados únicamente en base a TC. Debido a la baja sensibilidad de la RXT de forma independiente para discriminar la presencia de infección pulmonar⁶¹⁻⁶⁴, ésta ha sido considerada como una prueba de referencia de menor calidad.

Desde un punto de vista global el mayor problema se hace patente en lo concerniente al flujo de pacientes dentro de cada estudio, donde el principal inconveniente encontrado ha sido una aplicación desigual de la prueba de referencia, ya que en algunas ocasiones no se ha realizado la misma a todos los sujetos incluidos en el estudio (sesgo de verificación diferencial). Esto ha ocurrido en el 50% de los trabajos incluidos (Figura 13).

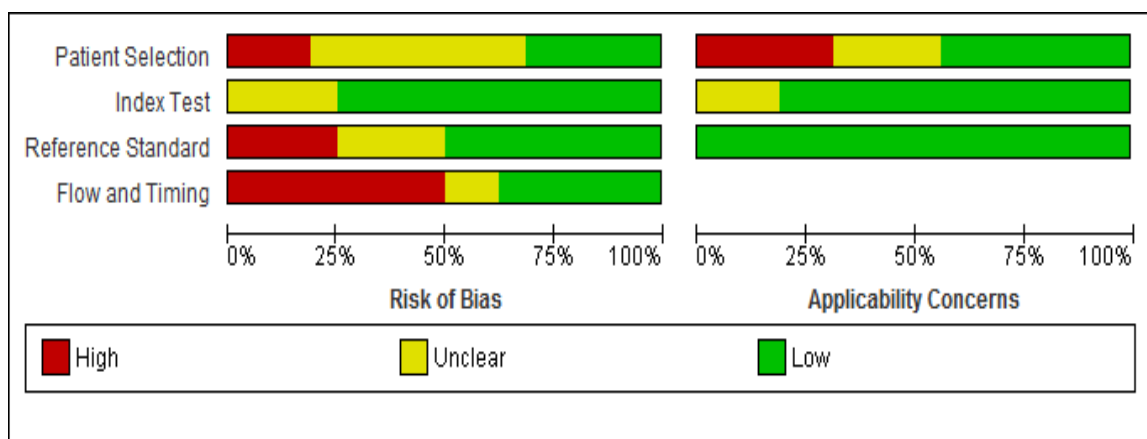
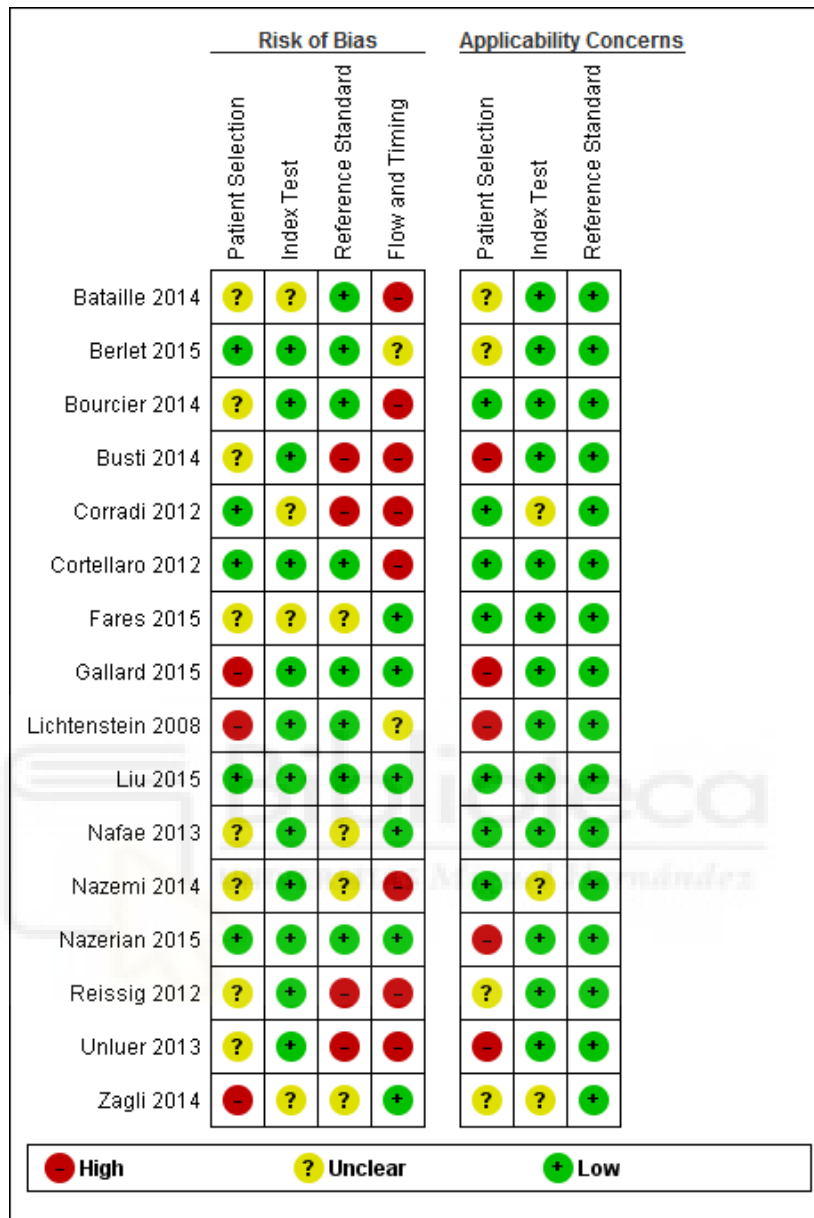


Figura 13. QUADAS-2. En los diferentes dominios se ilustran el número de estudios con riesgo de sesgo alto, bajo o dudoso, así como su porcentaje respecto al total. Igual para la aplicabilidad.

4.1.4.4. Meta-análisis.

Pese a la posibilidad de realizar mediciones, la ecografía no es una técnica cuantificable sino sujeta a la interpretación del observador, por lo tanto y asumiendo la desigualdad entre umbrales de positividad entre estudios, los datos de sensibilidad y especificidad se muestran por separado en cada estudio y no se realiza una ponderación global de acuerdo con las recomendaciones Cochrane para revisiones de exactitud diagnóstica⁹¹ (Figura 14). El test Q fue estadísticamente significativo, tanto para la sensibilidad como para la especificidad, lo que desaconsejaría, en principio, la estimación ponderada univariante de la sensibilidad y la especificidad. En cualquier caso, gráficamente no se observa un patrón de asociación negativa entre sensibilidad y especificidad. El intervalo de confianza al 95% del efecto global indicaría aproximadamente una sensibilidad del 80-90% y una especificidad del 70-90%.

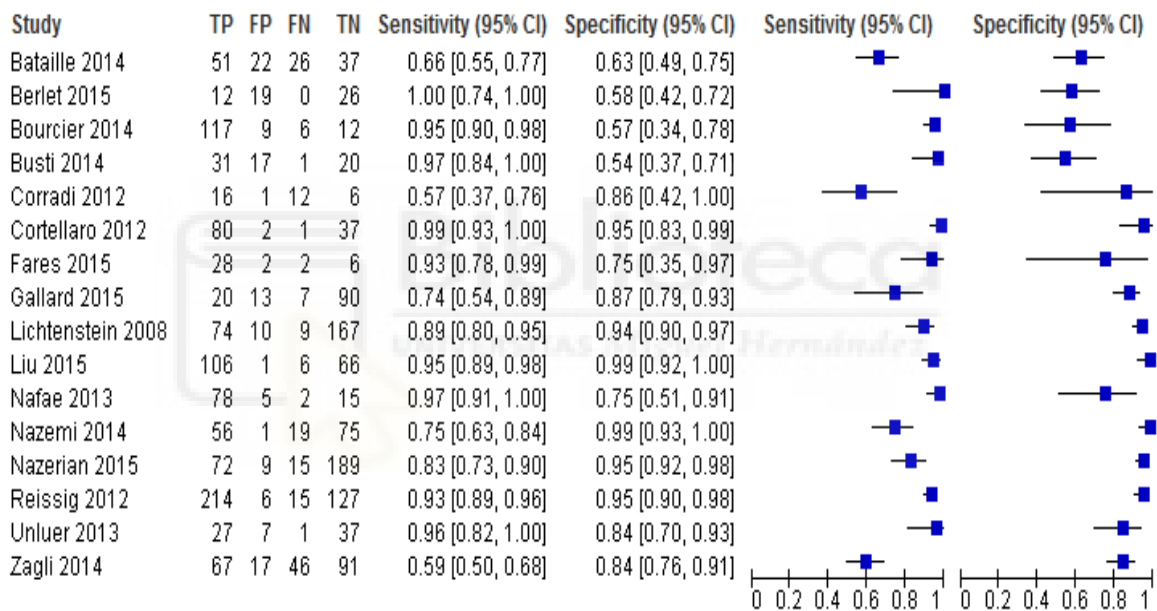


Figura 14. Forest plot de sensibilidad y especificidad de los distintos estudios. Abreviaturas: TP: verdaderos positivos; FP: falsos positivos; FN: falsos negativos; TN: verdaderos negativos.

La estimación global del área bajo la curva SROC de la ecografía pulmonar para el diagnóstico de neumonía fue de 0.93 (figura 4) y la ORD de 50 con un intervalo de confianza (IC) del 95% de (21, 120).

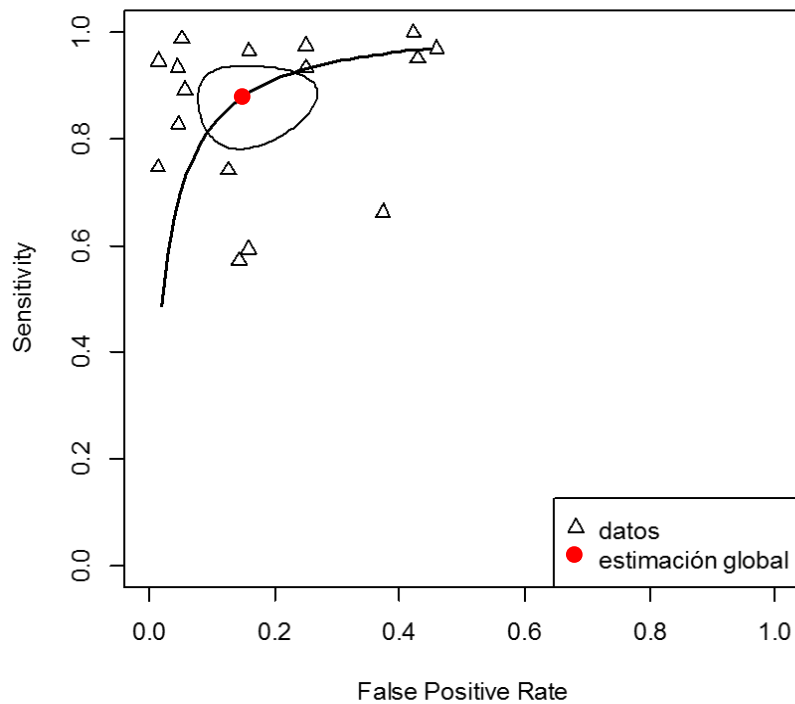


Figura 15. Curva SROC de la exactitud global de la ecografía pulmonar para el diagnóstico de neumonía y elipse de confianza alrededor del valor promedio. Los distintos estudios se representan con triángulos.

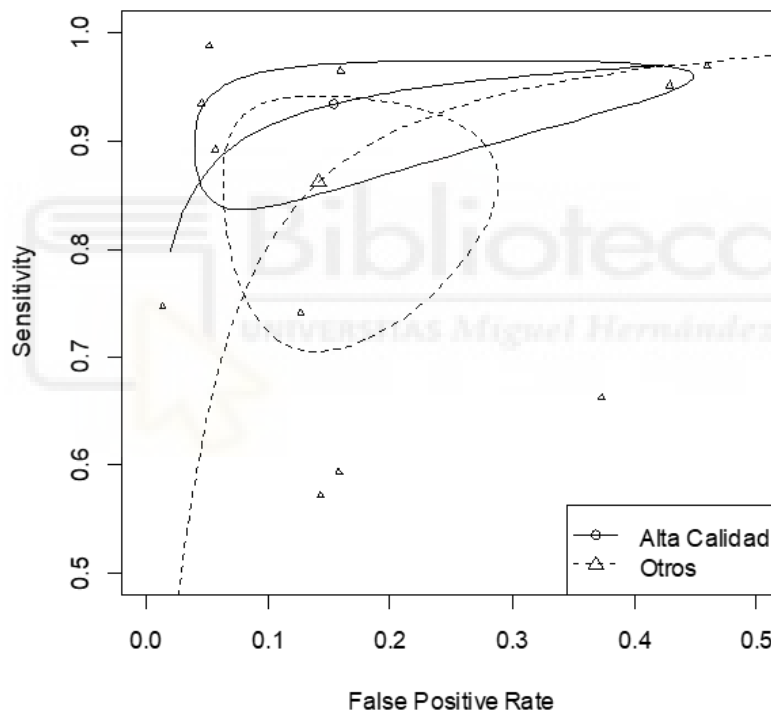
4.1.4.5. Exploración de las causas de heterogeneidad.

La heterogeneidad en los meta-análisis de estudios de diagnóstico suele ser mayor que la encontrada en los meta-análisis de estudios de intervención y muy frecuentemente mayor de lo que sería esperable por mera variabilidad aleatoria⁹¹. De hecho, la dispersión de los dieciséis estudios en el plano ROC y la amplitud del intervalo de confianza sugieren la presencia de una marcada heterogeneidad. En vistas a explorar las causas de dicha heterogeneidad, se llevó a cabo un análisis de subgrupos según la gravedad de los pacientes incluidos en cada estudio, la calidad global de los trabajos y el riesgo de sesgo en los dominios de gold standard y flujo de pacientes según la evaluación mediante los criterios QUADAS-2. No se analizó el subgrupo según riesgo de sesgo en la implementación de la prueba índice porque todos los estudios fueron considerados de riesgo bajo o dudoso. Tampoco se analizó el subgrupo de riesgo de sesgo en la selección de pacientes ya que únicamente hay dos estudios considerados con posibilidad de “alto riesgo” en esta categoría. De forma complementaria al análisis de subgrupos se llevó a cabo una metarregresión bivalente en la que se examina el impacto de las distintas variables definitorias de cada subgrupo sobre la sensibilidad y especificidad.

Definir la calidad global de los estudios incluidos es complejo y meritorio de múltiples y diversas matizaciones. Con la finalidad de simplificar su clasificación y en base al análisis de la calidad metodológica realizado aplicando el instrumento QUADAS-2, hemos considerado una “calidad alta” en aquellos estudios en los que ningún dominio ha sido calificado como de “alto riesgo de

sesgo”, mientras que aquellos trabajos con uno o más ítems clasificados como “susceptibles de alto riesgo de sesgo” pasan a componer el conjunto de estudios considerados de “baja calidad”.

La figura 16 aporta una doble información. Por un lado, al separarlos de los estudios de baja calidad, se observa que el IC al 95% de los estudios de alta calidad tiende a distribuirse a lo largo de la curva ROC, lo que sugiere la existencia de un efecto umbral (y desaconsejaría la estimación global de sensibilidad y especificidad). A pesar de que el análisis gráfico sugiere que los estudios de calidad alta tienen mejor sensibilidad y especificidad que los de baja calidad, la metarregresión bivalente muestra que los coeficientes de la variable “calidad del estudio” no son significativos ni para la sensibilidad ni para la especificidad y por lo tanto las diferencias observadas se podrían explicar por error aleatorio y no son necesariamente atribuibles a verdaderas diferencias entre los grupos (heterogeneidad). Lo mismo ocurre para las variables “gravidad”, “riesgo de sesgo asociado al gold standard” y “riesgo de sesgo asociado al flujo de pacientes”. Dichos análisis se detallan en el apéndice 2.



tsens 0.845 IC 95% (-0.405, 2.095) p= 0.185

tfpr 0.010 IC 95% (-1.329, 1.349) p= 0.988

Figura 16. Comparación de curvas ROC y elipse de confianza entre los estudios considerados de alta calidad y el resto de estudios (calidad considerada dudosa o baja). Tsens: sensibilidad transformada; tfpr: tasa de falsos positivos; IC 95%: intervalos de confianza al 95%.

4.1.4.6. APÉNDICES:

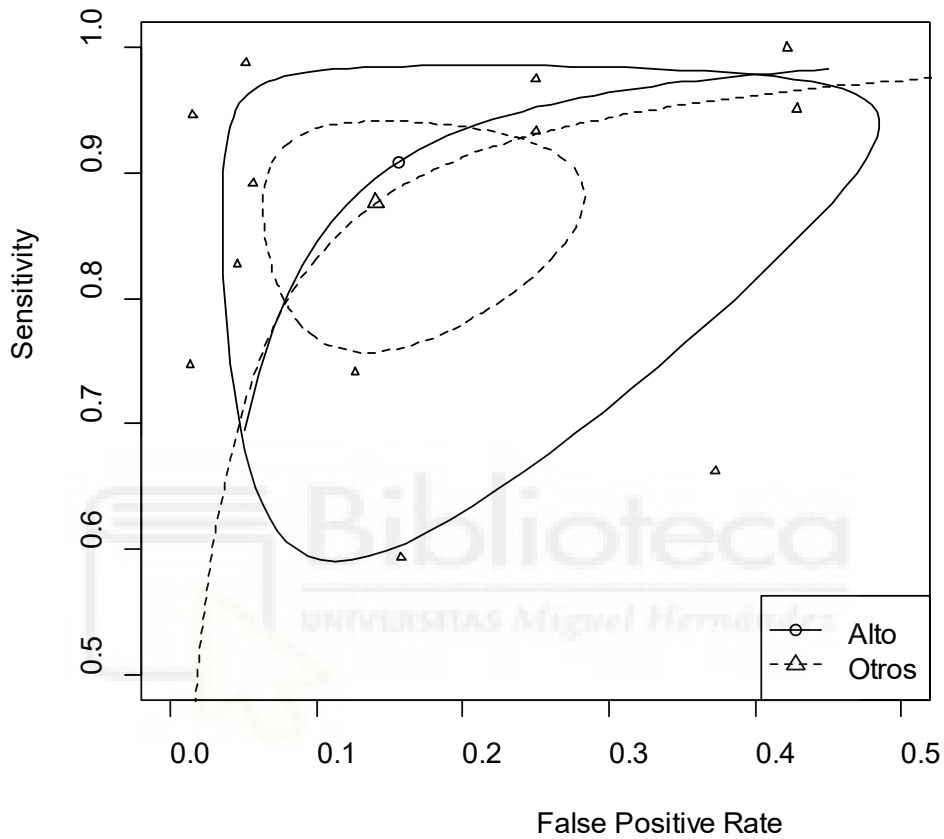
Apéndice 1. Características de ejecución de la ecografía pulmonar en los distintos estudios: ecógrafo y transductores o sondas, metodología y operador.

Estudio	Equipo técnico	Áreas pulmonares exploradas/Posición	Criterio de positividad	Ecografista
Bataille ⁷⁷	HP Sonos 5500. S: 2 -4 MHz.	Ant: 1 área. Lat: 3 áreas PSI	Consolidación alveolar ± BA ± DP o PAI	Dos investigadores (END)
Berlet ⁹⁰	SonoSite M-turbo; SonoSite, Bothell, WA. S: convex C-60 3.5 MHz.	Lateral: 4° EI LMA. Posterior:9°EI LAP PSI si posible.	Consolidación alveolar	4 Intensivistas (END)
Bourcier ⁸⁹	SonoSite M-turbo, Bothell, WA. S: convex 3.5 MHz.	8 áreas Posición sentada si es posible.	Patrón alveolar o intersticial ± DP	5 médicos Urg con 2 días de formación teórica y práctica.
Busti ⁷⁰	Ecógrafo portátil NE. S: convex 3.5-5 y 7-10 MHz.	Todos los EI. Decúbito supino y lateral.	CSP ± BA Patrón intersticial ± DP	Un médico experto (END)
Corradi ⁸⁴	LogiQ-E (GE Medical Systems), S:convex (MHz NE)	NE	Consolidación alveolar, patrón intersticial focal ± DP.	NE
Cortellaro ⁶⁶	Esaote Medical Systems, S: convex 3.5-5 MHz	2 áreas ant, 2 lat y 1 post cada hemitórax	CSP± BA. Patrón intersticial focal	Un médico experto (END)
Fares ⁸⁶	Sonoescape B5, S: convex 3-6 MHz.	6 áreas por hemitórax: ant, lat y post.	CSP ± BA± DP.	Un médico (END)
Gallard ⁷⁹	Sonosite Fujifilm M Turbo (Bothell, WA). S: 3-5 MHz.	3 regiones por hemitórax: ant, lat y post.	Consolidación ±DP	3 médicos Urg con competencias ACEP
Lichtenstein ⁸	Hitachi-405 (Hitachi Medical) S: microconvex 5 MHz.	6 áreas por hemitórax: ant lat y post. PSI, supino en intubados.	Consolidación, intersticial.	2 investigadores (END).
Liu ⁵⁰	SonoSite M-turbo, Bothell, WA. S: convex 3.5-5 MHz.	Todos los EI a nivel de LMC, LAA, LMA y línea paravertebral, desde ápex a diafragma.	Consolidación, patrón intersticial focal, lesiones subpleurales.	3 médicos Urg con formación según ACEP ≥50 ecografías.
Nafae ⁸⁸	Equipo NE, S: convex 3.5- 5 MHz.	5 áreas por hemitórax: Ant, lat y post.	Consolidación ± BA y/o fluido.PAI. ±DP.	Un médico (END)
Nazemi ⁸⁰	Medison V10, sonda de 5 MHz	NE	CSP ± BA ±DP.	Un radiólogo (END)
Nazerian ⁸⁷	MyLab30 Gold (Esaote) y HD7 (Philips). S: lineal 4-8 MHz, convex 3.5-5 MHz.	4 áreas por hemitórax: 2 Ant-lat y 2 post. Sentado, si no: PSI o supino.	CSP ± BA dinámico	4 médicos Urg y MIN con ≥5 años haciendo ecografía y ≥1 año en Urg.
Reissig ⁷¹	Equipo NE, S: convex 3.5- 5 MHz, lineal 7.5 MHz.	Todos los EI anteriores (supino) y posteriores (sentado).	Consolidación ± áreas necróticas, BA o fluido ±DP.	Médicos experimentados (>100 exploraciones ecográficas pulmonares)
Unluer ⁸⁵	M7 (Mindray Biomedical Electronics). S: microconvex 3.6 MHz.	4 áreas: 2 ant, 1 lat, 1 post.	Consolidación alveolar Patrón intersticial	3 médicos Urg formados durante 6 horas por un radiólogo experto.
Zagli ⁸¹	MyLab TM 30CV (Esaote) S: convexa multifrecuencia 3.5-5 MHz.	Todos los EI de base a ápex pulmonar. Posición: supino.	Consolidación alveolar	NE.

Abreviaturas: S: Sonda; MHz: megahertzios; Ant: anterior; Lat: lateral; Post: posterior; DP: derrame pleural; EI: espacio intercostal; PAI: Patrón alveolointersticial; BA: Broncograma aéreo; CSP: Consolidación subpleural; END: experiencia en ecografía pulmonar no descrita; LMA: línea medioaxilar; LAP: línea axilar posterior; NE: no especificado; LMC: línea medioclavicular; LAA: línea axilar anterior; PSI: Posición semiincorporada; ACEP: American College of Emergency ; Physicians; Urg: Urgencias; MIN: Medicina Intena.

Apéndice 2. Análisis de subgrupos, según el riesgo de sesgo alto VS dudoso o bajo en cuanto a la elección del gold standard (2.A), el flujo de pacientes dentro del estudio (2.B) y la gravedad de los pacientes incluidos en cada trabajo (2.C).

● **2.A) Riesgo de sesgo por gold standard.**



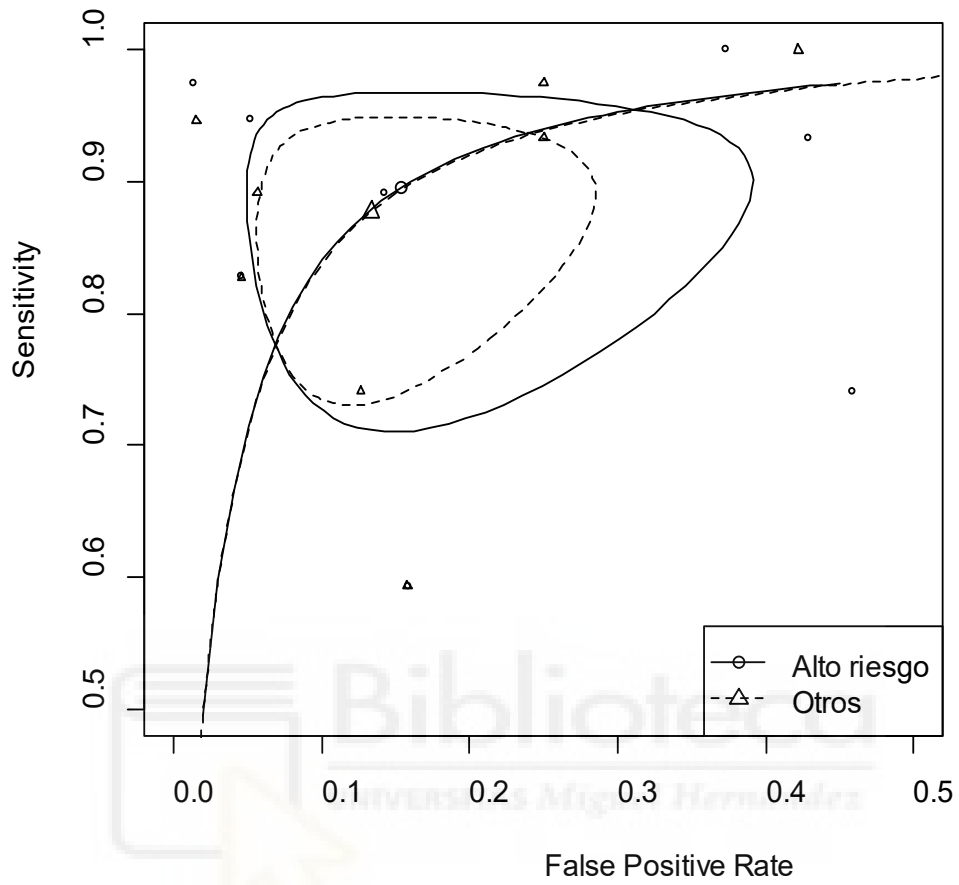
Riesgo alto versus Riesgo dudoso o bajo

tsens 0.133 IC 95% (-1.264, 1.530) p= 0.852

tfpr. 0.258 IC 95% (-1.165, 1.682) p= 0.722

Abreviaturas: tsens: sensibilidad transformada; tfpr: tasa de falsos positivos; IC 95%: intervalos de confianza al 95%.

● 2.B) Riesgo de sesgo por flujo de pacientes.



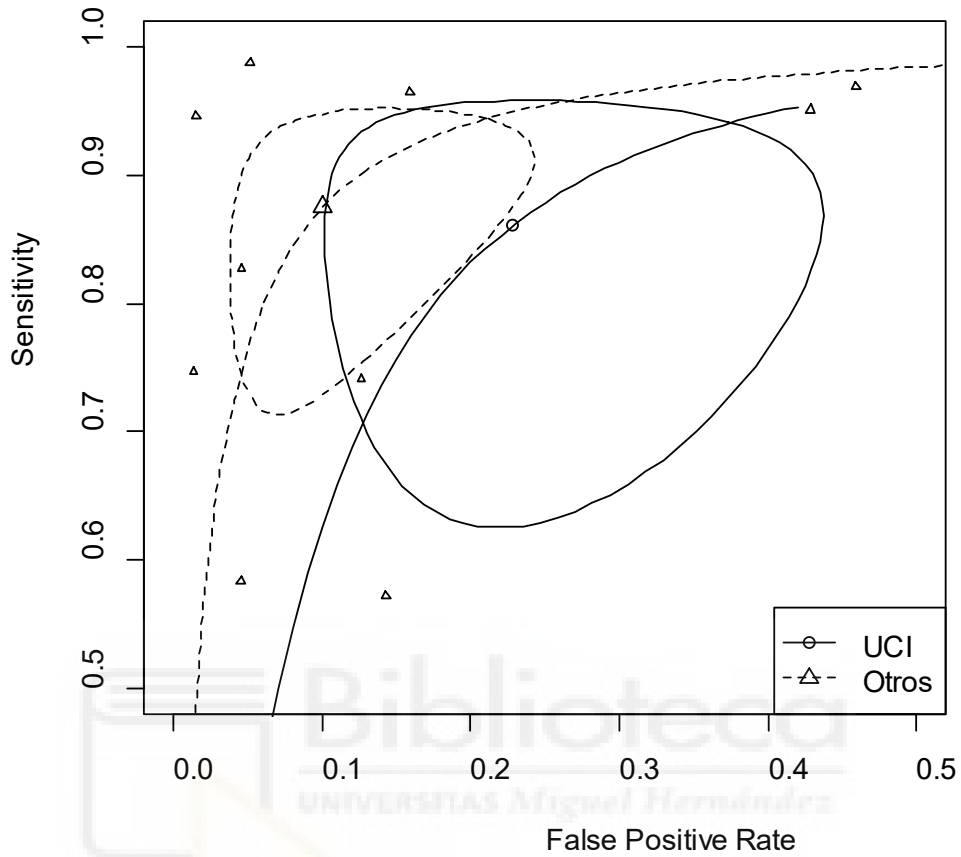
Riesgo alto versus Riesgo dudoso o bajo

tsens 0.063 IC 95% (-1.133, 1.259) p= 0.918

tfpr. 0.287 IC 95% (-0.930, 1.505) p= 0.644

Abreviaturas: tsens: sensibilidad transformada; tfpr: tasa de falsos positivos; IC 95%: intervalos de confianza al 95%.

2.C) Riesgo de sesgo por gravedad de los pacientes incluidos.



Pacientes de UCI versus Pacientes de Urgencias/Hospitalizados

tsens -0.307 IC 95% (-1.537, 0.923) p= 0.624

tfpr. 0.884 IC 95% (-0.300, 2.068) p= 0.143

Abreviaturas: tsens: sensibilidad transformada; tfpr: tasa de falsos positivos; IC 95%: intervalos de confianza al 95%.

4.2. ECOGRAFÍA DIAFRAGMÁTICA Y PULMONAR PARA PREDECIR EL RESULTADO DEL DESTETE: REVISIÓN SISTEMÁTICA Y META-ANÁLISIS.

4.2.1. RESUMEN.

Antecedentes: decidir el momento óptimo para la extubación de un paciente supone un reto, pues tanto un destete precoz como tardío se asocian a múltiples eventos adversos, además las herramientas tradicionales de predicción del resultado del destete (éxito o fracaso) son poco precisas. El objetivo de este estudio es evaluar la exactitud de la ecografía pulmonar y diafragmática para predecir el resultado del destete a través de una revisión sistemática y un meta-análisis.

Métodos: se realizaron búsquedas bibliográficas en Medline, The Cochrane Library, Web of Science, Scopus, LILACS, Teseo, Tesis doctorales en red y OpenGrey, y se revisaron las referencias bibliográficas de los estudios relevantes. Dos investigadores seleccionaron de forma independiente los estudios que cumplieron con los criterios de inclusión. La calidad de los estudios se evaluó de acuerdo con la herramienta QUADAS-2. Se calculó la curva ROC promedio (SROC) y la odds ratio diagnóstica (ORD) mediante un método bivalente de efectos aleatorios. Las fuentes de heterogeneidad fueron exploradas mediante un análisis de subgrupos predefinidos y una metarregresión bivalente.

Resultados: se incluyeron diecinueve estudios (1071 participantes). Para la Fracción de Engrosamiento del Diafragma (FEDI): área bajo la curva SROC (AUSROC) = 0,87, ORD= 21 (intervalo de confianza (IC) del 95%: 11, 40). Respecto a la Excursión Diafragmática (ED): Sensibilidad promedio= 75% (IC 95%: 65,85), Especificidad promedio= 75% (IC 95%: 60,85), ORD = 10 (IC 95%: 4, 24). En cuanto a LUS score (Lung Ultrasound Score) los resultados fueron: AUSROC = 0,77, ORD = 38 (IC 95%: 7,198). Al aplicar la metarregresión bivalente se observó una mayor especificidad de la FEDI en los estudios con problemas de aplicabilidad y también una mayor sensibilidad de la ED en estudios valorados como de menor calidad y con problemas de aplicabilidad.

Conclusión: la ecografía pulmonar y diafragmática puede proporcionar información valiosa para predecir el resultado del destete, pero su exactitud puede ser variable en función de la subpoblación a la que se aplique.

4.2.2. INTRODUCCIÓN Y OBJETIVOS.

Determinar el momento óptimo para extubar a un paciente en la Unidad de Cuidados Intensivos (UCI) puede suponer un reto, ya que una retirada prematura de la ventilación mecánica invasiva (VMI) implica un elevado riesgo de fracaso en el destete con necesidad de reintubación precoz, lo cual conlleva someter al paciente a un estrés respiratorio y hemodinámico innecesario⁹². Por otro lado, el retraso en la extubación aumenta el tiempo que el paciente es sometido a VMI con los riesgos que esto conlleva (desarrollo de neumonía asociada a VMI, barotrauma, lesiones traqueales, etc)⁹³, asociándose tanto la retirada precoz como la excesivamente tardía a un aumento de la mortalidad, estancia en UCI y coste económico^{94,95}.

Las recomendaciones actuales para la retirada de la VMI^{96,97} abogan por una estrategia en dos pasos: en primer lugar, se identifican aquellos pacientes que pueden estar preparados para el weaning mediante el cumplimiento de una serie de criterios de estabilidad clínica y hemodinámica, seguida de una prueba de respiración espontánea (PRE) como herramienta predictora de la probabilidad de éxito en el destete. La PRE puede realizarse tanto retirando por completo el soporte ventilatorio (“tubo en T”) como manteniendo niveles bajos de presión de soporte, si bien las recomendaciones más recientes se inclinan hacia esta última aproximación⁹⁸. No obstante, se ha descrito que aproximadamente entre el 13 y el 26% de los pacientes que son extubados tras pasar con éxito la PRE precisan ser reintubados durante las siguientes 48 horas⁹⁹⁻¹⁰². En los últimos años se han propuesto múltiples índices y parámetros como predictores del resultado del weaning, habiendo demostrado hasta la fecha tan sólo una modesta exactitud pronóstica¹⁰³⁻¹⁰⁶.

Se han descrito varios mecanismos que pueden influir en el desarrollo de fracaso respiratorio tras la extubación, incluyendo la disminución de aireación del parénquima pulmonar o desreclutamiento, relacionado con el desarrollo insuficiencia cardíaca inducida por la PRE¹⁰⁶ y con alteraciones de la complianza pulmonar^{107,108}, así como la disfunción diafragmática asociada a ventilación mecánica¹⁰⁹.

El uso de los ultrasonidos en la UCI es un área de interés creciente¹¹⁰ debido a su portabilidad, rapidez, inocuidad y los resultados alentadores obtenidos en su aplicación para el manejo de múltiples entidades¹¹¹⁻¹¹³, pues no sólo nos proporciona información morfológica útil para el diagnóstico, sino también información funcional a tiempo real. Por ello se postula que la ecografía podría ser de utilidad para valorar tanto el estado de aireación del parénquima pulmonar¹⁰⁷ como la situación funcional del diafragma¹¹⁴, y por lo tanto aportar información de carácter pronóstico sobre el éxito o fracaso de la retirada de la VMI. Los predictores sonográficos propuestos a nivel diafragmático son dos: por un lado la excursión diafragmática (ED)¹¹⁵, que mide la distancia que es capaz de desplazarse el diafragma durante el ciclo respiratorio, y por otro la fracción de engrosamiento diafragmático (FEDI)¹¹⁶ que refleja la variación en el grosor del diafragma durante el esfuerzo respiratorio y se calcula por la fórmula $FEDI = (\text{grosor al final de la inspiración} - \text{grosor al final de la espiración}) / \text{grosor al final de la espiración}$. A nivel pulmonar podemos detectar la disminución de la aireación del parénquima por ecografía, ya sea por causa cardíaca, respiratoria o diafragmática y una forma de cuantificarlo es mediante el denominado Lung Ultrasound Score (LUS score)¹¹⁷, una escala validada cuyo rango de valores oscila entre 0 y 36 puntos obtenidos del sumatorio de puntuaciones asignadas a los distintos patrones de aireación observados en cada área de exploración pulmonar^{54,118}.

La presente revisión sistemática de la literatura y meta-análisis tiene como objetivo valorar la exactitud de la ecografía pulmonar y diafragmática, en concreto la ED, FEDI y LUS score, para predecir el resultado del destete de la ventilación mecánica en pacientes adultos.

4.2.3. MATERIAL Y MÉTODOS.

4.2.3.1. Estrategia de búsqueda y selección de los estudios.

Se incluyeron aquellos estudios llevados a cabo sobre pacientes con edad igual o mayor a 18 años ingresados en una unidad de críticos y sometidos a VMI al menos durante 24 horas en los que se realice ecografía pulmonar y/o diafragmática y se aporten datos relativos al éxito o fracaso del destete de dicha VMI. Consideramos una definición amplia de fracaso en el destete, asumiendo como tal la necesidad de reintubación o ventilación mecánica no invasiva (VMNI) postextubación no programada o traqueostomía y/o muerte en las primeras 24-72 horas o fracaso en la PRE. Dado que éxito y fracaso en el destete con situaciones necesariamente excluyentes, definimos éxito en el destete en aquellas situaciones en las que no existan criterios de fracaso.

Se aplicaron los siguientes criterios de exclusión: estudios no primarios, datos insuficientes para calcular una tabla 2x2 de sensibilidad y especificidad, imposibilidad para obtener el trabajo a texto completo y/o su traducción, aquellos en los que la unidad de análisis no es el propio paciente (por ejemplo hemitórax o regiones pulmonares) y estudios con un tamaño muestral inferior a 20 pacientes.

Dos investigadores independientes (A. M. L-A. y E. M. T-L.) llevaron a cabo una búsqueda bibliográfica para identificar los estudios potencialmente relevantes en las bases de datos Medline, Cochrane Library, Web of Science, Scopus y LILACS, así como en repositorios de tesis doctorales (Teseo, Tesis doctorales en red) y revisión de literatura gris en OpenGrey. Además, se revisó la bibliografía de los estudios relevantes y en algunos casos se contactó personalmente con autores de ponencias en congresos para la obtención de datos no publicados o bien con otros autores para solicitar aclaraciones sobre sus trabajos en caso de duda, obteniendo respuesta aproximadamente en el 30% de las ocasiones. En el proceso de búsqueda se emplearon los descriptores “extubation”, “weaning”, “discontinuation of mechanical ventilation” “disconnect of mechanical ventilation”, “ultrasound”, “ultrasonography” y “echography”. No se impuso ninguna restricción temporal, incluyendo todos los trabajos publicados hasta la fecha de la búsqueda (noviembre de 2016) y se filtraron aquellos estudios realizados sobre seres humanos. Pese a que la presente revisión se centra en pacientes adultos, no se usó ningún filtro de edad, con la intención de hacer un cribado manual de los mismos y aumentar la exhaustividad de la búsqueda. Tampoco se limitó la búsqueda por tipo de diseño ni por idioma. Se activó la alerta de Google Scholar para recibir notificaciones sobre la publicación de estudios potencialmente relevantes entre noviembre de 2016 y abril de 2017. Las discrepancias entre ambos investigadores se consensuaron con un tercer investigador experto (J. L-P.).

Los resultados se presentan según las recomendaciones PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses)⁶⁹.

4.2.3.2. Extracción de los datos.

Dos investigadores extrajeron manualmente los datos necesarios para construir una tabla 2x2 de los estudios seleccionados, bien a partir de datos brutos de verdaderos positivos, falsos negativos, verdaderos negativos y falsos positivos, o bien a partir de datos de sensibilidad y especificidad y/o gráficas. En algunos estudios se aportan datos relativos a la predicción del éxito y en otros predicción del fracaso en el destete para el punto de corte óptimo calculado, por lo que para poder comparar los resultados se construyeron por consenso las tablas para la predicción de éxito en base a la naturaleza dicotómica de la variable “resultado del destete”. A diferencia del FEDI y ED, en el caso del LUS score valores de puntuación altos predicen fracaso, valores bajos “ausencia de fracaso” (igual a éxito) y valores intermedios tienen un significado incierto, es decir, los puntos de corte para predicción de éxito y fracaso son distintos, por lo que se extrajeron los datos para ambas situaciones (éxito y fracaso) a partir de los puntos de corte proporcionados.

En uno de los estudios¹¹⁹ se aportan datos de fallo primario de weaning (en las primeras 48 horas de respiración espontánea) y de fallo secundario (acontecido más allá de las primeras 48 horas), de los cuales hemos seleccionado los relativos al fallo primario porque se ajusta con mayor fidelidad a los criterios de inclusión de este trabajo.

Uno de los trabajos¹²⁰ proporciona los puntos de corte óptimos de FEDI para tres niveles de presión de soporte (PS): 15/5, 10/5 y 5/5, de los cuales hemos tenido en cuenta únicamente los referidos a PS 5/5 cm H₂O, ya que consideramos que es el valor más adecuado para una PRE⁹⁷. En este estudio además no se define claramente el concepto de éxito o fracaso en el destete, por lo que consensuamos considerar como éxito de weaning a los pacientes que permanecen extubados pasadas 48 horas de la PRE.

4.2.3.3. Valoración del riesgo de sesgo.

La valoración de la calidad metodológica de los estudios incluidos fue realizada aplicando el instrumento QUADAS-2¹²¹ (Quality Assessment of Diagnostic Accuracy Studies 2) de forma independiente por dos investigadores (A. M. L-A. y E. M. T-L.). La herramienta QUADAS-2 es un método validado que permite valorar el riesgo de sesgo en distintos aspectos de los estudios incluidos, así como factores que puedan afectar a la aplicabilidad. De forma complementaria se elaboró un protocolo con las definiciones operativas de riesgo de sesgo alto/dudoso/bajo en cada una de las áreas exploradas (ver anexos). Las eventuales discrepancias se resolvieron por consenso con un tercer investigador (J. L-P.).

4.2.3.4. Análisis de los datos.

La fase de análisis se ha desarrollado se acuerdo con las recomendaciones de la Eunetha (European network for Health Technology Assessment)¹²² mediante el uso del software R versión 3.3.2⁷⁴ y el paquete estadístico Mada incluido en el mismo, así como la aplicación Review Manager 5.3⁷⁶ desarrollada por la Colaboración Cochrane.

En el análisis univariante se construyó un diagrama de bosque para la sensibilidad y la especificidad y se calculó la odds ratio diagnóstica (ORD). La correlación entre la sensibilidad y

la tasa de falsos positivos se exploró gráficamente (diagramas de bosque) y estadísticamente mediante el examen de la correlación de Spearman y de sus intervalos de confianza al 95%. La heterogeneidad entre los estudios se analizó gráficamente y mediante I^2 y el test Q.

Los datos se representaron en el plano ROC y, dependiendo de la existencia de indicios gráficos y estadísticos de efecto umbral, se planteó el cálculo del área bajo la curva (variabilidad atribuible a efecto umbral) o bien una estimación ponderada de la sensibilidad y la especificidad (no indicios de efecto umbral).

Para ello se utilizó el modelo bivalente de Reitsma⁷³, equivalente al modelo jerárquico (hierarchical summary receiver operating characteristic o HSROC) de Rutter y Gatsonis⁷⁵. En caso de detectar un “outlier”, es decir, un estudio con valores de sensibilidad y/o especificidad ostensiblemente distintos al resto, se realizó un análisis de sensibilidad excluyendo dicho trabajo.

Las posibles causas de heterogeneidad entre los estudios se examinaron mediante el análisis de subgrupos predefinidos según criterios de calidad y aplicabilidad y una metarregresión bivalente.

4.2.4. RESULTADOS.

4.2.4.1. Selección de estudios.

En la búsqueda inicial se obtuvieron un total de 3833 referencias, de las cuales se seleccionaron 55 para su revisión a texto completo tras cribar por título y abstract. Fueron descartados 36 estudios por los criterios de exclusión descritos en el apartado previo y se seleccionaron finalmente un total de 19 trabajos para su inclusión en la fase de síntesis cuanti y cualitativa (Figura 17).

4.2.4.2. Características de los estudios incluidos.

Los diecinueve estudios seleccionados siguen un diseño de cohortes y han sido publicados entre los años 2004 y 2017. En algunos trabajos se mide más de un predictor, refiriéndose diez a la ED, diez a la FEDI y cinco a LUS score. La definición de “fracaso en el destete” no es homogénea, abarcando según cada estudio uno o más de los siguientes ítems en las primeras 48-72 horas: necesidad de reintubación, VMNI postextubación no programada, muerte, traqueostomía, extubación terminal, retraso en la extubación y/o fracaso en la PRE (tabla 4).

La mayoría de los trabajos se realizaron en UCIs polivalentes, incluyendo algunas de ellas pacientes postquirúrgicos. Tres estudios se llevaron a cabo en UCIs respiratorias¹²³⁻¹²⁵ y uno en una “Unidad de Alta Dependencia”¹²⁶. En general las ecografías se llevan a cabo sobre pacientes considerados como “preparados” para iniciar el destete antes de la PRE, pero en cuatro estudios^{108,124,125,127,128} únicamente consideran a aquellos individuos que superan con éxito dicha prueba (tabla 5). El total de pacientes incluidos en este meta-análisis es de 1071 pacientes.

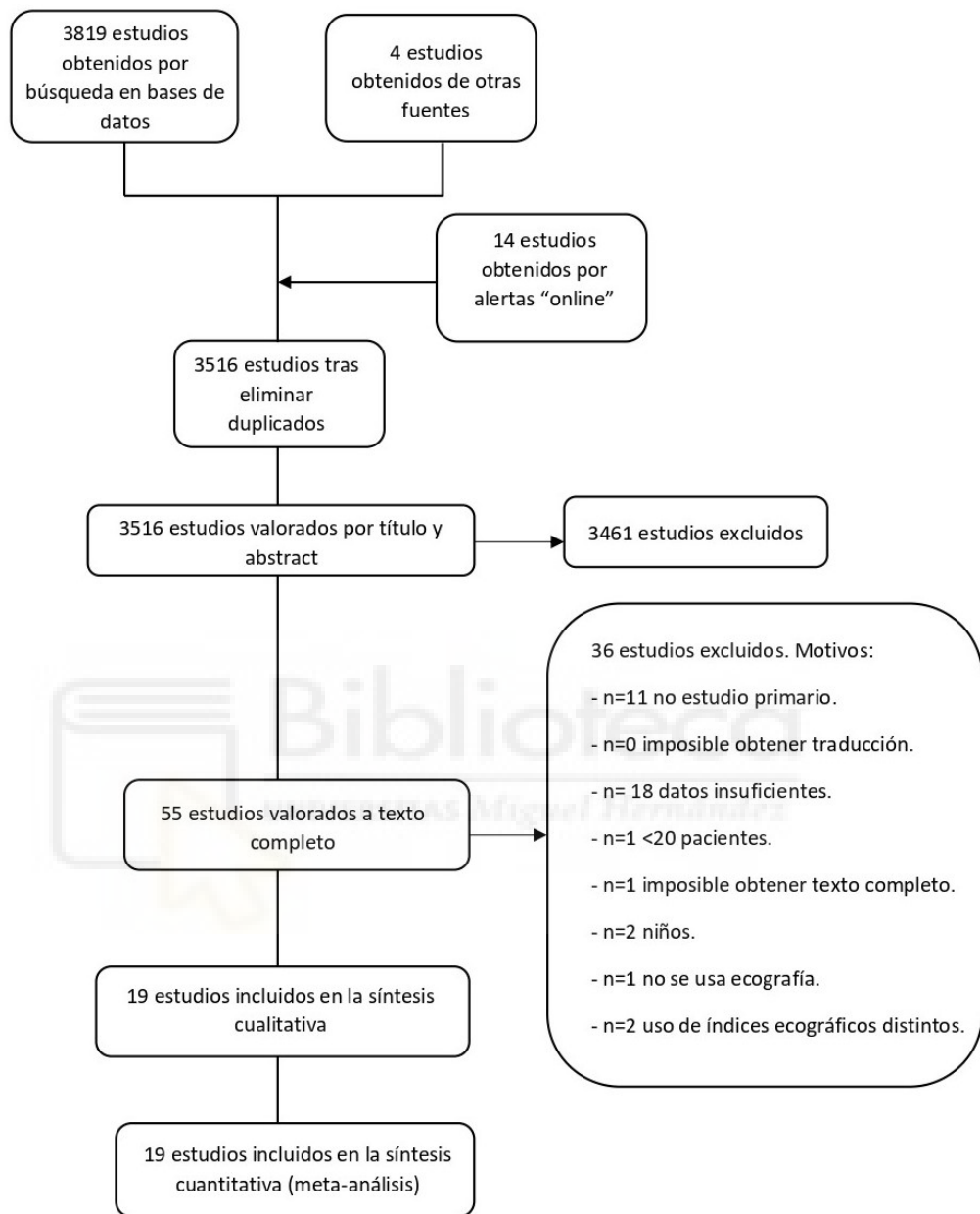


Figura 17. Diagrama de flujo de los estudios cribados según las recomendaciones PRISMA y motivos de exclusión.

Estudio y año	País	Diseño	Medidas y puntos de corte	Definición de fracaso de weaning
Ali 2016 ¹²⁹	Egipto	Cohortes	FEDI >30% ED >1.5 cm	MV durante las primeras 48 horas de respiración autónoma
Baess 2016 ¹²³	Egipto	Cohortes	FEDI ≥30% ED >1 cm	Reintubación o VMNI en las primeras 48 horas tras la extubación.
Binet 2014 ¹³⁰	Francia	Cohortes	LUS score ≤14 (≥19 para fracaso)	Reintubación, VMNI no programada o muerte en las primeras 48 horas post-extubación.
Blumhof 2016 ¹²⁰	EEUU	Cohortes	FEDI >20%	Reintubación or retraso en la extubación (>48 horas)
Carrie 2016 ¹³¹	Francia	Cohortes	ED >2.7 cm	Fallo de PRE o necesidad de VM o muerte en las primeras 48 horas tras la extubación.
Dinino 2014 ¹³²	EEUU	Cohortes	FEDI ≥ 30%	Reintubación en las primeras 48 horas o extubación terminal o traqueostomía.
Farghaly 2017 ¹²⁴	Egipto	Cohortes	FEDI ≥34.5% ED ≥1.5 cm	Incapacidad de mantener una ventilación autónoma sin ningún soporte ventilatorio en las primeras 48 horas post-extubación.
Fayed 2016 ¹³³	Egipto	Cohortes	FEDI >29%	Incapacidad de mantener una ventilación autónoma sin ningún soporte ventilatorio en las primeras 48 horas post-extubación.
Ferrari 2014 ¹²⁶	Italia	Cohortes	FEDI >36%	Fallo en la PRE
Flevari 2016 ¹³⁴	Grecia	Cohortes	ED ≥ 1 cm	Necesidad de soporte ventilatorio (invasivo o no invasivo) durante las primeras 48 horas tras la PRE.
Jiang 2004 ¹²⁸	Taiwan	Cohortes	ED ≥1.1 cm	Reintubación o VMNI en las primeras 72 horas tras la extubación
Jung 2016 ¹³⁵	Francia	Cohortes	FEDI ≥ 30%	Reintubación en las primeras 72 horas tras la extubación o traqueostomía.
Kim 2011 ¹¹⁹	Korea	Cohortes	ED <1.4 cm	Necesidad de VM en las primeras 48 horas de respiración autónoma.
Osman 2017 ¹³⁶	Egipto	Cohortes	FEDI ≥ 28% ED ≥ 1 cm LUS score <12	Reintubación en las primeras 48 horas.
Saeed 2016 ¹²⁵	Egipto	Cohortes	ED >1.1 cm	VM en las primeras 48 horas.
Shoair 2016 ¹³⁷	Egipto	Cohortes	LUS score ≤10 (>18 para fracaso)	Reintubación en las primeras 48 horas.
Soummer 2012 ¹⁰⁸	Francia	Cohortes	LUS score <13 (>17 para fracaso)	Necesidad de soporte ventilatorio (invasivo o no invasivo) durante las primeras 48 horas tras la extubación.
Spadaro 2016 ¹³⁸	Italia	Cohortes	ED >1.4 cm	Fallo en la PRE, reintubación o VMNI en las primeras 48 horas.
Tenza 2017 ¹³⁹	España	Cohortes	FEDI >24%	Fallo en la PRE, reintubación o VMNI en las primeras 48 horas.

Tabla 4. Características de los estudios incluidos. Las medidas y puntos de corte se refieren a la predicción del éxito del destete. Abreviaturas: FEDI: fracción de engrosamiento diafragmático; ED: excursión diafragmática; VM: ventilación mecánica; VMNI: ventilación mecánica no invasiva; LUS score: lung ultrasound score; PRE: prueba de respiración espontánea.

Estudio	N	Servicio hospitalario	Edad* (años)	Tipo de pacientes incluidos
Ali ¹²⁹	54	UCIP	54 ±11.23	No especificado
Baess ¹²³	30	UCI general y respiratoria	59.17±13.17	Preparados para el destete
Binet ¹³⁰	48	UCIP	59±16	Ventilados >48 horas
Blumhof ¹²⁰	52	UCI	62±17	Ventilados >24 horas
Carrie ¹³¹	67	1 UCI médica 1 UCI medicoquirúrgica	66 (58-74)	Primera PRE tras al menos 48 horas en VM.
Dinino ¹³²	63	2 ICU médicas	66±19	Preparados para la primera PRE
Farghaly ¹²⁴	54	UCI respiratoria	65 (55-67.8) SG 62.5 (55-70.7) FG	Enfermedad pulmonar subyacente y con insuficiencia respiratoria aguda que hayan superado con éxito la PRE.
Fayed ¹³³	112	UCI	62.61±12.04	EPOC
Ferrari ¹²⁶	46	“Unidad de alta dependencia”	64.6±12.1	Traqueostomizados que hayan fracasado en uno o más intentos previos de destete.
Flevari ¹³⁴	27	UCIP	65 (53-75)	Destete difícil o prolongado
Jiang ¹²⁸	55	UCI médica	67 (33, 84)	Pacientes que han superado con éxito la PRE.
Jung ¹³⁵	33	UCI medicoquirúrgica	58 (51-67)	Diagnosticados de polineuropatía del paciente crítico, en VM durante >48 horas a los que se les realiza PRE.
Kim ¹¹⁹	82	UCI médica	66	Ventilados >48 horas y preparados para PRE.
Osman ¹³⁶	68	UCI medicoquirúrgica	56 (45-65)	Preparados para PRE (la mayoría postoperados).
Saeed ¹²⁵	30	UCI respiratoria	59±6	Intubados por exacerbación de EPOC que han superado con éxito la PRE.
Shoair ¹³⁷	50	UCI	47.52±14.60 SG 51.89±14.58 FG	Ventilados >48 horas y preparados para la primera PRE.
Soummer ¹⁰⁸	86	2 UCIP	61±14	Ventilados >48 horas que han superado con éxito la primera PRE.
Spadaro ¹³⁸	51	UCI	65±13	Ventilados >48 horas y preparados para la PRE.
Tenza ¹³⁹	63	UCIP	63±16	Ventilados >48 horas y preparados para la PRE.

Tabla 5. Características de los pacientes incluidos en los distintos estudios.

Abreviaturas: N: número de pacientes incluidos; UCI: unidad de cuidados intensivos (no se especifica el tipo); UCIP: unidad de cuidados intensivos polivalente; VM: ventilación mecánica; PRE: prueba de respiración espontánea; EPOC: enfermedad pulmonar obstructiva crónica.

*La edad se expresa como media ± desviación estándar o mediana (rango intercuartílico), según los datos aportados en cada estudio.

4.2.4.3. Riesgo de sesgo y aplicabilidad.

La calidad metodológica de los estudios incluidos se evaluó mediante la herramienta QUADAS-2 (figuras 18, 19 y 20) en la cual se analizan distintos dominios como son la selección de pacientes, prueba índice, gold standard y flujo de pacientes, en cuanto al riesgo de sesgo y posibles problemas de aplicabilidad. Tras la valoración y discusión de estas áreas por el equipo investigador, se asigna de forma consensuada y mediante un código semafórico de colores la consideración de “alto”, “dudoso” o “bajo” riesgo de sesgo o de potenciales objeciones en cuanto a la aplicabilidad.

	<u>Risk of Bias</u>				<u>Applicability Concerns</u>		
	Patient Selection	Index Test	Reference Standard	Flow and Timing	Patient Selection	Index Test	Reference Standard
Ali 2016	⊖	⊖	?	⊖	?	+	+
Baess 2016	+	?	+	+	+	+	+
Blumhof 2016	?	+	+	+	⊖	+	+
Dinino 2014	+	+	+	?	+	+	+
Farghaly 2016	+	+	+	+	⊖	+	+
Fayed 2016	+	+	?	+	⊖	+	+
Ferrari 2014	+	+	⊖	+	⊖	+	⊖
Jung 2016	+	+	+	+	⊖	+	+
Osman 2017	+	+	?	+	⊖	+	+
Tenza 2017	+	+	+	+	+	+	+

⊖ High	⊙ Unclear	⊕ Low
--------	-----------	-------

Figura 18. Aplicación de QUADAS-2 a los estudios que miden FEDI.

	Risk of Bias				Applicability Concerns		
	Patient Selection	Index Test	Reference Standard	Flow and Timing	Patient Selection	Index Test	Reference Standard
Ali 2016	⊖	⊖	?	⊖	?	+	+
Baese 2016	+	+	+	+	+	+	+
Carrie 2017	+	+	+	+	+	+	+
Faghaly 2016	+	+	+	+	⊖	+	+
Flovani 2016	+	+	+	+	⊖	+	+
Jiang 2004	+	+	+	+	+	+	+
Kim 2011	+	+	+	+	+	+	+
Osman 2017	+	+	?	+	⊖	+	+
Saeed 2016	⊖	+	+	+	⊖	+	+
Spadaro 2016	+	+	+	+	+	+	+

⊖ High ? Unclear + Low

Figura 19. Aplicación de QUADAS-2 a los estudios que miden ED.

	Risk of Bias				Applicability Concerns		
	Patient Selection	Index Test	Reference Standard	Flow and Timing	Patient Selection	Index Test	Reference Standard
Binet 2014	+	?	+	+	+	+	+
Osman 2017	+	+	?	+	⊖	+	+
Shoaeir 2016	⊖	+	?	+	+	+	+
Soummer 2012	⊖	+	+	+	+	+	+
Tenza 2017	+	+	+	+	+	+	+

⊖ High ? Unclear + Low

Figura 20. Aplicación de QUADAS-2 a los estudios que miden LUS score.

En líneas generales el principal problema detectado al aplicar la herramienta QUADAS-2 es la presencia de aspectos que pueden comprometer la aplicabilidad en lo que respecta a la selección de los pacientes, pues algunos estudios se centran exclusivamente en pacientes con EPOC¹³³ o cuya causa de intubación es necesariamente respiratoria^{124,125,133}, enfermos con miopatía del paciente crítico¹³⁵, pacientes predominantemente postquirúrgicos¹³⁶ o pacientes que han fracasado en intentos previos de weaning^{126,134}.

En cuanto a la selección de pacientes, algunos estudios excluyen a tipos de pacientes concretos como por ejemplo aquellos con neuromiopatía severa del paciente crítico^{108,137}, EPOC severo¹⁰⁸, pacientes a los que se les haya administrado fármacos como corticoides o aminoglucósicos¹²⁹, intubación por un motivo distinto a EPOC¹²⁵ o pacientes con algún ingreso en UCI en los 12 meses previos¹²⁹.

En lo que respecta a la prueba índice, la ecografía, en general no se apreciaron problemas metodológicos, salvo en un estudio¹²⁹ en el cual ésta se realiza mientras el paciente está en ventilación mecánica invasiva, lo cual no es adecuado para valorar los esfuerzos respiratorios del paciente¹⁴⁰ en el contexto de la predicción del resultado del destete.

En esta revisión no existe un gold standard de comparación, por lo que hemos considerado el éxito o fracaso en el destete según la definición que dé cada autor en su estudio para comprobar si las mediciones ecográficas han sido capaces de predecirlo correctamente o no, siendo dichas definiciones heterogéneas entre los trabajos incluidos. En este sentido, únicamente un estudio¹²⁶ ha sido penalizado al considerar como fracaso en el destete el fallo en la PRE, ya que puede acontecer que un paciente requiera reintubación y/o reconexión a VMI en las horas posteriores a haber superado con éxito una PRE. En general los médicos encargados de tomar la decisión de la retirada de la VMI no conocen las mediciones realizadas en la ecografía (ciego), aunque dicha condición no se menciona en cuatro de las referencias^{129,133,136,137}.

En cuanto al flujo de pacientes, tan sólo en un trabajo¹²⁹ se ha considerado que pueda existir un sesgo potencial, ya que se excluyen del análisis seis pacientes que fallecen y no se especifica si el óbito se produce o no durante proceso de weaning.

4.2.4.4. Meta-análisis.

La ecografía es una técnica que permite hacer mediciones pero también está sujeta a la interpretación del observador, por lo que asumimos la diversidad de umbrales de positividad entre estudios. Por este motivo, de acuerdo con las recomendaciones Cochrane para revisiones de exactitud diagnóstica³¹, no realizamos de entrada una ponderación global y se muestran los valores de sensibilidad (S) y especificidad (E) por separado para cada trabajo.

● FEDI.

No se aprecia una correlación negativa entre sensibilidad y especificidad ni gráfica (figura 21) ni estadísticamente (Rho de Spearman= 0.3, IC 95% (-0.4, 0.7)). El test Q fue de 9.5 (p=0.386) y el I² del 6%. Se obtuvo una Odds ratio diagnóstica (ORD) de 21 con IC 95% (11, 40).

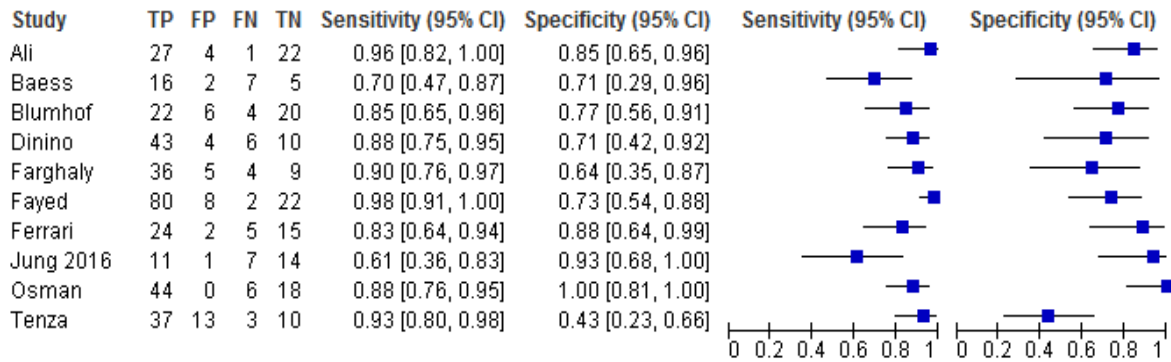


Figura 21. Forest plot de sensibilidad y especificidad de los estudios que miden FEDI. Abreviaturas: TP: verdaderos positivos; FP: falsos positivos; FN: falsos negativos; TN: verdaderos negativos.

Al representar los datos en el plano ROC (figura 22) se observa una dispersión moderada de los mismos y una elipse de confianza que tiende a proyectarse sobre la curva SROC, por lo se optó por el cálculo del AUC, que fue de 0.87, en lugar de una estimación ponderada de S y E.

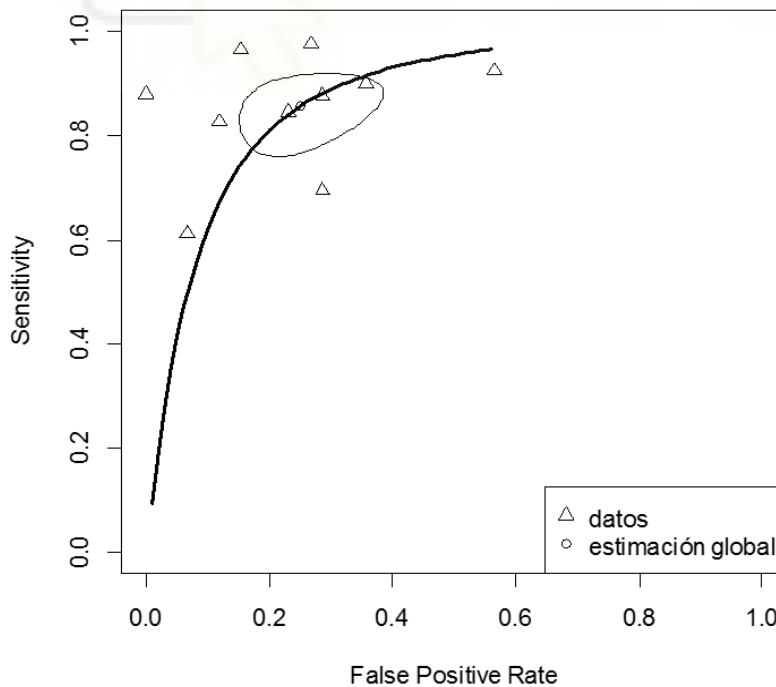


Figura 22. Curva SROC de la FEDI para predecir el éxito en el destete y elipse de confianza alrededor del valor promedio. Los distintos estudios se representan con triángulos.

Se identificó un “outlier” y se realizó un análisis de sensibilidad eliminando dicho trabajo¹³⁹ (e-Fig 1 de la publicación original en *Chest journal*), apreciándose un estrechamiento de la elipse de confianza en cuanto a E que nos permite estimarla de forma ponderada en torno al 80% con un IC 95% (65,90)

● ED.

No se aprecian indicios de correlación negativa entre S y E en el forest plot (figura 23) ni en aplicar la Rho de Spearman (Rho= -0.45, IC 95% (-0.84, 0.25)). La ORD calculada es de 10.6, con IC 95% (5, 24). El estadístico Q fue de 10.7 (p=0.29), lo cual apunta a la posibilidad de una heterogeneidad ligera-moderada entre estudios.

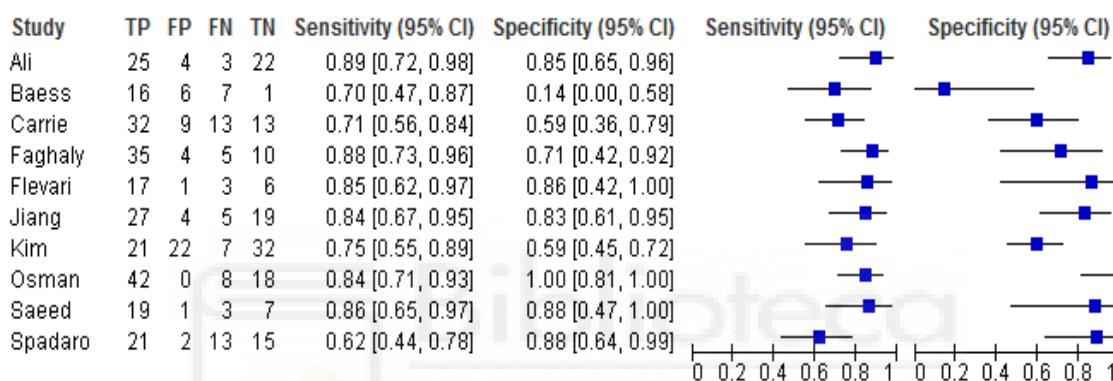


Figura 23. Forest plot de sensibilidad y especificidad de los estudios que miden ED. Abreviaturas: TP: verdaderos positivos; FP: falsos positivos; FN: falsos negativos; TN: verdaderos negativos.

Al representar los distintos estudios en el plano ROC (figura 24) no se aprecian indicios de efecto umbral. La sensibilidad de los distintos estudios se mueve en un rango relativamente estrecho, con un valor promedio en torno al 75% IC 95% (70, 85), sin embargo, las diferencias entre estudios en lo concerniente a E son importantes y por lo tanto no sería adecuado hacer una estimación ponderada.

Se realizó un análisis de sensibilidad por la presencia de un “outlier” de especificidad¹²³ (e-Fig 2 de la publicación original en *Chest journal*), obteniendo un rango menor de valores de especificidad con un promedio alrededor del 75%, IC 95% (60, 85)

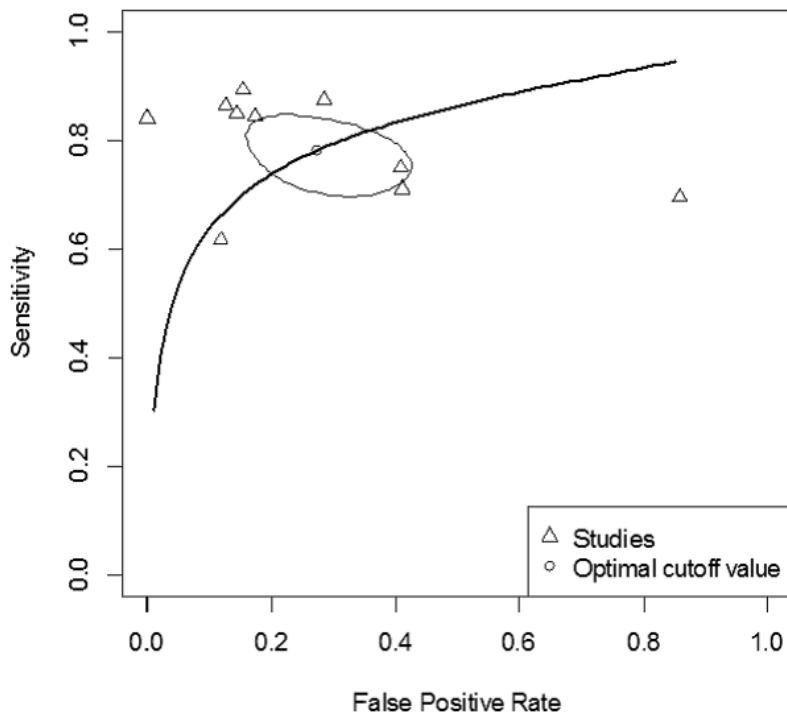


Figura 24. Curva SROC de la ED para predecir el éxito en el destete y elipse de confianza alrededor del valor promedio. Los distintos estudios se representan con triángulos.



● **LUS score.**

No parece existir una correlación negativa entre sensibilidad y especificidad (Rho de Spearman= 0.2, IC 95% (-0.8, 0.9) (figura 25). El test Q fue de 5.1 (p=0.272) y se obtuvo una ORD de 38 con IC 95% (7, 198).

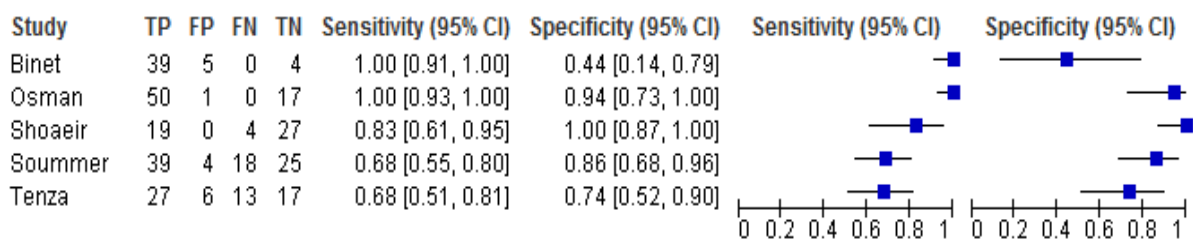


Figura 25. Forest plot de sensibilidad y especificidad de los estudios que miden LUS score. Abreviaturas: TP: verdaderos positivos; FP: falsos positivos; FN: falsos negativos; TN: verdaderos negativos.

Se obtuvo un AUC =0.77, sin embargo y debido al reducido número de estudios disponibles, su representación en el plano ROC implica una notable incertidumbre en cuanto a los resultados (e-Fig 3 de la publicación original en *Chest journal*).

4.2.4.5. Exploración de las posibles causas de heterogeneidad.

La heterogeneidad en los meta-análisis de diagnóstico suele ser mayor que la encontrada en los meta-análisis de estudios de intervención y muy frecuentemente mayor de lo que sería esperable por mera variabilidad aleatoria⁹¹. En vistas a explorar las causas potenciales de dicha heterogeneidad, se llevó a cabo un análisis de subgrupos según la calidad global de los trabajos y posibles problemas de aplicabilidad de acuerdo con las conclusiones obtenidas tras la evaluación con QUADAS-2. Este análisis se ejecutó únicamente para los predictores FEDI y ED, ya que para LUS score disponemos de muy pocos estudios. De forma complementaria se llevó a cabo una metarregresión bivalente en la que se evalúa el impacto de las distintas variables de cada subgrupo sobre la sensibilidad y especificidad.

Calidad de los estudios.

Definir la calidad de los distintos estudios implica una labor compleja de síntesis y discusión, que de forma implícita supone cierta subjetividad por parte de los evaluadores, siendo proclive a múltiples matizaciones. Con el objetivo de simplificar su clasificación y de forma consensuada, se ha definido como subgrupo de alta calidad el integrado por aquellos trabajos en los cuales ninguno de los dominios QUADAS-2 haya sido calificado como de “alto riesgo de sesgo” o “dudoso”, componiendo el resto de estudios el conjunto de estudios de “menor calidad”.

● Calidad de los estudios que miden FEDI.

En la metarregresión bivalente los coeficientes de la variable “calidad del estudio” no son significativos ni para sensibilidad ni para especificidad, por lo que probablemente las diferencias observadas sean secundarias a error intraestudio (Figura 26 A).

● Calidad de los estudios que miden ED.

En cuanto a los estudios que miden ED, al analizar por separado los considerados de calidad alta versus menor calidad, se observa gráficamente que aparentemente los estudios de menor calidad tienen una mayor capacidad para predecir el éxito en el destete (Figura 26 B). Sin embargo en la metarregresión se aprecia que dichas diferencias no fueron significativas, aunque rozando la significación estadística en el caso de la especificidad.

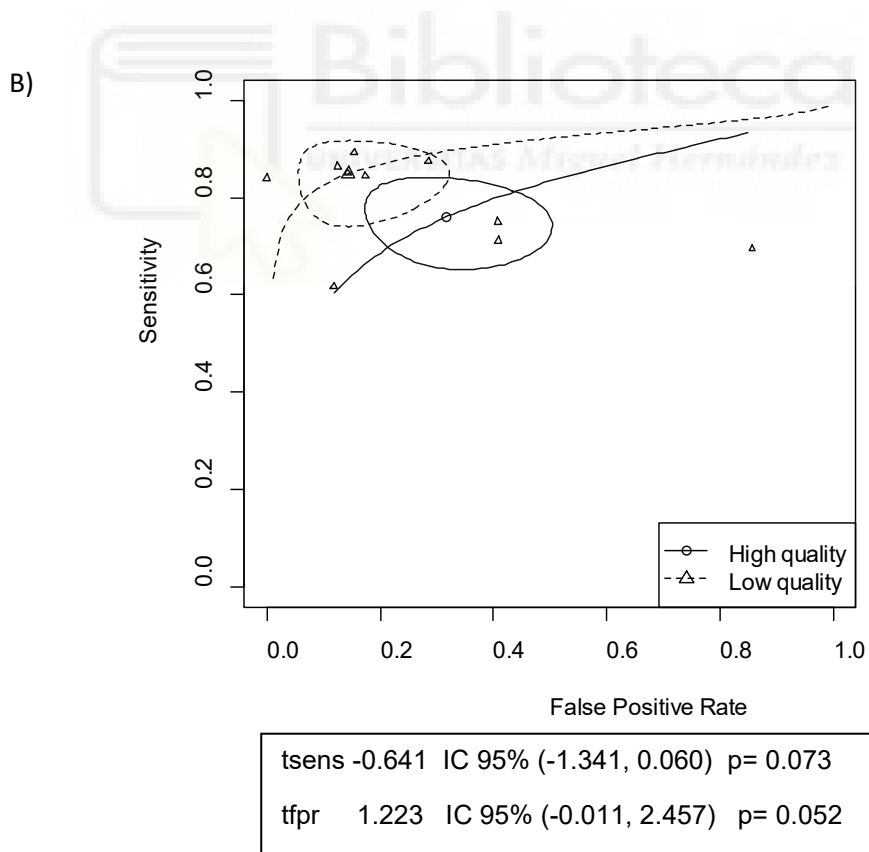
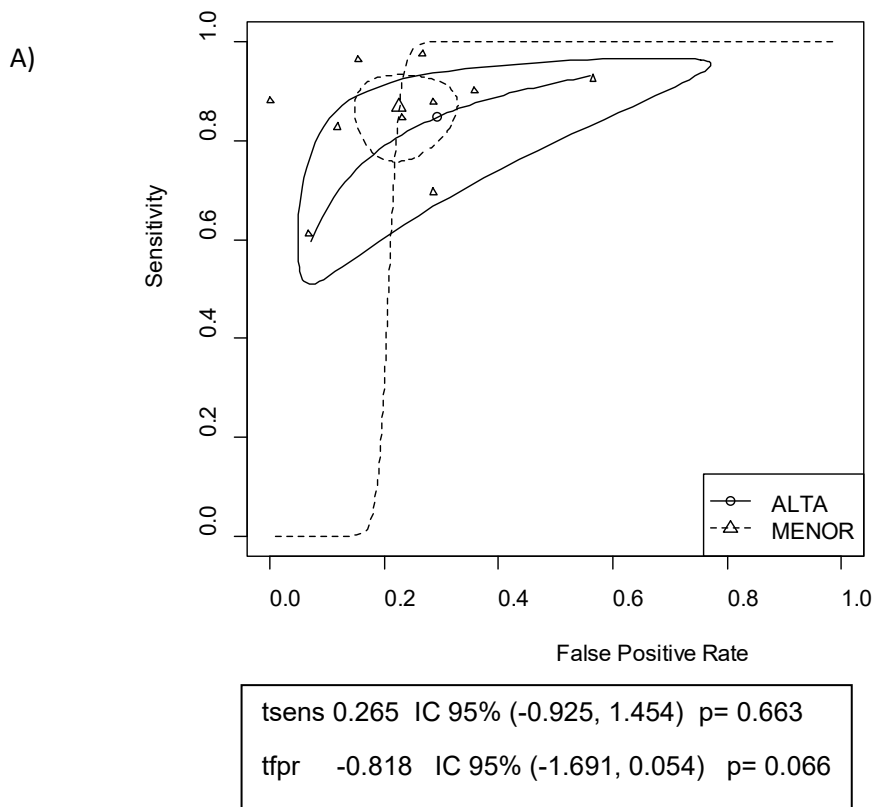


Figura 26. Análisis de subgrupos de calidad en estudios que miden FEDI (A) y ED (B). Comparación de curvas SROC entre los estudios considerados de alta calidad (círculos) y menor calidad (triángulos). Abreviaturas: tsens: sensibilidad transformada; tfpr: tasa de falsos positivos; IC 95%: intervalo de confianza al 95%.

Aplicabilidad de los estudios.

Según los aspectos desarrollados en el apartado 4.2.4.3, la aplicabilidad de algunos estudios puede verse comprometida por unos criterios de inclusión de pacientes que seleccionen subgrupos muy concretos de enfermos, lo cual afecta a la probabilidad pre-prueba y por lo tanto a la validez externa. Se han considerado estudios como “estudios con problemas de aplicabilidad” aquellos calificados de “alto riesgo” o “dudoso riesgo” en la evaluación QUADAS-2.

● **Aplicabilidad en estudios que miden FEDI.**

Se observa que en los estudios con problemas de aplicabilidad la S y E son mayores (figura 27), siendo estas diferencias significativas para la E pero no significativas para la S. Esto sugiere que la FEDI es un predictor de destete más específico en poblaciones concretas, como por ejemplo pacientes respiratorios, pero su exactitud en pacientes críticos generales es limitada.

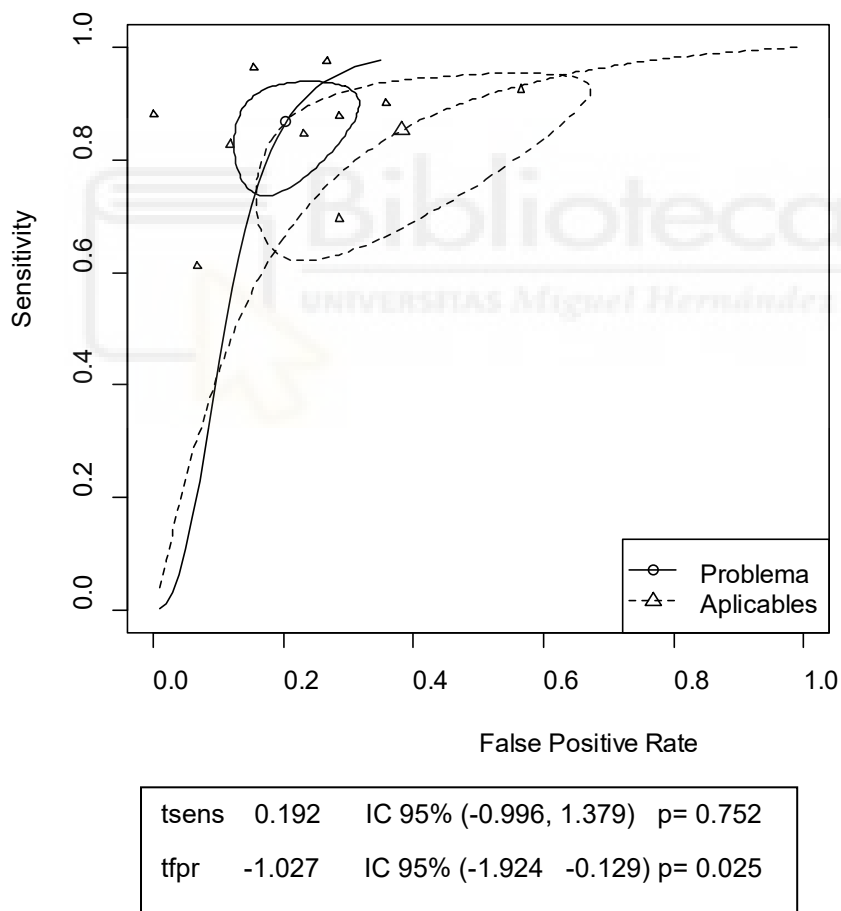


Figura 27. Análisis de subgrupos de aplicabilidad en estudios que miden FEDI. Comparación de curvas SROC y elipses de confianza entre los estudios considerados aplicables (triángulos) y aquellos que probablemente impliquen problemas en cuanto a la aplicabilidad (círculos). Abreviaturas: tsens: sensibilidad transformada; tfpr: tasa de falsos positivos; IC 95%: intervalo de confianza al 95%.

● **Aplicabilidad en estudios que miden ED.**

Las curvas SROC sugieren la S y E son mayores en los estudios con problemas de aplicabilidad (figura 28), siendo estas diferencias significativas para la S pero no significativas para la E. Esto indica que posiblemente la ED es un predictor de destete más sensible en poblaciones concretas de pacientes, pero su exactitud en pacientes críticos generales es limitada.

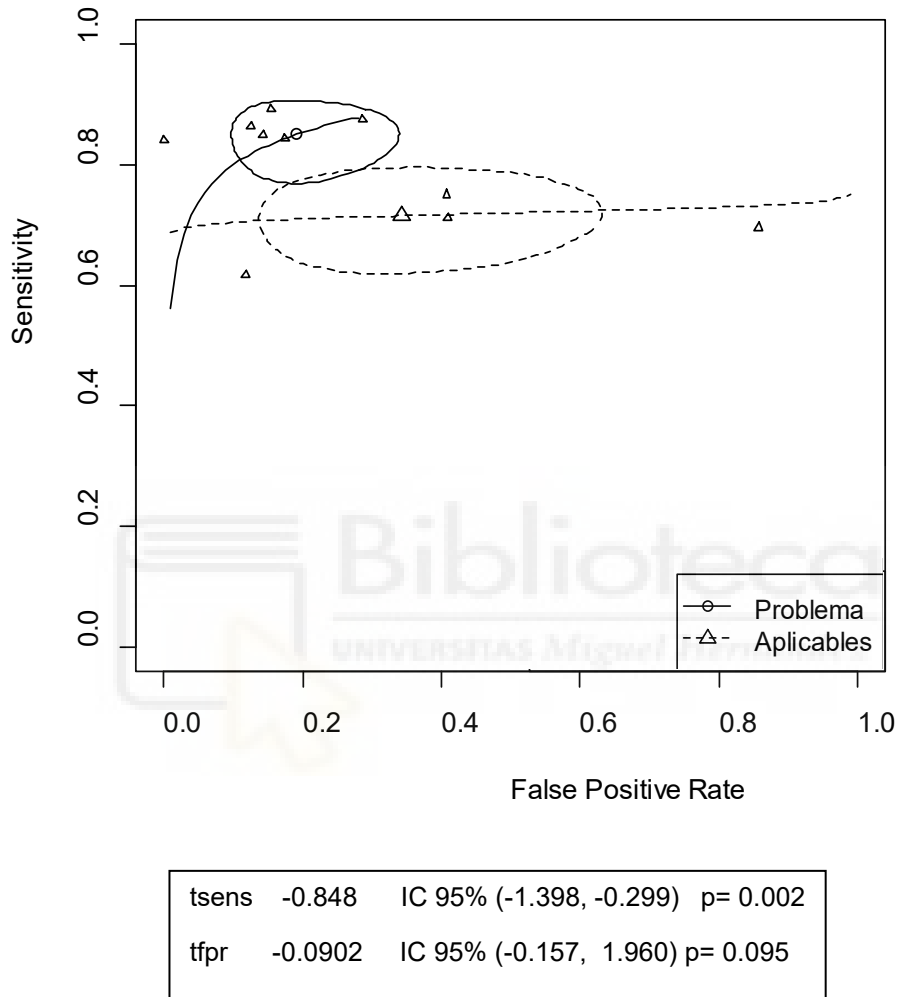


Figura 28. Análisis de subgrupos de aplicabilidad en estudios que miden ED. Comparación de curvas SROC y elipses de confianza entre los estudios considerados aplicables (triángulos) y aquellos que probablemente impliquen problemas en cuanto a la aplicabilidad (círculos). Abreviaturas: tsens: sensibilidad transformada; tfpr: tasa de falsos positivos; IC 95%: intervalo de confianza al 95%.

4.3. ECOGRAFÍA PULMONAR Y DIAFRAGMÁTICA COMO PREDICTORES DE ÉXITO EN LA RETIRADA DE LA VENTILACIÓN MECÁNICA.

4.3.1. RESUMEN

Antecedentes: La disfunción diafragmática inducida por la ventilación mecánica (VM) se asocia al fracaso en la extubación, con un aumento de los días en VM y de la mortalidad. Su evaluación se realiza con métodos invasivos, por lo que no se suele realizar en la práctica clínica habitual.

Objetivos: El objetivo de este estudio es valorar la exactitud y reproducibilidad de la ecografía pulmonar y diafragmática para la predicción de éxito del destete de la VM en pacientes críticos adultos.

Métodos: Llevamos a cabo un estudio de cohortes prospectivo en pacientes en VM candidatos en condiciones para iniciar el destete. Se evaluó el valor predictivo del Lung ultrasound score modificado (LUSm) y la fracción de engrosamiento diafragmático (FEDI) en el resultado de destete mediante el cálculo de las correspondientes curvas ROC. Además, se evaluó la concordancia interobservador para ambos índices ecográficos, mediante el coeficiente kappa ponderado para LUSm y el coeficiente de correlación intraclase (CCI) y el método de Bland-Altman para la FEDI.

Resultados: LUSm mostró un poder discriminativo moderado para predecir el éxito en el destete y la extubación, (Área bajo la curva (AUC)= 0.80 para éxito en el destete, sensibilidad y especificidad en el punto de corte óptimo 0.76 y 0.73 respectivamente; AUC para predecir el éxito en la extubación =0.78, sensibilidad óptima y especificidad de 0.76 y 0.47 respectivamente). La FEDI es más sensible pero menos específica que LUSm para predecir el éxito en el destete (AUC 0.71, sensibilidad y especificidad óptimas de 0.93 y 0.48, respectivamente) y el éxito en la extubación (AUC 0.76, sensibilidad y especificidad óptimas de 0.93 y 0.58). Se objetivó una adecuada concordancia interobservador tanto para LUSm (kappa ponderado 0,95) como para FEDI (ICC 0,78, diferencia según el análisis de Bland-Altman +/- 12,5%).

Conclusiones: Un valor bajo de FEDI y/o alto de LUSm indica elevado riesgo del fracaso de la retirada de la VM. La concordancia interobservador es excelente para el LUSm y moderada-buena para la FEDI.

4.3.2. INTRODUCCIÓN Y OBJETIVOS.

El uso generalizado de la VM en la unidad de cuidados intensivos (UCI) salva cientos de vidas diarias. Sin embargo, incluso cuando se usa correctamente, la VM prolongada puede provocar complicaciones como el desarrollo de neumonía asociada a ventilación mecánica (NAVVM) y disfunción diafragmática inducida por la ventilación mecánica (DDIVM)^{95,141,142}. Ello justifica que el destete deba considerarse tan pronto como sea posible. Por otro lado, la retirada prematura de la VM puede provocar el fracaso de la extubación, que se asocia con un aumento de la morbilidad y la mortalidad^{143,144}, por lo que identificar el momento óptimo para iniciar el destete es fundamental y varios índices se han desarrollado para este propósito. En un estudio de 500 pacientes¹⁴⁵, se evaluó el valor predictivo de varios índices predictores de resultado de destete (volumen minuto, frecuencia respiratoria (FR), presión parcial de oxígeno (PaO₂), índice de respiración superficial rápida (Rapid Shallow Breathing Index= RSBI); presión inspiratoria máxima (PIMax), presión espiratoria máxima; índice CROP (dynamic Compliance, Respiratory rate, Oxygenation, maximal inspiratory Pressure)) y se observó que ninguno de ellos era capaz de predecir con exactitud el resultado del destete.

En los últimos años se han desarrollado predictores basados en ecografía que evalúan la función del diafragma y de la aireación pulmonar con resultados alentadores, sin embargo existe de momento poca evidencia disponible sobre su exactitud y reproducibilidad. Bouhemad¹¹⁷ fue el primer autor en proponer un sistema de puntuación, el LUS score, para cuantificar evolutivamente la aireación pulmonar en pacientes con neumonía asociada a ventilación mecánica. En estudios posteriores^{108,130,136,137} el LUS score fue usado para predecir el resultado de destete con la obtención de resultados prometedores. A nivel diafragmático se han propuesto diversos parámetros ecográficos para evaluar el resultado del weaning, como el grosor del diafragma, movimiento o excursión del mismo durante el ciclo respiratorio y también engrosamiento o fracción de engrosamiento diafragmático (FEDI). De ellos, en un estudio reciente¹⁴⁶, únicamente la FEDI ha demostrado tener una exactitud similar al gold standard (la presión transdiafragmática o Pdi), para valorar la función diafragmática.

La evidencia actual sugiere que LUS score y FEDI pueden ser buenos predictores ecográficos del resultado del destete, sin embargo, existen pocos estudios al respecto, con un espectro muy variable de pacientes, una técnica ecográfica poco estandarizada y escasos datos sobre su reproducibilidad.

El objetivo del presente estudio es evaluar la exactitud y reproducibilidad de la ecografía pulmonar y diafragmática para predecir el éxito en el destete en pacientes críticos adultos sometidos a ventilación mecánica.

4.3.3. MATERIAL Y MÉTODOS.

4.3.3.1. Diseño.

Se realizó un estudio de cohortes prospectivo para evaluar la exactitud de la ecografía pulmonar y diafragmática en la predicción del éxito del destete y otro de concordancia entre dos observadores de ecografía pulmonar y diafragmática. En ambos estudios se consideraron los predictores Lung Ultrasound Score modificado (LUSm) y Fracción de Engrosamiento diafragmático (FEDI).

4.3.3.2. Población.

Para el estudio de exactitud pronóstica, se incluyeron consecutivamente a todos los pacientes ingresados en la UCI del Hospital General Universitario de Elche del 15 de enero de 2016 al 15 de abril de 2017 que cumplieran los siguientes criterios de inclusión: 1) mayores de 18 años; 2) más de 24 horas en MV; 3) listo para el destete según el criterio de su médico responsable; 4) firma del consentimiento informado.

Para el estudio de concordancia interobservador, se incluyeron 50 pacientes (con o sin VM) que ingresaron consecutivamente en la UCI del mismo hospital de diciembre de 2016 a febrero de 2017 y que requirieron un examen de ultrasonido torácico por razones clínicas.

Se aplicaron los mismos criterios de exclusión para ambos estudios: 1) lesión de la médula espinal por encima de T8; 2) arritmias e inestabilidad hemodinámica; 3) extubación terminal; 4) embarazo; 5) neumotórax, neumomediastino, toracostomía, tubo de tórax o lesiones en el tórax que impiden la adquisición de una correcta ventana ecográfica; 6) lesiones pleurales o pleurodesis.

4.3.3.3. Mediciones.

Dos médicos entrenados en ecografía pulmonar y diafragmática realizaron las mediciones. Se usó una sonda convexa de 2-4 MHz en el modo B y un ecógrafo Micromax® Sonosite.

Se exploraron las regiones pulmonares según los puntos recomendados en el protocolo BLUE⁸: zonas antero-superior, antero-inferior, lateral y postero-basal. En lugar de utilizar el LUS score original, se aplicó un procedimiento modificado (LUSm), evaluando cuatro regiones pulmonares en cada lado en lugar de las seis estándar, evitando con ello la movilización del paciente y sus potenciales complicaciones.

En cada región pulmonar evaluada pueden coexistir varios patrones de aireación del parénquima, por lo que por consenso se asignó la puntuación correspondiente al patrón más consolidativo detectado, según se detalla a continuación:

- Patrón normal (0 puntos): presencia de deslizamiento pleural o “lung-sliding” con líneas A, ocasionalmente alguna línea B aislada (2 o menos).
- Pérdida moderada del patrón aéreo (1 punto): 3 o más líneas B bien definidas e irregularmente separadas (cometas) que salen de la línea pleural o de una consolidación subpleural.
- Pérdida severa del patrón aéreo (2 puntos): múltiples líneas B confluentes que salen de la línea pleural o de una consolidación subpleural.
- Patrón de consolidación pulmonar (3 puntos): patrón tisular con imágenes puntiformes hiperecoicas que representan el broncograma aéreo (puede tratarse de neumonía o atelectasia).

Se sumaron todas las puntuaciones de ambos pulmones, obteniendo una puntuación total que oscilaría entre 0 a 24 puntos. Las puntuaciones bajas en el LUSm se relacionan con un patrón de aireación normal pulmonar y puntuaciones altas con un patrón alterado.

Para la ecografía diafragmática, se midió el grosor diafragmático según las recomendaciones publicadas^{140,147}. Ambos operadores midieron el grosor diafragmático usando un ecógrafo Micromax® Sonosite con una sonda lineal de 7-10 MHz. El hemidiafragma derecho se visualizó en la zona de aposición, en la línea axilar media entre los espacios intercostales octavo y décimo, con el paciente en posición semiincorporada (20º-40º). Se capturaron al menos tres imágenes durante la respiración espontánea del paciente, midiendo el grosor del diafragma al final de la inspiración y al final de la espiración. Se hizo la media de tres mediciones de FEDI, que se calculó según la siguiente fórmula: $[(\text{grosor diafragmático al final de la inspiración} - \text{grosor del diafragma al final de la espiración}) / \text{grosor del diafragma al final de la espiración}] \times 100$.

4.3.3.4. Procedimiento.

Los pacientes que cumplían las condiciones para iniciar el destete según los criterios de la última guía internacional de consenso⁵¹, fueron sometidos a una ventilación en presión de soporte de 8 cm H₂O y presión positiva al final de la espiración (PEEP) de 5 cm H₂O durante la realización de las ecografías pulmonar y diafragmática. Una vez finalizadas las exploraciones ecográficas, el paciente fue reconectado en el modo ventilatorio controlado que tuviese previamente, o directamente se inició la prueba de respiración espontánea (PRE) con presión de soporte ventilatorio o tubo en T, según la decisión de su médico responsable (Figura 29). Se recomendó realizar la PRE y la extubación en un periodo no superior a 4 horas tras realizar las ecografías, ya que periodos de tiempo más prolongados podrían suponer cambios en las condiciones clínicas del paciente y también del patrón ecográfico. El médico responsable juzgó qué pacientes pasaron con éxito la PRE, y los que lo hicieron fueron extubados. El equipo médico estaba cegado a los resultados de la ecografía, y el equipo de investigación no desempeñó ningún papel asistencial en el destete del paciente.

El fracaso de weaning se definió según los criterios de la conferencia de consenso internacional de 2007⁵¹ como el fracaso de la PRE (Tabla 6) o fracaso de la extubación, el cual viene condicionado por la necesidad de reintubación, soporte ventilatorio no invasivo o muerte durante las 48 horas posteriores a la extubación.

Para el estudio de concordancia interobservador, ambos operadores realizaron las mediciones ecográficas sobre la misma muestra de 50 pacientes de forma ciega, es decir, desconociendo los resultados obtenidos por el otro. En un mismo paciente, el tiempo transcurrido entre la exploración de ambos ecografistas fue inferior a 5 horas.

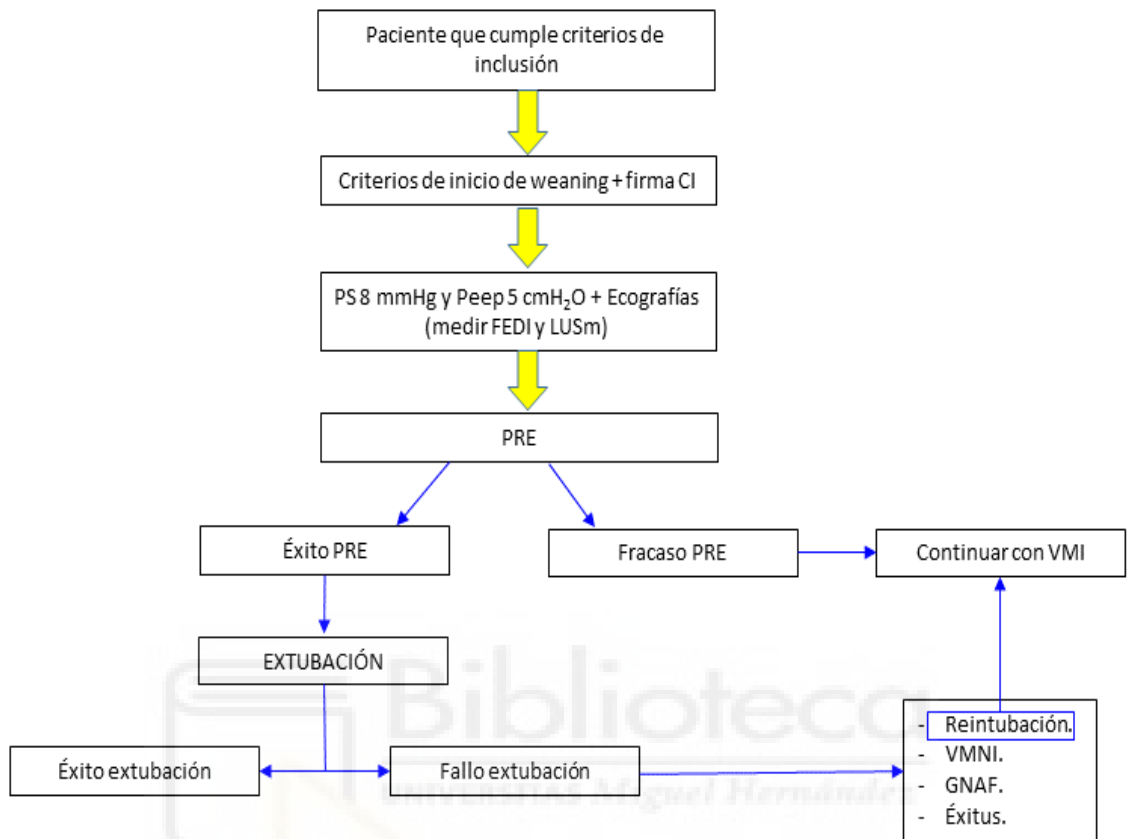


Figura 29. Esquema del proceso de evaluación de los pacientes incluidos en el estudio. Abreviaturas: CI: consentimiento informado; PS: presión de soporte; Peep: Presión positiva al final de la espiración (*positive end expiratory pressure*); FEDI: fracción de engrosamiento diafragmático; LUSm: *Lung ultrasound score* modificado; PRE: prueba de respiración espontánea; VMI: ventilación mecánica invsiva; VMNI: ventilación mecánica no invasiva; GNAF: gafas nasales de alto flujo.

CRITERIOS DE FRACASO DE LA PRUEBA DE RESPIRACIÓN ESPONTÁNEA.
Índices subjetivos
Causas neurológicas: Agitación o ansiedad. Bajo nivel de conciencia. Incremento del trabajo respiratorio: uso de musculatura accesoria, expresión facial de estrés, disnea, diaforesis
Índices objetivos
Hipoxemia: PaO ₂ < 60 mmHg o SatO ₂ < 90% con FiO ₂ ≥ 0.5 o SatO ₂ < 90%. Hipercapnia: PaCO ₂ > 50 mmHg o un incremento en la PaCO ₂ > 8 mmHg pH < 7.32 o un descenso en el pH mayor o igual a 0.07 Taquipnea: FR > 35 rpm Inestabilidad hemodinámica: FC > 140 lpm o > 20% respecto a la basal; TAS > 180 mmHg o > 20% respecto a la basal; TAS < 90 mmHg; arritmias cardíacas.

Tabla 6. Criterios de fracaso de la prueba de respiración espontánea. Abreviaturas: PaO₂: presión arterial de oxígeno; mmHg: milímetros de mercurio; PaCO₂: presión arterial de dióxido de carbono; FR: frecuencia respiratoria; FC: frecuencia cardíaca; TAS: tensión arterial sistólica.

4.3.3.5. Análisis estadístico

Los datos se expresaron como media y desviación standard o mediana y rango intercuartílicos, según su distribución era o no normal. Para el contraste de hipótesis de variables cuantitativas se utilizaron los tests T-Student, U de Mann-Whitney y Wilcoxon. Para la comparación de proporciones se empleó el test Chi-cuadrado o el test exacto de Fisher. Todos los contrastes fueron bilaterales, con un nivel de significación α del 5%.

Para evaluar la exactitud pronóstica de LUSm y FEDI se calculó el área bajo la curva ROC y el punto de corte óptimo, para el cual se determinó la sensibilidad, especificidad y cocientes de probabilidad de cada uno de los predictores ecográficos (FEDI y LUSm) con respecto a la variable desenlace éxito de weaning.

Para evaluar la concordancia interobservador del LUSm se calculó el Coeficiente Kappa ponderado por pesos cuadráticos, el cual es comparable al Coeficiente de Correlación Intraclass, (CCI) y para la FEDI se usó el CCI y también el método de Bland-Altman¹⁴⁸ para examinar la magnitud y distribución de las diferencias entre observadores y su intervalo de confianza al 95%. Los análisis se realizaron usando el paquete estadístico StatsDirect v3.0.194¹⁴⁹.

4.3.3.6. Aspectos éticos

La Comisión de Investigación y Ética del Hospital General Universitario de Elche aprobó la realización de este estudio. Todos los pacientes incluidos o sus familiares responsables firmaron el consentimiento informado.

4.3.4. RESULTADOS.

4.3.4.1. Exactitud para predecir el éxito en el weaning.

Durante el período de estudio, 139 pacientes se sometieron a MV, de los cuales 52 no cumplieron los criterios de inclusión (48 fallecimientos antes del inicio del destete, 2 autoextubaciones, 2 en VM durante menos de 24 horas) y 18 no fueron incluidos por motivos ajenos al control del equipo investigador (en 8 pacientes el médico responsable decidió realizar una reducción gradual del soporte ventilatorio en lugar de una PRE, 2 pacientes no firmaron el consentimiento informado, 2 pacientes fueron trasladados a otro hospital, en el caso de 5 pacientes elegibles el equipo médico de guardia no avisó a los investigadores, en un paciente no se pudieron realizar las ecografías por avería del ecógrafo). Finalmente se incluyeron 69 pacientes, cuyas características basales se describen en la Tabla 7.

Características basales de los pacientes (n=69)	
Edad años P50 (P25, P75)	66 (53, 78)
Sexo masculino (%)	43 (62,3)
Comorbilidad al ingreso (nº pacientes y % del total)	
<ul style="list-style-type: none"> • Cardiopatía 18 (26.1%) • Patología neurológica 18 (26.1%) • EPOC 13 (18.8%) • Diabetes mellitus 19 (27.5%) • Cáncer 8 (11.6%) • Insuficiencia renal crónica 12 (17.4%) • Hepatopatía 8 (11.6%) 	
Causas ingreso UCI (nº pacientes y % del total)	
<ul style="list-style-type: none"> • Patología neurológica 28 (40.6%) • Patología respiratoria 18 (26%) • Patología cardíaca 11 (15.9%) • Sepsis no respiratoria 7 (10.1%) • Patología digestiva 4 (5.6%) • Politraumatizado 1 (1.4%) 	
Escalas de gravedad	
<ul style="list-style-type: none"> • SAPS 3 al ingreso puntos P50 (P25, P75): 63 (48, 73) • APACHE II al ingreso puntos P50 (P25, P75): 13 (9, 17) • APACHE II durante la PRE puntos P50 (P25, P75): 4 (2, 6) 	

Tabla 7. Características basales de los pacientes incluidos en el estudio (n=69).

Se utilizó soporte de presión en el 49% de las PRE, el tubo en T en el 42% y ambos métodos en el 9%. 8 pacientes fracasaron en la PRE y 61 pacientes la superaron con éxito y por lo tanto fueron extubados. 17 pacientes fracasaron en la extubación (de ellos 9 precisaron ventilación mecánica no invasiva, 3 de ellos cánula nasal de alto flujo y en 5 pacientes fue precisa la reintubación) lo cual implica que un total de 25 pacientes fracasaron en el destete (Figura 30).

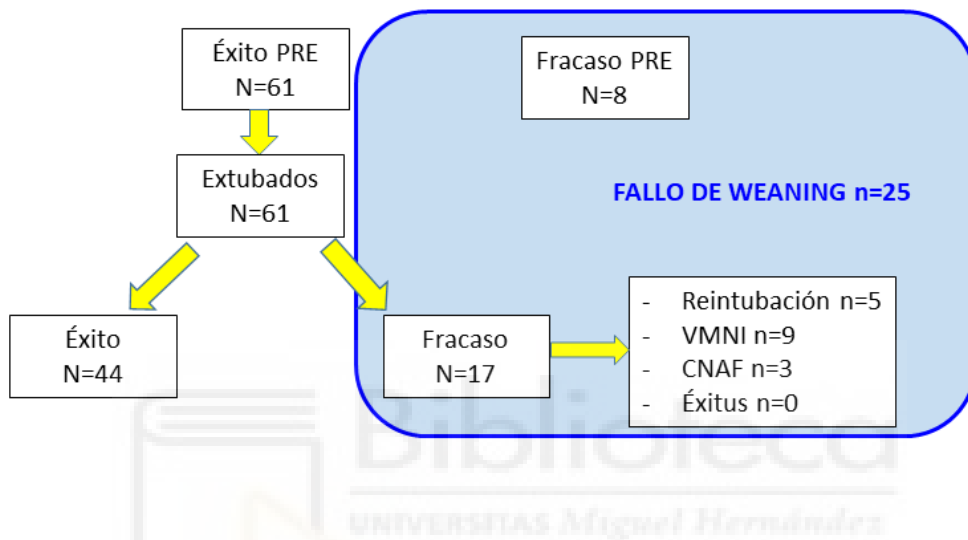


Figura 30. Resultado del destete y flujo de pacientes. Abreviaturas: PRE: prueba de respiración espontánea; VMNI: ventilación mecánica no invasiva; CNAF: cánula nasal de alto flujo.

Al comparar el grupo de pacientes de éxito y fracaso en el weaning (Tabla 8), se aprecia que éste último se asocia a mayor tiempo en VM, más casos de enfermedad pulmonar obstructiva crónica (EPOC), mayor puntuación de LUSm, menor FEDI y menor saturación periférica de oxígeno (SpO2). La mediana de diferencia en los predictores LUSm y FEDI en ambos grupos fue de 3 puntos y 11% respectivamente.

Variables	Éxito weaning (n=44)	Fallo weaning (n=25)	p
Sexo: Varones (%)	26 (63,4)	15 (62,5)	0,8
Edad años P50 (P25, P75)	65 (53, 78)	69 (64, 78)	0,37
Duración VM días P50 (P25, P75)	4 (2, 6)	5 (3, 9)	0,04
Nº pacientes con más 6 días VM (%)	9 (20,4)	10 (40)	0,09
Comorbilidad: nº pacientes (%) *			
Cardiopatía	14 (31,8)	4 (16)	0,25
Neurológicos	12 (27,3)	6 (24)	0,99
EPOC	5 (11,4)	8 (32)	0,05
Diabetes mellitus	12 (27,3)	7 (28)	0,99
Cáncer	4 (9,1)	4 (16)	0,45
Insuficiencia renal	6 (13,6)	6 (24)	0,33
Hepatopatía	3 (6,8)	5 (20)	0,36
Patología ingreso: nº pacientes (%)			
Respiratoria	16 (36,4)	8 (32)	0,7
Neurológica	18 (41)	11 (44)	
LUSm P50 (P25, P75)	5 (3, 7)	8 (7, 11)	0,0001
FEDI % P50 (P25, P75)	38 (31, 45)	27 (20, 40)	0,003
APACHEII-ingreso P50 (P25, P75)	12,5 (9, 17)	14 (11, 15)	0,85
APACHEII-PRE P50 (P25, P75)	4 (2, 6)	5 (3, 9)	0,07
SAPS 3 P50 (P25, P75)	61 (48, 69)	69 (54, 80)	0,09
VE L/min P50 (P25, P75)	8,25 (7,3, 9,8)	9 (8,1, 11,6)	0,13
Compliance ml/cm H2O P50 (P25, P75)	55,5 (43, 69)	59 (40,5, 67)	0,86
PIMax cm H2O P50 (P25, P75)	-25 (-25, -18)	-25 (-26, -24)	0,28
P0.1 cm H2O P50 (P25, P75)	1 (1, 3)	1,5 (1, 2,5)	0,47
RSBI resp/min/L P50 (P25, P75)	31 (20, 43)	37 (30, 54)	0,16
Frec Resp rpm P50 (P25, P75)	17 (14, 19)	19 (16, 22)	0,09
V tidal ml P50 (P25, P75)	508 (452, 572)	500 (440, 600)	0,71
FiO2 % P50 (P25, P75)	30 (28, 35)	30 (28, 35)	0,83
SpO2 P50 (P25, P75)	99 (97, 100)	97 (96, 99)	0,027
PaCO2 mm de Hg P50 (P25, P75)	41 (36, 45,6)	39,6 (37, 45)	0,63
PaO2 mm de Hg P50 (P25, P75)	96,5 (83, 117)	92(74, 115)	0,61
pH P50 (P25, P75)	7,4 (7,3, 7,5)	7,4 (7,4, 7,4)	0,93
Bicarbonato mmol/L P50 (P25, P75)	25,6 (23, 30,6)	26 (23, 28,7)	0,73
Lactato mmol/L P50 (P25, P75)	1,2 (1, 1,5)	1,3 (1,1, 2)	0,2
Mortalidad UCI nº pacientes (%)	1 (2,3)	6 (24)	0,005
Mortalidad hospital nº pacientes (%)	3 (6,8)	8 (32)	0,003
Estancia UCI días P50 (P25, P75)	8 (6, 15)	18 (11, 21)	0,002
Estancia Hospital días P50 (P25, P75)	17 (13,5, 31)	23 (17, 30)	0,08

Tabla 8. Comparativa de los grupos de éxito y fracaso en el destete. Las variables en las que las diferencias alcanzan la significación estadística ($p < 0.05$) se señalan en amarillo. Ver abreviaturas en el listado de la página 7. *Las categorías de comorbilidad no son excluyentes.

En el caso de LUSm como predictor de éxito de weaning se obtuvo un área bajo la curva ROC de 0,80 (IC 95%: 0,69 -0,91) (Figura 31). El punto de corte óptimo fue de 7 puntos, con una Sensibilidad de 0,76 (IC 95%: 0,55- 0,90), Especificidad de 0,73 (IC 95%: 0,57-0,85), Valor Predictivo Positivo de 45,5%, Valor Predictivo Negativo de 84,2%, Cociente de Probabilidad Positivo de 2,8 (IC95%: 1,7-4,8), Cociente de Probabilidad Negativo de 0,44 (IC95%: 0,15-0,62) y Odds Ratio diagnóstica de 8,44.

Para FEDI como predictor de éxito de weaning se obtuvo un área bajo la curva ROC de 0,71 (IC 95%: 0,58-0,84) (Figura 31). El punto de corte óptimo fue de 0,24, con una Sensibilidad de 0,93 (IC 95%: 0,81-0,98), Especificidad de 0,48 (IC 95%: 0,28-0,68), Valor Predictivo Positivo de 75,9%, Valor Predictivo Negativo de 80%, Cociente de Probabilidad Positivo de 1,79 (IC 95%: 1,3-2,8), Cociente de Probabilidad Negativo de 0,14 (IC95%: 0,05-0,42) y Odds Ratio diagnóstica de 12,6.

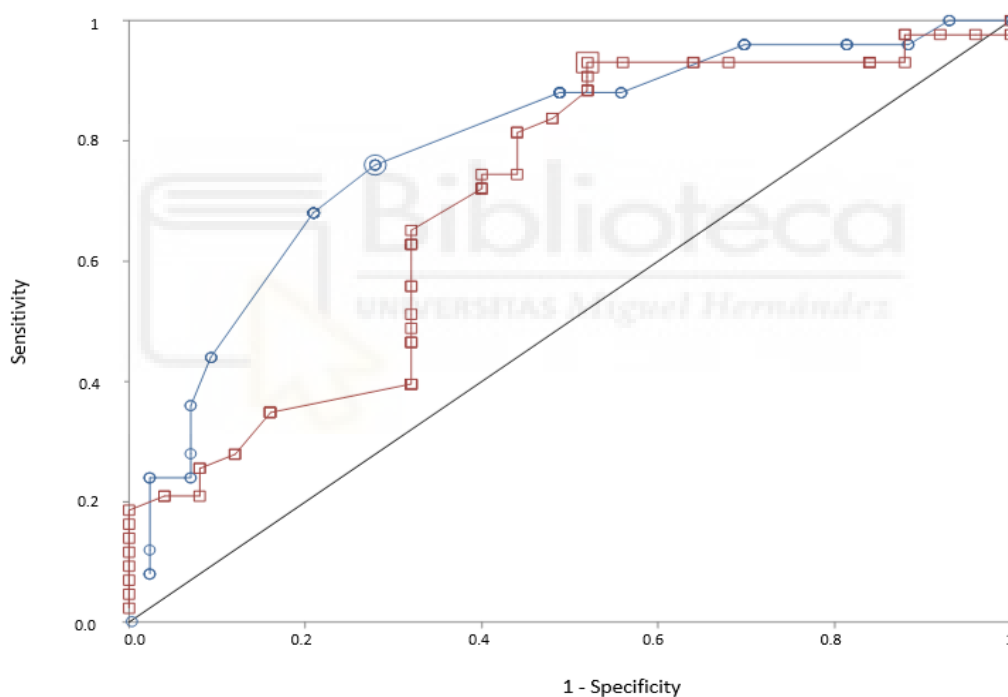


Figura 31. Curvas ROC de exactitud de los predictores FEDI (rojo) y LUSm (azul) para predecir el éxito en el destete. Se resaltan los puntos de corte óptimos para cada predictor.

4.3.4.2. Concordancia interobservador

El Coeficiente Kappa ponderado por pesos cuadráticos para LUSm fue de 0.95 (95% IC: 0.92, 0.98) y el CCI para FEDI fue de 0.78 (95% IC: 0.65, 0.87) y la diferencia en las medidas según el método de Bland-Altman fue de +/-12.5%.







5. DISCUSIÓN.

5.1. RESUMEN DE RESULTADOS.

5.1.1. Ecografía pulmonar para el diagnóstico de neumonía en adultos: revisión sistemática y meta-análisis.

Se seleccionaron 16 estudios, la mayoría de cohortes, en los que el gold estándar de comparación más frecuente es el diagnóstico final basado en datos clínicos, de laboratorio, microbiológicos, RxT y en algunos casos también ecocardiografía y/o TC. En cuanto a la calidad metodológica de dichos estudios, el principal problema encontrado ha sido una aplicación desigual de la prueba de referencia, ya que en algunas ocasiones no se ha realizado la misma a todos los sujetos incluidos en el estudio (sesgo de verificación diferencial).

Dada la presencia de indicios de efecto umbral, se optó por evitar la ponderación de la sensibilidad y especificidad. La estimación global del área bajo la curva SROC de la ecografía pulmonar para el diagnóstico de neumonía fue de 0.93 y la ORD de 50 con un intervalo de confianza (IC) del 95% de (21, 120).

5.1.2. Ecografía diafragmática y pulmonar para predecir el resultado del destete: revisión sistemática y meta-análisis.

Se incluyeron 19 estudios, todos ellos de cohortes, refiriéndose 10 a la ED, 10 a la FEDI y 5 a LUS score (en algunos trabajos se mide más de un predictor). En lo que respecta a la calidad metodológica de dichos estudios, el principal problema detectado es la presencia de aspectos que pueden comprometer la aplicabilidad en lo que respecta a la selección de los pacientes, pues algunos estudios se centran en poblaciones muy concretas, por ejemplo, en pacientes EPOC o con miopatía del paciente crítico.

Para los distintos predictores los resultados fueron:

- FEDI: Rho de Spearman= 0.3, IC 95% (-0.4, 0.7). El test Q fue de 9.5 ($p=0.386$) y se obtuvo una Odds ratio diagnóstica (ORD) de 21 con IC 95% (11, 40). Debido a la presencia de indicios de efecto umbral se optó por el cálculo del AUC, que fue de 0.87.
- ED: Rho de Spearman= -0.45, IC 95% (-0.84, 0.25). La ORD calculada es de 10.6, con IC 95% (5, 24). El estadístico Q fue de 10.7 ($p=0.29$), lo cual apunta a la posibilidad de una heterogeneidad ligera-moderada entre estudios. Dado que no se apreciaron indicios de efecto umbral al representar los estudios en el plano ROC, se ponderó la sensibilidad de los distintos estudios, cuyo valor promedio es del 75%, IC 95% (70, 85). Las diferencias entre estudios en lo concerniente a especificidad son importantes y por lo tanto no sería adecuado hacer una estimación ponderada.
- LUS score: Rho de Spearman= 0.2, IC 95% (-0.8, 0.9). El test Q fue de 5.1 ($p=0.272$) y se obtuvo una ORD de 38 con IC 95% (7, 198). Se estimó un AUC =0.77.

5.1.3. Ecografía pulmonar y diafragmática como predictores de éxito en la retirada de la ventilación mecánica.

En el caso de LUSm como predictor de éxito de weaning se obtuvo un área bajo la curva ROC de 0,80 (IC 95%: 0,69 -0,91). El punto de corte óptimo fue de 7 puntos, con una Sensibilidad de 0,76 (IC 95%: 0,55- 0,90), Especificidad de 0,73 (IC 95%: 0,57-0,85), Valor Predictivo Positivo de 45,5%, Valor Predictivo Negativo de 84,2%, Cociente de Probabilidad Positivo de 2,8 (IC95%: 1,7-4,8), Cociente de Probabilidad Negativo de 0,44 (IC95%: 0,15-0,62) y Odds Ratio diagnóstica de 8,44.

Para FEDI como predictor de éxito de weaning se obtuvo un área bajo la curva ROC de 0,71 (IC 95%: 0,58-0,84). El punto de corte óptimo fue de 0,24, con una Sensibilidad de 0,93 (IC 95%: 0,81-0,98), Especificidad de 0,48 (IC 95%: 0,28-0,68), Valor Predictivo Positivo de 75,9%, Valor Predictivo Negativo de 80%, Cociente de Probabilidad Positivo de 1,79 (IC 95%: 1,3-2,8), Cociente de Probabilidad Negativo de 0,14 (IC95%: 0,05-0,42) y Odds Ratio diagnóstica de 12,6.

El Coeficiente Kappa ponderado por pesos cuadráticos para LUSm fue de 0.95 (95% IC: 0.92, 0.98) y el CCI para FEDI fue de 0.78 (95% IC: 0.65, 0.87) y la diferencia en las medidas según el método de Bland-Altman fue de +/-12.5%.

5.2. VALIDEZ.

Los resultados de nuestro trabajo sugieren una excelente exactitud diagnóstica para discriminar la presencia de neumonía en pacientes adultos. Aunque la mayoría de los estudios tienen problemas de riesgo de sesgo, la excelente exactitud de la ecografía pulmonar no parece ser atribuible a la mala calidad de los estudios, ya que al estratificar por las distintas dimensiones de QUADAS-2, la exactitud es numéricamente mejor en los estudios de alta calidad. Revisiones previas señalan en el mismo sentido en cuanto a un perfil de exactitud diagnóstica conveniente aunque con datos aún más favorables. Chávez et al⁶⁷ obtuvieron una sensibilidad ponderada del 94% y especificidad ponderada del 96%, con un área bajo la curva SROC de 0.98. Las principales diferencias con nuestro trabajo son metodológicas en cuanto al análisis, aunque además se incluyeron estudios con pacientes menores de 18 años, así como un menor número de estudios (n=10) de los cuales el 70% implicaría un alto riesgo de sesgo por gold standard y flujo de pacientes. También se incluyó al menos un estudio en el que la unidad de análisis es el hemitórax¹⁵⁰. Hu et al¹⁵¹ obtuvieron una ORD de 509.99 (IC 95% 128.11, 2014.34) y área bajo la curva SROC de 0.99, aunque siete de los nueve estudios sometidos a análisis incluyeron niños e incluso neonatos, por lo que las muestras no son comparables. De hecho, diversos estudios pediátricos sugieren un mejor rendimiento diagnóstico de la ecografía torácica respecto a los adultos, lo cual posiblemente tenga que ver con que los niños tienen una pared torácica más fina y menor volumen de parénquima pulmonar. En la revisión de Ye et al¹⁵² la ORD fue de 151.19 (IC 38.50, 593.77) y el área bajo la curva SROC de 0.97, con el hándicap de un número de artículos muy reducido (n=5).

La FEDI se considera un buen indicador del esfuerzo inspiratorio diafragmático en pacientes con respiración espontánea¹⁴⁷ y valores bajos se asocian a un aumento en la duración de la VM, estancia en UCI y mortalidad¹⁴⁶. Nuestros datos sugieren que es además un buen predictor del resultado del destete, apreciándose en general consistencia entre los distintos estudios. En este sentido nuestro estudio con datos primarios en la UCI del Hospital General de Elche¹³⁹ se comporta como un "outlier", pues tiene una especificidad ostensiblemente menor al resto.

Analizando en detalle las posibles causas de este resultado, se aprecia que en la mayoría de estudios la ecografía se realiza durante la PRE, mientras que en nuestro trabajo la exploración ecográfica se lleva a cabo antes de someter al paciente a la PRE, durante un breve periodo de desconexión, lo cual sería compatible con la hipótesis de que la FEDI puede variar con la fatiga diafragmática conforme avanza la PRE, es decir, que un test demasiado precoz supondría una baja sensibilidad para predecir el fracaso en el destete, o lo que es lo mismo, una baja especificidad para predecir el éxito en el mismo. No existe de momento evidencia disponible que apoye esta posibilidad, sin embargo actualmente se está llevando a cabo un estudio que podría arrojar luz sobre esta cuestión¹⁵³. Otro punto que puede apoyar la baja especificidad en nuestro estudio es la elección de un punto de corte de FEDI relativamente bajo, es decir, muy sensible.

La ED se relaciona con el volumen inspiratorio¹⁵⁴ y recientemente se ha observado una correlación moderada con el cambio en la presión endotraqueal tras la estimulación del nervio frénico en pacientes con ventilación asistida en PS y ausencia de correlación con otros índices de esfuerzo respiratorio¹⁴⁷. Los datos de nuestra revisión sugieren una buena exactitud de la ED en términos generales para la predicción del éxito en el destete, aunque inferior a la FEDI y aparentemente con mayor heterogeneidad entre estudios. En este caso también encontramos un “outlier”¹²³ con una Especificidad manifiestamente inferior al resto de estudios, en lo cual podría haber influido el bajo número de eventos en el grupo de fracaso en el destete.

Además de la funcionalidad del diafragma, otras causas que pueden afectar al proceso de destete implican alteraciones en la aireación del parénquima pulmonar (edema pulmonar, exceso de secreciones, etc)¹⁵⁵, las cuales pueden ser evaluadas por LUS score. En nuestro estudio con datos primarios, obtuvimos un punto de corte óptimo inferior al resto de estudios de la revisión sistemática (7 puntos versus 10-14 puntos), lo cual puede explicarse porque con LUSm se evaluaron un total de ocho áreas pulmonares mientras que los otros estudios evaluaron 12 (LUS convencional). No obstante consideramos que esto no introduce realmente un sesgo si es tenido en cuenta, pues se evaluaron igualmente todas las áreas afectadas con más frecuencia en pacientes críticos⁸, al tiempo que simplificamos la técnica para evitar la movilización del paciente y sus posibles complicaciones. Por lo tanto, consideramos que LUSm podría constituir una nueva propuesta ventajosa tanto para el paciente como para el operador, lo cual sería conveniente estudiar con más detalle en estudios comparativos de ambos scores. Los datos aportados por nuestra revisión sistemática y meta-análisis sugieren una excelente exactitud del LUS score para predecir el resultado en el destete, no obstante esta información debe ser interpretada con cautela, pues los estudios disponibles hasta la fecha son escasos y existe de momento gran incertidumbre al respecto.

Hasta donde conocemos, la presente revisión sistemática y metaanálisis sobre ecografía pulmonar para el diagnóstico de neumonía en adultos, con 16 estudios incluidos y 2359 pacientes fue el más amplio ejecutado hasta su fecha de publicación (febrero de 2017). En octubre de 2018, Orso et al¹⁵⁶ publicaron una revisión sistemática y meta-análisis sobre el mismo tema incluyendo 17 estudios, aunque con importantes diferencias metodológicas respecto al nuestro ya que, por ejemplo, consideraron en el análisis un estudio en el que el test a evaluar está compuesto por datos ecocardiográficos, ecografía pulmonar y colapsabilidad de la vena cava inferior¹⁵⁷. En cuanto a ecografía diafragmática, la expuesta es la primera revisión sistemática en incluir un meta-análisis que valore la exactitud de la ecografía torácica no cardiológica en la predicción del resultado de la VM en pacientes adultos, aportando además

datos novedosos derivados del análisis de la heterogeneidad como la distinta capacidad predictiva de los ultrasonidos según el espectro de pacientes sobre el que se aplique.

Entre las fortalezas de las dos revisiones sistemáticas presentadas en este proyecto destaca su exhaustividad en la no restricción por idioma y el esfuerzo invertido en las traducciones, lo cual nos ha permitido tener acceso a datos publicados en idiomas “poco habituales” en investigación, como el chino, polaco y ruso, que generalmente son más proclives a publicar datos no tanto con un efecto negativo sino con una menor magnitud del mismo¹⁵⁸. Tampoco se han usado filtros temporales ni de edad, habiendo realizado un cribado manual de los estudios. Se ha invertido además un importante esfuerzo en el contacto con diversos autores de los estudios revisados y en la consecución de datos no publicados

En cuanto a las limitaciones de la revisión sobre neumonía, el principal escollo radica en que incluye un número relativamente bajo de estudios y existe entre ellos elevada heterogeneidad clínica. En vistas a abordar esta última cuestión, Volpicelli et al¹⁵⁹ publicaron en 2012 un documento elaborado por los expertos con mayor índice de publicación en la materia en el que se proporcionan una serie de recomendaciones basadas en la evidencia sobre la implementación, desarrollo y estandarización de la ecografía pulmonar en el ámbito de las Urgencias, planta de hospitalización y unidades de críticos.

La principal limitación de la revisión sobre ecografía diafragmática en el destete es el escaso número de estudios disponibles relativos a cada predictor ecográfico, lo cual se justifica por tratarse de aplicaciones incipientes y en pleno desarrollo de los ultrasonidos en el paciente crítico. Esto ha de tenerse en cuenta a la hora de interpretar el análisis de subgrupos, pues éstos son muy imprecisos aunque orientativos. No debemos perder de vista que la afectación diafragmática y el desrecultamiento pulmonar son dos de las múltiples causas potenciales de fracaso del destete y resulta imprescindible contextualizar la información obtenida con datos clínicos, de laboratorio y derivados de otras técnicas de imagen como la ecocardiografía¹⁶⁰.

5.3. APLICABILIDAD.

En contraposición a otros métodos para valorar el parénquima pulmonar y la función diafragmática, la ecografía nos aporta información dinámica y además resulta sencilla, se realiza a pie de cama y suele estar disponible en la mayoría de unidades de críticos. Además resulta inocua para el paciente y el personal sanitario, pues se evita cualquier exposición a radiaciones ionizantes y tiene un coste económico relativamente menor. En el caso de la ecografía pulmonar se evitarían los traslados a la sala de Radiodiagnóstico, durante los cuales la probabilidad de eventos adversos es considerable, así como la posibilidad de observar a tiempo real los cambios dinámicos en el patrón de aireación durante la PRE. No obstante la ecografía también presenta inconvenientes, como la dificultad de obtener una adecuada ventana acústica, principalmente en pacientes obesos o en casos de interposición aérea por ejemplo si hay enfisema subcutáneo, así como en casos de problemas de accesibilidad a las áreas de exploración por la presencia de vendajes u otras circunstancias, alteraciones cutáneas que contraindiquen la aplicación del transductor, etc. Otra desventaja de la ecografía pulmonar y diafragmática radica en su carácter observador-dependiente, pues implica la necesidad de ciertas habilidades y experiencia del operador en su manejo, sin embargo la curva de aprendizaje resulta relativamente rápida^{161,162} y finalmente se consigue una alta concordancia interobservador¹⁶³. No olvidemos de que se trata

de una prueba de imagen y sus hallazgos deben ser interpretados conjuntamente con datos clínicos, de laboratorio y microbiológicos en caso de disponer de los mismos.

Los datos presentados sugieren que la ecografía pulmonar es un test que presenta elevada exactitud para el diagnóstico de neumonía. Sugerimos un papel adyuvante con la RxT dado que ésta, pese a tener una menor sensibilidad y valor predictivo positivo para el diagnóstico de neumonía, aporta información complementaria importante sobre estructuras óseas, mediastino, vía aérea y partes blandas. No obstante es posible que la implementación de los ultrasonidos contribuya a minimizar el número de exploraciones radiológicas, lo cual podría repercutir positivamente además en la seguridad del paciente y del personal sanitario, en la optimización del tiempo y de los recursos económicos. En cuanto al papel de la TC en el diagnóstico de neumonía consideramos que en este momento es poco probable que su rol de prueba de referencia en casos dudosos o de mala evolución pese al tratamiento se vea modificado por la introducción de la ecografía como técnica de imagen habitual.

En lo que respecta a la ecografía diafragmática para predecir el resultado del destete, tenemos que tener en cuenta que en pacientes sometidos a VM la ED observada es el resultado del esfuerzo respiratorio del paciente y de la presión administrada por el respirador¹⁶⁴, sin embargo el cambio de grosor diafragmático reflejaría realmente la contracción activa del diafragma¹⁶⁵. La medición de la ED puede verse afectada por la posición del sujeto, siendo mayor en decúbito supino que en sedestación¹⁶⁶, así como en situaciones en que se modifique la presión abdominal y/o torácica, por ejemplo por la presencia de ascitis, visceromegalias, resección pulmonar, atelectasias, etc¹⁶⁷. Pese a presentar el hándicap de ser una técnica observador-dependiente, la evidencia disponible sugiere que tanto la FEDI como la ED son mediciones reproducibles¹⁶⁸⁻¹⁷⁰. En nuestro medio, la concordancia interobservador también es moderada-buena para FEDI y los límites de concordancia nos permiten acotar el error de medición (mínima diferencia detectable) según el método de Bland-Altman¹⁴⁸. Si comparamos las diferencias de medianas encontradas en el FEDI en el grupo de fallo y éxito de weaning (Tabla 8), observamos que la diferencia entre ambas es de 11%. Dado que el error de medición es mayor (+/- 12,5%), podemos afirmar que aunque la concordancia valorada con los criterios habituales se puede considerar como buena-moderada, las diferencias observadas son menores que la mínima diferencia detectable, lo que supone una importante limitación para su aplicación práctica. Probablemente sea debido a que el FEDI es una medida que tiene sus propias limitaciones en su cálculo, ya que una pequeña diferencia de pocos milímetros en la medición del grosor inspiratorio o espiratorio diafragmáticos tiene consecuencias extremas en el valor final. De ello se deduce que pequeñas variaciones de medida entre observadores pueden afectar sustancialmente al resultado¹⁷¹.

La concordancia interobservador resultó excelente para LUSm en nuestro trabajo, no habiendo encontrado hasta la fecha ningún estudio que valore la reproducibilidad del LUS score.

En contraste con la excelente exactitud de la FEDI obtenida en el análisis univariante, al estratificar según criterios de aplicabilidad, la especificidad es significativamente menor en pacientes críticos generales que en grupos de pacientes con mayor probabilidad pre-prueba de fracaso en el destete como son aquellos con EPOC¹⁷² o polineuropatía del paciente crítico¹⁷³. Respecto a la ED, en el análisis de subgrupos y de forma análoga a lo que ocurre con la FEDI, los estudios considerados de menor calidad y con posibles problemas de aplicabilidad, tienen una S significativamente mayor, lo cual sugiere que los aspectos relativos al menoscabo de la calidad y aplicabilidad (ver protocolo QUADAS en el estudio publicado en *Chest journal*) podrían estar condicionando una sobreestimación de la S de la ED en la predicción del éxito en el destete.

Estos hallazgos son concordantes con el estudio de Dubé et al, en el que en una muestra de 112 pacientes en PS la ED no fue capaz de predecir la disfunción diafragmática¹⁴⁶.





Biblioteca
UNIVERSITAS Miguel Hernández

6. CONCLUSIONES



6. CONCLUSIONES.

- 1) La exactitud de la ecografía pulmonar para el diagnóstico de neumonía en pacientes críticos es excelente.
- 2) La excelente exactitud de la ecografía pulmonar para el diagnóstico de neumonía no parece ser atribuible a deficiencias en la calidad ni a problemas de aplicabilidad en los estudios incluidos.
- 3) En cuanto a la ecografía diafragmática para la predicción del resultado del destete, consideramos que la FEDI es por sí misma un predictor modesto del resultado del destete en pacientes críticos generales. Respecto a la ED, no consideramos su uso sistemático, ya que su precisión es menor y su medición e interpretación conlleva varios inconvenientes. El LUS score parece ser un buen predictor, pero de momento existe elevada incertidumbre al respecto y son necesarios más estudios que lo evalúen.
- 4) Al estratificar según criterios de aplicabilidad en los estudios que miden FEDI, la especificidad es significativamente menor en pacientes críticos generales que en grupos de pacientes con mayor probabilidad pre-prueba de fracaso en el destete. Ello sugiere que los aspectos que afectan a la aplicabilidad de los estudios condicionan una sobreestimación de la especificidad de la FEDI en la predicción del éxito del destete.
- 5) Respecto a la ED, en el análisis de subgrupos los estudios considerados de menor calidad y con posibles problemas de aplicabilidad, tienen una sensibilidad significativamente mayor, lo cual sugiere que los aspectos relativos al menoscabo de la calidad y aplicabilidad podrían estar condicionando una sobreestimación de la sensibilidad de la ED en la predicción del éxito en el destete.
- 6) En nuestro medio, la concordancia interobservador fue excelente para LUSm y moderada-buena para FEDI.



7. REFERENCIAS BIBLIOGRÁFICAS.





1. Ruiz Villandiego I. et al. Bases físicas de la ecografía. Imágenes esenciales. Aplicaciones de la ecografía en Neumología. Rev Esp Patol Torac 2016; 28 (2) Suplemento 1: 4-14.
2. Rumack, C. M. et al. Physics of Ultrasound. In: Diagnostic ultrasound, 5th edition. Vol 1. Philadelphia, PA, USA: Elsevier; 2017.
3. García de Casasola G, Torres Macho J. Principios físicos de los ultrasonidos. In: Manual de ecografía clínica. Sociedad Española de Medicina Interna (SEMI). Madrid; 2014.
4. Kim K, Jang DM, Park JY, Yoo H, Kim HS, Choi WJ. Changes of diaphragmatic excursion and lung compliance during major laparoscopic pelvic surgery: A prospective observational study. PLoS ONE. 2018; 13(11): e0207841. <https://doi.org/10.1371/journal.pone.0207841>.
5. Schmidt G. Técnicas ecográficas. In: Ecografía. de la imagen al diagnóstico. Stuttgart, Alemania: Editorial médica Panamericana; 2007:3-8.
6. Cohen Hea. AIUM Practice Guideline for the performance of the Focused Assessment with Sonography for Trauma (FAST) examination. J Ultrasound Med 2014 Nov;33(11):2047-56 doi: 10.7863/ultra.33.11.2047.
7. Breitzkreutz R, Price S, Steiger HV, Seeger FH, Ilper H, Ackermann H, Rudolph M, Uddin S, Weigand MA, Müller E, Walcher F; Emergency Ultrasound Working Group of the Johann Wolfgang Goethe-University Hospital, Frankfurt am Main. Focused echocardiographic evaluation in life support and peri-resuscitation of emergency patients: A prospective trial. Resuscitation. 2010 nov;81(11):1527-33. doi: 10.1016/j.resuscitation.2010.07.013.
8. Lichtenstein D, Mezière G. Relevance of lung ultrasound in the diagnosis of acute respiratory failure: The BLUE protocol. Chest. 2008; 134: 117–125.
9. Lichtenstein D. Fluid administration limited by lung sonography: The place of lung ultrasound in assessment of acute circulatory failure. The FALLS-protocol. Expert rev respir med. 2012 apr;6(2):155-62. doi: 10.1586/ers.12.13.
10. Rudas M, Orde S, Nalos M. Bedside lung ultrasound in the care of the critically ill. Crit Care Resusc. 2017 dec;19(4):327-336.
11. Lichtenstein D, Menu Y. A bedside ultrasound sign ruling out pneumothorax in the critically ill: Lung sliding. Chest. 1995;108:1345–8.
12. Lichtenstein D. Diagnostic échographique de l'œdème pulmonaire. Rev im med. 1994;4:561–562.
13. Eibenberger KL, Dock WI, Ammann ME, R. Dorffner R, Hormann MF, Grabenwoger F. Quantification of pleural effusions: Sonography versus Radiography. Radiology, 191 (1994), pp. 681-684.
14. Mayo PH, Goltz HR, Tafreshi, M, Doelken, P. Safety of ultrasound-guided thoracentesis in patients receiving mechanical ventilation. Chest. 2004;125(3):1059.
15. Ding W, Shen Y, Yang J, He X, Zhang M. Diagnosis of pneumothorax by radiography and ultrasonography: a meta-analysis. Chest. 2011 Oct;140(4):859-866. doi: 10.1378/chest.10-2946.
16. Lichtenstein D, Mezière G, Biderman P, Gepner A. The “lung point”: An ultrasound sign specific to pneumothorax. intensive care med, 26 (2000), pp. 1434-1440.

17. Colmenero M, García-Delgado M, Navarrete I, López-Milena G. Utilidad de la ecografía pulmonar en la unidad de Medicina Intensiva. *Med Intensiva*. 2010;34(9):620–628. .
18. Bouhemad B, Brisson H, Le-Guen M, Arbelot C, Lu Q, Rouby JJ. Bedside ultrasound assessment of positive end-expiratory pressure-induced lung recruitment. *Am J Respir Crit Care Med*. 2011;183(3):341.
19. Bouhemad B, Liu ZH, Arbelot C, Zhang M, Ferarri F, Le-Guen M, Girard M, Lu Q, Rouby JJ. Ultrasound assessment of antibiotic-induced pulmonary re-aeration in ventilator-associated pneumonia. *Crit Care Med*. 2010 Jan;38(1):84-92. doi: 10.1097/CCM.0b013e3181b08cdb
20. Reissig A, Copetti R. Lung ultrasound in community-acquired pneumonia and in interstitial lung diseases. *Respiration*. 2014;87(3):179-189.
21. Koenig S, Chandra S, Alaverdian A, Dibello C, Mayo PH, Narasimhan M. Ultrasound assessment of pulmonary embolism in patients receiving CT pulmonary angiography. *Chest*. 2014 Apr;145(4):818-23.
22. Lichtenstein D, Mezière G, Lagoueyte JF, Biderman P, Goldstein I, Gepner A. A-lines and B-lines: Lung ultrasound as a bedside tool for predicting pulmonary artery occlusion pressure in the critically ill. *Chest*. 2009 Oct;136(4):1014-20.
23. Leblanc D, Bouvet C, Degiovanni F, Nedelcu C, Bouhours G, Rineau E, Ridereau-Zins C, Beydon L, Lasocki S. Early lung ultrasonography predicts the occurrence of acute respiratory distress syndrome in blunt trauma patients. *Intensive Care Med*. 2014 Oct;40(10):1468-74.
24. Criner GJ. Measuring diaphragm shortening using ultrasonography to predict extubation success. *Thorax* 2014 May;69(5):402-404.
25. Delgado-Ayala LY et al. El ultrasonido pulmonar como herramienta para predecir éxito a la extubación en pacientes con ventilación mecánica invasiva. *Rev Asoc Mex Med Crit Ter Int* 2016;30(3):153-160.
26. Chalmers I, Glasziou P. Avoidable waste in the production and reporting of research evidence. *Lancet* 2009; 374: 86–89.
27. Cooper NJ, Jones DR, Sutton A J. The use of systematic reviews when designing studies. *Clin Trials*. 2005;2:260-4.40. .
28. Latour-pérez J. investigación en el enfermo crítico. dificultades y perspectivas. *Med intensiva*. 2018;42(3):184-195. <https://Doi.org/10.1016/j.medin.2017.07.008>. .
29. Salman RA et al. Increasing value and reducing waste in biomedical research regulation and management. *Lancet*. 2014 Jan 11; 383(9912): 176–185. doi:10.1016/S0140-6736(13)62297-7. .
30. Moher D et al. Increasing value and reducing waste in bio-medical research: Who’s listening? *Lancet*. 2016;387:1573-86.4. .
31. Macaskill P, Gatsonis C, Deeks JJ, Harbord RM, Takwoingi Y. Chapter 10: Analysing and presenting results. in: *Cochrane handbook for systematic reviews of diagnostic test accuracy, version 1.0*. The Cochrane Collaboration, 2010. .
32. Deville WL, Buntinx F, Bouter LM, Montori VM, De Vet HC, Van der Windt DA, Bezemer PD. Conducting systematic reviews of diagnostic studies: Didactic guidelines. *BMC Med Res Methodol* 2002;2:9.

33. Zamora-Romero J, Plana MN, Abraira Santos, V. Estudios de evaluación de la validez de una prueba diagnóstica: Revisión sistemática y metanálisis. *Nefrología* 2009;29(sup. ext. 6):15-20.
34. Abraira V. índices de rendimiento de las pruebas diagnósticas. *SEMERGEN* 2008;28:93-194.
35. European network for health technology assessment, Eunethhta. Methodological guideline on "Meta-analysis of diagnostic test accuracy studies." Disponible en <http://www.eunethhta.eu/outputs/methodological-guideline-meta-analysis-diagnostic-test-accuracy-studies>
36. Lee J, Kim KW, Choi SH, Huh J, Park SH. Systematic review and meta-analysis of studies evaluating diagnostic test accuracy: A practical review for clinical researchers-part II. Statistical methods of meta-analysis. *Korean J Radiol.* 2015 Nov-Dec;16(6):1175-87. doi: 10.3348/kjr.2015.16.6.1175.
37. Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *J Clin Epidemiol* 2005;58:982-9. .
38. Rutter CM, Gatsonis CA. A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations. *Stat Med* 2001;20:2865-84.
39. Harbord RM, Deeks JJ, Egger M, Whiting P, Sterne JA. A unification of models for meta-analysis of diagnostic accuracy studies. *Biostatistics* 2007;8:239-51.
40. Moons KG, Altman DG, Vergouwe Y, Royston P. Prognosis and prognostic research: Application and impact of prognostic models in clinical practice. *BMJ* 2009;338:1487-90.
41. Debray TP, Damen JA, Snell K, Ensor J, Hooft L, Reitsma JB, Riley RD, Moons KG. A guide to systematic review and meta-analysis of prediction model performance. *BMJ.* 2017 Jan 5;356:i6460. doi: 10.1136/bmj.i6460. .
42. Brenner DJ, Hall EJ. Computed tomography, an increasing source of radiation exposure. *N engl J med.* 2007;357(22):2277-2284.
43. Mayo JR, Aldrich J, Muller NL. Radiation exposure at chest CT: A statement of the fleischner society. *Radiology.* 2003;228(1):15-21.
44. Albaum MN, Hill LC, Murphy M, et al. Interobserver reliability of the chest radiograph in community-acquired pneumonia. PORT investigators. *Chest.* 1996;110(2):343-350.
45. American Thoracic Society, Infectious Diseases Society of America. Guidelines for the management of adults with hospital-acquired, ventilator-associated, and healthcare-associated pneumonia. *Am J Respir Crit Care Med.* 2005 Feb 15;171(4):388-416.
46. Woodhead M, Blasi F, Ewig S, et al. Guidelines for the management of adult lower respiratory tract infections. *Clin Microbiol Infect.* 2011;17Suppl6:E1-59.
47. Nazerian P, Volpicelli G, Vanni S, et al. Accuracy of lung ultrasound for the diagnosis of consolidations when compared to chest computed tomography. *Am J Emerg Med.* 2015;33(5):620-625. .
48. Reissig A, Copetti R, Mathis G, Mempel C, Schuler A, Zechner P, et al. Lung ultrasound in the diagnosis and follow-up of community-acquired pneumonia: A prospective, multicenter, diagnostic accuracy study. *Chest* 2012 Oct;142(4):965-972.

49. Lichtenstein DA. Ultrasound examination of the lungs in the intensive care unit. *Pediatr Crit Care Med*. 2009;10(6):693-698. .
50. Liu XL, Lian R, Tao YK, Gu CD, Zhang GQ. Lung ultrasonography: An effective way to diagnose community-acquired pneumonia. *Emerg Med J* 2015 Jun;32(6):433-438.
51. Boles JM, Bion J, Connors A, Herridge M, Marsh B, Melot C, Pearl R, Silverman H, Stanchina M, Vieillard-Baron A, Welte T. Weaning from mechanical ventilation. *European Respiratory Journal* 2007 29: 1033-1056; doi: 10.1183/09031936.00010206. .
52. Jiang JR, Yen SY, Chien JY, Liu HC, Wu YL, Chen CH. Predicting weaning and extubation outcomes in long-term mechanically ventilated patients using the modified burns wean assessment program scores. *Respirology*. 2014 May;19(4):576-82. doi: 10.1111/resp.12266.
53. DiNino E, Gartman EJ, Sethi JM, McCool FD. Diaphragm ultrasound as a predictor of successful extubation from mechanical ventilation. *Thorax* 2014 May;69(5):423-427.
54. Mayo P, Volpicelli G, Lerolle N, Schreiber A, Doelken P, Vieillard-Baron A. Ultrasonography evaluation during the weaning process: The heart, the diaphragm, the pleura and the lung. *Intensive Care Med*. 2016 Jul;42(7):1107-17.
55. Dres et al. Reliability of diaphragmatic ultrasonography to detect diaphragm dysfunction in critically ill patients. *Intensive Care Medicine Experimental* 2015 3(suppl 1):A452.
56. Global Burden of Disease Study 2013 Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990-2013: A systematic analysis for the global burden of disease study 2013. *Lancet*. 2015;386(9995):743-800.
57. Torres A, Peetermans WE, Viegi G, Blasi F. Risk factors for community-acquired pneumonia in adults in europe: A literature review. *Thorax*. 2013;68(11):1057-1065.
58. Gutierrez F, Masiá M, Mirete C, et al. The influence of age and gender on the population-based incidence of community-acquired pneumonia caused by different microbial pathogens. *J Infect*. 2006;53(3):166-174.
59. Mandell LA, Wunderink RG, Anzueto A, Bartlett JG, Campbell GD, Dean NC, et al. Infectious Diseases Society of America. American Thoracic Society consensus guidelines on the management of community-acquired pneumonia in adults. *Clin Infect Dis* 2007 Mar 1;44 suppl 2:S27-72.
60. Wilkins TR, Wilkins RL. Clinical and radiographic evidence of pneumonia. *Radiol Technol*. 2005;77(2):106-110.
61. Hagaman JT, Rouan GW, Shipley RT, Panos RJ. Admission chest radiograph lacks sensitivity in the diagnosis of community-acquired pneumonia. *Am J Med Sci*. 2009;337(4):236-240.
62. Hayden GE, Wrenn KW. Chest radiograph vs computed tomography scan in the evaluation for pneumonia. *J Emerg Med*. 2009;36(3):266-270.
63. Maughan BC, Asselin N, Carey JL, Sucov A, Valente JH. False-negative chest radiographs in emergency department diagnosis of pneumonia. *R I Med J*. 2014;97(8):20-23.

64. Self WH, Courtney DM, McNaughton CD, Wunderink RG, Kline JA. High discordance of chest x-ray and computed tomography for detection of pulmonary opacities in ED patients: Implications for diagnosing pneumonia. *Am J Emerg Med.* 2013;31(2):401-405.
65. Sartori S, Tombesi P. Emerging roles for transthoracic ultrasonography in pulmonary diseases. *World J Radiol.* 2010;2(6):203-214.
66. Cortellaro F, Colombo S, Coen D, Duca PG. Lung ultrasound is an accurate diagnostic tool for the diagnosis of pneumonia in the emergency department. *Emerg Med J.* 2012;29(1):19-23.
67. Chavez MA, Shams N, Ellington LE, et al. Lung ultrasound for the diagnosis of pneumonia in adults: A systematic review and meta-analysis. *Respir Res.* 2014;15(1).
68. PROMPT. <http://es.pons.com/traducción2010>.
69. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Med.* 2009;6(7):e1000097.
70. Busti C, Agnelli G, Duranti M, Orlandi C, Marcucci M, Paciaroni M. Lung ultrasound in the diagnosis of stroke-associated pneumonia. *Intern Emerg Med.* 2014;9(2):173-178.
71. Reissig A, Copetti R, Mathis G, et al. Lung ultrasound in the diagnosis and follow-up of community-acquired pneumonia: A prospective, multicenter, diagnostic accuracy study. *Chest.* 2012;142(4):965-972.
72. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: A revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med.* 2011;155(8):529-536.
73. Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *J Clin Epidemiol.* 2005;58(10):982-990.
74. R Core Team (2016). R: A language and environment for statistical computing. R foundation for statistical computing, Vienna, Austria. URL <https://www.R-project.org/>.
75. Rutter CM, Gatsonis CA. A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations. *Stat Med.* 2001;20(19):2865-2884.
76. Review Manager (RevMan), version 5.3. Copenhagen. The nordic Cochrane Centre, The Cochrane Collaboration; 2014.
77. Bataille B, Riu B, Ferre F, et al. Integrated use of bedside lung ultrasound and echocardiography in acute respiratory failure: A prospective observational study in ICU. *Chest.* 2014;146(6):1586-1593.
78. Bourcier J, Paquet J, Seinger M, et al. Performance comparison of lung ultrasound and chest x-ray for the diagnosis of pneumonia in the ED. *Am J Emerg Med.* 2014;32(2):115-118.
79. Gallard E, Redonnet J, Bourcier J, et al. Diagnostic performance of cardiopulmonary ultrasound performed by the emergency physician in the management of acute dyspnea. *Am J Emerg Med.* 2014.
80. Nazemi S, Farokhnia M, Pirouzi M, Najarion O, Ahmadipour H. Accuracy of ultrasound as A diagnostic tool of pneumonia for admitted patients in infectious disease section of Afzalipour Hospital. *Jam J Sci Tech.* 2014;25:48-54.

81. Zagli G, Cozzolino M, Terreni A, Biagioli T, Caldini AL, Peris A. Diagnosis of ventilator-associated pneumonia: A pilot, exploratory analysis of a new score based on procalcitonin and chest echography. *Chest*. 2014;146(6):1578-1585.
82. Bourcier J, Paquet J, Seinger M, et al. Performance comparison of lung ultrasound and chest x-ray for the diagnosis of pneumonia in the ED. *Am J Emerg Med*. 2014;32(2):115-118.
83. Reissig A, Copetti R. Lung ultrasound in community-acquired pneumonia and in interstitial lung diseases. *Respiration*. 2014;87(3):179-189.
84. Corradi F, Ball L, Brusasco C, et al. Lung ultrasonography fails detection of non-subpleural community acquired pneumonia. *Intensive Care Med*. 2012;38:S238-S238.
85. Unluer EE, Karagoz A, Senturk GO, Karaman M, Olow KH, Bayata S. Bedside lung ultrasonography for diagnosis of pneumonia. *Hong Kong Journal of Emergency Medicine*. 2013;20(2):98-104.
86. Fares A, Abo-Naghl A, Moustafa Z. Role of transthoracic ultrasound in detection of pneumonia in ICU patients. *Med J Cairo Univ*. 2015;83(1):307-314.
87. Nazerian P, Volpicelli G, Vanni S, et al. Accuracy of lung ultrasound for the diagnosis of consolidations when compared to chest computed tomography. *Am J Emerg Med*. 2015;33(5):620-625.
88. Nafae R, Eman SR, Mohamad NA, El-Ghamry R, Ragheb AS. Adjuvant role of lung ultrasound in the diagnosis of pneumonia in intensive care unit-patients. *Egypt J Chest Dis Tuberc*. 2013;62(2):281-285.
89. Bourcier J, Paquet J, Seinger M, et al. Performance comparison of lung ultrasound and chest x-ray for the diagnosis of pneumonia in the ED. *Am J Emerg Med*. 2014;32(2):115-118.
90. Berlet T, Etter R, Fehr T, Berger D, Sendi P, Merz TM. Sonographic patterns of lung consolidation in mechanically ventilated patients with and without ventilator-associated pneumonia: A prospective cohort study. *J Crit Care*. 2015;30(2):327-333.
91. Plana, M. N., Zamora, J., & Abaira, V. Lectura crítica de revisiones sistemáticas y metanálisis en el diagnóstico por imagen. *Radiología*. 2015;57:23-30.
92. Lemaire F, Teboul JL, Cinotti L, et al. Acute left ventricular dysfunction during unsuccessful weaning from mechanical ventilation. *Anesthesiology*. 1988;69(2):171-179.
93. Caroleo S, Agnello F, Abdallah K, Santangelo E, Amantea B. Weaning from mechanical ventilation: An open issue. *Minerva Anesthesiol*. 2007;73(7-8):417-427.
94. Epstein SK, Ciubotaru RL, Wong JB. Effect of failed extubation on the outcome of mechanical ventilation. *Chest*. 1997;112(1):186-192.
95. Peñuelas O, Frutos-Vivar F, Fernandez C, et al. Characteristics and outcomes of ventilated patients according to time to liberation from mechanical ventilation. *Am J Respir Crit Care Med*. 2011;184(4):430-437.
96. MacIntyre NR, Cook DJ, Ely EW, et al. Evidence-based guidelines for weaning and discontinuing ventilatory support: A collective task force facilitated by the American College of Chest Physicians; The American Association for Respiratory Care; and the American College of Critical Care Medicine. *Chest*. 2001;120(6 Suppl):375S-95S.

97. Boles JM, Bion J, Connors A, et al. Weaning from mechanical ventilation. *Eur Respir J*. 2007;29(5):1033-1056.
98. Ouellette DR, Patel S, Girard TD, et al. Liberation from mechanical ventilation in critically ill adults: An official American College of Chest Physicians/American Thoracic Society clinical practice guideline: Inspiratory pressure augmentation during spontaneous breathing trials, protocols minimizing sedation, and noninvasive ventilation immediately after extubation. *Chest*. 2017;151(1):166-180.
99. Esteban A, Alia I, Tobin MJ, et al. Effect of spontaneous breathing trial duration on outcome of attempts to discontinue mechanical ventilation. spanish lung failure collaborative group. *Am J Respir Crit Care Med*. 1999;159(2):512-518.
100. Frutos-Vivar F, Ferguson ND, Esteban A, et al. Risk factors for extubation failure in patients following a successful spontaneous breathing trial. *Chest*. 2006;130(6):1664-1671.
101. Khamiees M, Raju P, DeGirolamo A, Amoateng-Adjepong Y, Manthous CA. Predictors of extubation outcome in patients who have successfully completed a spontaneous breathing trial. *Chest*. 2001;120(4):1262-1270.
102. Ouanes-Besbes L, Dachraoui F, Ouanes I, et al. NT-proBNP levels at spontaneous breathing trial help in the prediction of post-extubation respiratory distress. *Intensive Care Med*. 2012;38(5):788-795.
103. Boutou AK, Abatzidou F, Tryfon S, et al. Diagnostic accuracy of the rapid shallow breathing index to predict a successful spontaneous breathing trial outcome in mechanically ventilated patients with chronic obstructive pulmonary disease. *Heart Lung*. 2011;40(2):105-110.
104. Conti G, Montini L, Pennisi MA, et al. A prospective, blinded evaluation of indexes proposed to predict weaning from mechanical ventilation. *Intensive Care Med*. 2004;30(5):830-836.
105. Meade M, Guyatt G, Cook D, et al. Predicting success in weaning from mechanical ventilation. *Chest*. 2001;120(6 Suppl):400S-24S.
106. Teboul J, Monnet X, Richard C. Weaning failure of cardiac origin: Recent advances. *Critical Care*. 2010;14(2):211-211.
107. Bouhemad B, Brisson H, Le-Guen M, Arbelot C, Lu Q, Rouby JJ. Bedside ultrasound assessment of positive end-expiratory pressure-induced lung recruitment. *Am J Respir Crit Care Med*. 2011;183(3):341-347.
108. Soummer A, Perbet S, Brisson H, et al. Ultrasound assessment of lung aeration loss during a successful weaning trial predicts postextubation distress. *Crit Care Med*. 2012;40(7):2064-2072.
109. Martin DA, Smith BK, Gabrielli A. Mechanical ventilation, diaphragm weakness and weaning: A rehabilitation perspective. *Respir Physiol Neurobiol*. 2013;189(2):377-383.
110. Mayo P, Arntfield R, Balik M, et al. The ICM research agenda on critical care ultrasonography. *Intensive Care Med*. 2017.
111. Alrajab S, Youssef AM, Akkus NI, Caldito G. Pleural ultrasonography versus chest radiography for the diagnosis of pneumothorax: Review of the literature and meta-analysis. *Crit Care*. 2013;17(5):R208.

112. Lichtenstein DA. BLUE-protocol and FALLS-protocol: Two applications of lung ultrasound in the critically ill. *Chest*. 2015;147(6):1659-1670.
113. Llamas-Alvarez AM, Tenza-Lozano EM, Latour-Perez J. Accuracy of lung ultrasonography in the diagnosis of pneumonia in adults: Systematic review and meta-analysis. *Chest*. 2017;151(2):374-382.
114. Dube BP, Dres M. Diaphragm dysfunction: Diagnostic approaches and management strategies. *J Clin Med*. 2016;5(12):E113.
115. Boussuges A, Gole Y, Blanc P. Diaphragmatic motion studied by m-mode ultrasonography: Methods, reproducibility, and normal values. *Chest*. 2009;135(2):391-400.
116. Summerhill EM, El-Sameed YA, Glidden TJ, McCool FD. Monitoring recovery from diaphragm paralysis with ultrasound. *Chest*. 2008;133(3):737-743.
117. Bouhemad B, Liu ZH, Arbelot C, et al. Ultrasound assessment of antibiotic-induced pulmonary re-aeration in ventilator-associated pneumonia. *Crit Care Med*. 2010;38(1):84-92.
118. Via G, Storti E, Gulati G, Neri L, Mojoli F, Braschi A. Lung ultrasound in the ICU: From diagnostic instrument to respiratory monitoring tool. *Minerva Anestesiol*. 2012;78(11):1282-1296.
119. Kim WY, Suh HJ, Hong SB, Koh Y, Lim CM. Diaphragm dysfunction assessed by ultrasonography: Influence on weaning from mechanical ventilation. *Crit Care Med*. 2011;39(12):2627-2630.
120. Blumhof S, Wheeler D, Thomas K, McCool FD, Mora J. Change in diaphragmatic thickness during the respiratory cycle predicts extubation success at various levels of pressure support ventilation. *Lung*. 2016;194(4):519-525.
121. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: A revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med*. 2011;155(8):529-536.
122. European network for health technology assessment, eunetha. methodological guideline on "meta-analysis of diagnostic test accuracy studies." available at <http://www.eunetha.eu/outputs/methodological-guideline-meta-analysis-diagnostic-test-accuracy-studies>. Accessed march 13, 2017.
123. Baess AI, Abdallah TH, Emara DM, Hassan M. Diaphragmatic ultrasound as a predictor of successful extubation from mechanical ventilation: Thickness, displacement, or both? *Egypt J Bronchol* 2016 10:162±166.
124. Farghaly S, Hasan AA. Diaphragm ultrasound as a new method to predict extubation outcome in mechanically ventilated patients. *Aust Crit Care*. 2017;30(1):37-43.
125. Saeed AM, El Assal GI, Ali TM., Hendawy M.M. Role of ultrasound in assessment of diaphragmatic function in chronic obstructive pulmonary disease patients during weaning from mechanical ventilation. *Egypt J Broncho* 2016 10:167-172.
126. Ferrari G, De Filippi G, Elia F, Panero F, Volpicelli G, Apra F. Diaphragm ultrasound as a new index of discontinuation from mechanical ventilation. *Crit Ultrasound J*. 2014;6(1):8-7902-6-8. eCollection 2014.

127. Jiang J, Yen S, Chien J, Liu H, Wu Y, Chen C. Predicting weaning and extubation outcomes in long-term mechanically ventilated patients using the modified burns wean assessment program scores. *Respirology*. 2014;19(4):576-582.
128. Jiang JR, Tsai TH, Jerng JS, Yu CJ, Wu HD, Yang PC. Ultrasonographic evaluation of liver/spleen movements and extubation outcome. *Chest*. 2004;126(1):179-185.
129. Ali ER, Mohamad AM. Diaphragm ultrasound as a new functional and morphological index of outcome, prognosis and discontinuation from mechanical ventilation in critically ill patients and evaluating the possible protective indices against VIDD. *J Chest Dis Tuberc* (2016).
130. Binet C, Neron L, Rochart N, et al. Validation of a ultrasound index predictive of failure of weaning from mechanical ventilation. *Annales Francaises D Anesthesie Et De Reanimation*. 2014;33:A383-A383.
131. Carrie C, Gisbert-Mora C, Bonnardel E, et al. Ultrasonographic diaphragmatic excursion is inaccurate and not better than the MRC score for predicting weaning-failure in mechanically ventilated patients. *Anaesth Crit Care Pain Med*. 2017;36(1):9-14.
132. DiNino E, Gartman EJ, Sethi JM, McCool FD. Diaphragm ultrasound as a predictor of successful extubation from mechanical ventilation. *Thorax*. 2014;69(5):423-427.
133. Fayed AM, Abd El Hady MA , Shaaban MS, Fikry DM. Use of ultrasound to assess diaphragmatic thickness as a weaning parameter in invasively ventilated chronic obstructive pulmonary disease patients. *J Am Sci* 2016;12(6):96-105.
134. Flevari A, Lignos M, Konstantonis D, Armaganidis A. Diaphragmatic ultrasonography as an adjunct predictor tool of weaning success in patients with difficult and prolonged weaning. *Minerva Anesthesiol*. 2016;82(11):1149-1157.
135. Jung B, Moury PH, Mahul M, et al. Diaphragmatic dysfunction in patients with ICU-acquired weakness and its impact on extubation failure. *Intensive Care Med*. 2016;42(5):853-861.
136. Osman AM HR. Diaphragmatic and lung ultrasound application as new predictive indices for the weaning process in ICU patients. *Egypt J Radiol Nucl Med* (2017).
137. Shoaair M, Noeam K, Mahrous A, Alaa A. Lung aeration loss as a predictor of reintubation using lung ultrasound in mechanically ventilated patients. *Biolife* 4(3), pp 514-520. 2016.
138. Spadaro S, Grasso S, Mauri T, et al. Can diaphragmatic ultrasonography performed during the T-tube trial predict weaning failure? the role of diaphragmatic rapid shallow breathing index. *Critical Care*. 2016;20:305.
139. Tenza-lozano E, Llamas-álvarez A, Jaimez-navarro E, Fernández-Sánchez J. Lung and diaphragmatic ultrasound as predictors of success in weaning. *Crit Ultrasound J*. 2018; 10: 12. <https://doi.org/10.1186/s13089-018-0094-3>
140. Matamis D, Soilemezi E, Tsagourias M, et al. Sonographic evaluation of the diaphragm in critically ill patients. technique and clinical applications. *Intensive Care Med*. 2013;39(5):801-810.
141. Powers SK, Kavazis AN, Levine S. Prolonged mechanical ventilation alters diaphragmatic structure and function. *Crit Care Med*. 2009Oct;37(10suppl):S347-53.doi: 10.1097/CCM.0b013e3181b6e760. . .

142. Levine S, Nguyen T, Taylor N, et al. Rapid disuse atrophy of diaphragm fibers in mechanically ventilated humans. *N Engl J Med*. 2008;358(13):1327-1335.
143. Funk GC, Anders S, Breyer MK et al. Incidence and outcome of weaning from mechanical ventilation according to new categories. *Eur Respir J* 2010;35:88–94.
144. Béduneau G, Pham T, Schortgen F, et al. Epidemiology of weaning outcome according to a new definition. The WIND study. *Am J Respir Crit Care Med*. 2017;195(6):772-783.
145. Savi A, Teixeira C, Silva JM, Borges LG, Pereira PA, Pinto KB et al. Weaning predictors do not predict extubation failure in simple-to-wean patients. *J Crit Care* 2012,27(2):221-e1–221-e8. <https://doi.org/10.1016/j.jcrc.2011.07.079> . .
146. Dube BP, Dres M, Mayaux J, Demiri S, Similowski T, Demoule A. Ultrasound evaluation of diaphragm function in mechanically ventilated patients: Comparison to phrenic stimulation and prognostic implications. *Thorax*. 2017 Sep;72(9):811-818. doi: 10.1136/thoraxjnl-2016-209459.
147. Umbrello M, Formenti P, Longhi D, et al. Diaphragm ultrasound as indicator of respiratory effort in critically ill patients undergoing assisted mechanical ventilation: A pilot clinical study. *Crit Care*. 2015;19:161-015-0894-9.
148. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986 feb 8;1(8476):307-10.
149. StatsDirect Ltd. StatsDirect statistical software.<http://www.statsdirect.com>. England. StatsDirect Ltd. 2013.
150. Xirouchaki N, Magkanas E, Vaporidi K, et al. Lung ultrasound in critically ill patients: Comparison with bedside chest radiography. *Intensive Care Med*. 2011;37(9):1488-1493.
151. Hu Q, Shen Y, Jia L, et al. Diagnostic performance of lung ultrasound in the diagnosis of pneumonia: A bivariate meta-analysis. *Int J Clin Exp Med*. 2014;7(1):115-121.
152. Ye X, Xiao H, Chen B, Zhang S. Accuracy of lung ultrasonography versus chest radiography for the diagnosis of adult community-acquired pneumonia: Review of the literature and meta-analysis. *PLoS ONE*. 2015;10(6).
153. Zhou P, Zhang Z, Hong Y, et al the predictive value of serial changes in diaphragm function during the spontaneous breathing trial for weaning outcome: A study protocol. *BMJ open* 2017;7:e015043. doi: 10.1136/bmjopen-2016-015043. .
154. Cohen E, Mier A, Heywood P, Murphy K, Boulton J, Guz A. Excursion-volume relation of the right hemidiaphragm measured by ultrasonography and respiratory airflow measurements. *Thorax*. 1994;49(9):885-889.
155. Oropello J, Rahmanian M. Can chest sonography predict and facilitate successful ventilator weaning? *Crit Care Med*. 2013;41(8):2065-2067.
156. Orso D, Guglielmo N, Copetti R. Lung ultrasound in diagnosing pneumonia in the emergency department: A systematic review and meta-analysis. *Eur J emerg med*. 2018 oct;25(5):312-321. doi: 10.1097/MEJ.0000000000000517. .
157. Mantuani D, Frazee BW, Fahimi J, Nagdev A. point-of-care multi-organ ultrasound improves diagnostic accuracy in adults presenting to the emergency department with acute dyspnea.

western journal of emergency medicine. 2016;17(1):46-53.
doi:10.5811/westjem.2015.11.28525. .

158. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ*. 1997 sep 13;315(7109):629-34.

159. Volpicelli G, Elbarbary M, Blaivas M, et al. International evidence-based recommendations for point-of-care lung ultrasound. *Intensive Care Med*. 2012;38(4):577-591.

160. Mongodi S, Via G, Bouhemad B, Storti E, Mojoli F, Braschi A. Usefulness of combined bedside lung ultrasound and echocardiography to assess weaning failure from mechanical ventilation: A suggestive case. *Crit Care Med*. 2013;41(8):e182-5.

161. See KC, Ong V, Wong SH, et al. Lung ultrasound training: Curriculum implementation and learning trajectory among respiratory therapists. *Intensive Care Med*. 2016;42(1):63-71.

162. Lichtenstein D, Goldstein I, Mourgeon E, Cluzel P, Grenier P, Rouby JJ. Comparative diagnostic performances of auscultation, chest radiography, and lung ultrasonography in acute respiratory distress syndrome. *Anesthesiology*. 2004;100(1):9-15.

163. Gullett J, Donnelly JP, Sinert R, et al. Interobserver agreement in the evaluation of B-lines using bedside ultrasound. *J Crit Care*. 2015;30(6):1395-1399.

164. Matamis D, Soilemezi E, Tsagourias M, et al. Sonographic evaluation of the diaphragm in critically ill patients. technique and clinical applications. *Intensive Care Med*. 2013;39(5):801-810.

165. Goligher EC, Fan E, Herridge MS, et al. Evolution of diaphragm thickness during mechanical ventilation: Impact of inspiratory effort. *Am J Respir Crit Care Med*. 2015;192(9):1080-1088.

166. Houston JG, Angus RM, Cowan MD, McMillan NC, Thomson NC. Ultrasound assessment of normal hemidiaphragmatic movement: Relation to inspiratory volume. *Thorax*. 1994;49(5):500-503.

167. Haji K, Royse A, Green C, Botha J, Canty D, Royse C. Interpreting diaphragmatic movement with bedside imaging, review article. *J Crit Care*. 2016;34:56-65.

168. Francis CA, Hoffer JA, Reynolds S. Ultrasonographic evaluation of diaphragm thickness during mechanical ventilation in intensive care patients. *Am J Crit Care*. 2016;25(1):e1-8.

169. Bobbia X, Clement A, Claret PG, et al. Diaphragmatic excursion measurement in emergency patients with acute dyspnea: Toward a new diagnostic tool? *Am J Emerg Med*. 2016;34(8):1653-1657.

170. Boussuges A, Gole Y, Blanc P. Diaphragmatic motion studied by m-mode ultrasonography: Methods, reproducibility, and normal values. *Chest*. 2009;135(2):391-400.

171. Zambon M, Greco M, Bocchino S, Cabrini L, Beccaria PF, Zangrillo A. Assessment of diaphragmatic dysfunction in the critically ill patient with ultrasound: A systematic review. *Intensive Care Med*. 2017;43(1):29-38.

172. Lagonidis D. Cl. Physiologic determinants of prolonged mechanical ventilation and unweanable patients. In: Esquinas A., ed. *Noninvasive mechanical ventilation and difficult weaning in critical care. Key topics and practical approaches.*

. 1st ed. Springer; 2016:3-14. Accessed April 24, 2017. 10.1007/978-3-319-04259-6_1.

173. Dres M, Dube BP, Mayaux J, et al. Coexistence and impact of limb muscle and diaphragm weakness at time of liberation from mechanical ventilation in medical intensive care unit patients. *Am J Respir Crit Care Med*. 2017;195(1):57-66.



8. ANEXOS (PUBLICACIONES).

UNIVERSITAS Miguel Hernández



8.1. ANEXO 1.

Accuracy of lung ultrasound in the diagnosis of pneumonia in adults: systematic review and meta-analysis.

Original research.

Ana M. Llamas-Álvarez, MD, Eva M. Tenza-Lozano, MD, Jaime Latour-Pérez, MD, PhD.

Please cite this article as: Llamas-Álvarez AM, Tenza-Lozano EM, Latour-Pérez J, Accuracy of lung ultrasound in the diagnosis of pneumonia in adults: systematic review and meta-analysis, CHEST (2016), doi: 10.1016/j.chest.2016.10.039.

CONTENTS

- **Abbreviations.**
- **Abstract.**
- **Introduction.**
- **Methods.**
- **Results.**
- **Discussion.**
- **Conclusion.**
- **Tables.**
- **Figures.**
- **References.**
- **Supplemental material.**
- **Supplemental material's references.**
- **Acknowledgments**

**ACCURACY OF LUNG ULTRASOUND IN THE DIAGNOSIS OF PNEUMONIA IN ADULTS:
SYSTEMATIC REVIEW AND META-ANALYSIS.**

Ana M. Llamas-Álvarez, MD^a; Eva M. Tenza-Lozano, MD^a; Jaime Latour-Pérez, MD, PhD^{a,b}.

^aIntensive Care Unit, Elche General University Hospital, Elche, Spain

^bDepartment of Clinical Medicine, Miguel Hernández University, Sant Joan d'Alacant, Spain

CORRESPONDENCE TO: Ana Llamas-Álvarez, Department of Intensive Care Medicine, Elche General University Hospital, Camino de la Almazara 11, 03203, Elche, Spain; email: llamasalvarez@yahoo.es

CONFLICT OF INTEREST: The authors have no conflicts of interest to declare.

FUNDING/SUPPORT: Department of Clinical Medicine, Miguel Hernández University.

PRIOR ABSTRACT PUBLICATION/PRESENTATION: No.

IRB APPROVAL/ PATIENT CONSENT: not provided as this is a review of previously published studies. No patients were enrolled.

ABBREVIATIONS

CI: confidence interval.

CT: computed tomography.

CXR: chest X-ray.

DOR: diagnostic odds ratio

ER: emergency room

ICU: intensive care unit

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta- Analyses.

QUADAS-2: Quality Assessment of Diagnostic Accuracy Studies-2.

SROC: Summary Receiver Operating Characteristics.



ABSTRACT

Background: Some studies suggest that lung ultrasound could be useful for diagnosing pneumonia; moreover, it has a more favorable safety profile and lower cost than chest X-ray (CXR) and computed tomography (CT). The aim of this study is to assess the accuracy of bedside lung ultrasound for diagnosing pneumonia in adults through a systematic review and meta-analysis.

Methods: We searched MEDLINE, Scopus, The Cochrane Library, Web of Science, DARE, HTA Database, Google Scholar, LILACS, ClinicalTrials.gov, TESEO and OpenGrey. In addition, we reviewed the bibliographies of relevant studies. Two researchers independently selected studies that met the inclusion criteria. Quality of the studies was assessed in accordance with the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool. The summary receiver operating characteristics (SROC) curve and a pooled estimation of the diagnostic odds ratio (DOR) was estimated using a bivariate random-effects analysis. The sources of heterogeneity were explored using predefined subgroup analyses and bivariate meta-regression.

Results: Sixteen studies (2359 participants) were included. There was significant heterogeneity of both sensitivity and specificity according to Q test, without clear evidence of threshold effect. The area under the SROC curve was 0.93, with a DOR at the optimal cutpoint of 50 (95% confidence interval (CI): 21, 120). A tendency towards a higher area under the SROC curve in high quality studies was detected, however these differences were not significant after applying the bivariate meta-regression.

Conclusions: Lung ultrasound can help to accurately diagnose pneumonia, and it may be promising as an adjuvant resource to traditional approaches.

INTRODUCTION

Pneumonia ranks as the third cause of death worldwide, preceded only by ischemic heart disease and cerebrovascular diseases¹. It is the leading infectious cause of death and one of the most common reasons for emergency room (ER) visits and hospital admissions^{2,3}.

In patients with clinically suspected pneumonia, a correct differential diagnosis is essential for proper treatment orientation. Diagnostic and treatment delays and failure can entail a significant increase in mortality, while inappropriate use of antibiotics to treat respiratory symptoms that are not secondary to bacterial infection contributes to the development of antibiotic resistance⁴.

Traditionally, the diagnosis of pneumonia is based on three mainstays: clinical and laboratory data, imaging techniques, and microbiological studies⁴. Chest X-ray (CXR) currently constitutes the first approach in suspected pneumonia, where the presence of a new infiltrate is the characteristic radiographic finding⁵. However, several studies point to CXR's low sensitivity for the diagnosis of pneumonia⁶⁻⁸. Besides, image quality is lower in patients in decubitus position and when using portable devices; it implies patient mobilization, radiation exposure, and high interobserver variability⁹.

Chest computed tomography (CT) has greater sensitivity than CXR for the diagnosis of pneumonia⁸, but its use is reserved for more complex cases and where therapeutic failure occurs. Its main drawbacks are the need to transfer the patient to the X-ray unit, increased exposure to radiation, and high cost^{10,11}.

In this context, and although it is not considered in current clinical practice guidelines^{4,12,13}, bedside lung ultrasound for pleuropulmonary pathologies has attracted interest in recent years because it is portable and fast, does not use radiation, is easily reproducible, and allows real-time scanning¹⁴. Available data also suggest high diagnostic performance¹⁵⁻¹⁷.

This systematic literature review and meta-analysis aims to assess the accuracy of lung ultrasound for the diagnosis of pneumonia in adult patients.

METHODS

Search Strategy and Study Selection

We included studies in people aged 18 years and older with clinical suspicion or confirmed diagnosis of pneumonia, comparing lung ultrasound versus other diagnostic strategies including CXR and/or CT as imaging technique. We considered both community-acquired and nosocomial alveolar and interstitial pneumonia, including ventilator-associated pneumonia. Studies related to eosinophilic pneumonia were excluded given their different characteristics and etiology.

The following exclusion criteria were applied: not a primary study, impossibility of defining the diagnosis of pneumonia, reference standard does not include CXR or CT, inability to extract the necessary data for calculating a 2x2 table for sensitivity and specificity, inability to obtain the

full-text study and/or translation. We also excluded studies with fewer than 20 participants¹⁹ and studies in which the unit of analysis was not the participant (for example those considering hemithorax, pulmonary areas or lesions), assuming that such units cannot be independent of each other²⁰. In case of doubt, we contacted authors personally.

Two researchers (ALA and ETL) conducted an independent literature search to identify potentially relevant studies in MEDLINE, Scopus, The Cochrane Library, the Web of Science, DARE, HTA Database, Google Scholar, LILACS and ClinicalTrials.gov. We also searched for doctoral theses in the TESEO and Tesis Doctorals en Xarxa (TDX) databases, for conference papers in the Conference Proceedings section of the Web of Science, and for gray literature in OpenGrey. In addition, we reviewed the bibliographies of relevant studies.

Our search strategy employed the terms "pneumonia", "ultrasound", "ultrasonography", "ecograph *" and "sonograph *", and we did not impose any time restriction, including published studies from database inception until April 2016. In order to provide greater comprehensiveness to the review, we did not use any age filters despite the fact that potentially eligible studies had to be in adults; we reviewed the full text of any studies involving children to identify possible subgroups of adults. We did not limit the search based on study design nor language of publication. Discrepancies between the two researchers were resolved by consultation with a third expert researcher (JLP).

Data Extraction

For each included study, two researchers (ALA and ETL) manually extracted data concerning the number of true positives, false positives, true negatives, and false negatives. In cases where this information was not explicitly reflected, 2x2 tables for calculating the required data were constructed.

In two studies^{21,22}, there were ultrasound examinations classified as "inconclusive". In these cases, and in order to analyze all participants, we considered these explorations as "positive". However, one inconclusive ultrasound scan was also inconclusive when the reference standard was used²¹, so it was necessary to exclude that participant from the analysis.

Assessment of Risk of Bias

Two reviewers (ALA and ETL) independently performed the assessment of methodological quality of included articles according to QUADAS-2 criteria²³, resolving discrepancies through consultation with a third expert researcher (JLP).

Data Analysis

Data were presented in the ROC plane and analyzed according to the European Association for Technology Assessment recommendations²⁴.

In the univariate analysis, we constructed a forest plot for sensitivity, specificity and diagnostic odds ratio (DOR). The correlation between sensitivity and false positive rate was explored graphically using the corresponding forest plots and statistically by examining the Spearman correlation coefficient and 95% confidence interval (CI). We used the Mada application included

in the R statistical package²⁵ to apply the bivariate Reitsma²⁶ model, which in the absence of covariates is equivalent to the Rutter-Gatsonis hierarchical SROC method²⁷. We analyzed study heterogeneity graphically and through the Q test.

Possible causes of heterogeneity between studies were explored through prespecified subgroup analyses: study setting (ICU versus other), overall quality of studies (low versus high or unclear risk of bias) and individual dimensions of quality (QUADAS-2). The existence of interaction between these variables and diagnostic accuracy (sensitivity and specificity) was examined by using a bivariate meta-regression. The figures included were created using Mada and Review Manager²⁸.

The systematic review is presented according to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines¹⁸.

RESULTS

Characteristics of Included Studies

The initial literature search yielded a total of 2812 studies, 16 of which met the inclusion criteria and were included in the qualitative and quantitative synthesis (Figure 1).

The predominant design was the cohort study (Tables 1 and 2)²⁹⁻³¹, and all took place between 2008 and 2015. Most studies were European, and altogether they involve a total of 2359 participants. Although CT is traditionally considered the reference test for diagnosing pneumonia, it is not applied systematically due to its ethical implications. The most commonly proposed gold standard was “final diagnosis” or “diagnosis at discharge”^{16,32-38}, taking into account clinical, laboratory, microbiological, and CXR data, and in some cases also echocardiography³² and/or CT^{16,32,33,37} results.

Most studies describe blinding the professionals performing ultrasound to the results of the reference standard and vice versa. None of the studies report the absence of blinding, although some references do not clearly state whether blinding took place^{32,37-41}. Further details about how ultrasonography was conducted are described in e-Table 1.

Risk of Bias and Applicability

Regarding participant selection, we considered studies to be at high risk of bias where there were inadequate exclusions. However, we generally found no major conflicts that could compromise the applicability or confer a high risk of bias in relation to the index test.

Concerning the reference standard, we considered it to be of high quality if based on CT alone or where it consisted of a final diagnosis made by experts using an integrated synthesis of radiology, laboratory and/or microbiological data. Due to the low sensitivity of CXR to independently identify lung infection^{6-9,42}, we considered this to be a lower quality reference standard.

Overall, the highest risk of bias stemmed from the flow of patients within each study because of an uneven application of the reference test (differential verification bias) (Figures 2 and 3).

Meta-analysis

Ultrasonography is subject to observer interpretation. Thus, we assumed there would be unequal (more or less restrictive) thresholds of positivity among studies, and we present sensitivity and specificity data separately for each study, with no global weighting⁴³ (Figure 4). This decision was supported by the statistically significant result of the Q test. In any case, we did not observe a pattern of graphically negative association between sensitivity and specificity. The 95% CI of the overall effect indicates a sensitivity of approximately 80% to 90% and a specificity of 70% to 90%.

The overall estimation of the area under the SROC curve of lung ultrasound for the diagnosis of pneumonia is 0.93, and the calculated DOR is 50 (95% CI: 21, 120) (Figure 5).

Causes of Heterogeneity

Heterogeneity in a meta-analysis of diagnostic test accuracy is usually higher than that found in meta-analysis of intervention studies and very often greater than would be expected due to random variability alone⁴⁴. In fact, the dispersion of the studies in the ROC plane suggests the presence of marked heterogeneity.

In the subgroup analyses, we separated high and low quality studies (Figure 6), observing that high quality studies tended to be distributed along the ROC curve, which suggested the existence of a threshold effect. Although the graphic analysis suggests that high quality studies have a better sensitivity and specificity than those with low quality, bivariate meta-regression shows that the association between the variable "study quality" and the test accuracy (sensitivity and specificity) are not significant. The same goes for the variables of severity, risk of bias associated with the reference standard, and risk of bias associated with the flow of patients. These analyses are detailed in e-Appendix 1.

DISCUSSION

The results of this study suggest that bedside lung ultrasound has an excellent accuracy for the diagnosis of pneumonia in adults. Although most of the included studies have problems regarding risk of bias, the high accuracy of lung ultrasound does not appear to be attributable to the poor quality of the studies, as when stratifying by the different dimensions of QUADAS-2, test accuracy is numerically better in high quality studies.

Our findings are consistent with previous reviews showing even more favorable data on the diagnostic accuracy profile of the lung ultrasound. Chavez et al.¹⁵ obtained a weighted sensitivity and specificity of 94% and 96%, respectively, with an area under the SROC curve of 0.98. The main differences with our review are related to methodology and to the inclusion of fewer studies (n = 10), 70% of which would merit an assessment of high risk of bias according to our criteria. Chavez et al. also included at least one study in which the unit of analysis was not the participant⁴⁵. Hu et al.⁴⁶ obtained a DOR of 509.99 (95% CI: 128.11, 2014.34) and area under the SROC curve of 0.99, although seven of the nine analyzed studies included children and even

infants, so the samples are not comparable. In fact, several pediatric studies suggest superior diagnostic performance for chest ultrasound in children compared to adults^{47,48}, which may be related to the fact that children usually have a thinner chest wall and a smaller volume of lung parenchyma. In the review of Ye et al.⁴⁹, the DOR was 151.19 (CI 95%: 38.50, 593.77), and the area under the SROC curve was 0.97, with the handicap of large imprecision (only five included studies).

In addition to its diagnostic accuracy, lung ultrasound has other advantages over CXR and CT, as it can be done quickly at bedside, is easily reproducible, has a relatively low cost, and avoids any exposure to ionizing radiation. On the other hand, ultrasounds are of limited value in patients with subcutaneous emphysema and in obese people due to the thickness of the chest wall⁵⁰. Moreover, it is not possible where access to the patient's chest is limited by large bandages, prosthetic material, or skin disorders. Another disadvantage of lung ultrasound is its observer-dependent nature, as it implies the need for operators with certain skills and experience; however, the learning curve is relatively fast^{51,52}, and in the end the method can obtain high interobserver agreement⁵³.

To our knowledge, this meta-analysis of sixteen studies and 2359 participants is the largest one to date for ultrasound in the diagnosis of pneumonia in adults, including a review of gray literature and absence of language restrictions. With regard to the review's limitations, it is important to note the high clinical heterogeneity between included studies and substantial differences in the implementation of the ultrasonography and scanning technique. These last questions have been addressed by Volpicelli et al⁵⁴ in a consensus document providing expert recommendations in line with the GRADE⁵⁵ initiative on the implementation, development, and standardization of lung ultrasound.

Our data suggest that lung ultrasound may be a valuable resource to consider alongside clinical, laboratory, and microbiological data for the diagnosis of pneumonia in adults. Regarding imaging techniques, we suggest an adjuvant role for lung ultrasonography with CXR, since the latter provides important additional information on bone, mediastinal structures, airway and soft tissues. Nevertheless, it is possible that the implementation of lung ultrasonography helps to minimize the number of radiological examinations, which could also have a positive impact on patient and staff safety, optimizing the use of time and economic resources.

CONCLUSION

Lung ultrasonography has high sensitivity and specificity and appears to be a valuable complement to chest radiography for the diagnosis of pneumonia in adults.

TABLES.

Table 1. Characteristics of included studies.

Study	Country	Design	Reference standard	Blind in LUS	Blind in gold standard
Bataille 2014 ³²	France	Cohort	FD (CD, Lab, Mic, CXR. ± EC ± CT)	Unclear	Yes
Berlet 2015 ³⁴	Switzerland	Cohort	FD (CD, Lab, Mic ± CXR)	Yes	Yes
Bourcier 2014 ³³	France	Cohort	FD (CD, Lab, CXR. ± CT)	Yes	Yes
Busti 2014 ²¹	Italy	Cohort	CXR ± CT	Yes	Yes
Corradi 2012 ³⁹	Italy	Cohort	CXR ± CT	Unclear	Unclear
Cortellaro 2012 ¹⁶	Italy	Cohort	FD (CD, CXR, Lab, Mic ± CT)	Yes	Yes
Fares 2015 ⁴⁰	Egypt	Cross-sectional	CT	Yes	Unclear
Gallard 2015 ³⁵	France	Cohort	FD (CD, Lab, CXR)	Yes	Yes
Lichtenstein 2008 ³⁶	France	Cohort	FD (CD, Lab, CXR, Mic)	Yes	Yes
Liu 2015 ²⁹	China	Cohort	CT	Yes	Yes
Nafae 2013 ⁴¹	Egypt	Cross-sectional	CT	Yes	Unclear
Nazemi 2014 ³⁷	Iran	Cross-sectional	FD (CD, CXR, Mic. ± CT)	Yes	Unclear
Nazerian 2015 ³⁰	Italy	Cohort	CT	Yes	Yes
Reissig 2012 ²²	Germany	Cohort	CXR ± CT	Yes	Yes
Unluer 2013 ³¹	Turkey	Cohort	CXR ± CT	Yes	Yes
Zagli 2014 ³⁸	Italy	Case-control	FD (CD, CXR, Mic)	Unclear	Unclear

Abbreviations: CD: clinical data; CT: computed tomography; CXR: chest X-ray; EC: echocardiography; FD: final diagnosis (brackets indicate diagnostic methods); Lab: laboratory data; LUS: lung ultrasonography; Mic: microbiology data.

Table 2. Participant characteristics.

Study	N	Setting	Age* (years)	Inclusion	Type of pneumonia
Bataille 2014 ³²	136	ICU	68 ± 15	RF	Not specified
Berlet 2015 ³⁴	57	ICU	61.3 (47.9-71.3)	MV not for respiratory cause	VAP
Bourcier 2014 ³³	144	ER	77.6 ± 15.2	PS	CAP
Busti 2014 ²¹	69	Stroke Unit	77.6 ± 9.3	PS	Nosocomial
Corradi 2012 ³⁹	35	ER	67.09 ± 20.84	PS	CAP
Cortellaro 2012 ¹⁶	120	ER	69 ± 18	PS	CAP
Fares 2015 ⁴⁰	38	ICU	61.02 ± 11.24	PS	Not specified
Gallard 2015 ³⁵	130	ER	79.0 ± 11.1 no LVF 81.9 ± 10.2 LVF	Acute dyspnea	CAP
Lichtenstein 2008 ³⁶	260	ICU	68 ± 16	RF	CAP and nosocomial
Liu 2015 ²⁹	179	ER	71.5 (36-88)	PS	CAP
Nafae 2013 ⁴¹	100	ICU	≥ 50 at 76.25%	PS	Not specified
Nazem 2014 ³⁷	151	Ward	61.44 ± 17.40	PS	CAP
Nazerian 2015 ³⁰	285	ER	71 ± 14	Unexplained respiratory symptoms needing CT	CAP
Reissig 2012 ²²	362	ER and ward	63.8 (19-95)	PS	CAP
Unluer 2013 ³¹	72	ER	Women 68.4 ± 11 Men 64.2 ± 12.4	Dyspnea	CAP
Zagli 2014 ³⁸	221	ICU	56 ± 20.9	Cases of VAP, controls without VAP	VAP

Abbreviations: CAP: community acquired pneumonia; ER: emergency room; ICU: intensive care unit; LVF: left ventricular failure; MV: mechanical ventilation; N: number of patients included; PS: pneumonia suspected; RF: respiratory failure; VAP: ventilator associated pneumonia.

*Age is expressed according to data extracted from each study as media ±SD, median (interquartile range) or percentage within a group.

FIGURES.

Figure 1. PRISMA flow diagram for study identification and selection, with reasons for exclusion

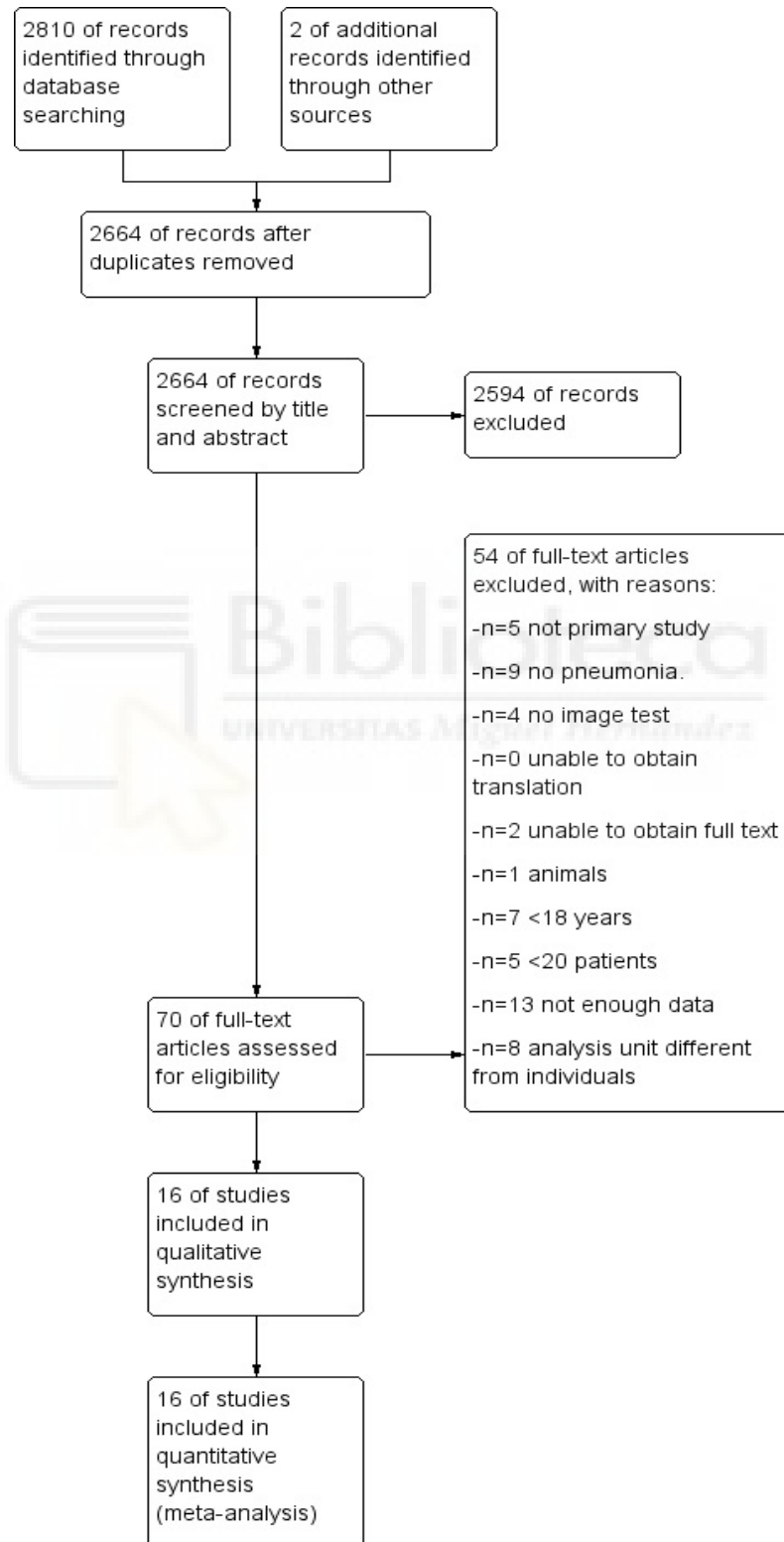


Figure 2. Quality assessment according to QUADAS-2. For each study, risk of bias and applicability concerns are classified as high risk, low risk, or unclear.

	<u>Risk of Bias</u>				<u>Applicability Concerns</u>		
	Patient Selection	Index Test	Reference Standard	Flow and Timing	Patient Selection	Index Test	Reference Standard
Bataille 2014	?	?	+	-	?	+	+
Berlet 2015	+	+	+	?	?	+	+
Bourcier 2014	?	+	+	-	+	+	+
Busti 2014	?	+	-	-	-	+	+
Corradi 2012	+	?	-	-	+	?	+
Cortellaro 2012	+	+	+	-	+	+	+
Fares 2015	?	?	?	+	+	+	+
Gallard 2015	-	+	+	+	-	+	+
Lichtenstein 2008	-	+	+	?	-	+	+
Liu 2015	+	+	+	+	+	+	+
Nafae 2013	?	+	?	+	+	+	+
Nazemi 2014	?	+	?	-	+	?	+
Nazerian 2015	+	+	+	+	-	+	+
Reissig 2012	?	+	-	-	?	+	+
Unluer 2013	?	+	-	-	-	+	+
Zagli 2014	-	?	?	+	?	?	+

- High
 ? Unclear
 + Low

Figure 3. QUADAS-2. Each domain is represented in a bar with the proportion of studies considered as high risk, low risk or unclear. The same applies to applicability concerns.

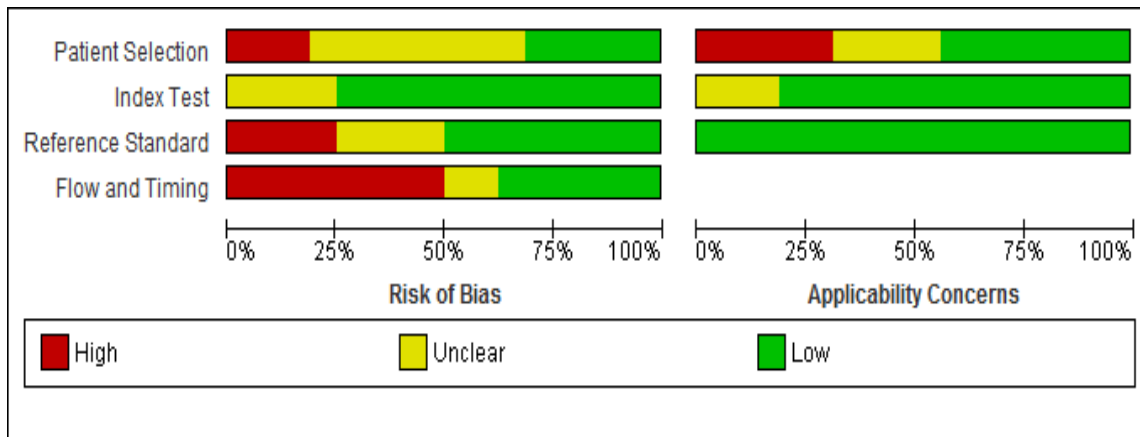
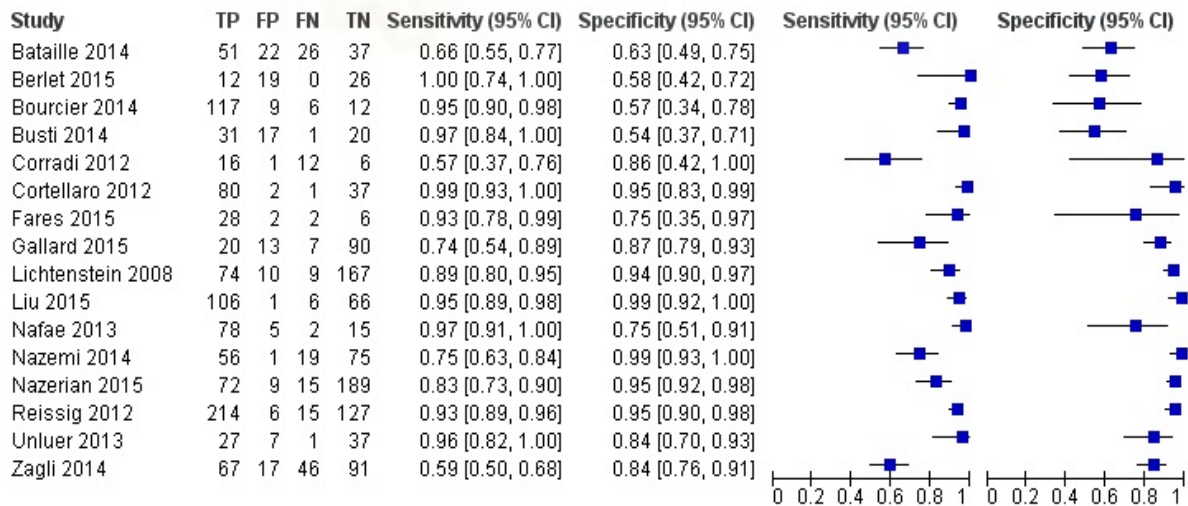


Figure 4. Forest plot of sensibility and specificity in each study.



Abbreviations: CI: confidence interval; TP: true positives; FP: false positives; FN: false negatives; TN: true negatives.

Figure 5. SROC curve of lung ultrasound for the diagnosis in pneumonia and confidence ellipse around the optimal cut-off value.

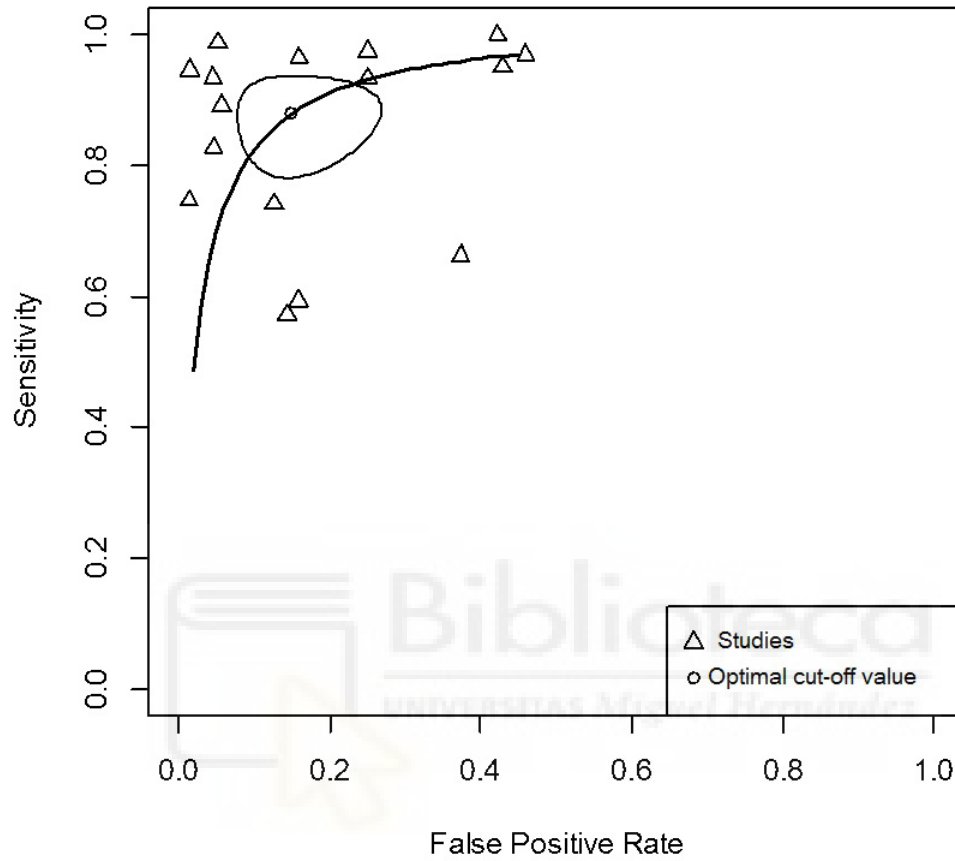
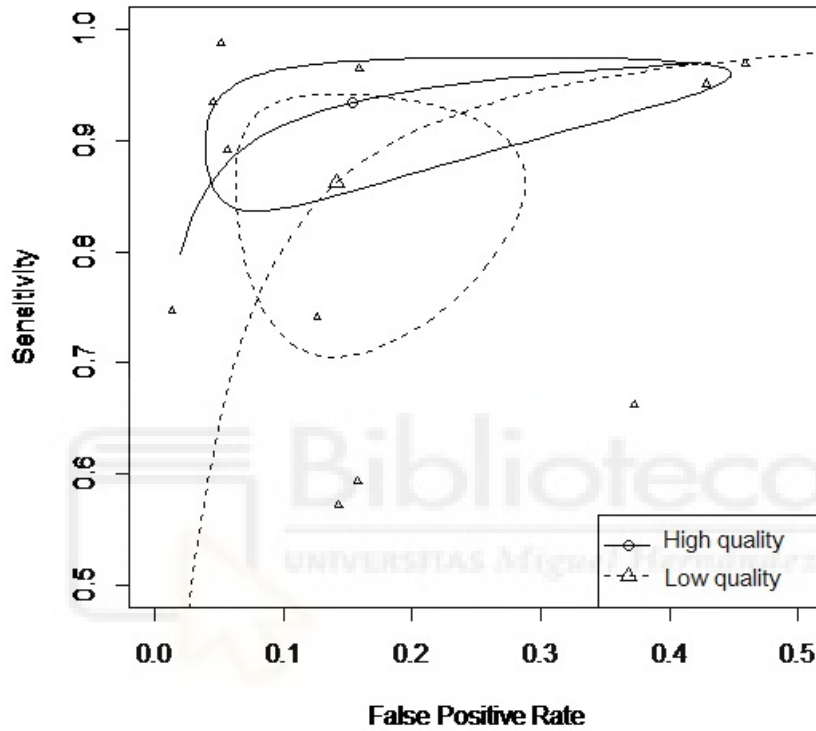


Figure 6. Comparison of ROC curves and confidence ellipse between studies of high and low quality. Based on the analysis of methodological quality using the QUADAS-2 instrument, we considered studies to be of high quality when no domain carried a high risk of bias, while studies at high risk of bias for one or more work items were classified as low quality.



Abbreviations: CI: confidence interval; p : p value, significance < 0.05 ; tfpr: transformed false positive rate; tsens: transformed sensibility.

REFERENCES.

1. Global Burden of Disease Study 2013 Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990-2013: A systematic analysis for the Global Burden of Disease study 2013. *Lancet*. 2015;386(9995):743-800.
2. Self WH, Grijalva CG, Zhu Y, et al. Rates of emergency department visits due to pneumonia in the United States, July 2006-June 2009. *Acad Emerg Med*. 2013;20(9):957-960.
3. Pfunter A, Wier LM, Stocks C. Most frequent conditions in U.S. hospitals, 2011: Statistical brief #162. In: *Healthcare cost and utilization project (HCUP) statistical briefs*. Rockville (MD): Agency for Healthcare Research and Quality; 2013.
4. Mandell LA, Wunderink RG, Anzueto A, et al. Infectious Diseases Society of America/American Thoracic Society consensus guidelines on the management of community-acquired pneumonia in adults. *Clin Infect Dis*. 2007;44 Suppl2:S27-72.
5. Wilkins TR, Wilkins RL. Clinical and radiographic evidence of pneumonia. *Radiol Technol*. 2005;77(2):106-110.
6. Hagaman JT, Rouan GW, Shipley RT, Panos RJ. Admission chest radiograph lacks sensitivity in the diagnosis of community-acquired pneumonia. *Am J Med Sci*. 2009;337(4):236-240.
7. Maughan BC, Asselin N, Carey JL, Sucov A, Valente JH. False-negative chest radiographs in emergency department diagnosis of pneumonia. *R I Med J*. 2014;97(8):20-23.
8. Hayden GE, Wrenn KW. Chest radiograph vs. computed tomography scan in the evaluation for pneumonia. *J Emerg Med*. 2009;36(3):266-270.
9. Albaum MN, Hill LC, Murphy M, et al. Interobserver reliability of the chest radiograph in community-acquired pneumonia. PORT investigators. *Chest*. 1996;110(2):343-350.
10. Brenner DJ, Hall EJ. Computed tomography--an increasing source of radiation exposure. *N Engl J Med*. 2007;357(22):2277-2284.
11. Mayo JR, Aldrich J, Muller NL. Radiation exposure at chest CT: A statement of the Fleischner Society. *Radiology*. 2003;228(1):15-21.
12. American Thoracic Society, Infectious Diseases Society of America. Guidelines for the management of adults with hospital-acquired, ventilator-associated, and healthcare-associated pneumonia. *Am J Respir Crit Care Med*. 2005;171(4):388-416.
13. Woodhead M, Blasi F, Ewig S, et al. Guidelines for the management of adult lower respiratory tract infections. *Clin Microbiol Infect*. 2011;17Suppl6:E1-59.
14. Sartori S, Tombesi P. Emerging roles for transthoracic ultrasonography in pulmonary diseases. *World J Radiol*. 2010;2(6):203-214.
15. Chavez MA, Shams N, Ellington LE, et al. Lung ultrasound for the diagnosis of pneumonia in adults: A systematic review and meta-analysis. *Respir Res*. 2014;15(1).
16. Cortellaro F, Colombo S, Coen D, Duca PG. Lung ultrasound is an accurate diagnostic tool for the diagnosis of pneumonia in the emergency department. *Emerg Med J*. 2012;29(1):19-23.

17. Lichtenstein DA. Ultrasound examination of the lungs in the intensive care unit. *Pediatr Crit Care Med*. 2009;10(6):693-698.
18. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLOS Med*. 2009;6(7):e1000097.
19. Button KS, Loannidis JP, Mokrysz C, et al. Power failure: Why small sample size undermines the reliability of neuroscience. *Nat Rev Neurosci*. 2013;14(5):365-376.
20. Obuchowski NA, Mazzone PJ, Dachman AH. Bias, underestimation of risk, and loss of statistical power in patient-level analyses of lesion detection. *Eur Radiol*. 2010;20(3):584-594.
21. Busti C, Agnelli G, Duranti M, Orlandi C, Marcucci M, Paciaroni M. Lung ultrasound in the diagnosis of stroke-associated pneumonia. *Intern Emerg Med*. 2014;9(2):173-178.
22. Reissig A, Copetti R, Mathis G, et al. Lung ultrasound in the diagnosis and follow-up of community-acquired pneumonia. *Chest*. 2012;142(4):965-972.
23. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: A revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med*. 2011;155(8):529-536.
24. European Network for Health Technology Assessment, Eunethhta. Methodological guideline on meta-analysis of diagnostic test accuracy studies; <http://www.eunethta.eu/outputs/methodological-guideline-meta-analysis-diagnostic-test-accuracy-studies>. Published November 2014. Accessed May 2016.
25. R Core Team. R: A language and environment for statistical computing. R foundation for statistical computing version 3.2.5. [Software]. 2016; Vienna, Austria. Available from <https://www.R-project.org/>.
26. Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *J Clin Epidemiol*. 2005;58(10):982-990.
27. Rutter CM, Gatsonis CA. A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations. *Stat Med*. 2001;20(19):2865-2884.
28. Review Manager (RevMan) [Computer program]. Version 5.3. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2014
29. Liu X, Lian R, Tao Y, Gu C, Zhang G. Lung ultrasonography: An effective way to diagnose community-acquired pneumonia. *Emerg Med J*. 2015;32(6):433-438.
30. Nazerian P, Volpicelli G, Vanni S, et al. Accuracy of lung ultrasound for the diagnosis of consolidations when compared to chest computed tomography. *Am J Emerg Med*. 2015;33(5):620-625.
31. Unluer EE, Karagoz A, Senturk GO, Karaman M, Olow KH, Bayata S. Bedside lung ultrasonography for diagnosis of pneumonia. *Hong Kong J Emerg Med*. 2013;20(2):98-104.
32. Bataille B, Riu B, Ferre F, et al. Integrated use of bedside lung ultrasound and echocardiography in acute respiratory failure: A prospective observational study in ICU. *Chest*. 2014;146(6):1586-1593.

33. Bourcier J, Paquet J, Seinger M, et al. Performance comparison of lung ultrasound and chest x-ray for the diagnosis of pneumonia in the ED. *Am J Emerg Med.* 2014;32(2):115-118.
34. Berlet T, Etter R, Fehr T, Berger D, Sendi P, Merz TM. Sonographic patterns of lung consolidation in mechanically ventilated patients with and without ventilator-associated pneumonia: A prospective cohort study. *J Crit Care.* 2015;30(2):327-333.
35. Gallard E, Redonnet J, Bourcier J, et al. Diagnostic performance of cardiopulmonary ultrasound performed by the emergency physician in the management of acute dyspnea. *Am J Emerg Med.* 2015;33(3):352-358.
36. Lichtenstein DA, Mezière GA. Relevance of lung ultrasound in the diagnosis of acute respiratory failure the BLUE protocol. *Chest.* 2008;134(1):117-125.
37. Nazemi S, Farokhnia M, Pirouzi M, Najarion O, Ahmadipour H. Accuracy of ultrasound as a diagnostic tool of pneumonia for admitted patients in Infectious Disease Section of Afzalipour Hospital. *Jam J Sci Tech.* 2014;25:48-54.
38. Zagli G, Cozzolino M, Terreni A, Biagioli T, Caldini AL, Peris A. Diagnosis of ventilator-associated pneumonia: A pilot, exploratory analysis of a new score based on procalcitonin and chest echography. *Chest.* 2014;146(6):1578-1585.
39. Corradi F, Ball L, Brusasco C, et al. Lung ultrasonography fails detection of non-subpleural community acquired pneumonia. *Intensive Care Med.* 2012;38:S238-S238.
40. Fares A, Abo-Naghl A, Moustafa Z. Role of transthoracic ultrasound in detection of pneumonia in ICU patients. *Med J Cairo Univ.* 2015;83(1):307-314.
41. Nafae R, Eman SR, Mohamad NA, El-Ghamry R, Ragheb AS. Adjuvant role of lung ultrasound in the diagnosis of pneumonia in intensive care unit-patients. *Egypt J Chest Dis Tuberc.* 2013;62(2):281-285.
42. Self WH, Courtney DM, McNaughton CD, Wunderink RG, Kline JA. High discordance of chest x-ray and computed tomography for detection of pulmonary opacities in ED patients: Implications for diagnosing pneumonia. *Am J Emerg Med.* 2013;31(2):401-405.
43. Macaskill P, Gatsonis C, Deeks JJ, Harbord RM, Takwoingi Y. Chapter 10: Analysing and presenting results. In: *Cochrane Handbook for Systematic Reviews of diagnostic test accuracy version 1.0.* The Cochrane Collaboration, 2010.
44. Plana MN, Zamora J, Abaira V. Lectura crítica de revisiones sistemáticas y metanálisis en el diagnóstico por imagen. *Radiología.* 2015;57:23-30.
45. Xirouchaki N, Magkanas E, Vaporidi K, et al. Lung ultrasound in critically ill patients: Comparison with bedside chest radiography. *Intensive Care Med.* 2011;37(9):1488-1493.
46. Hu Q, Shen Y, Jia L, et al. Diagnostic performance of lung ultrasound in the diagnosis of pneumonia: A bivariate meta-analysis. *Int J Clin Exp Med.* 2014;7(1):115-121.
47. Copetti R, Cattarossi L. Ultrasound diagnosis of pneumonia in children. *Radiol Med.* 2008;113(2):190-198.
48. Pereda MA, Chavez MA, Hooper-Miele CC, et al. Lung ultrasound for the diagnosis of pneumonia in children: A meta-analysis. *Pediatrics.* 2015;135(4):714-722.

49. Ye X, Xiao H, Chen B, Zhang S. Accuracy of lung ultrasonography versus chest radiography for the diagnosis of adult community-acquired pneumonia: Review of the literature and meta-analysis. *PLOS ONE*. 2015;10(6).
50. Bouhemad B, Zhang M, Lu Q, Rouby JJ. Clinical review: Bedside lung ultrasound in critical care practice. *Crit Care*. 2007;11(1):205.
51. See KC, Ong V, Wong SH, et al. Lung ultrasound training: Curriculum implementation and learning trajectory among respiratory therapists. *Intensive Care Med*. 2016;42(1):63-71.
52. Lichtenstein D, Goldstein I, Mourgeon E, Cluzel P, Grenier P, Rouby JJ. Comparative diagnostic performances of auscultation, chest radiography, and lung ultrasonography in acute respiratory distress syndrome. *Anesthesiology*. 2004;100(1):9-15.
53. Gullett J, Donnelly JP, Sinert R, et al. Interobserver agreement in the evaluation of B-lines using bedside ultrasound. *J Crit Care*. 2015;30(6):1395-1399.
54. Volpicelli G, Elbarbary M, Blaivas M, et al. International evidence-based recommendations for point-of-care lung ultrasound. *Intensive Care Med*. 2012;38(4):577-591.
55. Guyatt GH, Oxman AD, Vist GE, et al. GRADE: An emerging consensus on rating quality of evidence and strength of recommendations. *BMJ*. 2008;336(7650):924-926.



SUPPLEMENTAL MATERIAL

e-Table 1. Performance of lung ultrasonography in the included studies.

Study	Ultrasound equipment	Lung areas examined/Patient position (if indicated)	Diagnostic criteria	Performs ultrasonography	Time to complete US (minutes)
Bataille 2014¹	HP Sonos 5500. Probes 2-4 MHz.	Anterior:1 area Lateral: 3 areas Semirecumbent position	AC ± AB ± PE Alveolointerstitial syndrome	Two researchers (ENS)	9 ± 2
Berlet 2015²	SonoSite M-turbo; SonoSite, Bothell, WA. Convex transducer C-60 3.5 MHz.	Lateral: 4° intercostal space MAL. Posterior:9° intercostal space PAL Semirecumbent position	AC	Four Intensivists (ENS)	10 to 15
Bourcier 2014³	SonoSite M-turbo, Bothell, WA. Convex transducer 3.5 MHz.	8 areas Semirecumbent position	Alveolar or interstitial pattern ± PE	Five trained emergency physicians	NS
Busti 2014⁴	Portable, NS. Convex probe 3.5-5 MHz y 7-10 MHz.	All intercostal spaces. Supine and lateral.	Subpleural consolidations ± AB. Interstitial pattern ± PE	A expertise physician	10 to 15
Corradi 2012⁵	LogiQ-E (GE Medical Systems), Convex probe (MHz NS)	NS	AC, focal interstitial pattern ± PE	NS	NS
Cortellaro 2012⁶	Esaote Medical Systems, convex probe 3.5-5 MHz	2 anterior areas, 2 lateral, 1 posterior each hemithorax	Subpleural consolidation± AB. Focal interstitial syndrome	A expertise physician	<5
Fares 2015⁷	Sonoscape B5, convex transducer 3-6 MHz.	6 areas/hemithorax: anterior, lateral, posterior.	Subpleural consolidation ± AB± PE.	One physician (ENS)	NS
Gallard 2015⁸	Sonosite Fujifilm M Turbo (Bothell, WA). Probe 3-5 MHz.	3 regions/hemithorax: anterior, lateral, posterior.	AC ±PE	Three emergency physicians who met the ultrasound training requirements of the American College of Emergency Physicians	12 ± 3
Lichtens Tein 2008⁹	Hitachi-405 (Hitachi Medical) microconvex probe 5 MHz.	6 areas/hemithorax. Anterior, lateral, posterior. Semirecumbent position, supine in intubated patients.	AC, interstitial	Two researchers (ENS)	<3

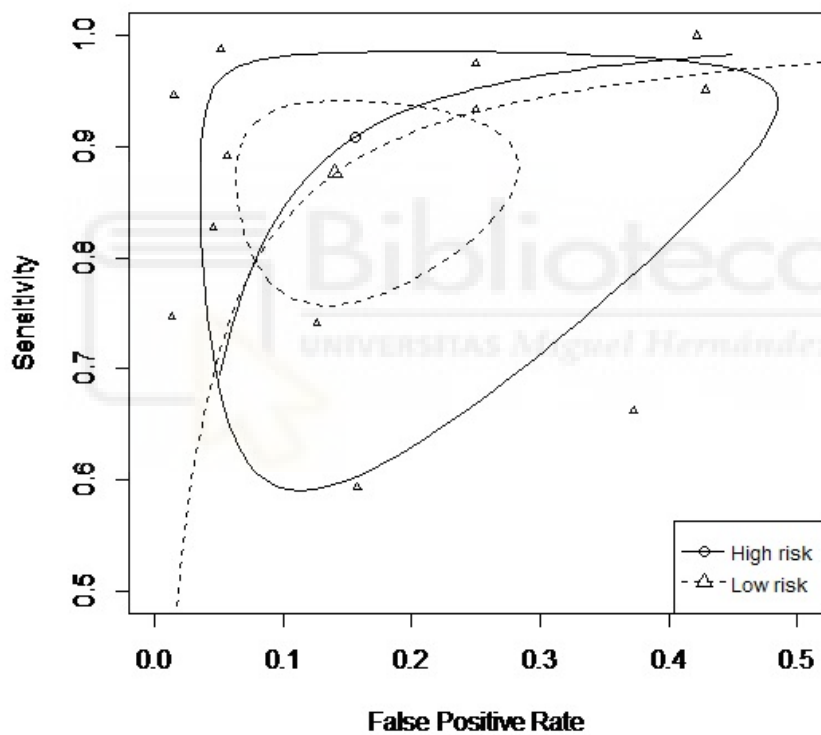
Liu 2015 ¹⁰	SonoSite M-turbo, Bothell, WA. Convex transducer 3.5-5 MHz.	All intercostal spaces (MCL, AAL, MAL and paravertebral line) from lung apex to the diaphragm.	AC, focal interstitial patten, subpleural lesions.	Three emergency physicians who met the ultrasound training requirements of the American College of Emergency Physicians	NS
Nafae 2013 ¹¹	Equipment NS, Convex transducer 3.5-5 MHz.	5 areas/hemithorax. Anterior, lateral, posterior.	AC ± AB and/or fluid bronchogram. Alveolointerstitial pattern±PE	One physician (ENS)	NS
Nazemi 2014 ¹²	Medison V10, probe 5 MHz	NS	Subpleural consolidation± AB ±PE	A radiologist (ENS)	NS
Nazerian 2015 ¹³	MyLab30 Gold (Esaote) and HD7 (Philips). Lineal probe 4-8 MHz , convex probe 3.5-5 MHz.	4 areas/ hemitórax: 2 Antero-lateral, 2 posterior. Sitting position. If not posible: semirecumbent or supine.	Subpleural consolidation± dynamic AB	4 Internal and Emergency medicine physicians with ≥ 5 years experience of POCUS and 4 resident physicians with ≥ 1 year training.	NS
Reissig 2012 ¹⁴	Equipment NS, convex transducer 3.5-5 MHz. Lineal probe 7.5 MHz.	All intercostal spaces anterior and posterior. Anterior examination: supine. Posterior examination: sitting position	AC ± necrotic areas, AB or fluid bronchogram ± PE.	Experienced physicians (have done > 100 chest ultrasonography procedures)	NS
Unluer 2013 ¹⁵	M7 (Mindray Biomedical Electronics) Microconvex transducer 3.6 MHz.	4 areas: 2 anterior, 1 lateral, 1 posterior.	AC. Interstitial pattern.	Three emergency physicians training taught by an experienced radiologist.	<10
Zagli 2014 ¹⁶	MyLab TM 30CV (Esaote) Convex probe 3.5-5 MHz.	All intercostal spaces from ápex to the base of the lung Supine.	AC	NS	NS

Abbreviations: AC: alveolar consolidation; AB: air bronchogram; AAL: anterior axillary line; ENS: expertise in lung ultrasound not specified; MAL: mild axillary line; MCL: medioclavícula line; NS: not stated; PAL: posterior axillary line; PE: pleural effusion. POCUS: point-of-care ultrasonography.

e-Appendix 1. Subgroup analyses

We did not perform a subgroup analysis related to the risk of bias in the implementation of the index test because none of the studies were considered to be at high risk, and only four were at unclear risk. Likewise, we did not perform a subgroup analysis for risk of bias in the patient selection domain because there were few studies (three) considered to be at high risk in this category.

- Risk of bias for reference standard



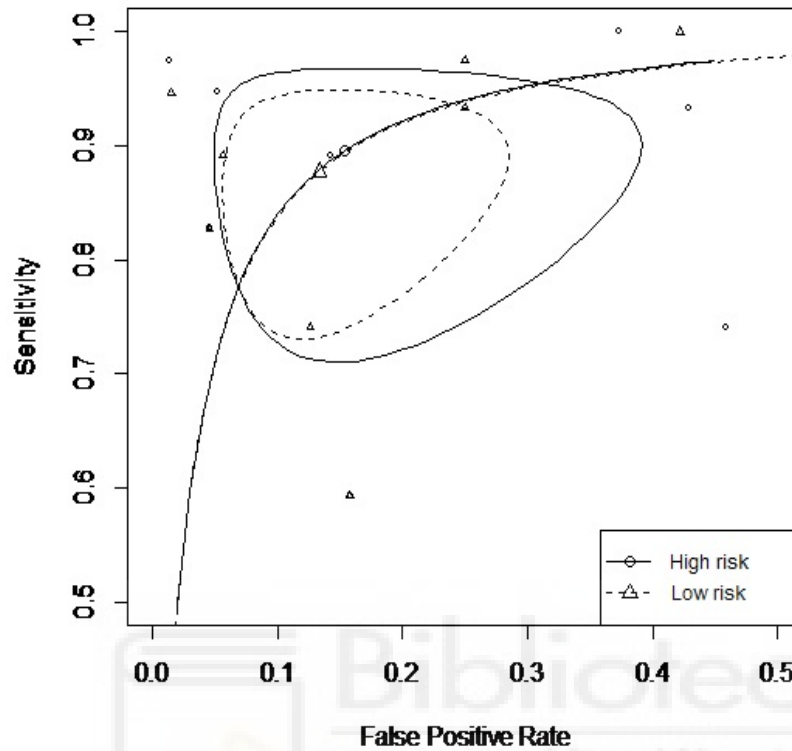
High risk versus low or unclear risk

tsens 0.133 (95% CI -1.264, 1.530), $p = 0.852$

tfpr 0.258 (95% CI -1.165, 1.682), $p = 0.722$

Abbreviations: CI: confidence interval; p : p value, significance < 0.05; tfpr: transformed false positive rate; tsens: transformed sensitivity.

- Risk of bias concerning patient flow and timing



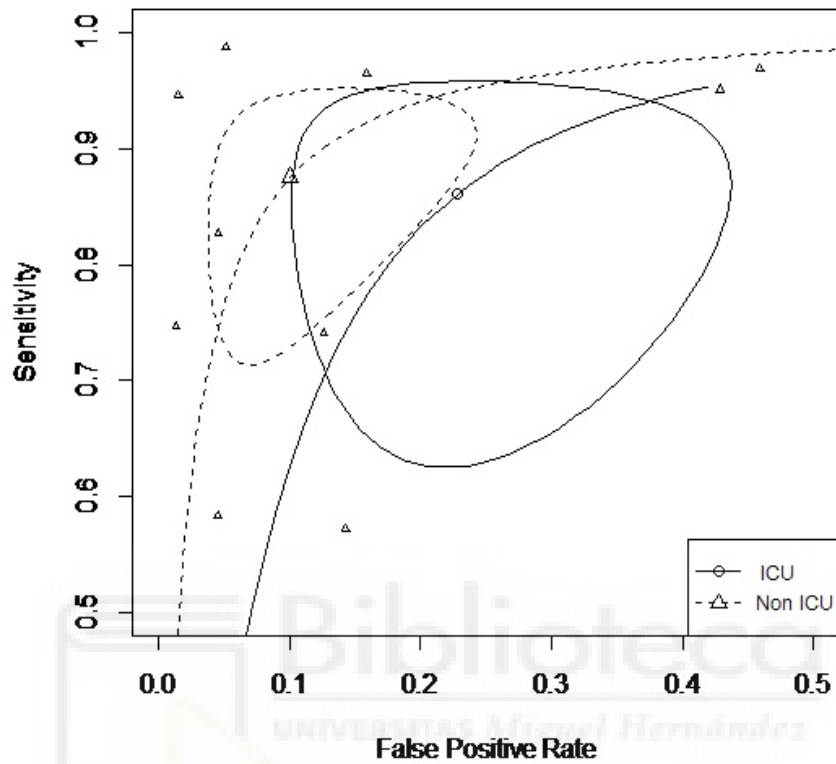
High risk versus low or unclear risk

tsens 0.063 (95% CI -1.133, 1.259), $p = 0.918$

tfpr 0.287 (95% CI -0.930, 1.505), $p = 0.644$

Abbreviations: CI: confidence interval; p : p value, significance < 0.05; tfpr: transformed false positive rate; tsens: transformed sensitivity.

- Illness severity.



ICU patients versus non-ICU patients

tsens -0.307 (95% CI $-1.537, 0.923$) $p = 0.624$

tfpr 0.884 (95% CI $-0.300, 2.068$) $p = 0.143$

Non-ICU patients: hospitalized or in the ER.

Abbreviations: CI: confidence interval; p : p value, significance < 0.05 ; tfpr: transformed false positive rate; tsens: transformed sensitivity.

SUPPLEMENTARY MATERIAL. REFERENCES.

1. Bataille B, Riu B, Ferre F, et al. Integrated use of bedside lung ultrasound and echocardiography in acute respiratory failure: A prospective observational study in ICU. *Chest*. 2014;146(6):1586-1593.
2. Berlet T, Etter R, Fehr T, Berger D, Sendi P, Merz TM. Sonographic patterns of lung consolidation in mechanically ventilated patients with and without ventilator-associated pneumonia: A prospective cohort study. *J Crit Care*. 2015;30(2):327-333.
3. Bourcier J, Paquet J, Seinger M, et al. Performance comparison of lung ultrasound and chest x-ray for the diagnosis of pneumonia in the ED. *Am J Emerg Med*. 2014;32(2):115-118.
4. Busti C, Agnelli G, Duranti M, Orlandi C, Marcucci M, Paciaroni M. Lung ultrasound in the diagnosis of stroke-associated pneumonia. *Intern Emerg Med*. 2014;9(2):173-178.
5. Corradi F, Ball L, Brusasco C, et al. Lung ultrasonography fails detection of non-subpleural community acquired pneumonia. *Intensive Care Med*. 2012;38:S238-S238.
6. Cortellaro F, Colombo S, Coen D, Duca PG. Lung ultrasound is an accurate diagnostic tool for the diagnosis of pneumonia in the emergency department. *Emerg Med J*. 2012;29(1):19-23.
7. Fares A, Abo-Naghl A, Moustafa Z. Role of transthoracic ultrasound in detection of pneumonia in ICU patients. *Med J Cairo Univ*. 2015;83(1):307-314.
8. Gallard E, Redonnet J, Bourcier J, et al. Diagnostic performance of cardiopulmonary ultrasound performed by the emergency physician in the management of acute dyspnea. *Am J Emerg Med*. 2015;33(3):352-358.
9. Lichtenstein DA, Mezière GA. Relevance of lung ultrasound in the diagnosis of acute respiratory failure the BLUE protocol. *Chest*. 2008;134(1):117-125.
10. Liu X, Lian R, Tao Y, Gu C, Zhang G. Lung ultrasonography: An effective way to diagnose community-acquired pneumonia. *Emerg Med J*. 2015;32(6):433-438.
11. Nafae R, Eman SR, Mohamad NA, El-Ghamry R, Ragheb AS. Adjuvant role of lung ultrasound in the diagnosis of pneumonia in intensive care unit-patients. *Egypt J Chest Dis Tuberc*. 2013;62(2):281-285.

12. Nazemi S, Farokhnia M, Pirouzi M, Najarian O, Ahmadipour H. Accuracy of ultrasound as a diagnostic tool of pneumonia for admitted patients in Infectious Disease Section of Afzalipour Hospital. *Jam J Sci Tech*. 2014;25:48-54.
13. Nazerian P, Volpicelli G, Vanni S, et al. Accuracy of lung ultrasound for the diagnosis of consolidations when compared to chest computed tomography. *Am J Emerg Med*. 2015;33(5):620-625.
14. Reissig A, Copetti R, Mathis G, et al. Lung ultrasound in the diagnosis and follow-up of community-acquired pneumonia. *Chest*. 2012;142(4):965-972.
15. Unluer EE, Karagoz A, Senturk GO, Karaman M, Olow KH, Bayata S. Bedside lung ultrasonography for diagnosis of pneumonia. *Hong Kong J Emerg Med*. 2013;20(2):98-104.
16. Zagli G, Cozzolino M, Terreni A, Biagioli T, Caldini AL, Peris A. Diagnosis of ventilator-associated pneumonia: A pilot, exploratory analysis of a new score based on procalcitonin and chest echography. *Chest*. 2014;146(6):1578-1585.



ACKNOWLEDGMENTS

Author contributions:

Dr Llamas-Álvarez 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, the literature search, data analysis, and drafting of the manuscript; and served as principal author.

Dr Tenza-Lozano contributed to the literature search.

Dr Latour-Pérez contributed to the study design, statistical analysis and revision of the manuscript.

All authors approved the final version of the manuscript.

Other contributions: The authors appreciate the assistance of Meggan Harris for her help with the edition and review of the manuscript.



8.2. ANEXO 2.

Lung ultrasound for the diagnosis of pneumonia.

Correspondence.

Jaime Latour-Pérez, MD, PhD, Ana M. Llamas-Álvarez, MD, Eva M. Tenza-Lozano, MD, PhD.

Please cite this article as: Latour-Perez J, Llamas-Álvarez A-M, Tenza-Lozano EM. Lung ultrasound for the diagnosis of pneumonia. *Eur J Emerg Med.* 2019 Apr;26(2):147-148. doi: 10.1097/MEJ.0000000000000567.



LUNG ULTRASOUND FOR THE DIAGNOSIS OF PNEUMONIA.

In a recent metaanalysis, Orso et al.¹ analyzed the diagnostic accuracy of lung ultrasound for the diagnosis of pneumonia. In this study, the authors compared their results with those of a systematic review carried out by our group², and attributed the differences to the non-inclusion of a large study (Zanobetti et al.³) and to the inclusion of studies conducted outside the emergency area, with the consequent heterogeneity.

First of all, we are surprised by the emphasis on the differences between the area under the ROC curve (AUC) of both studies, since they are quite similar (97% versus 93%) and do not justify a change in the conclusions.

Second, the study of Zanobetti et al. -published after our study- was not primarily designed to assess the diagnostic accuracy of ultrasound, and does not provide transparent data to construct the 2x2 diagnostic crosstable. Indeed, Orso et al. misinterpret the total number of patients with pneumonia (n= 1086) as the number of true positive; as a consequence, the sensitivity and specificity calculated by Orso et al. (100% and 99% respectively) does not correspond with that reported by Zanobetti et al. (88.5% and 91.6%).

Third, Orso et al. assume that the lower value of AUC of our study is mainly due to the inclusion of studies carried outside the Emergency department, and they regret that we do not communicate the heterogeneity of the study. The authors are right that we do not communicate I^2 , but this is due to the fact that I^2 in diagnostic accuracy metaanalysis is not recommended, since it does not differentiate between the real heterogeneity and that due to the threshold effect⁴. Anyway, the online supplement of our article includes the analysis of the AUC of the studies carried out in the ICU versus those performed outside the ICU, which showed no significant differences.

REFERENCES.

1. Orso D, Guglielmo N, Copetti R. Lung ultrasound in diagnosing pneumonia in the emergency department: a systematic review and meta-analysis. *Eur J Emerg Med* 2017.
2. Llamas-Alvarez AM, Tenza-Lozano EM, Latour-Perez J. Accuracy of Lung Ultrasonography in the Diagnosis of Pneumonia in Adults: Systematic Review and Meta-Analysis. *Chest* 2017; 151 (2):374-382.
3. Zanobetti M, Scorpiniti M, Gigli C, Nazerian P, Vanni S, Innocenti F, et al. Point-of-Care Ultrasonography for Evaluation of Acute Dyspnea in the ED. *Chest* 2017; 151 (6):1295-1301.
4. Macaskill P, Gatsonis C, Deeks J, Harbord R, Takwoingi Y. *Cochrane handbook for systematic reviews of diagnostic test accuracy*. Version 09 0 London: The Cochrane Collaboration 2010.





8.3. ANEXO 3.

Diaphragm and lung ultrasound to predict weaning outcome: systematic review and meta-analysis.

Original research.

Ana M. Llamas-Álvarez, MD, Eva M. Tenza-Lozano, MD, Jaime Latour-Pérez, MD, PhD.

Please cite this article as: Llamas-Álvarez AM, Tenza-Lozano EM, Latour-Pérez J. Diaphragm and lung ultrasound to predict weaning outcome: systematic review and meta-analysis. Chest. 2017 Aug 29. pii: S0012-3692(17)31482-4. doi: 10.1016/j.chest.2017.08.028

CONTENTS

- **Abbreviations.**
- **Abstract.**
- **Introduction.**
- **Methods.**
- **Results.**
- **Discussion.**
- **Conclusion.**
- **Tables.**
- **Figures.**
- **References.**
- **Supplemental material.**
- **Supplemental material's references.**
- **Acknowledgments**

DIAPHRAGM AND LUNG ULTRASOUND TO PREDICT WEANING OUTCOME: SYSTEMATIC REVIEW AND META-ANALYSIS.

Ana M. Llamas-Álvarez, MD^a; Eva M. Tenza-Lozano, MD^a; Jaime Latour-Pérez, MD, PhD^{a,b}.

^aIntensive Care Unit, Elche General University Hospital, Elche, Spain

^bDepartment of Clinical Medicine, Miguel Hernández University, Sant Joan d'Alacant, Spain

CORRESPONDENCE TO: Ana Llamas-Álvarez, Department of Intensive Care Medicine, Elche General University Hospital, Camino de la Almazara 11, 03203, Elche, Spain; email: llamasalvarez@yahoo.es

CONFLICT OF INTEREST: The authors have no conflicts of interest to declare.

FUNDING/SUPPORT: Department of Clinical Medicine, Miguel Hernández University.

PRIOR ABSTRACT PUBLICATION/PRESENTATION: No.

IRB APPROVAL/ PATIENT CONSENT: not provided as this is a review of previously published studies. No patients were enrolled.

ABBREVIATIONS.

AUSROC: area under the summary receiver operating characteristics curve.

CI: confidence interval.

COPD: chronic obstructive pulmonary disease

DE: diaphragmatic excursion.

DOR: diagnostic odds ratio.

DTF: diaphragm thickening fraction.

ICU: intensive care unit.

LUS: lung ultrasound.

MV: mechanical ventilation.

NIMV: non-invasive mechanical ventilation.

QUADAS-2: Quality Assessment of Diagnostic Accuracy Studies-2.

ROC: receiver operating characteristic.

SBT: spontaneous breathing trial.

SROC: summary receiver operating characteristic.



ABSTRACT.

Background: Deciding the optimal timing for extubation in patients who are mechanically ventilated can be challenging, and traditional weaning predictor tools are not very accurate. The aim of this systematic review and meta-analysis was to assess the accuracy of lung and diaphragm ultrasound for predicting weaning outcomes in critically ill adults.

Methods: MEDLINE, the Cochrane Library, Web of Science, Scopus, LILACS, Teseo, Tesis Doctorales en Red, and OpenGrey were searched, and the bibliographies of relevant studies were reviewed. Two researchers independently selected studies that met the inclusion criteria and assessed study quality in accordance with the Quality Assessment of Diagnostic Accuracy Studies-2 tool. The summary receiver-operating characteristic curve and pooled diagnostic OR (DOR) were estimated by using a bivariate random effects analysis. Sources of heterogeneity were explored by using predefined subgroup analyses and bivariate meta-regression.

Results: Nineteen studies involving 1,071 people were included in the study. For diaphragm thickening fraction, the area under the summary receiver-operating characteristic curve was 0.87, and DOR was 21 (95% CI, 11-40). Regarding diaphragmatic excursion, pooled sensitivity was 75% (95% CI, 65-85); pooled specificity, 75% (95% CI, 60-85); and DOR, 10 (95% CI, 4-24). For lung ultrasound, the area under the summary receiver-operating characteristic curve was 0.77, and DOR was 38 (95% CI, 7-198). Based on bivariate metaregression analysis, a significantly higher specificity for diaphragm thickening fraction and higher sensitivity for diaphragmatic excursion was detected in studies with applicability concerns.

Conclusions: Lung and diaphragm ultrasound can help predict weaning outcome, but its accuracy may vary depending on the patient subpopulation.

INTRODUCTION

Determining the optimal moment to extubate a critically ill patient remains a challenge, as premature removal of mechanical ventilation (MV) entails a high risk of weaning failure, prompting reintubation that subjects the patient to unnecessary haemodynamic and respiratory stress¹. On the other hand, delayed extubation increases the duration of MV and carries other risks (development of ventilator-associated pneumonia, tracheal injury, barotrauma, etc)². Thus, both early and delayed weaning are associated with increased mortality, stay in the intensive care unit (ICU), and economic cost³.

Current guidelines for weaning^{4,5} recommend the implementation of a spontaneous breathing trial (SBT) as a tool to predict weaning outcome. However, 13% to 26% of patients who are extubated following a successful SBT need to be reintubated within 48 hours^{6,7}. In the last few years, multiple indices and parameters have been proposed as predictors of weaning outcome, but none have shown more than modest prognostic accuracy^{8,9}.

Successful liberation from MV depends on several factors, as patients must be haemodynamically stable, have an adequate ventilation-perfusion ratio, have the ability to generate a strong cough and expectorate endotracheal secretions, and generate a reliable ventilator pattern. Diverse mechanisms can influence the development of respiratory distress after extubation, such as the decrease of aeration of the pulmonary parenchyma related to heart failure induced by SBT¹⁰, alterations in pulmonary compliance^{11,12}, and diaphragmatic dysfunction associated with MV¹³.

Ultrasound use in the ICU is an area of growing interest¹⁴ because of its portability, speed, safety, and the encouraging results obtained for managing multiple entities^{15,16}. As ultrasound provides both morphologic and functional information in real time, it may be useful to assess two important factors among those that can influence weaning: the status of aeration of the pulmonary parenchyma¹¹ and the functional status of the diaphragm¹⁷, eliciting clues on the probability of success when removing MV.

There are two proposed diaphragm sonographic predictors: the diaphragmatic excursion (DE)¹⁸, which measures the distance that the diaphragm is able to move during the respiratory cycle, and the diaphragm thickening fraction (DTF)¹⁹, which reflects variation in the thickness of the diaphragm during respiratory effort and is calculated as (thickness at end-inspiration – thickness at the end-expiration)/thickness at the end of the expiration. Ultrasound can detect the decrease of the aeration of the lung parenchyma due to cardiac, respiratory, or diaphragmatic origin. This can be quantified through the so-called lung ultrasound (LUS) score²⁰, a validated scale whose values range from 0 to 36 points, obtained from the sum of the grades assigned to different ventilation patterns observed in every area of the lung scan^{21,22}.

This systematic review and meta-analysis aims to assess the accuracy of the lung and diaphragm ultrasound, in particular the DE, DTF and LUS score, for predicting MV weaning outcomes in critically ill adults.

METHODS

Search strategy and selection criteria

In this systematic review and meta-analysis, we included any studies in participants aged 18 years and older, admitted to an ICU and subjected to invasive MV for at least 24 hours; patients underwent lung and/or diaphragmatic ultrasound and had available data on weaning outcome. Weaning failure was defined broadly as the need for reintubation with reconnection to invasive MV, unscheduled post-extubation noninvasive mechanical ventilation (NIMV), tracheostomy, death within the first 72 hours, or SBT failure. Since weaning success and failure are mutually exclusive, we defined weaning success as the absence of criteria for failure.

Our exclusion criteria were: non-primary studies, insufficient data to calculate a 2 × 2 table for sensitivity and specificity, studies with fewer than 20 participants and those in which the unit of analysis was not the patient (e.g. chest or lung regions).

Two researchers (AML-A and EMT-L) independently carried out a literature search to identify potentially relevant studies in MEDLINE, the Cochrane Library, the Web of Science, Scopus, and LILACS, as well as repositories of doctoral theses (Theseus, Tesis doctorales en red) and reviews of grey literature on Open Grey. We handsearched the bibliographies of relevant studies and in some cases attempted to contact authors of conference proceedings to obtain unpublished data or other authors to request clarifications.

Our search used the terms 'extubation', 'weaning', 'discontinuation of mechanical ventilation', 'disconnect of mechanical ventilation', 'ultrasound', 'ultrasonography' and 'echography'. We imposed no restrictions on the date of publication and included all papers published from database inception to November 2016. Only studies carried out in humans were eligible. Despite the fact that the present review focuses on adult patients, we did not apply an age filter to the search, intending instead to manually screen search results to increase comprehensiveness. Likewise, we did not limit the search by study design or language. We activated Google Scholar Alert to receive notifications about the publication of potentially relevant studies between November 2016 and April 2017. Discrepancies between researchers were resolved through consensus with a third expert researcher (JL-P).

We present results according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) recommendations²³.

Data analysis

Two researchers manually extracted the data necessary to build a 2 × 2 table, using raw data from each study on true positives, true negatives, false negatives, and false positives; reported sensitivity and specificity; and/or graphics. Given the dichotomous nature of the variable 'weaning outcome' and to facilitate comparison of data, we decided to build only tables regarding weaning success. Unlike with DTF and DE, high LUS scores predict weaning failure, low scores predict success and intermediate values indicate uncertainty, that is, different cutoff points are used for predicting success and failure, so data for both situations were extracted from the cutoff points provided.

Two researchers (AML-A and EMT-L) independently evaluated methodological quality using the QUADAS-2 tool²⁴, a validated method for assessing risk of bias across various domains as well as

factors that may affect applicability. We developed a complementary protocol with operational definitions of risk of bias in each of the explored areas (e-Table 1). We resolved discrepancies by consensus with a third researcher (JL-P).

We analysed data according to the European Network for Health Technology Assessment recommendations²⁵, using the Mada application in the R statistical package (version 3.3.2)²⁶ and Review Manager 5.3²⁷, developed by the Cochrane Collaboration. In the univariate analysis, we constructed a forest plot for sensitivity and specificity, calculating the diagnostic odds ratio (DOR). We explored the correlation between sensitivity and the false positive rate graphically (forest plot) and statistically by examining the Spearman correlation coefficient and its 95% confidence interval (CI). We analysed heterogeneity among studies graphically and through the I^2 statistic and the Q test. Data were presented in the Receiver Operating Characteristic (ROC) plane and, depending on the existence of threshold effect according to graphics and statistical evidence, we calculated the area under the curve or a pooled estimation of sensitivity and specificity. We used the bivariate Reitsma model²⁸, which in the absence of covariates is equivalent to the HSROC of Rutter and Gatsonis²⁹. In case of detecting an 'outlier', that is, a study with values for sensitivity and/or specificity ostensibly different from the rest, we performed a sensitivity analysis excluding that study. Possible causes of heterogeneity among studies were examined through an analysis of predefined subgroups, attending to quality and applicability criteria by using a bivariate meta-regression.

RESULTS

The initial database search yielded a total of 3819 references, and we identified 4 additional records from other sources. We also obtained 14 subsequent references from Google Scholar alerts. Therefore in total we obtained 3837 records. After removing duplicates and screening titles and abstracts, we considered the full text of 55 studies, of which we excluded 36 (reasons detailed in Fig 1). Thus, 19 references met the inclusion criteria and were included in qualitative and quantitative synthesis.

All the selected studies have a cohort design and took place between 2004 and 2017. Some studies measured more than one predictor: ten studies assessed DE; ten, DTF; and five, LUS score. The definition of 'weaning failure' is not standard, covering one or more of the following items in the first 48-72 hours: need for reintubation, non-scheduled post-extubation noninvasive mechanical ventilation (NIMV), death, tracheostomy, terminal extubation, extubation delay, and/or SBT failure (Table 1)^{12,30-47}. Most of the studies were carried out in polyvalent ICUs, but three studies^{31,36,44} took place in the respiratory ICU and one³⁸ in a high dependency unit. Ultrasounds were usually conducted on patients just before or during the SBT, but four studies^{12,36,40,44} only considered those individuals who successfully passed this test (Table 2).

Regarding the assessment of methodological quality (Fig 2), the main problem we detected was the presence of aspects that could compromise applicability due to patient selection, as different studies focused exclusively or primarily on patients with chronic obstructive pulmonary disease (COPD) or whose intubation was due to respiratory causes^{36,37,44}, on those with ICU-acquired weakness⁴¹, or on patients who had failed previous attempts of weaning^{38,39}. Some studies were at risk of selection bias as they excluded specific populations of patients, such as

those with severe ICU-acquired weakness^{12,45}, or those with any admission to ICU in the previous 12 months³⁰ (e-Table 2).

Overall, ultrasounds and their measurements were well conducted (e-Table 3), but one study³⁰ may have performed the procedure while patients were mechanically ventilated, which would not be suitable for assessing patients' respiratory efforts⁴⁸. In this review there is no gold standard test for comparison, so we considered the adequacy of the weaning failure definition made in each study to properly classify weaning outcome. In this sense, we considered only one study³⁸ to be at high risk of bias because its definition of weaning failure was just SBT failure, even though patients may require reintubation after a successful SBT. Generally, the doctors responsible for deciding to withdraw MV were blinded to ultrasound results, although four studies were unclear on this point^{30,37,43,45}. Concerning the flow of patients within each study, we considered one study³⁰ to be at high risk of bias because it excluded patients who died from the analysis but didn't clarify whether death occurred during weaning.

Given that ultrasound is subject to observer interpretation, we assume included studies used diverse thresholds for positivity. Thus we present sensitivity and specificity data separately for each study, with no global weighting⁴⁹. In any case we did not observe a pattern of graphically negative association between sensitivity and specificity in any predictor (Fig 3). Main results are shown in Table 3.

Regarding DTF, in the ROC plane studies showed moderate spread, and the confidence ellipse tended to be projected over the SROC curve, which supported the calculation of AUC at 0.87 (Figure 4A). There was one outlier in specificity⁴⁷, so we carried out a sensitivity analysis (e-Fig 1). For DE, there was no evidence of a threshold effect and low heterogeneity for sensitivity, with a pooled value of 75% (95% CI 65, 85). However, there were important differences among studies concerning specificity, so it was not appropriate to do a weighted estimation (Figure 4B). We performed a sensitivity analysis excluding an outlier in specificity³¹ (e-Fig 2). Concerning LUS, the AUC was 0.77; however, its representation in the ROC plane involves considerable uncertainty about results due to the small number of studies available (e-Fig 3).

In order to explore potential causes of heterogeneity, we conducted a subgroup analysis and bivariate meta-regression to consider the overall quality of the studies, and possible problems of applicability derived from patient selection observed in the QUADAS-2 evaluation (Fig 5). This analysis was performed for DTF and DE, since few studies reported on LUS. A 'high quality' subgroup comprises studies at low risk of bias in all QUADAS-2 domains. With regard to applicability, the 'applicable' group comprises those studies involving a general population of critically ill patients versus a 'non-applicable' group of studies performed on more specific subpopulations of critically ill patients (i.e., failed previous weaning attempts, COPD, etc), according to our QUADAS-2 evaluation. There were no significant differences between high and low quality subgroups for DTF or DE. Nevertheless, regarding applicability there was a significantly higher specificity for DTF and higher sensitivity for DE in the 'applicability concerns' subgroup. This finding suggests that DTF and DE perform better in determined subpopulations of individuals with a higher pre-test probability, compared to the general ICU population.

DISCUSSION

DTF is considered a good indicator of the diaphragmatic inhalation effort⁵⁰, and low values are associated with an increase in the duration of the MV, ICU stay and mortality⁵¹. Our data suggest that DTF is also a good predictor of weaning outcome, with overall consistency across studies, except for one outlier reporting lower specificity⁴⁷. This discordant result may be because investigators performed the ultrasound before the SBT, during a brief MV disconnection period, while most other studies conducted LUS during SBT. This would lend credence to the hypothesis that DTF varies as STB progresses due to diaphragmatic fatigue, that is, an early test would provide lower sensitivity to predict weaning failure. There is currently no available evidence that supports this possibility; however, one ongoing study could eventually help to clarify this issue⁵².

DE is associated with inspiratory volume⁵³ but does not correlate with other indexes of respiratory effort⁵⁰. Our data suggest a lower accuracy for DE compared to DTF in predicting weaning outcome and higher heterogeneity. Moreover, in patients undergoing MV, DTF reflects active diaphragm contraction⁵⁴, while DE comes from adding patients' respiratory effort to the pressure generated by the ventilator⁴⁸, so its use is only meaningful in the absence of ventilatory support. Moreover, DE may vary depending on posture, showing higher values when patients are supine versus seated, as well as when the abdominal and/or thoracic pressure is altered⁵⁵ (ascites, atelectasis, etc.).

Diaphragmatic ultrasound requires the use of high frequency probes with precision to a few millimeters, meaning that slight variations in measurement between observers can substantially affect the result. Despite being an observer-dependent technique, the available evidence suggests that both DTF and DE are reproducible measures^{56,57}.

In a recent systematic review, Zambon et al⁵⁸ assessed the usefulness of diaphragm ultrasound in ICU patients and, based on four studies, suggested a good performance as weaning indexes. Although overall data from this meta-analysis indicate good performance for both DTF and DE in predicting weaning outcomes, our exploration of heterogeneity and evaluation of applicability revealed a significantly higher specificity for DTF and higher sensitivity for DE in the 'applicability concerns' subgroup. This is a novel finding that contrasts with the enthusiasm shown in the current literature, suggesting that accuracy of DTF and DE may be overestimated due to a large number of studies performed in subpopulations with a higher pre-test probability of weaning failure, for example in COPD patients or those who have failed previous weaning attempts.

Other causes that may affect the weaning process involve alterations in the aeration of the pulmonary parenchyma (pulmonary oedema, pneumonia, atelectasis, etc.), which can be evaluated by LUS. Our findings suggest an excellent performance for LUS in predicting weaning outcome; however, readers should exercise caution when interpreting this information because there are few available studies, and great uncertainty still remains. We did not identify any studies assessing the reproducibility of LUS score.

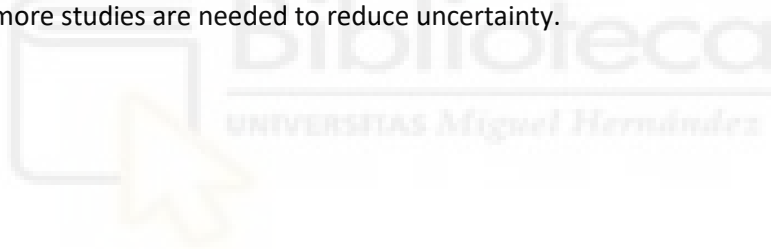
To our knowledge, this systematic review is the first to include a meta-analysis that assesses the accuracy of diaphragm and lung ultrasound to predict weaning outcome, providing novel data derived from an applicability assessment. Another strength of this study is its comprehensiveness, since the bibliographic search was not restricted by language, date of publication or participant age. We made a major effort to contact different authors to obtain unpublished data. On the other hand, the main limitation of this study is the small number of studies available on each ultrasound predictor, which is understandable because these are

incipient applications of ultrasound in the critical patient. This fact should be taken into account when interpreting subgroup analyses, as they are imprecise although indicative.

Our findings suggest that DTF, LUS score and to a lesser extent DE can provide valuable information for predicting weaning outcome, but taken alone, they may not perform as well as individual studies suggest. Diaphragmatic dysfunction and pulmonary aeration loss are two of the main potential causes of weaning failure, but they are not the only ones, and it is essential to contextualize the information obtained from ultrasound with clinical and laboratory data, as well as information derived from other image techniques such as echocardiography. Another aspect that should be taken into account is that all the reviewed studies implemented ultrasound only in patients previously classified as “ready to wean” by traditional assessment; however, it is unknown how many patients would meet this criterion according to ultrasound alone, while failing to fulfil traditional parameters. More high-quality studies that rigorously assess the ultimate role of diaphragm and lung ultrasound in critically ill patients – not only its accuracy and applicability – are needed.

CONCLUSION.

Our data suggest that DTF is by itself a modest predictor of weaning outcome in the general population of critically ill patients. We do not support the use of DE as its accuracy is lower and its measurement and interpretation entail several pitfalls. LUS score appears to be an accurate predictor, but more studies are needed to reduce uncertainty.



TABLES

Table 1. Characteristics of included studies. Measurements and cutoff values refer to weaning success prediction unless otherwise indicated.

Study/ Year	Country	Design	Measurements and cutoff value	Definition of weaning failure
Ali ³⁰ 2016	Egypt	Cohort	DTF >30% DE >1.5 cm	MV within 48 h of self-breathing
Baess ³¹ 2016	Egypt	Cohort	DTF ≥30% DE >1 cm	Reintubation or NIMV within 48 h
Binet ³² 2014	France	Cohort	LUS score ≤14 (≥19 for failure)	Reintubation, non-scheduled NIMV or death within 48 h
Blumhof ³³ 2016	USA	Cohort	DTF >20%	Reintubation or delayed extubation (>48 h)
Carrie ³⁴ 2017	France	Cohort	DE >2.7 cm	SBT failure or the need for MV or death within 48 h following extubation
Dinino ³⁵ 2014	USA	Cohort	DTF ≥ 30%	Reintubation within 48 h or terminal extubation or tracheostomy
Farghaly ³⁶ 2016	Egypt	Cohort	DTF ≥34.5% DE ≥1.5 cm	Inability to maintain spontaneous breathing without any ventilatory support within 48 h
Fayed ³⁷ 2016	Egypt	Cohort	DTF >29%	Inability to maintain spontaneous breathing without any ventilatory support within 48 h
Ferrari ³⁸ 2014	Italy	Cohort	DTF >36%	SBT failure
Flevari ³⁹ 2016	Greece	Cohort	DE ≥ 1 cm	Ventilatory support (non-invasive or invasive) within 48 h after SBT
Jiang ⁴⁰ 2004	Taiwan	Cohort	DE ≥1.1 cm	Reintubation or NIMV within 72 h
Jung ⁴¹ 2016	France	Cohort	DTF ≥ 30%	Reintubation within 72 h or tracheostomy
Kim ⁴² 2011	Korea	Cohort	DE <1.4 cm	MV within 48 h of self-breathing
Osman ⁴³ 2017	Egypt	Cohort	DTF ≥ 28% DE ≥ 1 cm LUS score <12	Reintubation within 48 h
Saeed ⁴⁴ 2016	Egypt	Cohort	DE >1.1 cm	MV within 48 h
Shoair ⁴⁵ 2016	Egypt	Cohort	LUS score ≤10 (>18 for failure)	Reintubation within 48 h
Soummer ¹² 2012	France	Cohort	LUS score <13 (>17 for failure)	Ventilatory support (either non-invasive or invasive ventilation) within 48 h after extubation

Spadaro ⁴⁶ 2016	Italy	Cohort	DE>1.4 cm	SBT failure, reintubation or NIMV within 48 h
Tenza ⁴⁷ 2017	Spain	Cohort	DTF>24%	SBT failure, reintubation or NIMV within 48 h

Abbreviations: DTF: diaphragmatic thickness fraction; DE: diaphragmatic excursion; MV: mechanical ventilation; NIMV: non-invasive mechanical ventilation; LUS: lung ultrasound; SBT: spontaneous breathing trial.



Table 2. Participant characteristics.

Study	N	Setting	Age* (years)	Inclusion
Ali ³⁰	54	MD ICU	54 ±11.23	Not specified
Baess ³¹	30	General and respiratory ICU	59.17±13.17	Patients who were planned for weaning
Binet ³²	48	MD ICU	59±16	Ventilated >48h
Blumhof ³³	52	ICU	62±17	Ventilated >24 h
Carrie ³⁴	67	1 Medical ICU 1 Medical and Surgical ICU	66 (58-74)	First SBT after 48 h of MV
Dinino ³⁵	63	2 Medical ICU	66±19	Ready for first SBT
Farghaly ³⁶	54	Respiratory ICU	65 (55-67.8) SG 62.5 (55-70.7) FG	Patients with underlying pulmonary disease causing ARF who had successfully passed the SBT
Fayed ³⁷	112	ICU	62.61±12.04	COPD patients
Ferrari ³⁸	46	High dependency unit	64.6±12.1	Patients with tracheostomy who had failed one or more weaning attempts
Flevari ³⁹	27	MD ICU	65 (53-75)	Difficult or prolonged weaning
Jiang ⁴⁰	55	Medical ICU	67 (33, 84)	Patients who had successfully passed the SBT
Jung ⁴¹	33	Medical and surgical ICU	58 (51-67)	Patients diagnosed for ICUAW with MV >48 h and undergoing a SBT
Kim ⁴²	82	Medical ICU	66	Patients ventilated >48 h and ready for a SBT
Osman ⁴³	68	Different ICU (Medical and Surgical)	56 (45-65)	Patients ready for SBT (most post-operatively)
Saeed ⁴⁴	30	Respiratory ICU	59±6	Patients intubated due to COPD who had successfully passed the SBT
Shoair ⁴⁵	50	ICU	47.52±14.60 SG 51.89±14.58 FG	Patients ventilated >48 h and ready for a first SBT

Soummer ¹²	86	2 MD ICU	61±14	Patients ventilated >48 h who successfully passed a first SBT
Spadaro ⁴⁶	51	ICU	65±13	Patients ventilated >48 h and ready for SBT
Tenza ⁴⁷	63	MD ICU	63±16	Patients ventilated >24 h and ready for SBT

Abbreviations: N: number of patients included; MD: multidisciplinary; ICU: intensive care unit; MV: mechanical ventilation; SBT: Spontaneous Breathing Trial; SG: success group; FG: failure group; ARF: acute respiratory failure; COPD: chronic obstructive pulmonary disease; ICUAW: intensive care unit acquired weakness.

*Age is expressed according to data extracted from each study as mean ± standard deviation or median (interquartile range).



Table 3. Main results regarding accuracy, correlation between sensitivity and specificity and heterogeneity.

		DTF	DE	LUS
Accuracy	Pooled sensitivity (95% CI)	NA	75% (65, 85%)	NA
	Pooled specificity (95% CI)	NA	NA	NA
	AUSROC	0.87	NA	0.77
	DOR (95% CI)	21 (11, 40)	10.6 (5, 24)	38 (7, 198)
Correlation sensitivity-specificity	Spearman's Rho (95% CI)	0.3 (-0.4, 0.7)	-0.45 (-0.84, 0.25)	0.2 (-0.8, 0.9)
Heterogeneity	Cochrane's Q (p value)	9.5 (p= 0.38)	10.7 (p= 0.29)	5.1 (p= 0.27)
	I²	6%	15.8%	22%

Abbreviations: DTF: diaphragm thickening fraction; DE: diaphragmatic excursion; LUS: lung ultrasound; 95% CI: 95% confidence interval; AUSROC: area under the summary operator receiver characteristic curve; NA: not applicable; DOR: diagnostic odds ratio.

FIGURES.

Figure 1. PRISMA Flow diagram for study identification and selection, with reasons for exclusion.

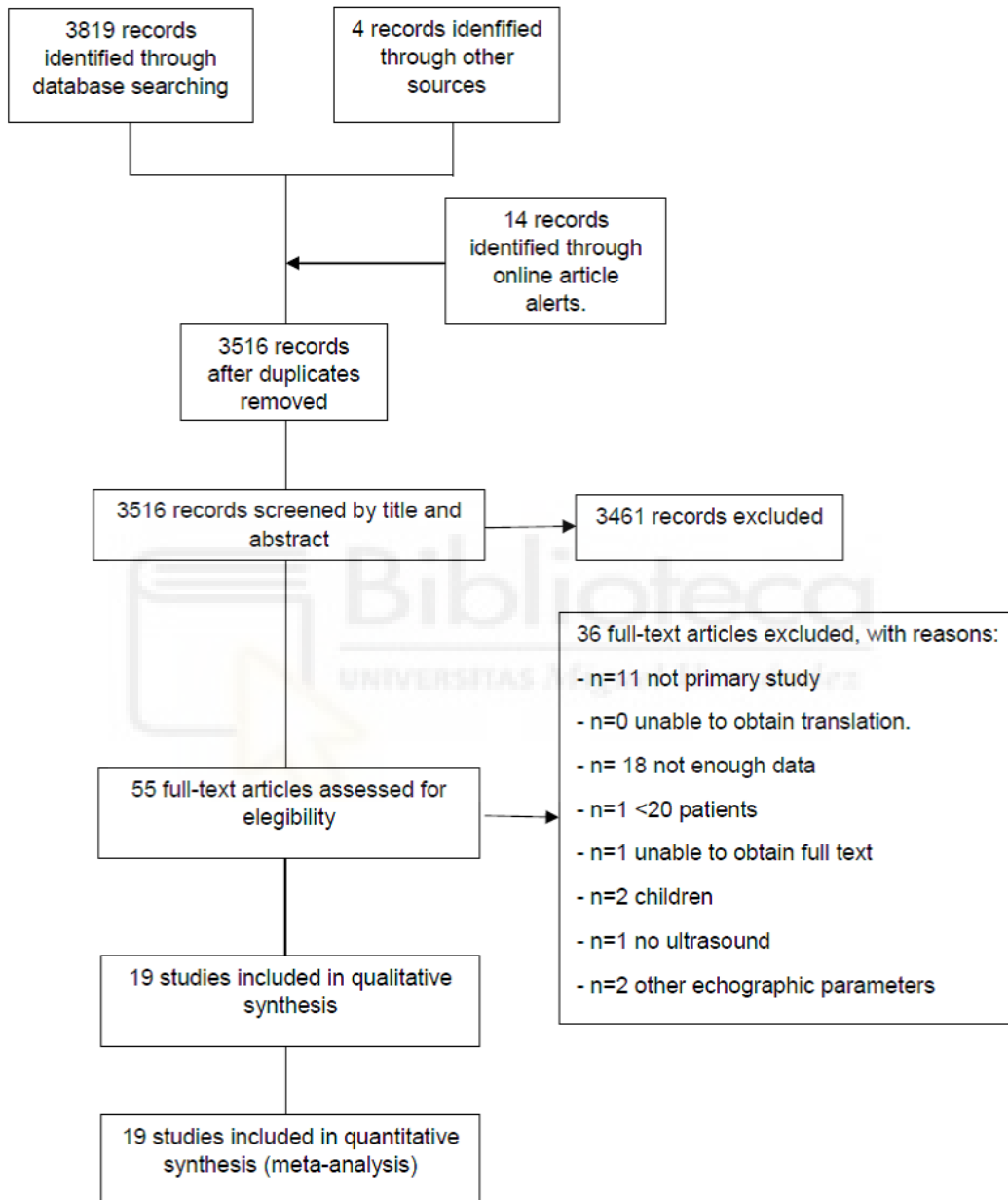


Figure 2. Methodological quality assessment according to QUADAS-2. For each study, risk of bias and applicability concerns are classified as high risk, low risk or unclear. A) Studies measuring DTF, B) Studies measuring DE, C) Studies measuring LUS score.

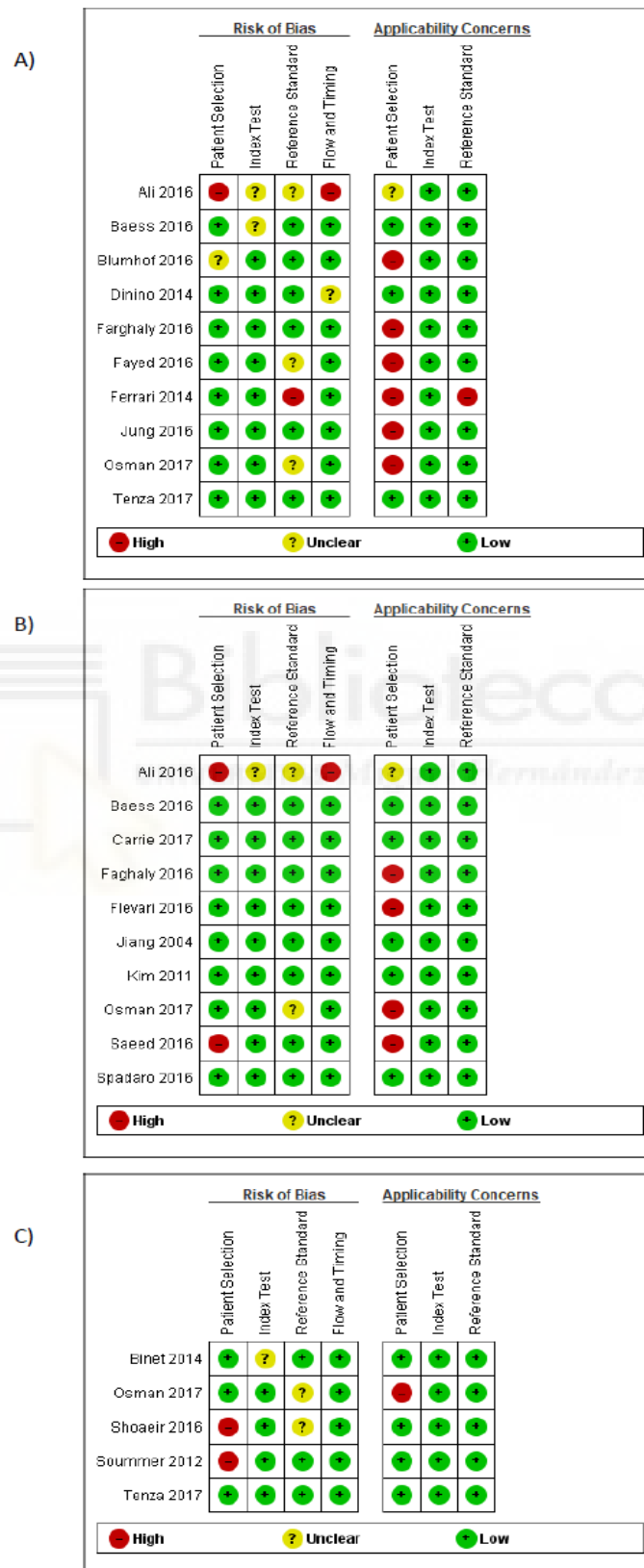
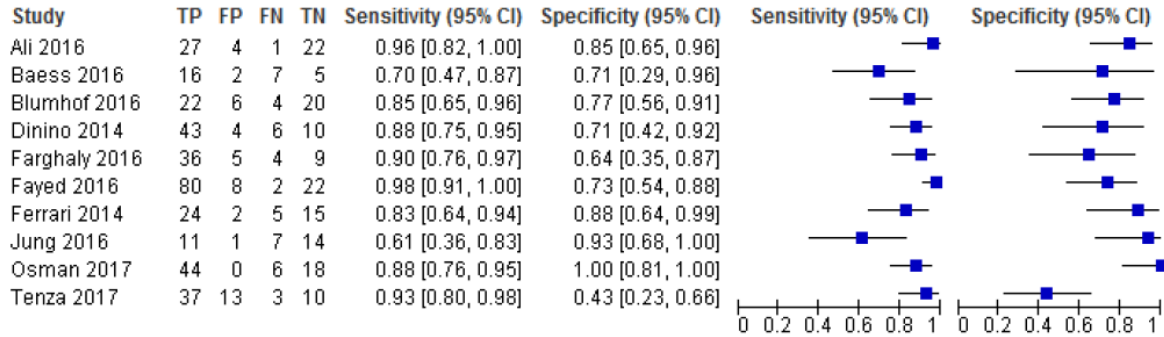
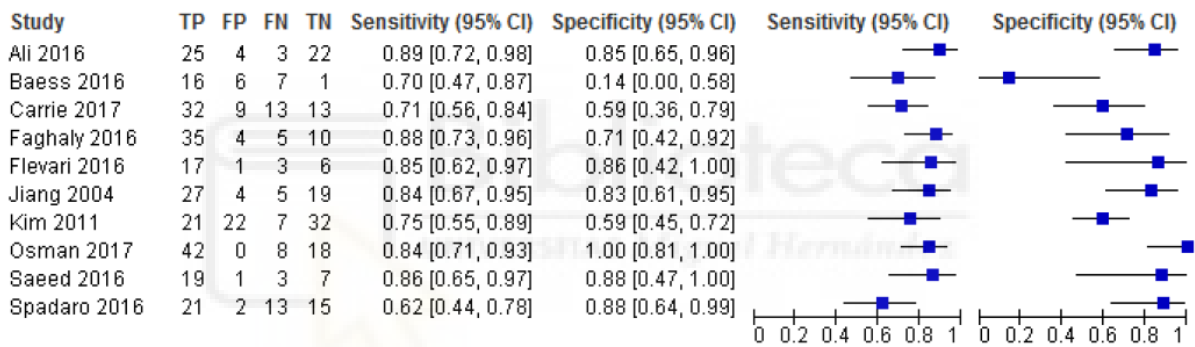


Figure 3. Forest plot of sensitivity and specificity in studies that measure A) diaphragm thickening fraction, B) diaphragm excursion, and C) LUS score.

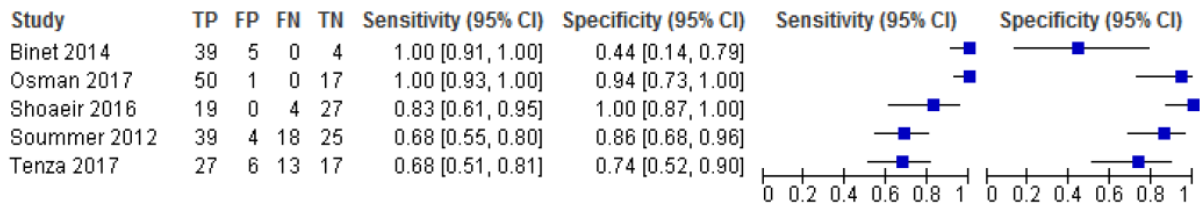
A) Diaphragm thickening fraction.



B) Diaphragm excursion.



C) LUS score.



Abbreviations: CI: confidence interval; LUS: lung ultrasound; TP: true positive; FP: false positive; TN: true negative; FN: false negative.

Figure 4. Summary receiver operating characteristic (SROC) curve of A) diaphragm thickening fraction and B) diaphragm excursion for weaning success prediction and confidence ellipse around the optimal cut-off value.

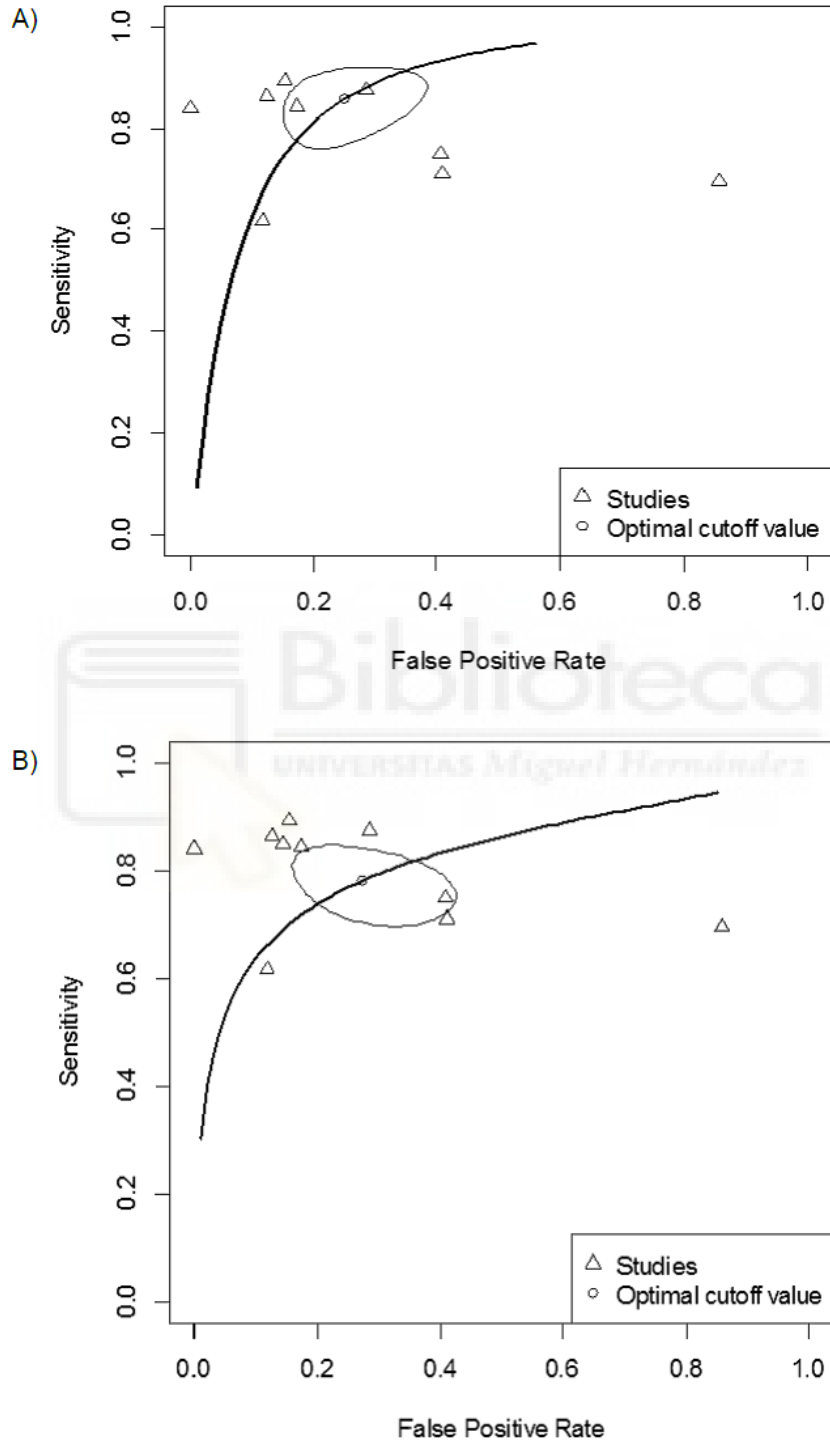
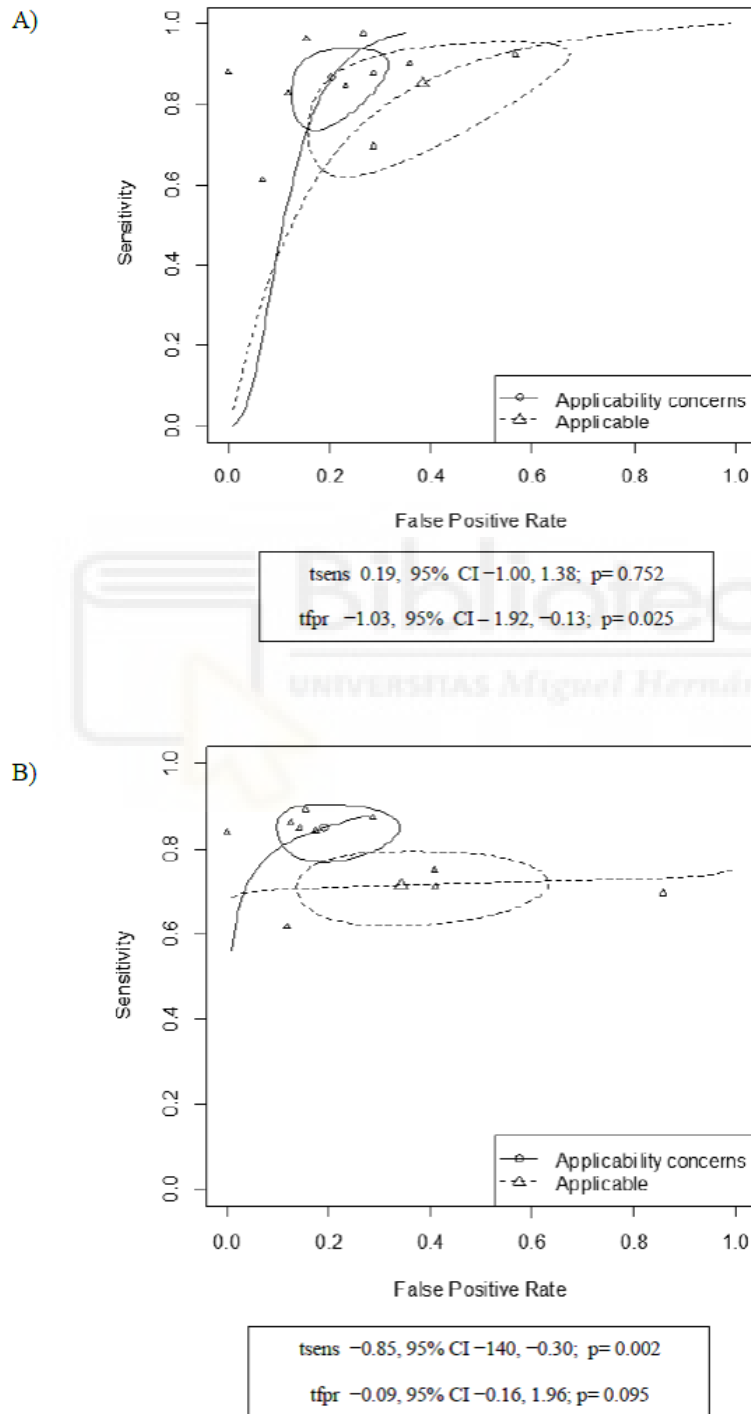


Figure 5. Subgroup analysis of applicability. Comparison of summary receiver operating characteristic (SROC) and confidence ellipses of applicable studies versus studies with applicability concerns, based on the analysis of methodological quality using the QUADAS-2 tool. Studies were considered applicable when there was low risk of concerns regarding applicability. A) Studies that measured DTF, B) Studies that measured DE. Statistical significance was set at $P < 0.05$.



Abbreviations: tsens: transformed sensitivity; tfpr: transformed false-positive rate; 95% CI: 95% confidence interval.

REFERENCES

1. Lemaire F, Teboul JL, Cinotti L, et al. Acute left ventricular dysfunction during unsuccessful weaning from mechanical ventilation. *Anesthesiology*. 1988;69(2):171-179.
2. Caroleo S, Agnello F, Abdallah K, Santangelo E, Amantea B. Weaning from mechanical ventilation: An open issue. *Minerva Anestesiol*. 2007;73(7-8):417-427.
3. Epstein SK, Ciubotaru RL, Wong JB. Effect of failed extubation on the outcome of mechanical ventilation. *Chest*. 1997;112(1):186-192.
4. Ouellette DR, Patel S, Girard TD, et al. Liberation from mechanical ventilation in critically ill adults: an Official American College of Chest physicians/American Thoracic Society clinical practice guideline: inspiratory pressure augmentation during spontaneous breathing trials, protocols minimizing sedation, and noninvasive ventilation immediately after extubation. *Chest*. 2017;151(1):166-180.
5. Boles JM, Bion J, Connors A, et al. Weaning from mechanical ventilation. *Eur Respir J*. 2007;29(5):1033-1056.
6. Frutos-Vivar F, Ferguson ND, Esteban A, et al. Risk factors for extubation failure in patients following a successful spontaneous breathing trial. *Chest*. 2006;130(6):1664-1671.
7. Ouanes-Besbes L, Dachraoui F, Ouanes I, et al. NT-proBNP levels at spontaneous breathing trial help in the prediction of post-extubation respiratory distress. *Intensive Care Med*. 2012;38(5):788-795.
8. Meade M, Guyatt G, Cook D, et al. Predicting success in weaning from mechanical ventilation. *Chest*. 2001;120(6 Suppl):400S-24S.
9. Conti G, Montini L, Pennisi MA, et al. A prospective, blinded evaluation of indexes proposed to predict weaning from mechanical ventilation. *Intensive Care Med*. 2004;30(5):830-836.
10. Teboul J, Monnet X, Richard C. Weaning failure of cardiac origin: Recent advances. *Critical Care*. 2010;14(2):211-211.
11. Bouhemad B, Brisson H, Le-Guen M, Arbelot C, Lu Q, Rouby JJ. Bedside ultrasound assessment of positive end-expiratory pressure-induced lung recruitment. *Am J Respir Crit Care Med*. 2011;183(3):341-347.
12. Soummer A, Perbet S, Brisson H, et al. Ultrasound assessment of lung aeration loss during a successful weaning trial predicts postextubation distress. *Crit Care Med*. 2012;40(7):2064-2072.
13. Martin AD, Smith B, Gabrielli A. Mechanical ventilation, diaphragm weakness and weaning: A rehabilitation perspective. *Respir Physiol Neurobiol*. 2013;189(2):377-383.
14. Mayo P, Arntfield R, Balik M, et al. The ICM research agenda on critical care ultrasonography. *Intensive Care Med*. 2017. doi: 10.1007/s00134-017-4734-z
15. Lichtenstein DA. BLUE-protocol and FALLS-protocol: Two applications of lung ultrasound in the critically ill. *Chest*. 2015;147(6):1659-1670.
16. Llamas-Álvarez AM, Tenza-Lozano EM, Latour-Pérez J. Accuracy of lung ultrasonography in the diagnosis of pneumonia in adults: Systematic review and meta-analysis. *Chest*. 2017;151(2):374-382.

17. Dube BP, Dres M. Diaphragm dysfunction: Diagnostic approaches and management strategies. *J Clin Med*. 2016;5(12):E113.
18. Boussuges A, Gole Y, Blanc P. Diaphragmatic motion studied by M-mode ultrasonography: Methods, reproducibility, and normal values. *Chest*. 2009;135(2):391-400.
19. Summerhill EM, El-Sameed YA, Glidden TJ, McCool FD. Monitoring recovery from diaphragm paralysis with ultrasound. *Chest*. 2008;133(3):737-743.
20. Bouhemad B, Liu Z-, Arbelot C, et al. Ultrasound assessment of antibiotic-induced pulmonary reaeration in ventilator-associated pneumonia. *Crit Care Med*. 2010;38(1):84-92.
21. Via G, Storti E, Gulati G, Neri L, Mojoli F, Braschi A. Lung ultrasound in the ICU: From diagnostic instrument to respiratory monitoring tool. *Minerva Anestesiol*. 2012;78(11):1282-1296.
22. Mayo P, Volpicelli G, Lerolle N, Schreiber A, Doelken P, Vieillard-Baron A. Ultrasonography evaluation during the weaning process: The heart, the diaphragm, the pleura and the lung. *Intensive Care Med*. 2016;42(7):1107-1117.
23. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLOS Med*. 2009;6(7):e1000097.
24. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: A revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med*. 2011;155(8):529-536.
25. European Network for Health Technology Assessment, Eunethhta. Methodological Guideline on "Meta-analysis of diagnostic test accuracy studies." Available at <http://www.eunethhta.eu/outputs/methodological-guideline-meta-analysis-diagnostic-test-accuracy-studies>. Accessed March 13, 2017.
26. R Core Team (2016). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
27. Review Manager (RevMan), version 5.3. (2014) Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration.
28. Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *J Clin Epidemiol*. 2005;58(10):982-990.
29. Rutter CM, Gatsonis CA. A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations. *Stat Med*. 2001;20(19):2865-2884.
30. Ali ER, Mohamad AM. Diaphragm ultrasound as a new functional and morphological index of outcome, prognosis and discontinuation from mechanical ventilation in critically ill patients and evaluating the possible protective indices against VIDD. *J Chest Dis Tuberc*. 2016. 66:339-351
31. Baess AI, Abdallah TH, Emara DM, Hassan M. Diaphragmatic ultrasound as a predictor of successful extubation from mechanical ventilation: Thickness, displacement, or both? *Egypt J Bronchol*. 2016;10:162±166.

32. Binet C, Neron L, Rochart N, et al. Validation d'un indice échographique prédictif d'échec de sevrage de la ventilation mécanique. *Annales Françaises D Anesthésie Et De Réanimation*. 2014;33:A383-A383.
33. Blumhof S, Wheeler D, Thomas K, McCool FD, Mora J. Change in diaphragmatic thickness during the respiratory cycle predicts extubation success at various levels of pressure support ventilation. *Lung*. 2016;194(4):519-525.
34. Carrie C, Gisbert-Mora C, Bonnardel E, et al. Ultrasonographic diaphragmatic excursion is inaccurate and not better than the MRC score for predicting weaning-failure in mechanically ventilated patients. *Anaesth Crit Care Pain Med*. 2017;36(1):9-14.
35. Dinino E, Gartman EJ, Sethi JM, McCool FD. Diaphragm ultrasound as a predictor of successful extubation from mechanical ventilation. *Thorax*. 2014;69(5):423-427.
36. Farghaly S, Hasan AA. Diaphragm ultrasound as a new method to predict extubation outcome in mechanically ventilated patients. *Aust Crit Care*. 2017;30(1):37-43.
37. Fayed AM, Abd El Hady MA, Shaaban MS, Fikry DM. Use of ultrasound to assess diaphragmatic thickness as a weaning parameter in invasively ventilated chronic obstructive pulmonary disease patients. *J Am Sci*. 2016;12(6):96-105.
38. Ferrari G, De Filippi G, Elia F, Panero F, Volpicelli G, Aprà F. Diaphragm ultrasound as a new index of discontinuation from mechanical ventilation. *Crit Ultrasound J*. 2014;6(1):8-7902-6-8.
39. Flevari A, Lignos M, Konstantonis D, Armaganidis A. Diaphragmatic ultrasonography as an adjunct predictor tool of weaning success in patients with difficult and prolonged weaning. *Minerva Anestesiologica*. 2016;82(11):1149-1157.
40. Jiang JR, Tsai TH, Jerng JS, Yu CJ, Wu HD, Yang PC. Ultrasonographic evaluation of liver/spleen movements and extubation outcome. *Chest*. 2004;126(1):179-185.
41. Jung B, Moury PH, Mahul M, et al. Diaphragmatic dysfunction in patients with ICU-acquired weakness and its impact on extubation failure. *Intensive Care Med*. 2016;42(5):853-861.
42. Kim WY, Suh HJ, Hong SB, Koh Y, Lim CM. Diaphragm dysfunction assessed by ultrasonography: Influence on weaning from mechanical ventilation. *Crit Care Med*. 2011;39(12):2627-2630.
43. Osman AM, Hashim RM. Diaphragmatic and lung ultrasound application as new predictive indices for the weaning process in ICU patients. *Egypt J Radiol Nucl Med*. 2017;48(1):61-66.
44. Saeed AM, El Assal GI, Ali TM, Hendawy MM. Role of ultrasound in assessment of diaphragmatic function in chronic obstructive pulmonary disease patients during weaning from mechanical ventilation. *Egypt J Bronchol*. 2016;10:167-172.
45. Shoaer M, Noeam K, Mahrous A, Alaa A. Lung aeration loss as a predictor of reintubation using lung ultrasound in mechanically ventilated patients. *Biolife*. 2016; 4(3), pp 514-520.
46. Spadaro S, Grasso S, Mauri T, et al. Can diaphragmatic ultrasonography performed during the T-tube trial predict weaning failure? the role of diaphragmatic rapid shallow breathing index. *Critical Care*. 2016;20:305.

47. Tenza-Lozano EM, Jaimez E, Latour-Pérez J. Ecografía pulmonar y diafragmática como predictores de weaning. (Material gathered from an unpublished draft by Eva María Tenza-Lozano). Forthcoming 2017.
48. Matamis D, Soilemezi E, Tsagourias M, et al. Sonographic evaluation of the diaphragm in critically ill patients. Technique and clinical applications. *Intensive Care Med.* 2013;39(5):801-810.
49. Macaskill P, Gatsonis C, Deeks JJ, Harbord RM, Takwoingi Y. Chapter 10: Analysing and presenting results. in: Deeks JJ, Bossuyt PM, Gatsonis C (editors), *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy* version 1.0. The Cochrane Collaboration, 2010.
50. Umbrello M, Formenti P, Longhi D, et al. Diaphragm ultrasound as indicator of respiratory effort in critically ill patients undergoing assisted mechanical ventilation: A pilot clinical study. *Crit Care.* 2015;19:161-015-0894-9.
51. Dubé BP, Dres M, Mayaux J, Demiri S, Similowski T, Demoule A. Ultrasound evaluation of diaphragm function in mechanically ventilated patients: Comparison to phrenic stimulation and prognostic implications. *Thorax.* 2017. Published Online First: 30 March 2017. doi:10.1136/thoraxjnl-2016-209459
52. Zhou P, Zhang Z, Hong Y, et al The predictive value of serial changes in diaphragm function during the spontaneous breathing trial for weaning outcome: a study protocol. *BMJ Open.* 2017;7:e015043. doi:10.1136/bmjopen-2016-015043
53. Cohen E, Mier A, Heywood P, Murphy K, Boulton J, Guz A. Excursion-volume relation of the right hemidiaphragm measured by ultrasonography and respiratory airflow measurements. *Thorax.* 1994;49(9):885-889.
54. Goligher EC, Fan E, Herridge MS, et al. Evolution of diaphragm thickness during mechanical ventilation: Impact of inspiratory effort. *Am J Respir Crit Care Med.* 2015;192(9):1080-1088.
55. Haji K, Royse A, Green C, Botha J, Canty D, Royse C. Interpreting diaphragmatic movement with bedside imaging, review article. *J Crit Care.* 2016;34:56-65.
56. Francis CA, Hoffer JA, Reynolds S. Ultrasonographic evaluation of diaphragm thickness during mechanical ventilation in intensive care patients. *Am J Crit Care.* 2016;25(1):e1-8.
57. Bobbia X, Clement A, Claret PG, et al. Diaphragmatic excursion measurement in emergency patients with acute dyspnea: Toward a new diagnostic tool? *Am J Emerg Med.* 2016;34(8):1653-1657.
58. Zambon M, Greco M, Bocchino S et al. Assessment of diaphragmatic dysfunction in the critically ill patient with ultrasound: a systematic review. *Intensive Care Med.* 2017;43:29. doi:10.1007/s00134-016-4524-z

SUPPLEMENTARY MATERIAL.

e-Table 1. Methodological assessment protocol using QUADAS-2 tool.

DOMAIN 1: PATIENT SELECTION	
<i>Risk of bias</i>	
Describe methods of patient selection: how patients were recruited, whether or not they are consecutive, where they come from (intensive care unit (ICU), Emergency room, etc.), exclusion criteria.	
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? Assess potential problems related to exclusion criteria.	
Yes (low risk of bias)	Excluded patients were highly unrepresentative of the general critically ill population. Examples: patients with neuromuscular diseases, impossible to obtain acoustic window (pneumothorax, subcutaneous emphysema, etc.), problems to access the anatomical area of exploration (bandages, skin disorders, wounds, etc.), intubation due to airway obstruction/stridor as the problem lies not in the diaphragm nor in the pulmonary parenchyma.
No (high risk of bias)	Excluding these patients could imply a selection bias. Examples: multi-pathological patients, those with ICU-acquired weakness, administration of common drugs as corticosteroids or aminoglycosids, etc.
Unclear	Insufficient data
Summary: could the selection of patients have introduced biases? Rating according to the above points. Risk: low/ high/ unclear.	
<i>Concerns regarding applicability</i>	
Describe included patients (prior testing, presentation, intended use of index test and setting): Assess potential problems related to inclusion criteria.	
Summary: Is there concern that the included patients do not match the review question? That is to say, general critically ill patients aged 18 years or older subjected to invasive mechanical ventilation during at least 24-48 hours.	
Low concern:	Included patients match the review question. For example, if only patients who have passed the test of spontaneous breathing (SBT) are included: it seems appropriate because it is not acceptable to extubate patients who do not pass the SBT.
High concern:	Included patients do not match the review question. For example, if there is an exclusive selection of patients who failed the SBT or who have already failed previous weaning attempts, or exclusively COPD, ICUAW, etc: it seems inappropriate since it would increase pre-test probability of weaning failure.
Unclear concern:	Insufficient data.
DOMAIN 2: INDEX TESTS	
Ultrasound measurement of diaphragm thickening fraction, diaphragm excursion and/or lung ultrasound (LUS) score. One form is filled in for each test.	
<i>Risk of bias</i>	
Describe the index test and how it was conducted and interpreted: note measurement and methodology of ultrasound, index calculation and timing of implementation.	
Were the index test results interpreted without knowledge of the results of the reference standard? Given the nature of the elected gold standard (see domain 3), the result of the gold standard measure could never be previously known, as ultrasound is performed first.	
If a threshold was used, was it pre-specified? Some studies provide a predefined threshold based on previously published references and others calculate the best cut-off value from the ROC curve. In this review this aspect does not especially affect validity.	

Could the conduct or interpretation of the index test have introduced bias?	
No (Low risk)	Ultrasound was performed during the SBT or just prior to extubation.
Yes (High risk)	Ultrasound was performed long time before the SBT or during assisted invasive mechanical ventilation.
Unclear risk	Use of a suboptimal probe (optimal probes are 3-5 MHz for lung ultrasound and 8-10 MHz for diaphragm ultrasound), insufficient data.
Concerns regarding applicability	
Is there concern that the index test, its conduct, or interpretation differ from the review question?	
Low concern	Ultrasound measurement of diaphragm thickening fraction, diaphragm excursion and/or LUS score are assessed according to current recommendations [1, 2].
High concern	Ultrasound measurement of diaphragm thickening fraction, diaphragm excursion and/or LUS score measurements substantially differ from current recommendations ^{1,2} .
Unclear	Insufficient data
DOMAIN 3: REFERENCE STANDARD	
<u>Risk of bias</u>	
Describe the reference standard and how it was conducted and interpreted: how the authors define weaning failure/success.	
Is the reference standard likely to correctly classify the target condition? i.e., is the provided definition of weaning failure/success able to properly classify weaning outcome?	
Yes (low risk of bias)	The provided definition must necessarily consider as weaning failure the need of reintubation/reconnection to invasive mechanical ventilation in the first 48 hours. In addition to this, other items can be included, as the need of non-invasive mechanical ventilation, tracheostomy, death, SBT failure, etc.
No (high risk of bias)	The provided definition only considers other situations different from the need of reintubation/reconnection to invasive mechanical ventilation in the first 48 hours.
Unclear	Insufficient data
Were the reference standard results interpreted without knowledge of the results of the index test? I.e., are physicians who take the decision to extubate aware of the information provided by the ultrasound? Yes/No/Unclear	
Could the reference standard, its conduct, or its interpretation have introduced bias? I. e., could the provided definition of weaning success/failure in this study have introduced bias? Rating according to the above points.	
Risk: low/ high/ uncertain.	
Concerns regarding applicability	
Is there concern that the target condition as defined by the reference standard does not match the review question? Potential problems with reference standard in this review may imply risk of bias but not applicability concerns, as any weaning failure/success definition (optimal or not) are accepted. Thus, this question seems irrelevant in this review.	
Concern: low /high/unclear	
DOMAIN 4: FLOW AND TIMING	
<u>Risk of bias</u>	
Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2 x 2 table (refer to flow diagram): exclusions for reasons other than the prespecified exclusion criteria, e.g. patients that don't receive ultrasound (index test) or are not classified as weaning failure or success.	
Describe the time interval and any interventions between index test(s) and reference standard: Time elapsed between ultrasound performance and weaning, and any intervention during the ultrasound.	
Was there an appropriate interval between index test(s) and reference standard?	

Yes (low risk of bias)	It is optimal when ultrasound is performed immediately prior to extubation, as sonographic measured parameters (diaphragm thickening fraction, diaphragmatic excursion and/or LUS score) are dynamic.
No (high risk of bias)	Long time intervals. It is not possible to establish a priori a threshold above which is considered an inappropriate time interval due to the absence of evidence about the change in the values of diaphragm thickening fraction, diaphragmatic excursion and LUS score at intervals less than 24 hours.
Unclear	Insufficient data
Did all patients receive a reference standard? Did patients receive the same reference standard? Both questions refer to patients that may not be finally classified as weaning failure or success.	
Yes (low risk of bias)	All patients were finally classified as weaning failure or success and the same definition of weaning failure/success is applied to all the participants.
No (high risk of bias)	All patients were not finally classified as weaning failure or success and the same definition of weaning failure/success is not applied to all the participants.
Unclear	Insufficient data
Could the patient flow have introduced bias? Decide based on that previously answered.	

Abbreviations: SBT: spontaneous breathing trial; COPD: chronic obstructive pulmonary disease; ICUAW: intensive care unit acquired weakness.



e- Table 2. Inclusion and exclusion criteria of the selected studies and identified causes of weaning failure.

Ali ³ 2016	Patients who were invasively mechanically ventilated and admitted to the ICU. Selection of patients was based on the likelihood of prolonged (>72 h) MV.	Patients with age <18 years, history of neuromuscular disease or known anatomical malformation of the diaphragm, use of any muscle paralyzing agent, use of aminoglycosides and corticosteroids, use of non-invasive ventilation before the start of MV, hemodynamic instability, the presence of a current tracheostomy, morbidly obese patients and an admission to an ICU within 12 months prior to this inclusion, pneumothorax or pneumomediastinum, or respiratory rate ≥ 30 breaths/minute.	DNA
Baess ⁴ 2016	Age above 18 years, hemodynamic stability without need for vasopressors, and a duration of continuous MV of 30 days or greater.	Patients were excluded if the cause for intubation was upper airway obstruction, pregnancy, dressing on the right lower rib cage hindering sonographic access, recent intra-abdominal surgery, a history of empyema, or pleurodesis.	DNA
Binet ⁵ 2014	51 patients invasively mechanically ventilated during >48 h.	Not specified.	DNA
Blumhof ⁶ 2016	Age greater than 18 years old, intubated and mechanically ventilated for greater than 24 h.	Ischemic or hemorrhagic stroke, spinal cord injury, neuromuscular disease, pregnancy, or any pathology that would obscure visualization of the right hemidiaphragm.	DNA
Carrie ⁷ 2017	First SBT after at least 48 h of MV, resolution of the underlying acute illness, respiratory rate <35 breaths/ min, PaO ₂ /FIO ₂ ratio >150 mmHg on FIO ₂ $\leq 40\%$ and PEEP ≤ 8 cmH ₂ O, haemodynamic stability in the absence of or with minimal doses of vasopressors, no sedation and a Glasgow coma score > 14.	Pregnancy, age < 18 years, history of hemiplegia or unilateral diaphragmatic palsy. Uncooperative or nonunderstanding patient or the absence of ultrasonographic windows allowing adequate visualization of diaphragmatic motion.	<ul style="list-style-type: none"> - Primary respiratory failure, n= 11 - Cardiogenic-induced weaning failure, n= 5 - Airway secretion build-up, n= 4 - New sepsis, n= 1 - Acute neurological impairment, n= 1
Dinino ⁸ 2014	No specific disease process was required for entry into the study other than the presence of respiratory failure. FIO ₂ <50%, haemodynamically stable in the absence of vasopressors, and alert.	Pregnant women, age <18 years and surgical dressings over the right lower rib cage which would preclude ultrasound exam.	<ul style="list-style-type: none"> - Congestive heart failure, n= 1 - Fever and acute change in mental status, n= 1. -Mucous plugging, n=1 - Aspiration, n= 2 - Respiratory pump failure, n=8 - Fluid overload, n=1
Farghaly ⁹ 2016	Adult patients who had underlying pulmonary disease causing acute respiratory failure and necessitating MV.	Suspicious diaphragm paralysis (raised copula in chest X-ray) or known neuromuscular diseases (as Gillian Barre, myasthenia gravis), pneumothorax or pleural effusion, previous cardiothoracic surgery or pleurodesis, and patients who presented with stridor (due to upperairway involvement) as a cause of extubation failure.	<ul style="list-style-type: none"> -Respiratory distress, n: DNA -Hypoventilation and acidosis, n: DNA. - Disturbed conscious level, n: DNA.
Fayed ¹⁰ 2016	COPD patients who were invasively mechanically ventilated, ≥ 18 years. Evidence for some reversal of the underlying cause for respiratory failure. Adequate oxygenation (PaO ₂ /FIO ₂ ratio ≥ 150 , FIO ₂ ≤ 0.40 , SpO ₂ $\geq 90\%$, PaO ₂ ≥ 60 and PEEP ≤ 8 cmH ₂ O. PH ≥ 7.30 . MIP < -20 to -30 cmH ₂ O. Respiratory	History of diaphragmatic or neuromuscular disease. Morbidly obese patients (BMI > 40).	<ul style="list-style-type: none"> -Fluid overload, fever and mucous plugging, n= 12. - 'Respiratory pump' failure, n=18.

	rate \leq 35 bpm. Heart rate \leq 140 bpm. SBP 90 – 180 mmHg. Hemodynamic stability; no or minimal vasopressors or inotropes. Appropriate level of consciousness. No continuous sedation infusion nor neuromuscular blocking agents. Afebrile. Hemoglobin \geq 8 g/dl and/or no evidence of hemorrhage.		
Ferrari ¹¹ 2014	Invasively mechanically ventilated in pressure support through tracheostomy, after failing one or more attempts of weaning. FiO ₂ < 0.5, PEEP \leq 5 cmH ₂ O, PaO ₂ /FiO ₂ > 200, respiratory rate < 30 breaths per minute, absence of fever, alert and cooperative, and hemodynamic stability in the absence of any vasoactive therapy support.	Diaphragm paralysis or neuromuscular diseases.	DNA
Flevari ¹² 2016	Patients which have had either more than one episode of failed SBT but successful extubation within a week, or 2 repeated failed episodes of SBT in more than a week time period. Age \geq 18 years; Glasgow Coma Scale \geq 14; hemodynamic stability; FiO ₂ \leq 0.5; pressure support < 12 cmH ₂ O and PEEP < 6 cmH ₂ O; RR < 35/min, Vt > 5 ml/kg IBW. The use of sedatives and/or analgesics was not contraindicated as long as the patient remained calm, cooperative and easily arousable (Ramsey score 2-3).	Lack of cooperation, incapability to initiate an inspiratory effort, intubation due to elective surgery or upper airway obstruction, ventilation via tracheostomy tube, hemodynamic instability.	DNA
Jiang ¹³ 2004	Intubated patients receiving MV, age \geq 20 years, FiO ₂ \leq 0.4, PEEP \leq 5 cm H ₂ O, pressure support \leq 8 cm H ₂ O, and a duration of continuous mechanical ventilation of \leq 30 days.	Intubation for elective surgery or due to upper airway obstruction, unstable hemodynamic status, mechanical ventilation via a tracheostomy tube, history of peritonitis, intraabdominal operation, empyema, or pleurodesis.	-Respiratory distress, n= 6. - Profuse airway secretion, n= 4. - Altered mental status, n=3. - Hemodynamic Instability, n=4.
Jung ¹⁴ 2016	Patients diagnosed for ICUAW (defined by a Medical Research Council Score < 48 and mechanically ventilated for at least 48 h, and were undergoing a SBT.	Contraindication to magnetic stimulation of the phrenic nerves (cardiac pacemaker or implanted defibrillator, cervical implants), pre-existing neuromuscular disorders, cervical spine injury; bihemispheric or brain stem lesions; impossibility to assess muscular strength in a limb because of amputation or immobilization.	DNA
Kim ¹⁵ 2011	Age \geq 18 years, FiO ₂ < 50%, PEEP \leq 5 cmH ₂ O, RR \leq 30 breaths/min, PaO ₂ /FiO ₂ ratio > 200 mm Hg, Glasgow coma score \geq 14, and hemodynamic stability in the absence of vasopressors. Analgesic-sedative was allowed, if any, to a minimal dose such that patients were calm and easily arousable (Ramsay score 3). The use of any muscle-paralyzing agent was discontinued at least 2 days before the study. Aminoglycosides were not used in any of the patients.	History of diaphragmatic palsy, cervical spine injury, or neuromuscular disease (myasthenia gravis, Guillain-Barre syndrome, amyotrophic lateral sclerosis); a current thoracostomy, pneumothorax, or pneumomediastinum.	DNA
Osman ¹⁶ 2017	Mechanically ventilated patients through a tracheostomy or endotracheal tube. PaO ₂ > 60 mmHg, PaCO ₂ < 50 mmHg, FiO ₂ < 0.5, PaO ₂ /FiO ₂ > 200, PEEP < 5 cmH ₂ O, < 30-35 breath/min, Vt > 5 ml/kg, Vital capacity > 10 ml/kg, MIP < -15 to -30 cmH ₂ O, Ve 4-10 L/min, RSBI < 100 breath/min/L, P _{0.1} < 2 cmH ₂ O.	Patients < 18 years old, known neuromuscular disorder, primary ultrasound revealed unilateral/bilateral absent diaphragmatic mobility, post esophageal or thoracic surgeries due to intra-operative diaphragmatic manipulation.	DNA
Saeed ¹⁷ 2016	Mechanically ventilated COPD patients admitted to the respiratory ICU, prepared for extubation after spontaneous breathing trials.	Unstable hemodynamics, disturbed conscious level, uncontrolled comorbid disease affecting weaning, intubation due to surgical or other medical problem other than COPD. Presence of ascites, colonic	DNA

Shoaeir ¹⁸ 2016	Patients intubated for ≥ 48 h and eligible for SBT according to Statement of the Sixth International Consensus Conference on Intensive Care Medicine.	distension, lung collapse, fibrosis or pleural effusion, mass or any mechanical factor in the chest or abdomen interfering with diaphragmatic mobility. Patients aged <18 years, patients with tracheostomy, paraplegia with a medullar level above T8, cardiac arrhythmias, severe ICU-acquired neuromyopathy, patients with planned prophylactic noninvasive ventilation after extubation, and patients who had previously failed an SBT.	DNA
Soumerai ¹⁹ 2012	Consecutive patients mechanically ventilated for more than 48 h were included when the underlying respiratory disease that had required intubation was considered by the attending physician as reversed, rendering the patient eligible to a first 1 hour SBT.	Patients aged <18 years, patients with tracheostomy, paraplegia with medullar level above T8, cardiac arrhythmias, severe ICU-acquired neuromyopathy, COPD with forced expiratory volume <50% of the theoretical predicted value, patients with planned prophylactic noninvasive ventilation after extubation, and patients who had previously failed a SBT.	DNA
Spadaro ²⁰ 2016	Patients intubated and mechanically ventilated for more than 48 h, eligible for a first SBT. Clinical improvement of the underlying acute cause of respiratory failure, adequate cough reflex, absence of excessive and/or purulent tracheobronchial secretion; stable cardiovascular status (i.e., heart rate <120 beats/min; SBP 90–160 mmHg; and no or minimal vasopressor use, i.e., dopamine or dobutamine <5 µg/kg/min or noradrenaline <0.05 µg/kg/min). Stable metabolic status (i.e., electrolytes and glycemia within normal range, body temperature <38 °C, hemoglobinemia ≥8–10 g/dL). Adequate oxygenation (i.e., SaO ₂ >92 % with FiO ₂ ≤0.5 or PaO ₂ /FiO ₂ ≥150 mmHg, both with PEEP ≤8 cmH ₂ O). Adequate pulmonary function (i.e., RR ≤30 breaths/min with Vt ≥5 mL/kg IBW and no significant respiratory acidosis. Richmond Agitation and Sedation Scale score ranging between –1 and +1.	Age <18 years, pregnancy, presence of thoracostomy, pneumothorax, or pneumomediastinum, presence of flail chest or rib fractures, neuromuscular disease, use of muscle-paralyzing agents within 48 h before the study, and history or new detection of paralysis or paradoxical movement of a single hemidiaphragm on diaphragmatic ultrasonography.	-Evident dyspnea, n=6. -Cardiovascular instability, n=2. - Desaturation, n=1. -Excessive tracheobronchial secretion, n=5. - Progressive hypertension and desaturation, n=1. - Progressive dyspnea associated with desaturation in the presence of a stable cardiovascular status, n= 3.
Tenza ²¹ 2017	Patients aged ≥ 18 years admitted to the ICU and invasively mechanically ventilated during at least 24 h. Ready for weaning, according to the attending physician criteria.	Medular injury in a level above T8, arrhythmias with hemodynamic instability, terminal extubation, pregnancy, pneumothorax, pneumomediastinum, thoracostomy, chest tube. Thoracic wounds which would preclude ultrasound exam. Pleural lesions, pleurodesis.	- Tachypnea and respiratory fatigue, n=8. - Hypoxemia, n= 8. - Respiratory acidosis, n=2. - Neurologic impairment, n= 3. - Excessive tracheobronchial secretion, n=2.

Abbreviations: ICU: intensive care unit; h: hours; MV: mechanical ventilation; DNA: data not available; SBT: spontaneous breathing trial; PaO₂: arterial oxygen partial pressure; FiO₂: fractional concentration of oxygen in the inspired air; PEEP: positive end-expiratory pressure; N: number of patients; COPD: chronic obstructive pulmonary disease; SpO₂: peripheral blood oxygen saturation; MIP: maximal inspiratory pressure; bpm: beats per minute; SBP: systolic blood pressure; BMI: body mass index; RR: respiratory rate; Vt: tidal volume; IBW: ideal body weight; ICUAW: intensive care unit-acquired weakness; PCO₂: partial pressure of carbon dioxide; Ve: minute ventilation; RSBI: rapid shallow breathing index; P0.1: airway occlusion pressure.

e- Table 3. Performance of diaphragm and lung ultrasound in the included studies.

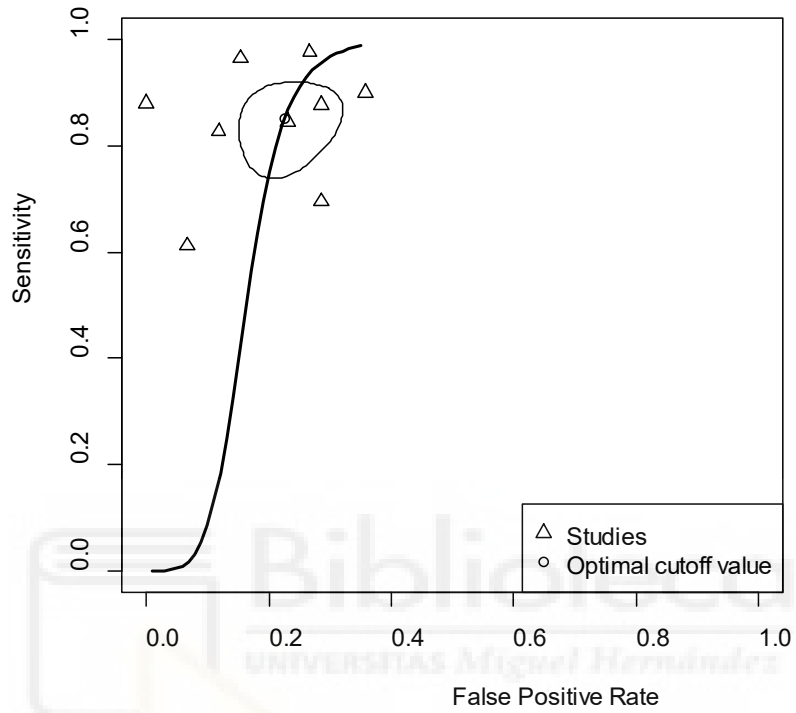
Time			
Ali ³ 2016	Echo Blaster 128 Kit. 10 MHz linear transducer for DTF and 4MHz probe for DE	Daily, during MV	Uncertain
Baess ⁴ 2016	Equipment NS. 2-4 MHz phased-array probe (Philips Healthcare, Andover, Massachusetts, USA)	During SBT	Yes
Binet ⁵ 2014	NS	After a 1 hour SBT	Yes
Blumhof ⁶ 2016	Equipment NS. Linear 7–10 mHz Probe (Sonosite M-Turbo, FujiFilm, Japan)	During PS weaning	Yes
Carrie ⁷ 2017	Vivid S5TM (GE Healthcare, Wauwatosa, WI, USA). 4 MHz cardiac probe.	Before the start of the first SBT, after a brief disconnection from the ventilator	Yes
Dinino ⁸ 2014	Equipment NS. 7–10 MHz Linear probe (LOGIQ Book, GE Healthcare, Waukesha, Wisconsin, USA)	Within the first 5 minutes of the SB or the PS trial	Yes
Farghaly ⁹ 2016	Samsung Medison Sono Ace R3 (Samsung company, Seoul, South Korea). For DTF: 7 MHz transducer, for DE a 3.5 MHz probe.	During SBT	Yes
Fayed ¹⁰ 2016	Mindray digital, model DP-20.	At the beginning of SBT	Uncertain
Ferrari ¹¹ 2014	10 MHz linear transducer Esaote MyLab 40 (Esaote, Genova, Italy). 10 MHz linear probe.	During SBT	Yes
Flevari ¹² 2016	Equipment NS. Multifrequency 5-MHz vector transducer.	During unassisted breathing	Yes
Jiang ¹³ 2004	Aloka Echo Camera SSD-1400 (Aloka; Zug, Switzerland). 3.5 MHz sonar probe.	During SBT	Yes
Jung ¹⁴ 2016	Equipment NS. 7.5 to 10 MHz probe.	During SBT	Yes

Kim ¹⁵ 2011	Esaote MyLab 30CV (Genoa, Italy). 3.5 MHz probe.	At the beginning of SBT	Yes
Osman ¹⁶ 2017	Logic E9 (GE) and Honda electronics HS-2100 Portable. 3.5 MHz convex probe for DE, 9-11 MHz linear probe for DTF. Probe NS for LUS score.	During SBT	Uncertain
Saeed ¹⁷ 2016	Mindray DP 1100. 2-5 MHz convex probe.	During SBT	Yes
Shoaeir ¹⁸ 2016	Mindray DP-3300. 2-4 MHz convex probe.	Before weaning, during SBT and 6 hours after extubation	Uncertain
Soummer ¹⁹ 2012	Siemens Acuson CV70 or a Philips Envisor. 2-4 MHz convex probe.	Before and at the end of the SBT, and 4-6 hours after extubation.	Yes
Spadaro ²⁰ 2016	M-Turbo (SonoSite, Inc., Bothell, USA). 3.5 to 5 MHz convex probe.	After 30 minutes from the beginning of the SBT, or immediately before reconnecting the patient to the ventilator in the case of SBT failure occurring before.	Yes
Tenza ²¹ 2017	Micromaxx Sonosite. Lineal 7-10MHz probe for DTF and 2-4 MHz for MLUS score	At least 4 hours before the start of the first SBT, after a brief period with a low PS.	Yes

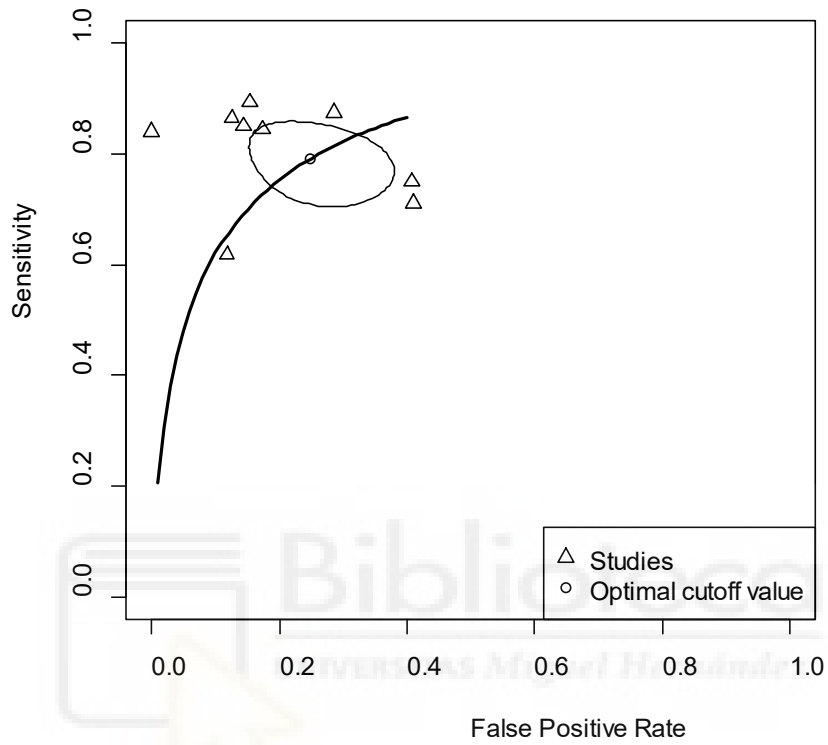
*“Blinded decisioner” means that clinical and weaning-related decisions were taken by clinicians without knowledge of ultrasound results

Abbreviations: US: ultrasonography; MHz: Megahertz; MV: mechanical ventilation; SBT: spontaneous breathing trial; PS: pressure support; LUS: lung ultrasound; MLUS score: modified lung ultrasound score.

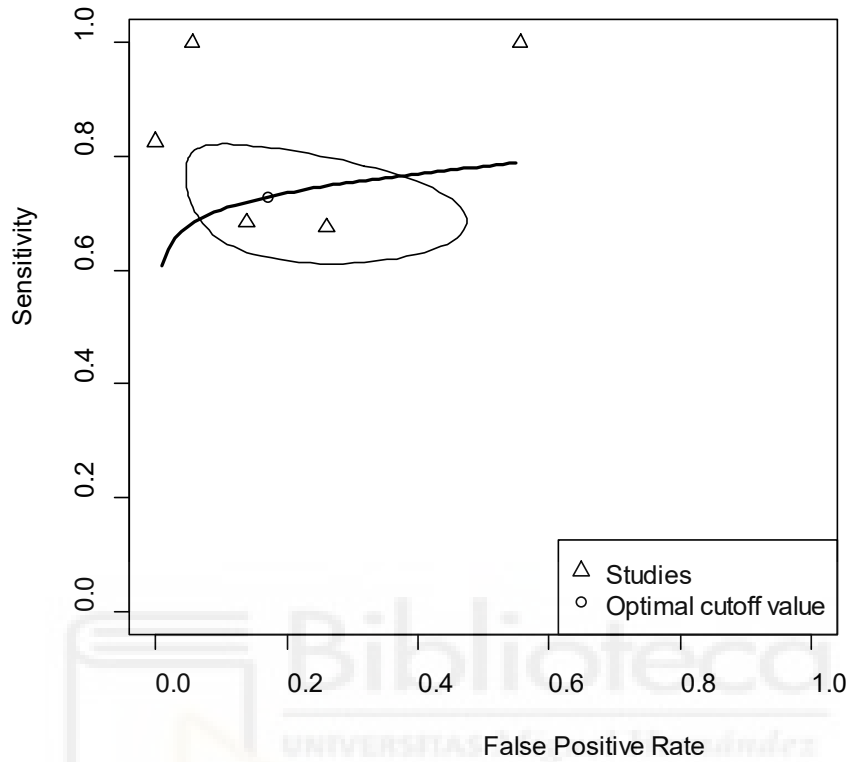
e- Figure 1. Sensitivity analysis of studies that measure diaphragm thickening fraction (DTF) without the specificity outlier²¹. We could then graphically estimate a pooled specificity of 80% with 95% CI (65, 90) and a pooled sensitivity of 85% with 95% CI (75, 90).



e- Figure 2. Sensitivity analysis of studies that measure diaphragmatic excursion (DE) without the specificity outlier⁴. We could then graphically estimate a pooled specificity of 75% with 95% CI (60, 85) and a pooled sensitivity of 80% with 95% CI (70, 85).



e- Figure 3. Summary receiver operating characteristic (SROC) curve of LUS score for weaning success prediction and confidence ellipse around the optimal cutoff value.



REFERENCES:

1. Matamis D, et al. Sonographic evaluation of the diaphragm in critically ill patients. Technique and clinical applications. *Intensive Care Med.* 2013; 39(5):801-10. doi: 10.1007/s00134-013-2823-1.
2. Bouhemad B, et al. Ultrasound assessment of antibiotic-induced pulmonary reaeration in ventilator-associated pneumonia. *Crit Care Med.* 2010; 38(1):84-92. doi: 10.1097/CCM.0b013e3181b08cdb.
3. Ali ER, Mohamad AM. Diaphragm ultrasound as a new functional and morphological index of outcome, prognosis and discontinuation from mechanical ventilation in critically ill patients and evaluating the possible protective indices against VIDD. *J Chest Dis Tuberc.* 2016. 66:339-351.
4. Baess AI, Abdallah TH, Emara DM, Hassan M. Diaphragmatic ultrasound as a predictor of successful extubation from mechanical ventilation: Thickness, displacement, or both? *Egypt J Bronchol.* 2016; 10:162±166.
5. Binet C, Neron L, Rochart N, et al. Validation d'un indice échographique prédictif d'échec de sevrage de la ventilation mécanique. *Annales Françaises D Anesthésie Et De Réanimation.* 2014;33:A383-A383.
6. Blumhof S, Wheeler D, Thomas K, McCool FD, Mora J. Change in diaphragmatic thickness during the respiratory cycle predicts extubation success at various levels of pressure support ventilation. *Lung.* 2016;194(4):519-525.
7. Carrie C, Gisbert-Mora C, Bonnardel E, et al. Ultrasonographic diaphragmatic excursion is inaccurate and not better than the MRC score for predicting weaning-failure in mechanically ventilated patients. *Anaesth Crit Care Pain Med.* 2017;36(1):9-14.
8. Dinino E, Gartman EJ, Sethi JM, McCool FD. Diaphragm ultrasound as a predictor of successful extubation from mechanical ventilation. *Thorax.* 2014;69(5):423-427.
9. Farghaly S, Hasan AA. Diaphragm ultrasound as a new method to predict extubation outcome in mechanically ventilated patients. *Aust Crit Care.* 2017;30(1):37-43.
10. Fayed AM, Abd El Hady MA, Shaaban MS, Fikry DM. Use of ultrasound to assess diaphragmatic thickness as a weaning parameter in invasively ventilated chronic obstructive pulmonary disease patients. *J Am Sci.* 2016;12(6):96-105.
11. Ferrari G, De Filippi G, Elia F, Panero F, Volpicelli G, Aprà F. Diaphragm ultrasound as a new index of discontinuation from mechanical ventilation. *Crit Ultrasound J.* 2014;6(1):8-7902-6-8.
12. Flevari A, Lignos M, Konstantonis D, Armaganidis A. Diaphragmatic ultrasonography as an adjunct predictor tool of weaning success in patients with difficult and prolonged weaning. *Minerva Anesthesiol.* 2016;82(11):1149-1157.
13. Jiang JR, Tsai TH, Jerng JS, Yu CJ, Wu HD, Yang PC. Ultrasonographic evaluation of liver/spleen movements and extubation outcome. *Chest.* 2004;126(1):179-185.
14. Jung B, Moury PH, Mahul M, et al. Diaphragmatic dysfunction in patients with ICU-acquired weakness and its impact on extubation failure. *Intensive Care Med.* 2016;42(5):853-861.

15. Kim WY, Suh HJ, Hong SB, Koh Y, Lim CM. Diaphragm dysfunction assessed by ultrasonography: Influence on weaning from mechanical ventilation. *Crit Care Med.* 2011;39(12):2627-2630.
16. Osman AM, Hashim RM. Diaphragmatic and lung ultrasound application as new predictive indices for the weaning process in ICU patients. *Egypt J Radiol Nucl Med.* 2017; 48(1):61-66.
17. Saeed AM, El Assal GI, Ali TM, Hendawy MM. Role of ultrasound in assessment of diaphragmatic function in chronic obstructive pulmonary disease patients during weaning from mechanical ventilation. *Egypt J Bronchol.* 2016;10:167-172.
18. Shoaier M, Noeam K, Mahrous A, Alaa A. Lung aeration loss as a predictor of reintubation using lung ultrasound in mechanically ventilated patients. *Biolife.* 2016; 4(3), pp 514-520.
19. Soummer A, Perbet S, Brisson H, et al. Ultrasound assessment of lung aeration loss during a successful weaning trial predicts postextubation distress. *Crit Care Med.* 2012;40(7):2064-2072.
20. Spadaro S, Grasso S, Mauri T, et al. Can diaphragmatic ultrasonography performed during the T-tube trial predict weaning failure? the role of diaphragmatic rapid shallow breathing index. *Critical Care.* 2016;20:305.
21. Tenza-Lozano EM, Jaimez E, Latour-Pérez J. Ecografía pulmonar y diafragmática como predictores de weaning. (Material gathered from an unpublished draft by Eva María Tenza-Lozano). Forthcoming 2017.



ACKNOWLEDGMENTS

Author contributions

AML-A 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, the literature search, data collection and analysis, methodological quality assessment, figures, translation and drafting of the manuscript.

EMT-L contributed to the literature search, data collection and methodological quality assessment.

JL-P contributed to the study design, statistical analysis and revision of the manuscript.

All authors approved the final version of the manuscript.

Other contributions.

The authors appreciate the assistance of Meggan Harris for her help with the edition and review of the manuscript, and the assistance of Pascual Gascón-Sansano for his help with the full-text articles obtention.

Financial/ nonfinancial disclosures: The Department of Clinical Medicine of Miguel Hernández University provided non-financial support supplying translation assistance.

Role of the sponsor

The sponsor did not participate in study design, execution, or in the preparation of the manuscript.

8.4. ANEXO 4.

Lung and diaphragmatic ultrasound as predictors of success in weaning.

Original research.

Eva M. Tenza-Lozano, MD, PhD, Ana M. Llamas-Álvarez, MD, Jaime Latour-Pérez, MD, PhD.

Please cite this article as: Tenza-Lozano E, Llamas Álvarez A, Jaimez Navarro E, Fernández Sanchez J. Lung and diaphragmatic ultrasound as predictors of success in weaning. Crit Ultrasound J. 2018; 10: 12. doi: 10.1186/s13089-018-0094-3

CONTENTS

- **Abbreviations.**
- **Abstract.**
- **Introduction.**
- **Methods.**
- **Results.**
- **Discussion.**
- **Conclusion.**
- **Tables.**
- **Figures.**
- **References.**
- **Supplementary material.**
- **Acknowledgments**

LUNG AND DIAPHRAGM ULTRASOUND AS PREDICTORS OF SUCCESS IN WEANING FROM MECHANICAL VENTILATION.

Eva Tenza-Lozano^a, PhD; Ana Llamas-Alvarez^a, MD; Enrique Jaimez-Navarro^a, MD;
Javier Fernández-Sánchez^b, PhD.

^a Intensive Care Unit, Elche General University Hospital, Elche, Spain

^b Department of Clinical Medicine, Miguel Hernandez University, Sant Joan d'Alacant,
Spain

CORRESPONDENCE TO: Eva Tenza-Lozano, Department of Intensive Care
Medicine, Elche General University Hospital, Camino de la Almazara 11, 03203, Elche,
Spain; email: evitatl@hotmail.com

Ana Llamas-Alvarez: llamasalvarez@yahoo.es

Enrique Jaimez-Navarro: enriquejnb@gmail.com

Javier Fernandez-Sanchez: fj.fernandez@umh.es

CONFLICT OF INTEREST: The authors have no conflicts of interest to declare.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE: The Research Ethics Committee of Elche
General University Hospital approved our work.

PRIOR ABSTRACT PUBLICATION/PRESENTATION: No

ABBREVIATIONS.

MV: mechanical ventilation

ICU: intensive care unit

LUS: lung ultrasound score

LUSm: modified lung ultrasound score

TI: diaphragm thickening index

ROC curves: receiver operating characteristic curves

ICC: intraclass correlation coefficient

AUC: area under the ROC curve

SBT: spontaneous breathing

PEEP: positive-end expiratory pressure

IQR: interquartile ranges

LR+: positive likelihood ratio

LR-: negative likelihood ratio

CI: confidence interval

NIV: non-invasive-ventilation

HFNC: high-flow nasal cannula

SW: successfully weaned

FW: failed weaning

COPD: chronic obstructive pulmonary disease

VE: minute ventilation

PIMax: maximal inspiratory pressure

P0.1: airway occlusion pressure

RR: respiratory rate

RSBI: rapid shallow breathing index

FiO₂: fraction of inspired oxygen

SpO₂: oxygen saturation

PaCO₂: partial pressure of carbon dioxide

PaO₂: partial pressure of oxygen;

APACHE II: acute physiology and chronic health evaluation II.



ABSTRACT.

Background: Lung and diaphragm ultrasound methods have recently been introduced to predict the outcome of weaning from mechanical ventilation (MV). The aim of this study is to assess the reliability and accuracy of these techniques for predicting successful weaning in critically ill adults.

Methods: We conducted two studies: a cross-sectional interobserver agreement study between two sonographers and a prospective cohort study to assess the accuracy of lung and diaphragm ultrasound for predicting weaning and extubation outcome. For the interobserver agreement study, we included 50 general critical care patients who were consecutively admitted to the ICU. For the predictive accuracy study, we included consecutively 69 patients on MV who were ready for weaning. We assessed interobserver agreement of ultrasound measurements, using the weighted kappa coefficient for LUSm score (modified lung ultrasound score) and the intraclass correlation coefficient (ICC) and Bland–Altman method for TI (diaphragm thickening index). We assessed the predictive value of LUSm and TI in weaning outcome by plotting the corresponding ROC curves.

Results: We found adequate interobserver agreement for both LUSm (weighted kappa 0.95) and TI (ICC 0.78, difference according to Bland–Altman analysis \pm 12.5%). LUSm showed good-moderate discriminative power for successful weaning and extubation (area under the ROC curve (AUC) for successful weaning 0.80, and sensitivity and specificity at optimal cut-off point 0.76 and 0.73, respectively; AUC for successful extubation 0.78, and optimal sensitivity and specificity 0.76 and 0.47, respectively. TI was more sensitive but less specific for predicting successful weaning (AUC 0.71, optimal sensitivity and specificity 0.93 and 0.48) and successful extubation (AUC 0.76, optimal sensitivity and specificity 0.93 and 0.58). The area under the ROC curve for predicting weaning success was 0.83 for both ultrasound measurements together.

Conclusions: Interobserver agreement was excellent for LUSm and moderate-good for TI. A low TI value or high LUSm value indicates high risk of weaning failure.

INTRODUCTION.

Although widespread use of MV in the intensive care unit (ICU) saves hundreds of lives daily, prolonged MV can lead to increased mortality and morbidity [1–3]. On the one hand, therefore, weaning should be considered as early as possible. On the other hand, however, premature withdrawal can result in extubation failure, which is also associated with increased morbidity and mortality [1, 4, 5].

Several ventilatory indices have been developed for identifying the right time to extubate the patient who has required endotracheal intubation and MV, but none met the criteria required to provide suitably accurate success rates [6]. More recently, lung and diaphragm ultrasound methods have been introduced, assessing pulmonary airway patterns and diaphragm function. Bouhemad [7] was the first author to propose the LUS score for calculating lung aeration patterns in patients with ventilator-associated pneumonia. In later articles, this score was used to predict weaning outcome [8–11], with promising results. Several parameters measured through diaphragm ultrasound have been proposed for the same purpose [11–21]. These include diaphragm thickness, diaphragm movement or excursion during the respiratory cycle [22], and diaphragm thickening or thickening fraction (TI). Although some studies have shown diaphragm excursion and thickness to be of low predictive value in the assessment of diaphragm function [18, 19, 23], a recent meta-analysis corroborates this and the best use of TI to weaning outcome [24].

The data suggest that TI and LUS are good non-invasive indicators of weaning outcome. However, the reliability and accuracy of these studies is limited, mainly due to small sample sizes, inadequate spectra of patients and study heterogeneity [24]. The aim of this study is to assess the reliability and accuracy of lung and diaphragm ultrasound for predicting successful weaning in general critical care patients on mechanical ventilation.

METHODS.

Design

We performed two independent studies: a cross-sectional concordance study between two sonographers (interobserver agreement study) and a prospective cohort study to assess the accuracy of lung and diaphragm ultrasound for predicting weaning and extubation outcome (predictive accuracy study).

Population

For the interobserver agreement study, we included 50 patients (with or without MV), who were consecutively admitted to the ICU of our hospital from December 2016 to February 2017, and who required a thoracic ultrasound examination for clinical reasons. For the predictive accuracy study, we included consecutively all patients on MV admitted to the ICU from 15 January 2016 to 15 April 2017 who have signed the informed consent (Additional file 1) and met the following inclusion criteria (1) over 18 years of age; (2) more than 24 h on MV; (3) ready for weaning.

We applied the same exclusion criteria for both studies: (1) spinal cord injury higher than T8; (2) arrhythmias and haemodynamic instability; (3) terminal extubation; (4) pregnancy; (5) pneumothorax, pneumomediastinum, thoracostomy, chest tube or chest injuries that prevent

ultrasound; (6) pleural lesions or pleurodesis.

Measurements/procedures

Ultrasound technique

Two sonographers trained in lung and diaphragm ultrasound, according to international recommendations [25] (Additional files 1- 3), performed the ultrasound measurements. They used a 2–4 MHz convex probe in B mode, as described in other studies [7, 8]. The scoring system adopted distinguishes four ventilation patterns as follows: normal aeration (N; presence of lung sliding with A lines and fewer than two isolated B lines), moderate loss of pulmonary ventilation (B1; more than two well defined B lines), severe loss of pulmonary ventilation (B2; multiple coalescing B lines) and pulmonary consolidation (C; presence of a tissue pattern). Scores of 0–3 were respectively attributed to the four categories (0 point for N, 1 point for B1, 2 points for B2 and 3 points for C), and for each region the worst visible pattern was recorded.

Rather than using the original LUS score, in our study, we applied a modified procedure (LUSm), evaluating four lung regions on each side instead of the standard six. Our intention in making this modification was to avoid having to move the critical patient, thus preventing the associated complications and facilitating the examination process for the operators. We assessed four areas: anterior-superior, anterior-inferior, lateral and postero-basal. The postero-basal area is where most of the pathology of the critical patient according to Lichtenstein [26] occurs. The total LUSm score for all areas ranged from 0 to 24 points.

In the diaphragm ultrasound examination, the sonographers measured diaphragm thickness using a 7–10 MHz linear probe in B mode (Micromax[®] Sonosite) following the technique described in other studies [11–18, 20, 27]. The right hemidiaphragm was visualised in the zone of apposition, on the midaxillary line between the 8th and 10th intercostal spaces, with the patient in a semi-decubitus position (20°–40°). The diaphragm was viewed in M-mode as a hypoechoic structure between two echoic lines (the diaphragmatic pleura and the peritoneal membrane).

The sonographers captured almost three images in M-mode during spontaneous patient breathing, measuring diaphragm thickness at the end of expiration and at the end of inspiration. We made the average of three TI measurements using the following formula: (end inspiratory diaphragm thickness – end expiratory diaphragm thickness)/end inspiratory diaphragm thickness.

Interobserver agreement study

Both sonographers took TI and LUSm measurements in the same sample of 50 stable patients, with a time difference of less than 5 h between the two operators. This sample was different from the predictive accuracy study.

Predictive accuracy study

In the predictive accuracy study, the patients who were ready to start weaning, according to the international consensus conference criteria [28], the respirator was selected with 8 cm H₂O pressure support (PS) and 5 cm H₂O positive end-expiratory pressure (PEEP) and ultrasound and ventilatory measurements were made. The ventilatory measurements are made automatically by the respirator (model: GE DATEX- OHMEDA Engström Carestation). Afterwards, SBT was continued with a T-tube or with 8 cm H₂O PS and 5 PEEP, depending on the decision of the responsible physician, who evaluated which of the patients successfully passed the SBT, and those who did were extubated. The medical team was blinded to the ultrasound results, and the research team played no role in the patient's weaning. Weaning failure according to the 2007 international consensus conference [28] is defined as either failure of SBT or failure of extubation. Extubation failure is defined as the occurrence of reintubation, non-invasive ventilatory support or death within 48 h following extubation.

Statistical analysis

Data were expressed as medians and interquartile ranges (IQR) or percentages. To compare continuous variables, we used the unpaired Student's t test, Mann–Whitney U test and Wilcoxon test. For categorical variables, we applied the Chi-square or Fisher's exact test.

To evaluate interobserver agreement for LUSm, we calculated the quadratic-weighted kappa coefficient (which is comparable to ICC) and for the TI variable we used ICC and the Bland–Altman method.

In the predictive accuracy study, we calculated the AUCs and their corresponding sensitivities, specificities and likelihood ratios (LR + and LR–) at the optimal cut-off points, to determine the predictive value of TI and LUSm for weaning and extubation success. We developed a predictive model using binary logistic regression, with the ultrasound measures (LUSm and TI) as independent variables to predict successful weaning. We used the StatsDirect v3.0.194 package to perform the statistical analysis.

Ethical aspects

The research ethics committee of Elche General University Hospital approved the study and all enrolled patients gave their informed consent.

RESULTS.

In the interobserver agreement study, the quadratic weighted kappa value for LUSm was 0.95 (95% CI 0.92, 0.98), which shows almost perfect interobserver agreement. For the TI variable, we calculated an ICC value of 0.78 (95% CI 0.65, 0.87), showing moderate to good interobserver agreement, and a difference in measurements according to the Bland–Altman method of \pm 12.5% (Fig. 1).

Over the study period, 139 patients underwent MV, of whom 52 did not meet the inclusion criteria (48 deaths before attempted weaning, 2 self-extubations, 2 on MV for less than 24 h) and 17 were not included for reasons beyond the research team's control (eight withdrawn more gradually from MV, two with no informed consent, two transferred to another hospital, four eligible patients of whom the investigator was not notified, one case of a non-functioning ultrasound scanner). The baseline characteristics of the 69 patients recruited are shown in Table 1. Pressure support was used in 49% of SBTs, a T-tube in 42% and both methods in 9%. Eight patients failed SBT and 61 were extubated, of whom 17 failed extubation. This means that a total of 25 patients failed weaning. Most patients who failed extubation recovered with non-invasive-ventilation (NIV) and high-flow nasal cannula (HFNC); only five patients (8.2%) were reintubated (Fig. 2).

If we compare the group that was successfully weaned (SW) with those who failed weaning (FW) (Table 1), we observe that the FW group was associated with more time on MV, more cases of chronic obstructive pulmonary disease (COPD), higher LUSm and mortality, and lower TI and SpO₂. The median difference in TI and LUSm between the SW and FW groups was 11% and 3 points, respectively.

The area under the ROC curve for predicting weaning success was 0.80 for LUSm (95% CI 0.69, 0.91), 0.71 for TI (95% CI 0.58, 0.84) (Fig. 3) and 0.83 for both (Fig. 4). Table 2 shows the sensitivity, specificity and likelihood ratios at the optimal cut-off points for successful weaning. The area under the ROC curve for predicting extubation success was 0.78 for LUSm (95% CI 0.64, 0.91) and 0.76 for TI (95% CI, 0.61–0.9) (Fig. 3). Table 3 shows the sensitivity, specificity and likelihood ratios at the same cut-off points as those shown in the previous two tables.

DISCUSSION.

According to our data, the reproducibility of lung ultrasound is excellent for the variable LUSm and moderate to good for TI. Regarding the prognostic accuracy of ultrasound for weaning outcome, we found that if TI is below 24% or LUSm is greater than 7 points, the patient has a high risk of weaning failure, with an AUC of 0.8 for LUSm and 0.71 for TI. We found similar values for extubation outcome.

Mean time on MV [29, 30], mortality of the patients included in the study (16%) [4], and SBT failure rate (11.6%) [31] was consistent with previously published results. Extubation failure occurred in 24.6% of the patients, a slightly higher proportion than the 10% to 20% reported in a number of other studies [1, 5, 32–35].

Of the 17 patients who failed extubation, only five (8.2%) required reintubation, a lower rate than reported in other studies [36, 37]. This shows that NIV plays a decisive role in reducing the need for reintubation without increasing morbidity or mortality [37, 38]. As such, through conceptually defined as a criteria of weaning failure [28], we consider that recovery with NIV in fact constitutes a success for the patient's clinical situation. We observed that FW patients were more likely to have COPD. This is a logical finding, as COPD is a risk factor for extubation failure [39, 40]. Of the standard predictors of weaning assessed in our study (PIMax, RSBI, P0.1), we found none to be useful. In a study with the largest number of patients performed for the study of weaning predictors [6], about 500 patients, to assess the predictability of many indices as possible weaning predictors (minute volume, respiratory rate, PaO₂, RSBI, PIMax, Maximum

expiratory pressure, dynamic respiratory compliance, CROP index), it is observed that none of them has value as a predictor of weaning.

The results obtained for LUSm are consistent with previously reported data [8–11]. Our cut-off point is lower (7 LUSm points) because we assessed eight lung areas, whereas the other studies assessed 12. Our aim was to assess all the areas normally affected in critical patients [26], while simplifying the technique so that the patient did not have to be moved, and the associated complications could be avoided. We therefore consider LUSm to constitute a useful new proposal that is beneficial for both the patient and the operator. For TI, we found a cut-off point of 24% for predicting successful weaning, within the range of values reported in other studies (20–36%) [11–20]. The LUS and TI variables tested in those studies showed a higher predictive value for weaning success (Table 2), but a number of factors may have influenced this result. In those studies, patient selection was in some cases very strict, resulting in a homogeneous sample with specific characteristics (patients with COPD [15], tracheotomised patients with prior weaning failure [13], patients with ICU-acquired weakness [17]). Since these samples already had a higher probability of weaning failure, their results cannot be generalised to the whole population of critical care patients. Other limitations in the reviewed studies included elimination of deceased patients [20], diaphragm ultrasound when the patient was on MV rather than during SBT [20], use of a non-validated probe for diaphragm measurement [18], periods of up to 36 h between the ultrasound scan and extubation [12], and using STB failure rather than extubation failure as an endpoint [13]. Only one of the reviewed studies included a reproducibility study comparing TI measurement by two observers, with results very similar to ours (ICC 0.81) [14].

We decided not to select specific populations of patients for our study, in order to obtain a heterogeneous sample and produce generalizable results. The ultrasound measurements were taken within the first minutes of pressure support ventilation, the physicians were blinded to the ultrasound data, no patients were lost to follow-up, and the median time between SBT and extubation was 120 min.

One limitation of our study is the small sample size, which led to imprecise results with broad confidence intervals, especially for TI. Further studies with larger samples of patients are required to establish the true predictive power of these ultrasound techniques.

It should be noted that the interobserver difference in TI measurement ($\pm 12.5\%$) was greater than the median difference in TI between the SW/FW groups (11%). In our sample, therefore, a degree of uncertainty was associated with this parameter. This is probably because the formula for calculating TI is such that a difference of a tenth of a millimeter in the measurement of inspiratory or expiratory diaphragm thickness has a considerable effect on the result. None of the studies reviewed considers this possibility. We believe the problem can be overcome with better ultrasound equipment and a more effective TI measuring technique, so that the true value of this parameter can be reported.

In the future, in order that results can be compared across studies, we believe a number of items should be standardised: the ultrasound technique, the definitions of weaning failure and extubation failure, the time between ultrasound and extubation, blinding of ultrasound findings, and the protocol. If results are to be applied to the whole population of general critical care patients on MV, a heterogeneous sample should be used. Studies in this domain should also assess reproducibility by measuring interobserver agreement in ultrasound measurements. In addition, it may be useful to consider other parameters, such as time on MV, disease severity,

comorbidities, etc., alongside lung and diaphragm ultrasound measurements, to better predict weaning success.

CONCLUSIONS.

In our study, interobserver agreement was excellent in LUSm measurements and moderate to good in TI measurements. The TI variable showed a degree of uncertainty for predicting weaning outcome, but overall its predictive value was found to be acceptable. LUSm produced stronger results in this regard. Lung and diaphragm ultrasound are promising techniques for predicting weaning outcome, but more studies are required to verify their reproducibility. These studies should have a standardised design and should assess interobserver agreement of ultrasound techniques. Using a non-specific sample would ensure that results can be generalized to all patients on MV.



TABLES.

Table 1. Characteristics of patients included in the study.

Variables	All patients (69)	SW (44)	FW (25)	P value
Sex (men) ^a	43 (62.3)	26 (63.4)	15 (62.5)	0.8
Age (years) ^b	66 (53, 78)	65 (53, 78)	69 (64, 78)	0.37
Time on MV (days) ^b	4 (3, 7)	4 (2, 6)	5 (3, 9)	0.04
Time SBT-extubation (minutes) ^b	120 (30, 120)	120 (30, 120)	60 (30, 120)	0.15
Mode of STB				
T-T	29 (42%)	17 (38.6%)	12 (48%)	0.14
PS 8 and PEEP 5	34 (49%)	21 (47.7%)	13 (52%)	
T-T + PS 8 PEEP 5	6 (9%)	6 (13.6%)	0	
Comorbidity ^a				
Chronic heart disease	18 (26)	14 (31.8)	4 (16)	0.25
Neurological disease	18 (26)	12 (27.3)	6 (24)	0.99
COPD	13 (18.8)	5 (11.4)	8 (32)	0.05
Diabetes mellitus	19 (27.5)	12 (27.3)	7 (28)	0.99
Cancer	8 (11.6)	4 (9.1)	4 (16)	0.45
Chronic kidney failure	12 (17.4)	6 (13.6)	6 (24)	0.33
Liver disease	8 (11.6)	3 (6.8)	5 (20)	0.36
Pathology at ICU admission ^a				
Neurological disease	29 (42)	18 (41)	11 (44)	
Respiratory disease	24 (34, 8)	16 (36.4)	8 (32)	
Cardiovascular disease	10 (14.5)	8 (18.2)	2 (8)	0.7
Sepsis	7 (10.1)	5 (11.4)	2 (8)	
Digestive pathology	4 (5.8)	1 (2.3)	3 (12)	
Polytrauma	1 (1.4)	1 (2.3)	0	0.7
LUSm ^b	6 (4, 8)	5 (3, 7)	8 (7, 11)	0.0001
TI (%) ^b	36 (27, 41)	38 (31, 45)	27 (20, 40)	0.003
APACHE II on SBT day ^b	4 (2, 6)	4 (2, 6)	5 (3, 9)	0.07
VE (L/min) ^b	8.6 (7.5, 10)	8.25 (7.3, 9.8)	9 (8.1, 11.6)	0.13
Compliance (mL/cm H ₂ O) ^b	56 (41, 67)	55.5 (43, 69)	59 (40.5, 67)	0.86
PI _{Max} (cm H ₂ O) ^b	-25 (-23, -25)	-25 (-25, -18)	-25 (-26, -24)	0.28
P0.1 (cm H ₂ O) ^b	1 (1, 3)	1 (1, 3)	1.5 (1, 2.5)	0.47
RSBI (breaths/min/L) ^b	35 (20, 50)	31 (20, 43)	37 (30, 54)	0.16
RR (breaths/min) ^b	17 (15, 20)	17 (14, 19)	19 (16, 22)	0.09
Tidal volume (ml) ^b	400 (450, 585)	508 (452, 572)	500 (440, 600)	0.71
FiO ₂ (%) ^b	30 (28, 35)	30 (28, 35)	30 (28, 35)	0.83
SpO ₂ (%) ^b	98 (97, 100)	99 (97, 100)	97 (96, 99)	0.027
PaCO ₂ (mm Hg) ^b	40 (36, 46)	41 (36, 45.6)	39.6 (37, 45)	0.63
PaO ₂ (mm Hg) ^b	93 (79, 115)	96.5 (83, 117)	92 (74, 115)	0.61
pH ^b	7.42 (7.32, 7.47)	7.4 (7.3, 7.5)	7.4 (7.4, 7.4)	0.93
Lactate (mmol/L) ^b	1.2 (1, 1.7)	1.2 (1, 1.5)	1.3 (1.1, 2)	0.2
ICU mortality ^a	7 (14)	1 (2.3)	6 (24)	0.005
Hospital mortality ^a	11 (16)	3 (6.8)	8 (32)	0.003
ICU stay (days) ^b	11 (7, 19)	8 (6, 15)	18 (11, 21)	0.002
Hospital stay (days) ^b	19 (14, 30)	17 (13.5, 31)	23 (17, 30)	0.08

SW, successful weaning; FW, failed weaning; MV, mechanical ventilation; T-T, disconnection with T-tube with oxygen; PS 8 and PEEP, 5 disconnection with pressure support and 5 cm H₂O positive end-expiratory pressure; ICU, intensive care unit; LUSm, modified lung ultrasound score; TI, diaphragm thickness index; APACHE II, acute physiology and chronic health evaluation II; SBT, spontaneous breathing trial; VE, minute ventilation; PI_{max}, maximal inspiratory pressure; P0.1, airway occlusion

Table 2. Comparison with other studies of predictive value of TI and LUS for successful weaning.

Study (ref)—variable	n	AUC	Sensitivity	Specificity	LR+	LR–	Cut-off point (%)
Binet [9]—LUS	48	0.89	1	0.44	1.80	0	14
Osman [11]—LUS	68	0.94	1	0.94	18	0	12
Shoaeir [10]—LUS	50	0.95	0.83	1		0.17	10
Soummer [8]—LUS	86	0.87	0.68	0.86	4.96	0.37	13
Tenza—LUSm	69	0.80	0.76	0.73	2.8	0.44	7
Ali [16]—TI	54	–	0.96	0.85	6.27	0.04	30 ^a
Baess [18]—TI	30	0.65	0.70	0.71	2.43	0.42	30 ^a
Blumhof(14)—TI	33	0.86	0.85	0.77	3.67	0.20	20 ^a
DiNino [12]—TI	63	0.79	0.88	0.71	3.07	0.17	30 ^a
Farghaly [16]—TI	54	0.71	0.9	0.64	2.52	0.16	34.5 ^a
Fayed [15]—TI	112	0.93	0.98	0.73	3.66	0.04	29 ^a ; 24 ^b
Ferrari [13]—TI	46	0.95	0.83	0.88	7.03	0.20	36 ^a
Jung [17]—TI	33	–	0.61	0.93	9.17	0.42	20 ^a
Osman [11]—TI	68	0.89	0.88	1	–	0.12	28 ^a
Tenza—TI	69	0.71	0.93	0.48	1.8	0.14	24 ^a

AUC, area under the ROC curve; LR+, positive likelihood ratio; LR–, negative likelihood ratio; LUS, lung ultrasound score; LUSm, modified lung ultrasound score; TI, diaphragm thickness index

^a Right hemidiaphragm; ^b Left hemidiaphragm

**Table 3.** Predictive value of LUSm and TI for successful extubation.

Variable	AUC	Sensitivity	Specificity	LR+	LR–	Cut-off point (%)
LUSm	0.78	0.76	0.72	2.74	0.33	7
TI	0.76	0.93	0.58	2.26	0.12	24
LUSm + TI	0.83	0.86	0.56	1.97	0.24	

AUC, area under the ROC curve; LR+, positive likelihood ratio; LR–, negative likelihood ratio; LUSm, modified lung ultrasound score; TI, diaphragm thickness index

FIGURES.

Figure 1. Bland-Altman method for interobserver difference in TI measurement.

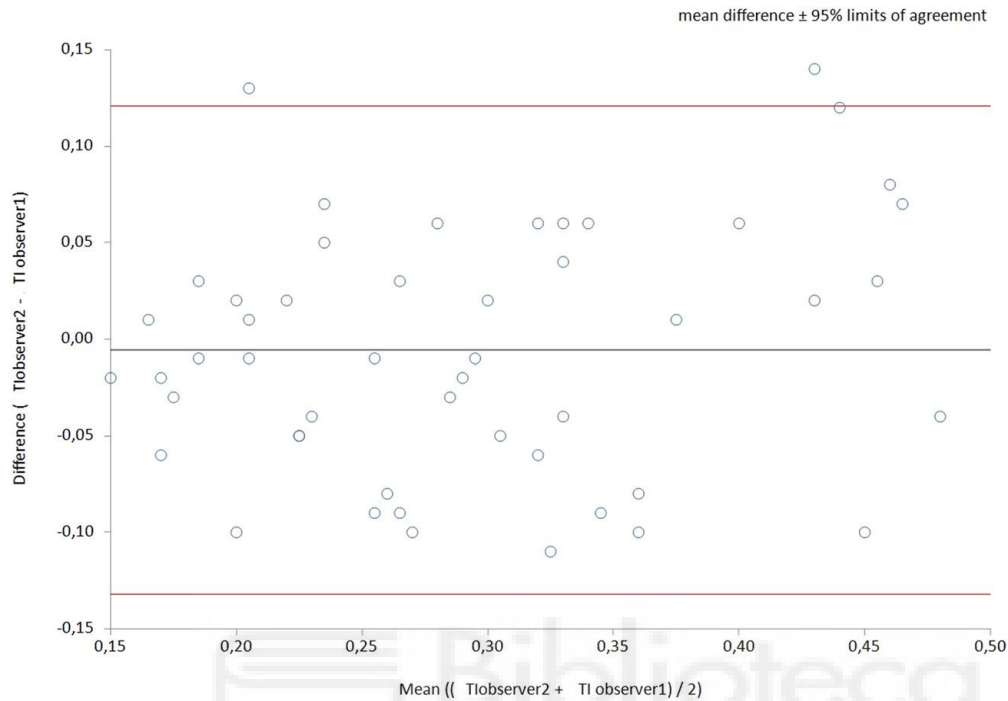


Figure 2. Flowchart of patients.

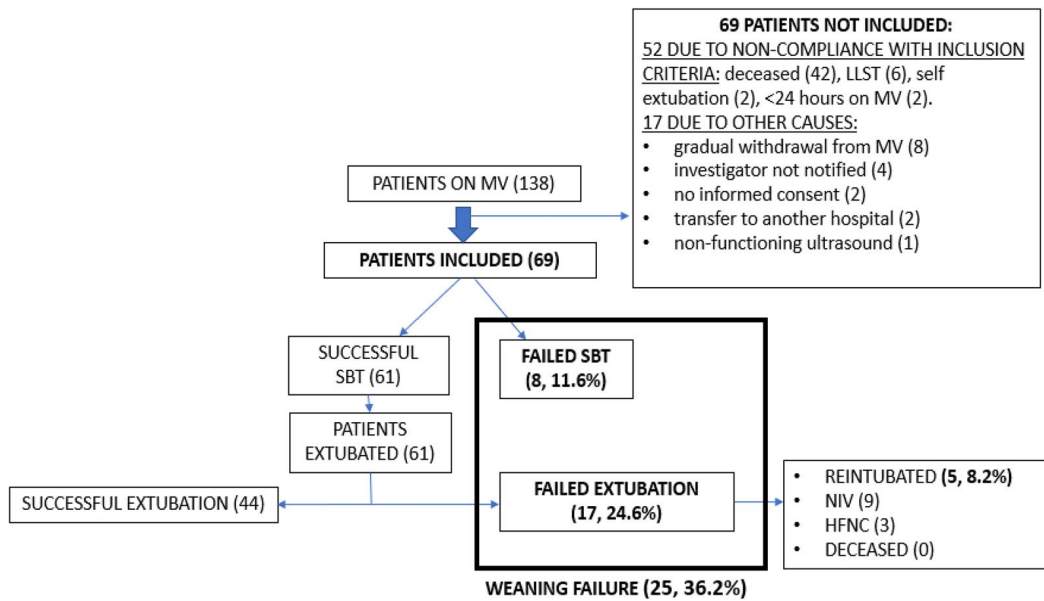


Figure 3: ROC curves for predictive value of TI in successful weaning (SW) (+), in successful extubation (SE) (□) and LUSm in SW (o) and SE (Δ). In SW: LUSm AUC 0.8; TI AUC 0.71. In SE: LUSm AUC 0.78; TI AUC 0.76.

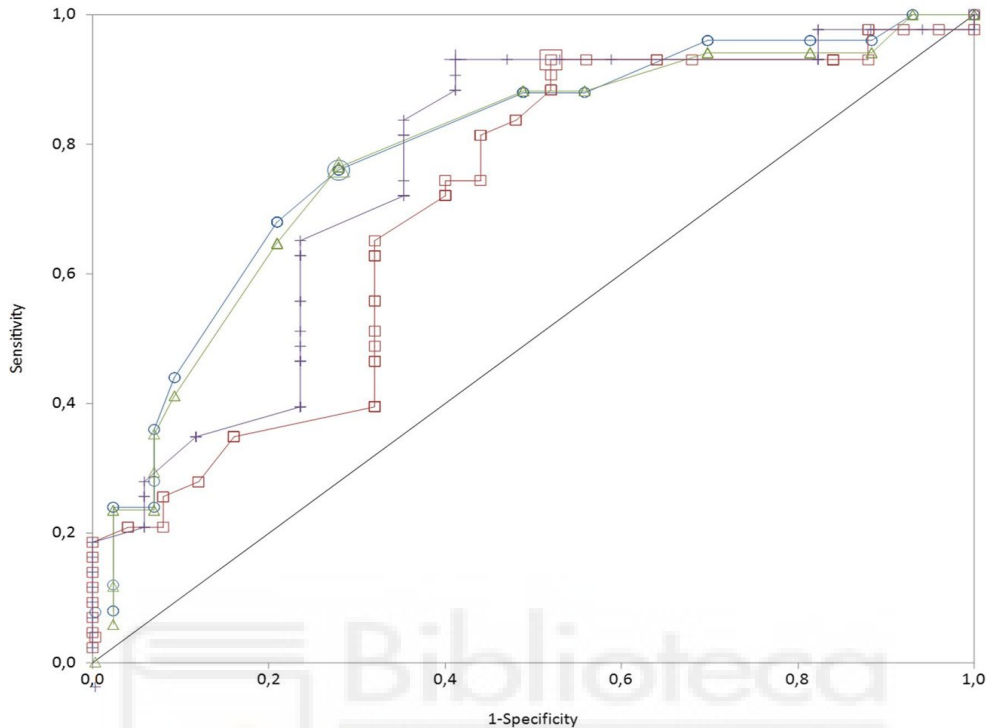
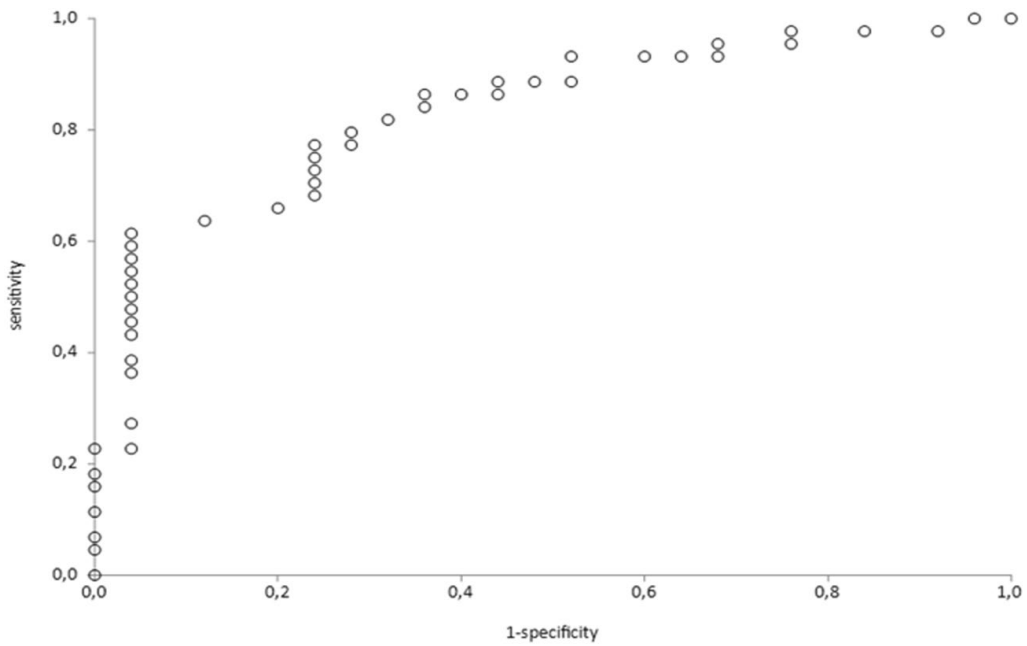


Figure 4: ROC curve for predictive value of TI plus LUSm in successful weaning (SW). AUC 0.83.



REFERENCES.

1. Peñuelas O, Frutos-Vivar F, Fernández C, Anzueto A, Epstein SK, Apezteguía C et al (2011) Characteristics and outcomes of ventilated patients according to time to liberation from mechanical ventilation. *Am J Respir Crit Care Med* 184(4):430–437
2. Powers SK, Kavazis ANLS (2009) Prolonged mechanical ventilation alters diaphragmatic structure and function. *Crit Care Med* 37(10 suppl):S347–S353
3. Levine S, Nguyen T, Taylor N, Friscia ME, Budak MT, Rothenberg P, Zhu J, Sachdeva R, Sonnad S, Kaiser LR, Rubinstein NA, Powers SKSJ (2008) Rapid disuse atrophy of diaphragm fibers in mechanically ventilated humans. *N Engl J Med* 358(13):1327–1335
4. Béduneau G, Pham T, Schortgen F, Piquilloud L, Zogheib E, Jonas M, Grelon F, Runge I, Terzi N, Grangé S, Barberet G, Guitard PG, Frat JP, Constan A, Chretien JM, Mancebo J, Mer BL (2016) Epidemiology of weaning outcome according to a new definition. The WIND study. *Am J Respir Crit Care Med* 195(6):772–783
5. Funk GC, Anders S, Breyer MK, Burghuber OC, Edelmann G, Heindl W et al (2010) Incidence and outcome of weaning from mechanical ventilation according to new categories. *Eur Respir J* 35(1):88–94
6. Savi A, Teixeira C, Silva JM, Borges LG, Pereira PA, Pinto KB et al (2012) Weaning predictors do not predict extubation failure in simple-to-wean patients. *J Crit Care* 27(2):221-e1–221-e8. <https://doi.org/10.1016/j.jcrc.2011.07.079>
7. Bouhemad B, Liu ZH, Arbelot C, Zhang M, Ferarri F, Le-Guen M, Girard M, Lu QRJ (2010) Ultrasound assessment of antibiotic-induced pulmonary reaeration in ventilator-associated pneumonia. *Crit Care Med* 38(1):84–92
8. Soummer A, Perbet S, Brisson H, Arbelot C, Constantin J-M, Lu Q et al (2012) Ultrasound assessment of lung aeration loss during a successful weaning trial predicts postextubation distress. *Crit Care Med* 40(7):2064–2072
9. Binet C, Neron L, Rochart N, Cousson J, Floch T, Charbit B et al (2014) Validation d'un indice échographique prédictif d'échec de sevrage de la ventilation mécanique. *Ann Fr Anesth Reanim* 33:A383
10. Shoaer M, Noeam KMA (2016) Lung aeration loss as a predictor of reintubation using lung ultrasound in mechanically ventilated patients. *Biolife* 4(3):514–520
11. Osman AM, Hashim RM (2017) Diaphragmatic and lung ultrasound application as new predictive indices for the weaning process in ICU patients. *Egypt J Radiol Nucl Med* 48:61–66
12. DiNino E, Gartman EJEJ, Sethi JMJM, McCool FDD (2014) Diaphragm ultrasound as a predictor of successful extubation from mechanical ventilation. *Thorax* 69(5):423–427. <https://doi.org/10.1136/thorax.xjnl-2013-20411.1>
13. Ferrari G, De Filippi G, Elia F, Panero F, Volpicelli GAF (2014) Diaphragm ultrasound as a new index of discontinuation from mechanical ventilation. *Crit Ultrasound J.* 6(1):8
14. Blumhof S, Wheeler D, Thomas K, McCool FD, Mora J (2016) Change in diaphragmatic thickness during the respiratory cycle predicts extubation success at various levels of pressure support ventilation. *Lung* 194(4):519–525

15. Fayed AM, Abd El Hady MA, Shaaban MSFD (2016) Use of ultrasound to assess diaphragmatic thickness as a weaning parameter in invasively ventilated chronic obstructive pulmonary disease patients. *J Am Sci.* 12(6):96–105
16. Farghaly S, Hasan AA (2015) Diaphragm ultrasound as a new method to predict extubation outcome in mechanically ventilated patients. *Aust Crit Care.* <https://doi.org/10.1016/j.aucc.2016.03.004>
17. Jung B, Moury PH, Mahul M, de Jong A, Galia F, Prades A et al (2016) Diaphragmatic dysfunction in patients with ICU-acquired weakness and its impact on extubation failure. *Intensive Care Med* 42(5):853–861
18. Baess AI, Abdallah TH, Emara DM, Hassan M (2016) Diaphragmatic ultrasound as a predictor of successful extubation from mechanical ventilation: thickness, displacement, or both? *Egypt J Bronchiol* 10(2):162–166
19. Umbrello M, Formenti P (2016) Ultrasonographic assessment of diaphragm function in critically ill subjects. *Respir Care* 61(4):542–555
20. Ali ER, Mohamad AM (2016) Diaphragm ultrasound as a new functional and morphological index of outcome, prognosis and discontinuation from mechanical ventilation in critically ill patients and evaluating the possible protective indices against VIDD. *Egypt J Chest Dis Tuberc.* <https://doi.org/10.1016/j.ejcdt.2016.10.006>
21. Zambon M, Greco M, Bocchino S, Cabrini L, Beccaria PF, Zangrillo A (2016) Assessment of diaphragmatic dysfunction in the critically ill patient with ultrasound: a systematic review. *Intensive Care Med* 43(1):29–38
22. Palkar A, Narasimhan M, Greenberg H, Singh K, Koenig S, Mayo P et al (2018) Diaphragm excursion-time index: a new parameter using ultrasonography to predict extubation outcome. *Chest* 153(5):1213–1220. <https://doi.org/10.1016/j.chest.2018.01.007>
23. Dubé B-P, Dres M, Mayaux J, Demiri S, Similowski T, Demoule A (2017) Ultrasound evaluation of diaphragm function in mechanically ventilated patients: comparison to phrenic stimulation and prognostic implications. *Thorax.* <https://doi.org/10.1136/thorax-2016-209459>
24. Llamas-Alvarez AM, Tenza-Lozano EM, Latour-Perez J (2017) Diaphragm and lung ultrasound to predict weaning outcome: systematic review and meta-analysis. *Chest* 152(6):1140–1150
25. Bouhemad B, Zhang M, Lu QRJ (2007) Clinical review: bedside lung ultrasound in critical care practice. *Crit Care* 11(1):205. <https://doi.org/10.1186/cc5668>
26. Lichtenstein DAMG (2008) Relevance of lung ultrasound in the diagnosis of acute respiratory failure: the BLUE protocol. *Chest* 134(1):117–125
27. Matamis D, Soilemezi E, Tsagourias M, Akoumianaki E, Dimassi S, Boroli F et al (2013) Sonographic evaluation of the diaphragm in critically ill patients. Technique and clinical applications. *Intensive Care Med* 39(5):801–810. <https://doi.org/10.1007/s00134-013-2823-1>
28. Boles J-M, Bion J, Connors A, Herridge M, Marsh B, Melot C et al (2007) Weaning from mechanical ventilation. *Eur Respir J* 29(5):1033–1056. <https://doi.org/10.1183/09031936.00010206>

29. Esteban A, Anzueto A, Frutos F, Alía I, Brochard L, Stewart TE, Benito S, Epstein SK, Apezteguía C, Nightingale P, Arroliga AC, Tobin MJ (2002) Characteristics and outcomes in adult patients receiving mechanical ventilation: a 28-day international study. *JAMA* 287(3):345
30. Esteban A, Frutos-Vivar F, Muriel A, Ferguson ND, Peñuelas O, Abraira V et al (2013) Evolution of mortality over time in patients receiving mechanical ventilation. *Am J Respir Crit Care Med* 188(2):220–230
31. Figueroa-Casas JB, Broukhim A, Vargas A, Milam L, Montoya R (2014) Interobserver agreement of spontaneous breathing trial outcome. *Respir Care* 59(9):1324–1328
32. Esteban A, Alía I, Tobin MTA et al (1999) Effect of spontaneous breathing trial duration on outcome of attempts to discontinue mechanical ventilation. *Am J Respir Crit Care Med* 159:512–518
33. Frutos-Vivar F, Esteban A, Apezteguia C, González M, Arabi Y, Restrepo MI et al (2011) Outcome of reintubated patients after scheduled extubation. *J Crit Care* 26(5):502–509. <https://doi.org/10.1016/j.jcrc.2010.12.015>
34. Tonnelier A, Tonnelier JM, Nowak E, Gut-Gobert C, Prat G, Renault A et al (2011) Clinical relevance of classification according to weaning difficulty. *Respir Care* 56(5):583–590
35. Sellares J, Ferrer M, Cano E, Loureiro H, Valencia M, Torres A (2011) Predictors of prolonged weaning and survival during ventilator weaning in a respiratory ICU. *Intensive Care Med* 37(5):775–784
36. Thille AW, Boissier F, Ben-Ghezala H, Razazi K, Mekontso-Dessap A, Brun-Buisson C et al (2016) Easily identified at-risk patients for extubation failure may benefit from noninvasive ventilation: a prospective before-after study. *Crit Care* 20(1):48
37. Ferrer M, Esquinas A, Arancibia F, Thomas Bauer T, Gonzalez G, Carrillo A et al (2003) Noninvasive ventilation during persistent weaning failure: a randomized controlled trial. *Am J Respir Crit Care Med* 168(1):70–76
38. Nava S, Gregoretti C, Fanfulla F, Squadrone E, Grassi M, Carlucci A, Beltrame FNP (2005) Noninvasive ventilation to prevent respiratory failure after extubation in high-risk patients. *Crit Care Med* 33(11):2465–2470
39. Esteban A, Alía I, Gordo FFR et al (1997) Extubation outcome after spontaneous breathing trials with T-tube or pressure support ventilation. *Am J Respir Crit Care Med* 156:459–465
40. Thille AW, Harrois A, Schortgen F, Brun-Buisson C, Brochard L (2011) Outcomes of extubation failure in medical intensive care unit patients. *Crit Care Med* 39(12):2612–2618.

SUPPLEMENTARY MATERIAL.

1. Approbation report of the Research Ethics Committee of Elche General University Hospital.



INFORME DE LA COMISIÓN DE INVESTIGACIÓN

En relación con el proyecto de investigación, titulado: "**Ecografía pulmonar y diafragmática como predictores de éxito en la retirada de la ventilación mecánica**" presentado por la **Dra. Eva María Tenza Lozano** del Servicio de **Medicina intensiva**, y evaluado por la Comisión de Investigación del Hospital de Elche, queremos resaltar los siguientes puntos:

1. Desde el punto de vista científico, la Comisión considera los objetivos del proyecto como de gran interés y potencial utilidad.
2. El diseño del estudio es correcto y adecuado para responder a los objetivos planteados.
3. El equipo tiene una gran experiencia profesional y solvencia como para asegurar una adecuada ejecución del proyecto.

Nota al Investigador:

Enhorabuena por el proyecto, que por supuesto es siempre bienvenido porque contribuye al crecimiento del hospital, aunque la labor de los miembros de la Unidad de Investigación sea precisamente destacar sus potenciales debilidades. Sólo dos puntualizaciones: no se habla de *outcome* primario combinado, sino de un *outcome* primario compuesto por 5 variables que tienen diferente significado clínico y para el que, según detallan los investigadores, la presencia de al menos una cualquiera de ellas define dicho *outcome*. La segunda es que los investigadores exponen con claridad y de forma convincente las razones para la selección de la población de estudio; sin embargo, otro aspecto que habría que tener también en cuenta es la capacidad para poder responder de forma satisfactoria a la pregunta de investigación.

Felicitaciones de nuevo, esta Comisión informa favorablemente dicho proyecto.

Elche, a 31 de marzo de 2016



Fdo: Dr. Andrés Navarro Ruiz
Secretario de la Comisión de Investigación
del Hospital General Universitario de Elche

2. Informed consent for participants.

**“ECOGRAFÍA PULMONAR Y DIAFRAGMÁTICA COMO PREDICTORES DE ÉXITO EN LA
RETIRADA DE LA VENTILACIÓN MECÁNICA”**

D/Dña. con DNI nº
..... como representante de D/Dña
.....manifiesta que:

1º he sido informado/a sobre los riesgos y beneficios del estudio por:
..... (nombre del investigador) 2º

Comprendo que la participación al estudio es voluntaria.

3º Comprendo que los resultados de la ecografía pulmonar y diafragmática no serán tenidos en cuenta para retirada de la ventilación mecánica de mi familiar, y que ésta última se realizará a criterio del médico responsable.

4º He sido también informado/a de que los datos personales serán protegidos e incluidos en un fichero que deberá estar sometido a y con las garantías de la ley 15/1999 de 13 de diciembre.

Tomando ello en consideración, DOY mi CONSENTIMIENTO a que se le incluya en este estudio y sus datos clínicos puedan utilizarse para cubrir los objetivos especificados en el proyecto.

Fecha Firma del representante

Fecha Firma del investigador

__/__/201_

ACKNOWLEDGMENTS.

ET-L is guarantor for the entire manuscript. ET-L had full access to all the data in the study; takes responsibility for the integrity of the data and the accuracy of the data analysis; contributed to the study concept and design, data analysis and drafting of the manuscript; and served as principal author.

AML-A contributed to the literature search and data analysis.

ET-L and EJ-N contributed to the ultrasound measurements.

JF-S contributed to the study design and revision of the manuscript.

All authors read and approved the final manuscript.

Other contributions:

The authors appreciate the assistance of Jaime Latour-Pérez, PhD, for his contribution to the study design, drafting of the manuscript and statistical analysis, and Julia Turner, MSc, for her help with the language editing and review of the manuscript.

