



SERIE EN MEDICINA INTENSIVA: ACTUALIZACIÓN EN METODOLOGÍA EN MEDICINA INTENSIVA

Investigación en el enfermo crítico. Dificultades y perspectivas



J. Latour-Pérez^{a,b}

^a Servicio de Medicina Intensiva, Hospital General Universitario de Elche, Elche, España

^b Departamento de Medicina Clínica, Universidad Miguel Hernández, Sant Joan d'Alacant, España

Recibido el 25 de mayo de 2017; aceptado el 27 de julio de 2017

Disponible en Internet el 21 de septiembre de 2017

PALABRAS CLAVE

Cuidados intensivos;
Ensayos clínicos;
Difusión de la
innovación;
Medicina basada en la
evidencia;
Medicina de precisión

KEYWORDS

Critical care;
Clinical trials;
Diffusion of
innovation;
Evidence-based
medicine;
Precision medicine

Resumen En el ámbito de la medicina intensiva, el aumento de la supervivencia ha venido de la mano de la mejora de los cuidados, la detección precoz del deterioro clínico y la prevención de la iatrogenia, mientras que la investigación de nuevos tratamientos se ha seguido de una abrumadora serie de decepciones. Las raíces de estos fracasos hay que buscarlas en la conjunción de problemas metodológicos —comunes a otras disciplinas— y las particularidades de los pacientes críticos. En este artículo se exploran ambos aspectos y se sugieren algunas vías de progreso.

© 2017 Elsevier España, S.L.U. y SEMICYUC. Todos los derechos reservados.

Clinical research in critical care. Difficulties and perspectives

Abstract In the field of Intensive Care Medicine, improved survival has resulted from better patient care, the early detection of clinical deterioration, and the prevention of iatrogenic complications, while research on new treatments has been followed by an overwhelming number of disappointments. The origins of these fiascos must be sought in the conjunction of methodological problems - common to other disciplines - and the particularities of critically ill patients. The present article discusses both aspects and suggests some options for progress.

© 2017 Elsevier España, S.L.U. y SEMICYUC. All rights reserved.

Correo electrónico: jlatour@coma.es

<https://doi.org/10.1016/j.medin.2017.07.008>

0210-5691/© 2017 Elsevier España, S.L.U. y SEMICYUC. Todos los derechos reservados.

Introducción

La finalidad última de la investigación clínica es la mejora del estado de salud de las personas. Para conseguir este objetivo, los estudios sobre intervenciones terapéuticas y preventivas deberían idealmente¹ abordar un objetivo relevante para la toma de decisiones clínicas²; realizarse de acuerdo con una metodología apropiada que minimice tanto el error aleatorio como el error sistemático³; difundirse de forma íntegra y en un formato adecuado para el decisor, y aplicarse de forma juiciosa en la atención al paciente (fig. 1)⁴. La rotura de alguno de estos 4 eslabones de la cadena supone un desperdicio de investigación¹⁻³.

De acuerdo con los datos disponibles, el volumen de esta investigación desperdiciada es considerable: entre un 30 y un 50% de los ensayos aleatorizados presentan fallos metodológicos importantes^{1,4}; la tasa de no replicabilidad de los estudios es superior al 50%⁵⁻⁸; gran parte de la investigación disponible es inservible^{1,9-11}, y un 40% de los pacientes en un momento dado no reciben tratamientos reconocidos como eficaces con el estado actual de la ciencia^{2,12,13}.

En el ámbito de la medicina intensiva se han producido importantes logros, como la reducción de la mortalidad del síndrome de distrés respiratorio agudo (SDRA)¹⁴ o la sepsis¹⁵. Sin embargo, estos avances han venido de la mano de la mejora general de los cuidados, la detección precoz del deterioro clínico y la prevención de la iatrogenia^{16,17}, mientras que un gran número de ensayos aleatorizados

esperanzadores han concluido con resultados negativos o con un aumento inesperado de la mortalidad¹⁸⁻²³ (tabla 1). Esta situación ha dado pie a que algunos influyentes investigadores hayan cuestionado la normativa europea sobre ensayos clínicos²⁴ e incluso la propia idoneidad de los ensayos aleatorizados en cuidados intensivos^{17,25-28}.

En este primer artículo de la serie *Metodología de investigación en el enfermo crítico* se abordan las deficiencias de la «cadena de la investigación» que podrían explicar —al menos en parte— los fracasos de los *estudios de efectividad comparada* (EEC) en cuidados intensivos; a continuación se examinan con más detenimiento las singularidades de la investigación clínica en pacientes críticos que podrían contribuir a estos fracasos; finalmente, se apuntan posibles vías de progreso. Otros problemas metodológicos, la investigación con *big data*²⁹ y los aspectos ético-legales de la investigación, serán tratados en otros artículos de la serie.

Desenlaces relevantes

La investigación preclínica es esencial para la comprensión de la fisiopatología y el desarrollo de tratamientos eficaces para el manejo del enfermo crítico³⁰. Por ejemplo, la demostración de la lesión pulmonar aguda en modelos animales ventilados con volúmenes altos³¹ se pudo trasladar con éxito a la clínica y fue la base de la actual ventilación protectora¹⁴. Sin embargo, hasta un 75-90% de los

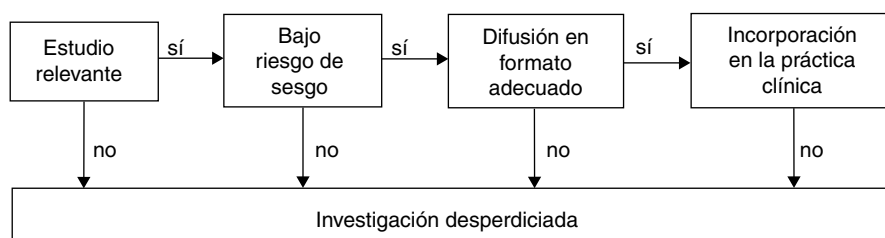


Figura 1 La cadena de la investigación clínica.

Tabla 1 Estudios en pacientes críticos que han reportado aumento de la mortalidad en el grupo experimental

| Estudio (año) | Intervención | Pacientes | RR de mortalidad |
|---------------------------------------|--|---|--|
| Hayes et al. (1994) ¹³ | Aumento aporte de oxígeno | Pacientes críticos generales | Mortalidad hospitalaria RR 1,58 (IC del 95% 1,01, 2,56) |
| Takala et al. (1999) ¹⁹ | Hormona del crecimiento | Pacientes críticos generales | Mortalidad hospitalaria (subestudio multinacional) RR 2,4 (IC del 95% 1,6 a 3,5) |
| Finfer et al. (2009) ²³ | Control estricto de glucemia | Pacientes críticos generales | Mortalidad a los 90 días OR 1,14 (IC del 95% 1,02 a 1,28) |
| Gao Smith et al. (2012) ²⁰ | Agonistas β -2 | SDRA bajo ventilación mecánica | Mortalidad a los 28 días RR 1,47 (IC del 95% 1,03 a 2,08) |
| Perner et al. (2012) ²² | Hidroxiethylalmidón | Sepsis grave | Mortalidad a los 90 días RR 1,17 (IC del 95% 1,01 a 1,36) |
| Ferguson et al. (2013) ²¹ | Ventilación oscilatoria de alta frecuencia | SDRA moderado-grave | Mortalidad hospitalaria RR 1,33 (IC del 95% 1,09 a 1,64) |
| Heyland et al. (2013) ⁷³ | Glutamina | Pacientes críticos con ventilación mecánica y fallo multiorgánico | Mortalidad ajustada a los 28 días OR 1,28 (IC del 95% 1,0 a 1,64) |

IC del 95%: intervalo de confianza del 95%; OR: odds ratio; RR: riesgo relativo; SDRA: síndrome de distrés respiratorio agudo.

resultados de la investigación preclínica publicados en revistas de alto perfil —habitualmente evidencias de tipo etiológico o fisiopatológico— no son reproducibles⁷ y, hasta la fecha, solo una mínima parte se ha traducido en innovaciones sustanciales para su uso a la cabecera del paciente³².

Al igual que la investigación preclínica, la investigación clínica tradicional se centra en el análisis de manifestaciones biológicas de la enfermedad (v.g. el gasto cardiaco o los niveles plasmáticos de citocinas). El problema es que la literatura científica ha mostrado reiteradamente una disociación entre estos resultados fisiopatológicos (de interés para el investigador) y los resultados clínicos (que realmente importan al paciente). El fracaso de la estrategia del aporte supranormal de oxígeno³³ o el más reciente de la reanimación hemodinámica precoz orientada a objetivos^{34,35} son solo 2 de los más conocidos ejemplos.

Desde la perspectiva de una investigación clínica centrada en el paciente³⁶, estos desenlaces deben ser considerados como *desenlaces subrogados (o evidencia indirecta)* de los desenlaces clínicos verdaderamente importantes para la toma de decisiones, como la supervivencia, la extubación tras la ventilación mecánica, la calidad de vida tras el alta hospitalaria o los costes³⁷. Si esa evidencia indirecta no se traduce en una mejora de los desenlaces clínicos importantes ello supondrá un desperdicio de investigación.

Otra forma de investigación inútil ocurre cuando un estudio se plantea una pregunta de investigación con desenlaces clínicos importantes, pero sobre la que ya existe evidencia científica satisfactoria para contestarla^{1,38}. En efecto, se ha estimado que el 50% de los estudios se diseñan sin referencia a revisiones sistemáticas previas³⁹, con el consiguiente riesgo de generar publicaciones redundantes. Por ejemplo, Fergusson et al.⁴⁰ identificaron mediante un metaanálisis acumulativo 64 ensayos publicados entre 1987 y 2002, demostrando que la efectividad de la aprotinina quedó claramente establecida en 1992 tras la realización del decimosegundo ensayo; los siguientes 52 ensayos, por tanto, se habrían podido evitar con una adecuada revisión sistemática. Esta investigación redundante, además de innecesaria, es éticamente cuestionable y ha generado una movilización de la sociedad civil y de algunas agencias financiadoras para exigir la realización de revisiones sistemáticas sobre el tema antes de iniciar un nuevo ensayo clínico^{3,41,42}.

Se ha debatido extensamente acerca de cuál es el desenlace clínico más adecuado para los ensayos clínicos en pacientes críticos^{18,43-46}. La mortalidad por cualquier causa es el desenlace relevante por excelencia. En cualquier caso, la mortalidad debería referirse a un plazo adecuado a la patología en estudio (p. ej., mortalidad a los 28 días o a los 90 días)⁴⁵ y tener en cuenta el posible impacto de la limitación de los tratamientos de soporte vital.

Una de las dificultades de la mortalidad es su carácter dicotómico, que aumenta considerablemente las necesidades de tamaño muestral, especialmente en situaciones de baja incidencia. Este problema se ha intentado paliar con el desarrollo de resultados combinados (*composite endpoints*), como la aparición de «muerte, infarto o reperfusión urgente» o «muerte o disfunción orgánica»⁴⁷. Sin embargo, los desenlaces combinados son a veces de difícil interpretación, son fácilmente manipulables por el investigador y favorecen una visión demasiado optimista de los resultados, por lo que su interpretación debe ser cautelosa^{48,49}. Por

ejemplo, la utilización del desenlace combinado «muerte, infarto o reperfusión urgente» tendría poco sentido si las diferencias observadas se debieran exclusivamente a las diferencias de reperfusión.

La mortalidad no es el único desenlace relevante para el paciente. Aspectos como la duración de la estancia en la UCI, el número de días sin ventilación mecánica o la calidad de vida tras el alta hospitalaria⁵⁰ son otros desenlaces importantes que deberían ser tenidos en cuenta por las guías de práctica clínica para el establecimiento de recomendaciones.

En los estudios comparativos de antimicrobianos en pacientes críticos, los desenlaces clínicos tradicionales presentan serias limitaciones. Por ejemplo, la mortalidad puede deberse solo parcialmente al proceso infeccioso. El cálculo de la mortalidad atribuible puede ser más útil que la mortalidad cruda cuando coexisten varias causas de muerte y comorbilidad⁵¹. Sin embargo, hay que recordar que el riesgo atribuible no mide la fuerza de la asociación (como el riesgo relativo o la reducción relativa del riesgo), sino que es una medida de impacto, cuyo valor depende del riesgo basal de los pacientes. Otros desenlaces habituales, como la tasa de curación clínica o la duración de la infección, conllevan un importante grado de subjetividad, ya que los síntomas pueden ser debidos a otros eventos intercurrentes. Por otra parte, los estudios comparativos de estrategias antimicrobianas requieren la valoración conjunta de desenlaces relevantes para el enfermo (mortalidad, curación, efectos adversos) con otros relevantes para la sociedad (duración del tratamiento, aparición de resistencias). La metodología Desirability Of Outcome Ranking/Response Adjusted Days of Antibiotic Risk (DOOR/RADAR) permite comparar la distribución de estos desenlaces de forma conjunta, lo que obviaría algunos de los problemas de los estudios de no inferioridad, habituales en este campo^{52,53}. Aunque ha sido propuesta como metodología de elección para el análisis coprimario de estudios de mejora de estrategias antibióticas, esta metodología comparte las limitaciones de los estudios con resultados combinados, su interpretación es susceptible de manipulación por parte del investigador y un resultado positivo no siempre descarta la no inferioridad del tratamiento experimental^{54,55}.

Estudios pequeños

La falta de potencia estadística para detectar efectos razonables podría explicar buena parte de los resultados negativos obtenidos en los estudios realizados en pacientes críticos^{56,57}.

Harhay et al.⁵⁷ revisaron los ensayos aleatorizados realizados en UCI publicados en 16 revistas de alto impacto. De los 40 estudios que analizaban la mortalidad, 36 fueron negativos y solo 15 tenían tamaño suficiente para detectar una reducción absoluta de la mortalidad del 10%. Uno de los factores determinantes de estos resultados fue un cálculo del tamaño muestral inadecuado basado en una estimación del efecto demasiado optimista. Ello suscita un problema ético, en la medida en que se somete a pacientes a una experimentación que no cuenta con los recursos suficientes para cumplir sus objetivos⁵⁸. Por otro lado, una parte de estos resultados negativos podría tratarse de «falsos negativos»,

lo que podría desanimar la realización de futuros estudios en un campo prometedor.

Paradójicamente, además de propiciar el error de tipo II (falsos negativos), los estudios pequeños favorecen la obtención de falsos positivos que posteriormente no son confirmados en ulteriores estudios. En efecto, cuando el número de pacientes es pequeño (p. ej., al inicio de la fase de reclutamiento de un ensayo) la estimación del efecto es muy inestable y produce frecuentemente resultados espurios (falsos positivos y falsos negativos)^{59,60}.

Típicamente estos falsos positivos («máximos aleatorios») coinciden con una estimación de efecto implausible⁶¹.

Los máximos aleatorios característicos de los estudios pequeños podrían explicar gran parte de los fracasos de la investigación en cuidados intensivos. Por ejemplo, en un ensayo aleatorizado realizado en pacientes postoperados de pancreatitis grave se asignó a 30 pacientes a nutrición yeyunal precoz (dentro de las 12 h de la cirugía), mientras que los 30 restantes sirvieron de grupo control. La mortalidad fue del 3,3% en el grupo experimental y del 23% en el grupo control (riesgo relativo 0,14; intervalo de confianza del 95% [IC del 95%] 0,02-0,81; $p = 0,03$)⁶². Al margen de posibles sesgos, estos resultados nos deberían alertar en un doble sentido. Primero, parece poco verosímil que el inicio de la nutrición enteral en las primeras 12 h reduzca la mortalidad en un 86% (riesgo relativo de 0,14). Segundo, el tamaño muestral calculado para detectar un efecto «razonable» del 30% (con un nivel de significación del 5% y un poder estadístico del 80%) es de 518 pacientes por grupo. El bajo tamaño muestral (por debajo del tamaño óptimo de información⁶³) y la magnitud implausible del efecto indican fuertemente que estamos ante un máximo aleatorio. Esta impresión fue corroborada por un metaanálisis posterior⁶⁴ basado en 12 ensayos y 662 pacientes que comparaban nutrición enteral precoz frente a tardía en pacientes críticos generales que no encontró diferencias de mortalidad significativas (riesgo relativo 0,76, IC del 95% 0,52-1,11).

Por último, los estudios pequeños con frecuencia presentan problemas metodológicos que los hacen más propensos a los sesgos, lo que podría contribuir a la sobreestimación del efecto.

Existe evidencia empírica de que esto ocurre en el ámbito de la investigación en el enfermo crítico. Así, Zhang et al.⁶⁵ examinaron 27 metaanálisis (317 estudios) de ensayos realizados en pacientes críticos que analizaban la mortalidad y comprobaron que los estudios pequeños (con menos de 100 pacientes por brazo) eran de peor calidad en aspectos como la generación y la ocultación de la secuencia de aleatorización, pérdidas en el seguimiento, análisis por intención de tratar o enmascaramiento y, considerados globalmente, sobrestimaban el efecto en un 40%.

El problema de los estudios pequeños está íntimamente ligado al de la interrupción precoz del ensayo por beneficio. Frecuentemente, durante la realización de un ensayo clínico se realizan análisis intermedios para prevenir el reclutamiento de nuevos enfermos cuando existe evidencia clara de superioridad, toxicidad o futilidad. Esta práctica abre la posibilidad de detener el estudio ante un máximo aleatorio (con la consiguiente sobreestimación del efecto)^{66,67}. Un ejemplo ilustrativo es el estudio OPTIMIST⁶⁸ en el que se analizaba el efecto de tifacogin en pacientes con sepsis grave y

coagulopatía. Un primer análisis intermedio con los primeros 722 pacientes reclutados detectó una reducción significativa de la mortalidad en el grupo experimental (29,1% frente al 38,9%; $p = 0,006$), que desapareció en el análisis final tras reclutar a 1.754 pacientes (34,2% frente a 33,9%; $p = 0,88$).

Subgrupos

Los ensayos proporcionan resultados promedio de la efectividad del tratamiento sobre el conjunto de los enfermos incluidos en el estudio. Sin embargo, no todos los pacientes incluidos en el estudio se benefician del tratamiento: algunos pacientes presentan los efectos adversos del tratamiento («perjudicados»), otros presentan el evento a pesar del tratamiento («condenados») y otros, finalmente, no habrían presentado el evento en ningún caso, por lo que el tratamiento habría sido superfluo («inmunes») ⁶⁹. El análisis de los resultados estratificados por datos demográficos, niveles de gravedad o comorbilidad facilita la identificación de los subgrupos que más se pueden beneficiar o sufrir daño como consecuencia del tratamiento.

Con todo, la distribución homogénea de las variables pronósticas entre los distintos brazos del ensayo dentro de cada subgrupo no está garantizada y el análisis de subgrupos conlleva importantes riesgos⁷⁰. En un reciente estudio metaepidemiológico⁷⁰ solo 46 de los 117 subgrupos declarados estaban apoyados en sus propios datos; solo 5 de estos 46 hallazgos de subgrupos tuvieron al menos un intento de corroboración y ninguno de los intentos de corroboración alcanzó una p significativa en el test de interacción.

Otros problemas metodológicos de los ensayos en pacientes críticos —como la generación y ocultación de la secuencia de aleatorización, el sesgo de desgaste o el enmascaramiento— se analizan detenidamente en otros artículos de la serie.

Difusión selectiva o inadecuada de resultados

Para que la investigación de efectividad comparada sea útil no basta con que sea relevante y esté bien diseñada y ejecutada. Es preciso además que la evidencia generada llegue al decisor clínico y que este la aplique en sus pacientes.

Se estima que un 50% de los estudios realizados no llegan a publicarse como artículo¹ y la parte que finalmente se publica sobrerrepresenta los estudios con resultados positivos e infrarrepresenta los efectos adversos de los tratamientos (*sesgo de publicación*). Por ejemplo, los suplementos de glutamina fueron recomendados en su día por algunas guías de práctica clínica en pacientes traumatizados y quemados⁷¹; posteriormente, un metaanálisis Cochrane en pacientes críticos generales detectó indicios de sesgo de publicación⁷² y un ensayo amplio bien diseñado encontró un aumento de la mortalidad en pacientes críticos con fallo multiorgánico⁷³.

Las raíces del sesgo de publicación habría que buscarlas en la resistencia de los promotores a publicar resultados desfavorables, amparados en una normativa permisiva con la ocultación de los resultados de los estudios⁷⁴, junto con la tendencia de los editores de revistas a publicar preferentemente estudios con resultados positivos. Sin embargo,

diversos estudios indican que los propios investigadores ejercen una especie de autocensura que favorece la no publicación de los resultados negativos⁷⁴. La gravedad del problema ha propiciado algunas iniciativas —como AllTrials⁷⁵ o James Lind⁴¹— que favorecen la publicación de toda la evidencia disponible y animan a los pacientes a que no participen en ensayos que no garanticen la publicación total de los resultados^{3,75}.

Por otro lado, no todos los desenlaces que se estudian son los que finalmente se publican, lo que genera una especie de «sesgo de publicación intraestudio». Chan et al.⁷⁶ comunicaron que en el 62% de los estudios cuyo protocolo había sido previamente registrado el artículo finalmente publicado presentaba al menos un desenlace primario cambiado, introducido u omitido; y además, los desenlaces estadísticamente significativos tenían mayor probabilidad de ser comunicados que los no significativos. Por ejemplo, en un estudio comparativo de fluidoterapia con salino frente a hidroxietilalmidón (HES) en un grupo de pacientes traumatizados⁷⁷ se concluyó que «el HES genera menos daño renal y un aclaramiento de lactato significativamente mejor que el suero salino». En el registro previo del ensayo (<http://www.isrctn.com/ISRCTN42061860>) los autores habían señalado 2 *outcomes* primarios y 7 desenlaces secundarios. Sin embargo, como señalan Reinhart y Hartog⁷⁸, 3 de los desenlaces publicados no figuraban en el registro, entre ellos el daño renal agudo y el aclaramiento del lactato, que constituyen los principales resultados del artículo. Esta comunicación selectiva de los resultados tiende a aumentar los falsos positivos y contribuye a la sobreestimación del beneficio de los tratamientos que se publican en la literatura médica.

Por último, aunque se publiquen, los resultados pueden presentarse de forma manipulada o poco transparente, lo que dificulta su replicación, y complica la valoración de su validez y aplicabilidad, convirtiéndolos en inservibles⁷⁹. Es bien conocido que los ensayos promovidos por la industria farmacéutica sobreestiman de forma muy importante la efectividad del tratamiento experimental en comparación con los estudios promovidos por entidades sin ánimo de lucro («sesgo de promotor»)⁸⁰. El problema se ha agravado en los últimos años debido a la proliferación de las llamadas «publicaciones depredadoras», un modelo de negocio basado en revistas de acceso gratuito que se financian con las tasas de los autores sin una adecuada revisión por pares ni otras medidas de control de la calidad⁸¹⁻⁸³.

La cada vez más frecuente solicitud a los autores por parte de las revistas de los datos crudos para que puedan ser revisados por los expertos estadísticos y la propuesta del Instituto de Medicina (IOM)⁸⁴ de facilitar el análisis compartido de los datos primarios de los ensayos aleatorizados parecen marcar un camino prometedor que podría abordar algunas de las deficiencias actuales y reducir el desperdicio de la investigación. Esta última propuesta ha sido recientemente recogida por el Comité Internacional de Editores de Revistas Médicas (ICMJE)⁸⁵, si bien solo la cumple una minoría de las revistas de medicina intensiva (fig. 2).

Por último, uno de los avances más importantes de los últimos años en la difusión de los resultados de la investigación clínica es la iniciativa EQUATOR⁸⁶, que ha generado guías para la publicación de diversos tipos de estudio, como CONSORT (para ensayos clínicos), PRISMA

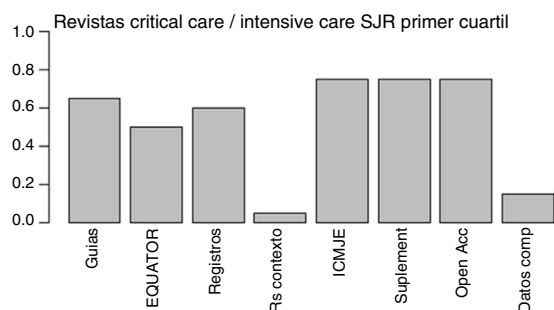


Figura 2 Iniciativas de las revistas de cuidados intensivos para mejorar la cadena de la investigación.

Medidas para reducir el desperdicio de la investigación incluidas en las normas para los autores de las revistas de cuidados intensivos indexadas por SJR dentro del primer cuartil (año 2015). De izquierda a derecha: menciona alguna guía de presentación de los resultados; menciona algunas de las guías de EQUATOR; menciona el registro previo de ensayos o revisiones sistemáticas; sugiere revisiones sistemáticas para contextualizar artículos originales; recomienda consultar la web del ICMJE; permite la publicación online de material suplementario; cita alguna forma de acceso gratuito; formula alguna política para favorecer el uso compartido de datos.

Datos comp: datos compartidos; EQUATOR: Enhancing the Quality and Transparency Of health Research; ICMJE: International Committee of Medical Journals Editors; Open Acc: Open Access; RS: revisiones sistemáticas; Supplement: material suplementario; SJR: Scimago Journal Reports.

(para metaanálisis) o STARD (para estudios de diagnóstico), que favorecen una mayor transparencia de los resultados, permitiendo al lector valorar por sí mismo la credibilidad de los resultados⁸⁶. Desafortunadamente, el número de revistas del campo de la medicina intensiva que recomiendan el uso de estas guías es todavía escaso (fig. 2).

Aplicación de los resultados

El cuarto eslabón de la cadena de la investigación lo constituye la decisión de aplicar o no los resultados de la evidencia científica. Se trata del salto desde la evidencia obtenida a partir de grupos de pacientes seleccionados (de los que frecuentemente se ha excluido a los pacientes con comorbilidad) bajo condiciones asistenciales óptimas, hasta la decisión clínica en el paciente individual (con su propia comorbilidad y preferencias) en un entorno asistencial distinto^{87,88}.

Las guías de práctica clínica basadas en la evidencia proporcionan una ayuda para establecer recomendaciones individuales. Actualmente, la metodología de la aproximación Grading of Recommendations Assessment, Development and Evaluation (GRADE) está considerada como la principal referencia para establecer el grado de evidencia y la fuerza de la recomendación^{89,90}. Sin embargo, no todas las guías de práctica clínica son igualmente confiables. La aplicación acrítica de antiguas jerarquías de la evidencia basadas en un solo desenlace clínico, que equiparaban de forma automática ensayo aleatorizado con evidencia de calidad alta, llevó a considerar que la hemofiltración de alto flujo en la sepsis estaba respaldada por evidencia de alta calidad,

llegándose a afirmar la obligatoriedad de su aplicación en cualquier paciente con fracaso renal agudo⁹¹. Habría que esperar algunos años hasta que estudios ulteriores acabaran demostrando que la terapia de alto flujo era fútil⁹².

Dificultades añadidas de los estudios de efectividad comparada en el paciente crítico

A las deficiencias anteriores, comunes a otras disciplinas, se añaden las peculiaridades de la investigación en el paciente crítico, que complican en gran medida la realización e interpretación de los estudios.

El primer rasgo destacable es la gran heterogeneidad de pacientes incluidos en los ensayos. Los diagnósticos habituales de los pacientes críticos (como SDRA o sepsis) son en realidad síndromes, necesitados de redefiniciones periódicas^{93,94}, en los que se incluye una gran variedad de patologías. Como muestra, menos de la mitad de los diagnosticados de SDRA con los criterios de la definición de Berlín⁹³ tienen en la autopsia lesión pulmonar difusa, el clásico correlato anatomopatológico del SDRA⁹⁵. A esta heterogeneidad entre pacientes se añade la fisiopatología rápidamente cambiante de la enfermedad dentro del mismo paciente, como ocurre en la sepsis. En estas condiciones parece poco razonable esperar que un único tratamiento sea efectivo para distintas formas de SDRA o que sea igualmente efectivo en una sepsis en fase hiperinflamatoria que en una sepsis en estadio de déficit inmunitario^{26,27}.

En segundo lugar, en el paciente crítico concurren una gran variedad de tratamientos y una alta prevalencia de comorbilidad, con la consiguiente probabilidad de interacciones y de aplicación de procedimientos diagnósticos o terapéuticos adicionales (cointervención), a veces en forma sutil⁹⁶⁻⁹⁹.

Por último, la vigilancia intensiva y los ajustes continuos de los tratamientos inherentes a los cuidados intensivos suponen una dificultad adicional para diseñar un grupo control adecuado, que refleje la práctica clínica habitual. En efecto, los tratamientos en UCI se dosifican frecuentemente de forma flexible según la situación del paciente (v.g. según los cambios hemodinámicos o de la mecánica pulmonar). Si estas intervenciones se fijan mediante un protocolo estricto se pueden crear en ambas ramas del ensayo subgrupos de pacientes que reciben dosis de tratamiento inconsistentes con la práctica habitual. Este desajuste con la práctica clínica habitual (*practice misalignment*)^{25,100}, además de comprometer la validez externa del estudio, genera nuevos problemas de interpretación y podría perjudicar a algunos enfermos.

Por ejemplo, en el ensayo Acute Respiratory Distress Syndrome (ARDS) Clinical Trials Network low-tidal-volume (VT) trial (ARMA)¹⁴ una serie de pacientes con SDRA se aleatorizaron a una ventilación con un volumen corriente (VC) de 6 ml/kg vs. 12 ml/kg. En el momento en que se realizó el estudio, la práctica habitual era reducir el VC conforme bajaba la distensibilidad pulmonar y aumentaban las presiones de la vía aérea. Por lo tanto, la aleatorización generó un subgrupo de pacientes más graves en los que un VC de 12 ml/kg resultaba excesivo, y un subgrupo de pacientes menos graves en los que un VC de 6 ml/kg era inferior a la

práctica habitual. Estudios posteriores mostraron que en los pacientes cuya distensibilidad prealeatorización era menor, el aumento del VC se asociaba a una mayor mortalidad en comparación con los pacientes tratados con volúmenes bajos (42% vs. 29%). Por el contrario, en los pacientes con mayor distensibilidad, la reducción del VC aumentaba la mortalidad (37% vs. 21%)¹⁰⁰.

En estas condiciones, el efecto del tratamiento no se compara con los resultados de la práctica clínica real, sino con unas condiciones experimentales artificiales y no reflejan adecuadamente la efectividad real del tratamiento.

Personalización de la medicina crítica y medicina de precisión

La medicina personalizada no es un descubrimiento reciente. La estratificación del riesgo y el desarrollo de reglas de predicción clínica tienen una larga tradición en la medicina intensiva y han facilitado la identificación de grupos relativamente homogéneos de pacientes y el establecimiento de recomendaciones ajustadas al riesgo-beneficio individual^{87,101}. En los últimos años, la personalización del tratamiento se ha beneficiado del desarrollo de marcadores biológicos, tanto pronósticos como de respuesta al tratamiento^{102,103}.

Un grado más de profundización en la personalización del tratamiento sería la prescripción de forma «precisa» sobre la base de las características genéticas, proteómicas y metabólicas individuales («medicina de precisión»)¹⁰⁴. Esta corriente ha tenido éxitos notables en algunas disciplinas, especialmente la oncología (v.g. cáncer de mama her-2 positivo, cáncer de pulmón EGFR positivo, etc.). Estos avances han generado expectativas de que la medicina de precisión permita identificar dianas terapéuticas y obtener grupos homogéneos de pacientes críticos, evitando así los falsos negativos debidos a la inclusión de pacientes que no se pueden beneficiar del tratamiento¹⁰⁴.

Esta idea ha sido recibida con entusiasmo por algunos investigadores que consideran que la raíz de los fracasos de los EEC en medicina intensiva está en el carácter proteiforme de la patología crítica y la heterogeneidad de la respuesta a los tratamientos, y proponen una reorientación de la investigación hacia la medicina de precisión. Con frecuencia esta propuesta se acompaña de una defensa de la supremacía del razonamiento fisiopatológico y de la generación de nueva evidencia a partir de los datos asistenciales (*big data*) frente a la evidencia procedente de ensayos aleatorizados^{17,25,27}.

En este punto hay que hacer algunas matizaciones. En primer lugar, la oposición entre una medicina de precisión —basada en la fisiopatología y en las características individuales del paciente— y una medicina basada en la evidencia —obtenida a partir de grupos de pacientes— es solo aparente¹⁰⁵, ya que ambas se necesitan mutuamente. En efecto, entre los hallazgos prometedores de la investigación preclínica y su integración en la práctica clínica se requiere una valoración de la validez (analítica y clínica) y de la utilidad clínica de las posibles aplicaciones (test o tratamientos)¹⁰⁶. Este proceso traslacional, que va más allá del paradigma *bench-to bedside*, incluye necesariamente la realización de ensayos clínicos aleatorizados, investigación de síntesis (revisiones sistemáticas, metaanálisis, análisis

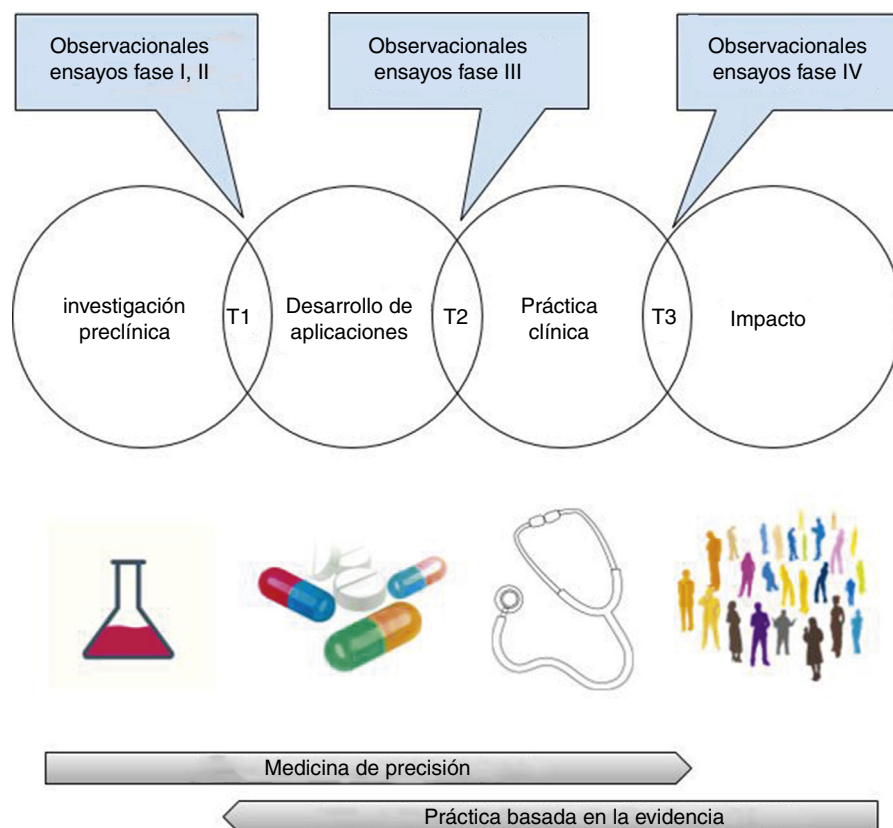


Figura 3 Investigación traslacional: del laboratorio al impacto en salud.

T1-T3: investigación traslacional tipo 1-tipo 3.

Imágenes obtenidas bajo licencia CC0 a través de pixabay (<https://pixabay.com>).

de decisiones) y elaboración de guías de práctica clínica basadas en la evidencia¹⁰⁷⁻¹¹⁰ (fig. 3). El conocimiento de la fisiopatología es crucial para un correcto diseño de los ensayos aleatorizados, así como para valorar la plausibilidad biológica de los hallazgos de la investigación clínica^{27,111,112}; sin embargo, hay que recordar que su uso exclusivo para la toma de decisiones ha generado importantes daños en los pacientes, que solo se han puesto de manifiesto con ensayos aleatorizados amplios^{19-23,73,113} (tabla 1).

En segundo lugar, la era de la medicina de precisión no ha llegado todavía a la medicina intensiva³² y aún tendría que superar algunos problemas antes de ser aceptada como una propuesta viable^{104,114,115}. Las dificultades para obtener biomarcadores predictivos de respuesta al tratamiento en la práctica clínica son notables, especialmente en las condiciones fisiopatológicas rápidamente cambiantes del enfermo crítico. Por ejemplo, a pesar de las más de 1.000 publicaciones existentes sobre polimorfismos genéticos en la sepsis no se ha encontrado un test diagnóstico suficientemente válido y disponible a pie de cama que empareje un genotipo específico con su terapia correspondiente¹⁰⁴. Por otra parte, el cuerpo de evidencia disponible en este campo presenta importantes debilidades: gran parte de los estudios se basan en el hallazgo de valores de *p* significativos sin hipótesis previas; frecuentemente utilizan muestras de conveniencia con alto riesgo de sesgo de selección, y adolecen de un insuficiente control de variables de confusión y una escasa tasa de reproducibilidad⁶. Respecto a su viabilidad,

la infraestructura informática y la base legal ineludible para gestionar los datos necesarios para una medicina crítica de precisión es a fecha de hoy claramente insuficiente^{29,104} y existen serias dudas sobre su sostenibilidad económica; prueba de ello es la reciente decisión del Medicare y Medicaid de no reembolsar la prescripción farmacogenómica de warfarina, a pesar del beneficio aportado a un pequeño grupo de pacientes con la variante genómica¹¹⁶. Finalmente, lejos de obviar los problemas logísticos de los macroensayos, la necesidad inherente a la medicina de precisión de evaluar tratamientos individualizados, supone estrechar cada vez más los criterios de inclusión, con la consiguiente ralentización de los tiempos de reclutamiento y un aumento de los costes de ejecución¹⁰⁴.

Por todo ello, el desarrollo de la medicina de precisión en cuidados intensivos no debería paralizar la realización de nuevos y mejores ensayos clínicos^{117,118}.

Mejores estudios clínicos

Los ensayos aleatorizados —y de forma especial los realizados en pacientes críticos— tienen limitaciones como su alejamiento del «mundo real» (debido a los criterios estrictos de selección, el entorno académico y la protocolización del tratamiento) y las dificultades (legales y logísticas) para reclutar una muestra suficiente^{17,119}.

Tabla 2 Características de los estudios de eficacia frente a estudios de efectividad

| | Ensayo explicativo | Ensayo pragmático |
|------------------------|-----------------------------------|--|
| Objetivo | Determinar eficacia | Determinar efectividad |
| Contexto | Investigadores profesionales | Práctica clínica habitual |
| Criterios de selección | Estrictos (prima validez interna) | Amplios (prima validez externa) |
| Tamaño muestral | Generalmente medio-bajo | Alto (miles de pacientes) |
| Desenlaces | Frecuentemente intermedios | Desenlaces relevantes para el paciente |

El alejamiento de la práctica clínica real de los ensayos clínicos genera un problema de validez externa, que se ha intentado paliar mediante el diseño de ensayos pragmáticos. A diferencia de los ensayos explicativos (o ensayos de eficacia), orientados a maximizar las posibilidades de que un tratamiento demuestre su efecto, los estudios pragmáticos (o ensayos de efectividad) se orientan a evaluar el efecto de la intervención en condiciones de la práctica clínica, lo que les dota de una mayor representatividad y cierta información sobre la efectividad real de la intervención (tabla 2)¹²⁰.

Por su parte, los metaanálisis de datos individuales ofrecen una oportunidad para explorar subgrupos homogéneos de pacientes con mayor potencia estadística sin renunciar a las técnicas de ajuste tradicionales³⁵.

Cuando no existen ensayos aleatorizados, los ensayos disponible son de baja calidad¹²¹ o carecen del poder estadístico suficiente (p. ej., para detectar eventos raros), los estudios observacionales basados en análisis secundarios de registros pueden ser adecuados; sin embargo, estos estudios están limitados por la calidad y el número de variables recogidas¹²². Por otro lado, los estudios observacionales adolecen del problema de la confusión por indicación (los pacientes pertenecientes a los grupos experimental y control tienen pronósticos de partida diferentes debido a variables de confusión conocidas o desconocidas). El ajuste estadístico mediante puntuación de propensión al tratamiento proporciona en general resultados consistentes con los de los ensayos aleatorizados; sin embargo, en ocasiones, los resultados son contradictorios, sin que exista ninguna forma fiable de predecir estos desacuerdos¹²³. Por todo ello, los estudios observacionales no parecen adecuados para evaluar intervenciones con un efecto pequeño^{124,125}.

Recientemente se han desarrollado nuevas modalidades de ensayos clínicos, que en ciertos casos podrían paliar estas limitaciones. Aquí hay que citar los ensayos cruzados binarios con datos agregados (*cluster randomized crossover trials*), con un gran potencial para la evaluación de intervenciones habituales en cuidados intensivos para reducir la mortalidad¹²⁶⁻¹²⁹.

Por su parte, el ensayo clínico con asignación aleatoria de cohorte múltiple (*cohort multiple randomised controlled trial*) combina las fortalezas del ensayo aleatorizado —validez interna— y las del registro observacional —muestreo consecutivo representativo, comparación con la práctica clínica real, ampliación del reclutamiento, reducción de costes^{125,130}. Este tipo de diseño podría beneficiarse, además, de la recogida de datos en tiempo real a través de

Tabla 3 Iniciativas para mejorar la relevancia, la calidad y la difusión de la investigación clínica

| Iniciativa/recursos | Objetivo |
|---|--|
| James Lind ⁴¹ https://www.ncbi.nlm.nih.gov/pmc/articles/PMC539653/ | Promover mejores ensayos clínicos: nuevos ensayos en contexto de revisiones sistemáticas; participación de pacientes en la priorización; participación en ensayos relevantes |
| https://clinicaltrials.gov/ https://www.clinicaltrialsregister.eu/ http://www.who.int/ictrp/en/ http://www.hkuctr.com/ https://www.crd.york.ac.uk/PROSPERO/ | Registros de ensayos y revisiones sistemáticas |
| CASP International CASP España http://www.redcaspe.org/ | Formación en lectura crítica de la literatura biomédica |
| GRADE http://www.gradeworkinggroup.org/ | Desarrollo de metodología rigurosa para la elaboración de guías de práctica clínica |
| Alltrials ⁷⁵ http://www.alltrials.net/ | Promover la publicación de toda la investigación clínica |
| EQUATOR http://www.equator-network.org/ | Guías para la publicación transparente de la investigación |
| Evidence Alerts https://plus.mcmaster.ca/EvidenceAlerts/ | Difusión selectiva de artículos válidos y relevantes |
| Antimicrobial Stewardships Programs https://www.cdc.gov/getsmart/healthcare/implementation/core-elements.html | Mejora de la utilización de antimicrobianos |

sistemas de vigilancia electrónica («sniffers») —una tecnología que notifica rápidamente al investigador la existencia de un paciente con criterios de inclusión, lo que ha permitido doblar la velocidad de reclutamiento en un ensayo de sepsis¹³¹.

Por último, los nuevos diseños adaptativos podrían ser especialmente útiles si se sospecha un efecto de subgrupo (marcadores de susceptibilidad)¹³². Este diseño se está utilizando en un amplio estudio europeo de la neumonía adquirida en la comunidad (AD-SCAP) para probar diversas combinaciones de antibióticos, esteroides y estrategias de ventilación (ClinicalTrials.gov NCT02735707). Mediante un procedimiento complejo basado en grandes cantidades de datos, una plataforma informática asigna de forma adaptativa a los pacientes con rasgos distintivos de enfermedad a las terapias más prometedoras para su perfil. Estas características indican que los ensayos de plataforma podrían servir como un primer paso hacia la medicina de precisión, ya que se inscriben en un grupo heterogéneo de pacientes, pero permiten el análisis de subtipos específicos en tiempo real, disminuyendo los falsos negativos debidos a los efectos diferenciales del tratamiento.

Corolario

Los resultados de la investigación clínica en el paciente crítico en las últimas décadas distan mucho de las expectativas que habían generado. Ello se debe tanto a las singularidades del entorno del paciente crítico como a deficiencias mejorables en la cadena de la investigación. El futuro desarrollo de una hipotética medicina crítica de precisión es esperanzador, pero entre tanto se avanza en esta línea, se debería tomar medidas para no seguir malgastando los esfuerzos de la investigación^{9,10,133-135} (tabla 3).

Conflictos de intereses

J. Latour es miembro de CASP-España.

Agradecimientos

A Juan B. Cabello, Vicente Gómez-Tello, Ana Llamas, Eva de Miguel, Juan Simó y José Antonio Viedma por la lectura crítica de una versión anterior del manuscrito. Procede la eximente habitual.

Bibliografía

- Chalmers I, Glasziou P. Avoidable waste in the production and reporting of research evidence. *Lancet*. 2009;374:86–9.
- Glasziou P, Haynes B. The paths from research to improved health outcomes. *Evid Based Nurs*. 2005;8:36–8.
- Moher D, Glasziou P, Chalmers I, Nasser M, Bossuyt PM, Korevaar DA, et al. Increasing value and reducing waste in biomedical research: Who's listening? *Lancet*. 2016;387:1573–86.
- Scott IA, Glasziou PP. Improving the effectiveness of clinical medicine: The need for better science. *Med J Aust*. 2012;196:304–8.
- Ioannidis JP. Why most published research findings are false. *PLoS Med*. 2005;2:e124.
- Dal-Re R, Bernad A, Garesse R. La reproducibilidad de las investigaciones biomédicas: ¿quo vadis? *Med Clin (Barc)*. 2016;146:408–12.
- Begley CG, Ioannidis JP. Reproducibility in science: Improving the standard for basic and preclinical research. *Circ Res*. 2015;116:116–26.
- Fernandez Mondejar E. Considerations on the low adherence to clinical practice guidelines. *Med Intensiva*. 2017;41:265–6.
- Al-Shahi Salman R, Beller E, Kagan J, Hemminki E, Phillips RS, Savulescu J, et al. Increasing value and reducing waste in biomedical research regulation and management. *Lancet*. 2014;383:176–85.
- Glasziou P, Altman DG, Bossuyt P, Boutron I, Clarke M, Julious S, et al. Reducing waste from incomplete or unusable reports of biomedical research. *Lancet*. 2014;383:267–76.
- Macleod MR, Michie S, Roberts I, Dirnagl U, Chalmers I, Ioannidis JP, et al. Biomedical research: Increasing value, reducing waste. *Lancet*. 2014;383:101–4.
- Scott IA, Glasziou PP. Improving effectiveness of clinical medicine: The need for better translation of science into practice. *Med J Aust*. 2012;197:374–8.
- Runciman WB, Hunt TD, Hannaford NA, Hibbert PD, Westbrook JI, Coiera EW, et al. CareTrack: Assessing the appropriateness of health care delivery in Australia. *Med J Aust*. 2012;197:100–5.
- ARDS Net. Ventilation with lower tidal volumes as compared with traditional tidal volumes for acute lung injury and the acute respiratory distress syndrome. The Acute Respiratory Distress Syndrome Network. *N Engl J Med*. 2000;342:1301–1308.
- Kaukonen KM, Bailey M, Suzuki S, Pilcher D, Bellomo R. Mortality related to severe sepsis and septic shock among critically ill patients in Australia and New Zealand, 2000–2012. *JAMA*. 2014;311:1308–16.
- Singer M, Azoulay E, Chiche JD. A new dawn for critical care research. *Intensive Care Med Exp*. 2013;1:1.
- Vincent JL. Improved survival in critically ill patients: Are large RCTs more useful than personalized medicine? No. *Intensive Care Med*. 2016;42:1778–80.
- Ospina-Tascon GA, Buchele GL, Vincent JL. Multicenter, randomized, controlled trials evaluating mortality in intensive care: Doomed to fail? *Crit Care Med*. 2008;36:1311–22.
- Takala J, Ruokonen E, Webster NR, Nielsen MS, Zandstra DF, Vundelinckx G, et al. Increased mortality associated with growth hormone treatment in critically ill adults. *N Engl J Med*. 1999;341:785–92.
- Gao Smith F, Perkins GD, Gates S, Young D, McAuley DF, Tunnicliffe W, et al. Effect of intravenous beta-2 agonist treatment on clinical outcomes in acute respiratory distress syndrome (BALTI-2): A multicentre, randomised controlled trial. *Lancet*. 2012;379:229–35.
- Ferguson ND, Cook DJ, Guyatt GH, Mehta S, Hand L, Austin P, et al. High-frequency oscillation in early acute respiratory distress syndrome. *N Engl J Med*. 2013;368:795–805.
- Perner A, Haase N, Guttormsen AB, Tenhunen J, Klemenzson G, Aneman A, et al. Hydroxyethyl starch 130/0.42 versus Ringer's acetate in severe sepsis. *N Engl J Med*. 2012;367:124–34.
- Finfer S, Chittock DR, Su SY, Blair D, Foster D, Dhingra V, et al. Intensive versus conventional glucose control in critically ill patients. *N Engl J Med*. 2009;360:1283–97.
- Blanch L, Annane D, Antonelli M, Chiche JD, Cunat J, Girard TD, et al. The future of intensive care medicine. *Med Intensiva*. 2013;37:91–8.
- Takala J. Better conduct of clinical trials: The control group in critical care trials. *Crit Care Med*. 2009;37:S80–90.

26. Vincent JL, Singer M. Critical care: Advances and future perspectives. *Lancet*. 2010;376:1354–61.
27. Pinsky MR. Clinical Trials without conceptual foundation may produce flawed results for the management of fluid therapy in the critically ill. *Br J Anaesth*. 2014;113:737–9.
28. Vincent JL. We should abandon randomized controlled trials in the intensive care unit. *Crit Care Med*. 2010;38:5534–8.
29. Docherty AB, Lone NI. Exploiting big data for critical care research. *Curr Opin Crit Care*. 2015;21:467–72.
30. Bihari S, Maiden M, Deane A, Fuchs R, Fraser J, Bersten AD, et al. Preclinical research in critical care—the Australasian perspective. *Crit Care Resusc*. 2015;17:151–2.
31. Dreyfuss D, Soler P, Basset G, Saumon G. High inflation pressure pulmonary edema. Respective effects of high airway pressure, high tidal volume, and positive end-expiratory pressure. *Am Rev Respir Dis*. 1988;137:1159–64.
32. Pene F, Ait-Oufella H, Taccone FS, Monneret G, Sharshar T, Tamion F, et al. Insights and limits of translational research in critical care medicine. *Ann Intensive Care*. 2015;5:8.
33. Shoemaker WC, Appel PL, Kram HB. Hemodynamic and oxygen transport responses in survivors and nonsurvivors of high-risk surgery. *Crit Care Med*. 1993;21:977–90.
34. Rivers E, Nguyen B, Havstad S, Ressler J, Muzzin A, Knoblich B, et al. Early goal-directed therapy in the treatment of severe sepsis and septic shock. *N Engl J Med*. 2001;345:1368–77.
35. The PRISM investigators. Early, goal-directed therapy for septic shock - A Patient-level meta-analysis. *N Engl J Med*. 2017;376:2223–2234.
36. Rubenfeld GD, Angus DC, Pinsky MR, Curtis JR, Connors AF Jr, Bernard GR. Outcomes research in critical care: Results of the American Thoracic Society Critical Care Assembly Workshop on Outcomes Research. The Members of the Outcomes Research Workshop. *Am J Respir Crit Care Med*. 1999;160:358–67.
37. Guyatt GH, Oxman AD, Kunz R, Atkins D, Brozek J, Vist G, et al. GRADE guidelines: 2. Framing the question and deciding on important outcomes. *J Clin Epidemiol*. 2011;64:395–400.
38. Young C, Horton R. Putting clinical trials into context. *Lancet*. 2005;366:107–8.
39. Cooper NJ, Jones DR, Sutton AJ. The use of systematic reviews when designing studies. *Clin Trials*. 2005;2:260–4.
40. Fergusson D, Glass KC, Hutton B, Shapiro S. Randomized controlled trials of aprotinin in cardiac surgery: Could clinical equipoise have stopped the bleeding? *Clin Trials*. 2005;2:218–29.
41. Chalmers I. The James Lind Initiative. *J R Soc Med*. 2003;96:575–6.
42. Partridge N, Scadding J. The James Lind Alliance: Patients and clinicians should jointly identify their priorities for clinical trials. *The Lancet*. 2004;364:1923–4.
43. Vincent JL. Endpoints in sepsis trials: More than just 28-day mortality? *Crit Care Med*. 2004;32:S209–13.
44. Ferguson ND. Using QALYS in critical care trials [consultado 18 Mar 2017]. Disponible en: https://criticalcarecanada.com/presentations/2012/icu.length.of.stay.new.approaches.to_a_challenging_outcome.pdf.
45. Taori G, Ho KM, George C, Bellomo R, Webb SA, Hart GK, et al. Landmark survival as an end-point for trials in critically ill patients—comparison of alternative durations of follow-up: An exploratory analysis. *Crit Care*. 2009;13:R128.
46. Petros AJ, Marshall JC, van Saene HKF. Should morbidity replace mortality as an endpoint for clinical trials in intensive care? *The Lancet*. 1995;345:369–71.
47. Mebazaa A, Laterre PF, Russell JA, Bergmann A, Gattinoni L, Gayat E, et al. Designing phase 3 sepsis trials: Application of learned experiences from critical care trials in acute heart failure. *J Intensive Care*. 2016;4:24.
48. Cordoba G, Schwartz L, Woloshin S, Bae H, Gotzsche PC. Definition reporting interpretation of composite outcomes in clinical trials: Systematic review. *BMJ*. 2010;341:c3920.
49. Latour-Perez J. Sobre el uso de resultados clínicos agregados en los ensayos clínicos [consultado 5 Jul 2017]. REMI. 2002;5. Disponible en: <http://remi.uninet.edu/2002/05/REMIED10.htm>.
50. Herridge MS, Cheung AM, Tansey CM, Matte-Martyn A, Diaz-Granados N, Al-Saidi F, et al. One-year outcomes in survivors of the acute respiratory distress syndrome. *N Engl J Med*. 2003;348:683–93.
51. Timsit JF, de Kraker MEA, Sommer H, Weiss E, Bettiol E, Wolke-witz M, et al. Appropriate endpoints for evaluation of new antibiotic therapies for severe infections: A perspective from COMBACTE's STAT-Net. *Intensive Care Med*. 2017;43:1002–12.
52. Molina J, Cisneros JM. Editorial commentary: A chance to change the paradigm of outcome assessment of antimicrobial stewardship programs. *Clin Infect Dis*. 2015;61:807–8.
53. Evans SR, Rubin D, Follmann D, Pennello G, Huskins WC, Powers JH, et al. Desirability of Outcome Ranking (DOOR) and Response Adjusted for Duration of Antibiotic Risk (RADAR). *Clin Infect Dis*. 2015;61:800–6.
54. Evans S, Follmann D, Schoenfeld D, Fowler VG Jr, Chambers HF. Reply to Phillips Morris, and Walker. *Clin Infect Dis*. 2016;62:815–6.
55. Phillips PP, Morris TP, Walker AS. DOOR/RADAR: A gateway into the unknown? *Clin Infect Dis*. 2016;62:814–5.
56. Aberegg SK, Richards DR, O'Brien JM. Delta inflation: A bias in the design of randomized controlled trials in critical care medicine. *Crit Care*. 2010;14:R77.
57. Harhay MO, Wagner J, Ratcliffe SJ, Bronheim RS, Gopal A, Green S, et al. Outcomes and statistical power in adult critical care randomized trials. *Am J Respir Crit Care Med*. 2014;189:1469–78.
58. Halpern SD, Karlawish JH, Berlin JA. The continuing unethical conduct of underpowered clinical trials. *JAMA*. 2002;288:358–62.
59. Pogue J, Yusuf S. Overcoming the limitations of current meta-analysis of randomised controlled trials. *Lancet*. 1998;351:47–52.
60. Ioannidis J, Lau J. Evolution of treatment effects over time: Empirical insight from recursive cumulative metaanalyses. *Proc Natl Acad Sci U S A*. 2001;98:831–6.
61. Nagendran M, Pereira TV, Kiew G, Altman DG, Maruthappu M, Ioannidis JP, et al. Very large treatment effects in randomised trials as an empirical marker to indicate whether subsequent trials are necessary: Meta-epidemiological assessment. *BMJ*. 2016;355:i5432.
62. Pupelis G, Selga G, Austrums E, Kaminski A. Jejunal feeding, even when instituted late, improves outcomes in patients with severe pancreatitis and peritonitis. *Nutrition*. 2001;17:91–4.
63. Guyatt GH, Oxman AD, Kunz R, Brozek J, Alonso-Coello P, Rind D, et al. GRADE guidelines 6. Rating the quality of evidence—imprecision. *J Clin Epidemiol*. 2011;64:1283–93.
64. Reintam Blaser A, Starkopf J, Alhazzani W, Berger MM, Casaer MP, Deane AM, et al. Early enteral nutrition in critically ill patients: ESICM clinical practice guidelines. *Intensive Care Med*. 2017;43:380–98.
65. Zhang Z, Xu X, Ni H. Small studies may overestimate the effect sizes in critical care meta-analyses: A meta-epidemiological study. *Crit Care*. 2013;17:R2.
66. Latour-Perez J, Cabello-López JB. Interrupción precoz de los ensayos clínicos ¿Demasiado bueno para ser cierto? *Med Intensiva*. 2007;31:518–20.
67. Montori VM, Devreux PJ, Adhikari NK, Burns KE, Eggert CH, Briel M, et al. Randomized trials stopped early for benefit: A systematic review. *JAMA*. 2005;294:2203–9.
68. Abraham E, Reinhart K, Opal S, Demeyer I, Doig C, Rodriguez AL, et al. Efficacy and safety of tifacogin (recombinant

- tissue factor pathway inhibitor) in severe sepsis: A randomized controlled trial. *JAMA*. 2003;290:238–47.
69. Dahabreh IJ, Hayward R, Kent DM. Using group data to treat individuals: Understanding heterogeneous treatment effects in the age of precision medicine and patient-centred evidence. *Int J Epidemiol*. 2017;45:2184–93.
 70. Wallach JD, Sullivan PG, Trepanowski JF, Sainani KL, Steyerberg EW, Ioannidis JA. Evaluation of evidence of statistical support and corroboration of subgroup claims in randomized clinical trials. *JAMA Intern Med*. 2017;177:554–60.
 71. Kreyman KG, Berger MM, Deutz NE, Hiesmayr M, Jolliet P, Kazandjiev G, et al. ESPEN Guidelines on Enteral Nutrition: Intensive care. *Clin Nutr*. 2006;25:210–23.
 72. Tao KM, Li XQ, Yang LQ, Yu WF, Lu ZJ, Sun YM, et al. Glutamine supplementation for critically ill adults [consultado 5 Jul 2017]. *Cochrane Database Syst Rev* CD010050.pub2 Disponible en: <http://www.ncbi.nlm.nih.gov/pubmed/25199493>.
 73. Heyland D, Muscedere J, Wischmeyer PE, Cook D, Jones G, Albert M, et al. A randomized trial of glutamine and antioxidants in critically ill patients. *N Engl J Med*. 2013;368:1489–97.
 74. Gøtzsche PC. *Medicamentos que matan y crimen organizado*. Barcelona: Libros del Lince; 2014.
 75. Alltrials. All Trials Registered | All Results Reported [consultado 21 Mar 2017]. Disponible en: www.alltrials.net.
 76. Chan AW, Hrobjartsson A, Haahr MT, Gotzsche PC, Altman DG. Empirical evidence for selective reporting of outcomes in randomized trials: Comparison of protocols to published articles. *JAMA*. 2004;291:2457–65.
 77. James MF, Michell WL, Joubert IA, Nicol AJ, Navsaria PH, Gillespie RS. Resuscitation with hydroxyethyl starch improves renal function and lactate clearance in penetrating trauma in a randomized controlled study: the FIRST trial (Fluids in Resuscitation of Severe Trauma). *Br J Anaesth*. 2011;107:693–702.
 78. Reinhart K, Hartog CS. Hydroxyethyl starch in patients with trauma. *Br J Anaesth*. 2012;108:321–2 [author reply 2–4].
 79. Delaney A, Bagshaw SM, Ferland A, Manns B, Laupland KB, Doig CJ. A systematic evaluation of the quality of meta-analyses in the critical care literature. *Crit Care*. 2005;9:R575–82.
 80. Lexchin J, Light DW. Commercial influence and the content of medical journals. *BMJ*. 2006;332:1444–7.
 81. Barnés HG. «Fake journals»: dentro del mercado negro de las investigaciones científicas falsas [consultado 5 Jul 2017]. *El Confidencial* 2017. Disponible en: http://www.elconfidencial.com/alma-corazon-vida/2017-02-13/fake-journals-mercado-negro-investigaciones-cientificas_1329506.
 82. Sorokowski P, Kulczycki E, Sorokowska A, Pisanski K. Predatory journals recruit fake. *Nature*. 2017;543:481–3.
 83. Predatory open access publishing [consultado 15 Abril 2017]. Disponible en: https://en.wikipedia.org/wiki/Predatory_open_access_publishing.
 84. Institute of Medicine. Discussion framework for clinical trial data sharing: Guiding principles, elements, and activities. 2015 [consultado 15 Abril 2017]. Disponible en: <http://www.iom.edu/Reports/2014/Discussion-Framework-for-Clinical-Trial-Data-Sharing.aspx>.
 85. Taichman DB, Backus J, Baethge C, Bauchner H, de Leeuw PW, Drazen JM, et al. Sharing clinical trial data: A proposal from the International Committee of Medical Journal. *N Z Med J*. 2016;129:7–9.
 86. EQUATOR [consultado 21 Mar 2017]. Disponible en: <http://www.equator-network.org/>.
 87. Dans AL, Dans LF, Agoritsas T, Guyatt G. Applying results to individual patients. En: Guyatt G, Rennie D, Meade MO, Cook DJ, editores. *User's guides to the medical literature: A manual for evidence-based clinical practice*. 3th ed. New York: JAMA evidence; 2015. p. 235–47.
 88. McCartney M, Treadwell J, Maskrey N, Lehman R. Making evidence based medicine work for individual patients. *BMJ*. 2016;353:i2452.
 89. Guyatt G, Gutterman D, Baumann MH, Addrizzo-Harris D, Hylek EM, Phillips B, et al. Grading strength of recommendations and quality of evidence in clinical guidelines: Report from an american college of chest physicians task force. *Chest*. 2006;129:174–81.
 90. Guyatt GH, Oxman AD, Kunz R, Falck-Ytter Y, Vist GE, Liberati A, et al. Going from evidence to recommendations. *BMJ*. 2008;336:1049–51.
 91. Honore PM, Joannes-Boyau O, Collin V, Boer W, Jennes S. Continuous hemofiltration in 2009: What is new for clinicians regarding pathophysiology, preferred technique and recommended dose? *Blood Purif*. 2009;28:135–43.
 92. Joannes-Boyau O, Honore PM, Perez P, Bagshaw SM, Grand H, Canivet JL, et al. High-volume versus standard-volume haemofiltration for septic shock patients with acute kidney injury (IVOIRE study): A multicentre randomized controlled trial. *Intensive Care Med*. 2013;39:1535–46.
 93. Ranieri VM, Rubenfeld GD, Thompson BT, Ferguson ND, Caldwell E, Fan E, et al. Acute respiratory distress syndrome: The Berlin Definition. *JAMA*. 2012;307:2526–33.
 94. Singer M, Deutschman CS, Seymour CW, Shankar-Hari M, Annane D, Bauer M, et al. The third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). *JAMA*. 2016;315:801–10.
 95. Thille AW, Esteban A, Fernandez-Segoviano P, Rodriguez JM, Aramburu JA, Penuelas O, et al. Comparison of the Berlin definition for acute respiratory distress syndrome with autopsy. *Am J Respir Crit Care Med*. 2013;187:761–7.
 96. Hurley JC. Individual patient data meta-analysis in intensive care medicine and contextual effects. *Intensive Care Med*. 2010;36:903–4 [author reply 5].
 97. Hurley JC. Topical antibiotics as a major contextual hazard toward bacteremia within selective digestive decontamination studies: A meta-analysis. *BMC Infect Dis*. 2014;14:714.
 98. Hurley JC. ICU-acquired candidemia within selective digestive decontamination studies: A meta-analysis. *Intensive Care Med*. 2015;41:1877–85.
 99. Hurley JC. SDD and contextual effect. *Br J Anaesth*. 2015;115:479–80.
 100. Deans KJ, Minneci PC, Danner RL, Eichacker PQ, Natanson C. Practice misalignments in randomized controlled trials: Identification, impact, and potential solutions. *Anesth Analg*. 2010;111:444–50.
 101. Garber AM, Tunis SR. Does comparative-effectiveness research threaten personalized medicine? *N Engl J Med*. 2009;360:1925–7.
 102. Rello J, Perez A. Precision medicine for the treatment of severe pneumonia in intensive care. *Expert Rev Respir Med*. 2016;10:297–316.
 103. Calfee CS, Delucchi K, Parsons PE, Thompson BT, Ware LB, Matthay MA. Subphenotypes in acute respiratory distress syndrome: Latent class analysis of data from two randomised controlled trials. *Lancet Respir Med*. 2014;2:611–20.
 104. Maslove DM, Lamontagne F, Marshall JC, Heyland DK. A path to precision in the ICU. *Crit Care*. 2017;21:79.
 105. Mokhles S, Takkenberg JJ, Treasure T. Evidence-based and personalized medicine. It's [AND] not [OR]. *Ann Thorac Surg*. 2017;103:351–60.
 106. Teutsch SM, Bradley LA, Palomaki GE, Haddow JE, Piper M, Calonge N, et al., EGAPP Working Group. The Evaluation of Genomic Applications in Practice and Prevention (EGAPP) Initiative: Methods of the EGAPP Working Group. *Genet Med*. 2009;11:3–14.
 107. Khoury MJ, Gwinn M, Yoon PW, Dowling N, Moore CA, Bradley L. The continuum of translation research in genomic medicine:

- How can we accelerate the appropriate integration of human genome discoveries into health care and disease prevention? *Genet Med.* 2007;9:665–74.
108. Khoury MJ, Gwinn M, Dotson WD, Schully SD. Knowledge integration at the center of genomic medicine. *Genet Med.* 2012;14:643–7.
109. Clyne M, Schully SD, Dotson WD, Douglas MP, Gwinn M, Kolor K, et al. Horizon scanning for translational genomic research beyond bench to bedside. *Genet Med.* 2014;16:535–8.
110. Schully SD, Lam TK, Dotson WD, Chang CQ, Aronson N, Birke-land ML, et al. Evidence synthesis and guideline development in genomic medicine: Current status and future prospects. *Genet Med.* 2015;17:63–7.
111. Tobin MJ. The role of a journal in a scientific controversy. *Am J Respir Crit Care Med.* 2003;168:511.
112. Brochard L, Mancebo J, Tobin M. Searching for evidence: Don't forget the foundations. *Intensive Care Med.* 2003;29:2109–11.
113. Hayes MA, Timmins AC, Yau EH, Palazzo M, Hinds CJ, Watson D. Elevation of systemic oxygen delivery in the treatment of critically ill patients. *N Engl J Med.* 1994;330:1717–22.
114. Joyner MJ, Paneth N. Seven questions for personalized medicine. *JAMA.* 2015;314:999–1000.
115. Ahasic AM, Christiani DC. Personalized critical care medicine: How far away are we? *Semin Respir Crit Care Med.* 2015;36:809–22.
116. Scott SA, Lubitz SA. Warfarin pharmacogenetic trials: Is there a future for pharmacogenetic-guided dosing? *Pharmacogenomics.* 2014;15:719–22.
117. Bellomo R, Landoni G, Young P. Improved survival in critically ill patients: Are large RCTs more useful than personalized medicine? Yes. *Intensive Care Med.* 2016;42:1775–7.
118. Gattinoni L, Tonetti T, Quintel M. Improved survival in critically ill patients: are large RCTs more useful than personalized medicine? We are not sure. *Intensive Care Med.* 2016;42:1781–3.
119. Berg RM, Moller K, Rossel PJ. European legislation impedes critical care research and fails to protect patients' rights. *Crit Care.* 2011;15:148.
120. Carcas Sansuán AJ. Investigación comparativa de la efectividad. Integración de la investigación y la práctica clínica. En: Dal-Re R, editor. *Luces y sombras en la investigación clínica.* Madrid: Triascastela; 2014. p. 371–97.
121. Seymour CW, Kahn JM. Resolving conflicting comparative effectiveness research in critical care. *Crit Care Med.* 2012;40:3090–2.
122. Cooke CR, Iwashyna TJ. Using existing data to address important clinical questions in critical care. *Crit Care Med.* 2013;41:886–96.
123. Kitsios GD, Dahabreh IJ, Callahan S, Paulus JK, Campagna AC, Dargin JM. Can we trust observational studies using propensity scores in the critical care literature? A systematic comparison with randomized clinical trials. *Crit Care Med.* 2015;43:1870–9.
124. Albert RK. "Lies, damned lies..." and observational studies in comparative effectiveness research. *Am J Respir Crit Care Med.* 2013;187:1173–7.
125. James S, Rao SV, Granger CB. Registry-based randomized clinical trials —a new clinical trial paradigm. *Nat Rev Cardiol.* 2015;12:312–6.
126. Crespi CM. Improved designs for cluster randomized trials. *Annu Rev Public Health.* 2016;37:1–16.
127. Forbes AB, Akram M, Pilcher D, Cooper J, Bellomo R. Cluster randomised crossover trials with binary data and unbalanced cluster sizes: Application to studies of near-universal interventions in intensive care. *Clin Trials.* 2015;12:34–44.
128. Seppelt IM. Trial designs for old problems in a new era. *Crit Care Med.* 2011;39:1561–2.
129. Van Duijn PJ, Bonten MJ. Antibiotic rotation strategies to reduce antimicrobial resistance in Gram-negative bacteria in European intensive care units: Study protocol for a cluster-randomized crossover controlled trial. *Trials.* 2014;15:277.
130. Relton C, Torgerson D, O'Cathain A, Nicholl J. Rethinking pragmatic randomised controlled trials: introducing the "cohort multiple randomised controlled trial" design. *BMJ.* 2010;340:c1066.
131. Herasevich V, Pieper MS, Pulido J, Gajic O. Enrollment into a time sensitive clinical study in the critical care setting: Results from computerized septic shock sniffer implementation. *J Am Med Inform Assoc.* 2011;18:639–44.
132. Bhatt DL, Mehta C. Adaptive designs for clinical trials. *N Engl J Med.* 2016;375:65–74.
133. Chalmers I, Bracken MB, Djulbegovic B, Garattini S, Grant J, Gulmezoglu AM, et al. How to increase value and reduce waste when research priorities are set. *Lancet.* 2014;383:156–65.
134. Chan AW, Song F, Vickers A, Jefferson T, Dickersin K, Gotzsche PC, et al. Increasing value and reducing waste: Addressing inaccessible research. *Lancet.* 2014;383:257–66.
135. Ioannidis JP, Greenland S, Hlatky MA, Khoury MJ, Macleod MR, Moher D, et al. Increasing value and reducing waste in research design, conduct, and analysis. *Lancet.* 2014;383:166–75.