



Planteamientos de Investigación Clínica y sus soluciones

ESCUELA SEMI DE INVESTIGACIÓN CLÍNICA



Área de Formación Online SEMI-FEMI
Sociedad Española de Medicina Interna (SEMI)
Fundación Española de Medicina Interna (FEMI)

© Sociedad Española de Medicina Interna (SEMI)
© Fundación Española de Medicina Interna (FEMI)

Todos los textos, imágenes y documentos presentes en esta publicación son propiedad intelectual de SEMI.
Se autoriza a visualizar e imprimir estos materiales mientras sean respetadas las siguientes condiciones:

1. Los textos, imágenes y documentos solo pueden ser utilizados con fines informativos.
2. Los textos, imágenes y documentos no pueden ser utilizados para propósitos comerciales.
3. Cualquier copia de estos textos, imágenes y documentos, o de parte de los mismos, deberá incluir esta advertencia de derechos reservados y el reconocimiento de la autoría.

15 de febrero de 2021

ISBN: 978-84-09-27653-0

Edita: Multimédica Proyectos, S.L.

C/. Menéndez Pidal, 27 bajo. 28036 Madrid

campussemi@multimedica proyectos.com

www.multimedica proyectos.com

Dr. Pablo Pérez Martínez

Director Científico del Instituto Maimónides
de Investigación Biomédica de Córdoba (IMIBIC)
Catedrático de Medicina, Universidad de Córdoba
Especialista de Medicina Interna
Hospital Universitario Reina Sofía (Córdoba)

Dr. Emilio Casariego Vales

Jefe de Servicio de Medicina Interna
Hospital Universitario Lucus Augusti (Lugo)
Expresidente de la Sociedad Española de Medicina Interna
y de la Fundación Española de Medicina Interna



directores
editores

prólogo

La visión global de la persona enferma define a la especialidad de Medicina Interna. Para los internistas no es posible atender bien a un enfermo sin tener en consideración todos sus problemas de salud, sus características únicas como persona y su interacción con el medio social y natural.

De la misma forma, los internistas tienen una visión integral de la medicina. Un buen internista atiende correctamente a sus enfermos, se preocupa por aprender y mantenerse continuamente formado, enseña a sus compañeros, a los médicos en formación y a los estudiantes de las profesiones sanitarias, y se preocupa por generar nuevo conocimiento que redunde en el beneficio de sus pacientes.

Este libro está dedicado a estas dos últimas tareas del internista. Un grupo de profesionales, muchos de ellos internistas, ha desarrollado un curso de iniciación a la investigación. En este curso enseñan metodología para ayudar a diseñar, desarrollar e interpretar los estudios de investigación. Han utilizado el método del caso, una aproximación docente que facilita la aplicación, la comprensión y el aprendizaje. Este método se basa en la presentación de uno o varios escenarios con diferentes resultados posibles. De esta forma se discute la aplicación de las técnicas estadísticas, la interpretación de los resultados y la forma más adecuada de presentarlos.

Los internistas y otros profesionales que sigan este curso y lean este libro van a desarrollar competencias en investigación que les van a permitir ensanchar los límites de la ciencia estableciendo nuevas hipótesis, desarrollando nuevos estudios y aumentando el conocimiento sobre las enfermedades, su pronóstico y su tratamiento.

Desde la Sociedad Española de Medicina Interna (SEMI) damos la bienvenida a iniciativas como esta que presento y agradecemos el esfuerzo de los directores y coordinadores del curso y de los profesores y docentes para llevarla a cabo. Nuestro agradecimiento se extiende a la empresa Boehringer Ingelheim, que ha creído en este proyecto y ha contribuido a su financiación. También quiero agradecer a Multimédica Proyectos la organización logística y de la plataforma docente en la que se han desarrollado las actividades docentes.

Finalmente, quiero manifestar mi esperanza de que iniciativas tan importantes y necesarias como esta tengan desarrollo y continuidad en el futuro.

Dr. Jesús Díez Manglano

Presidente

Sociedad Española de Medicina Interna (SEMI)

Fundación Española de Medicina Interna (FEMI)



presentaciones

Satisfacer necesidades médicas y asistenciales no cubiertas para mejorar la calidad de vida de las personas es la misión principal de Boehringer Ingelheim. Por ello, hace más de 130 años que aportamos valor a través de la innovación en diversas áreas terapéuticas y a través del compromiso con el entorno que nos rodea, con una firme apuesta por la sostenibilidad. Los pacientes y su bienestar son el motor de nuestra actividad y, por ello, nuestro objetivo es buscar soluciones que mejoren su calidad de vida.

Boehringer Ingelheim dedica gran parte de sus esfuerzos y recursos a la investigación y el desarrollo de nuevos tratamientos que mejoren la calidad de vida de los pacientes; desde la búsqueda y creación de nuevos principios activos, hasta su testeo en los ensayos clínicos.

En este sentido, cabe destacar la importante contribución de España a nivel mundial en los ensayos clínicos promovidos por la compañía en términos de pacientes reclutados. Y esto es gracias a la colaboración de investigadores de prestigio que llevan a cabo investigaciones con los más altos estándares de calidad y al compromiso de Boehringer Ingelheim en impulsar los ensayos clínicos como la principal fuente de avance del conocimiento en Medicina.

Boehringer Ingelheim y la Sociedad Española de Medicina Interna (SEMI) mantienen una estrecha e histórica colaboración en cuanto a la formación continuada de los internistas se refiere, siendo socios estratégicos desde 2015 en el campo del paciente pluripatológico y de edad avanzada. Sabemos que la esperanza de vida será cada vez mayor en España, aparecen nuevas patologías, los pacientes necesitan soluciones más globales y los profesionales sanitarios se enfrentan a situaciones cada vez más complejas que requieren de más competencias, como la investigación clínica.

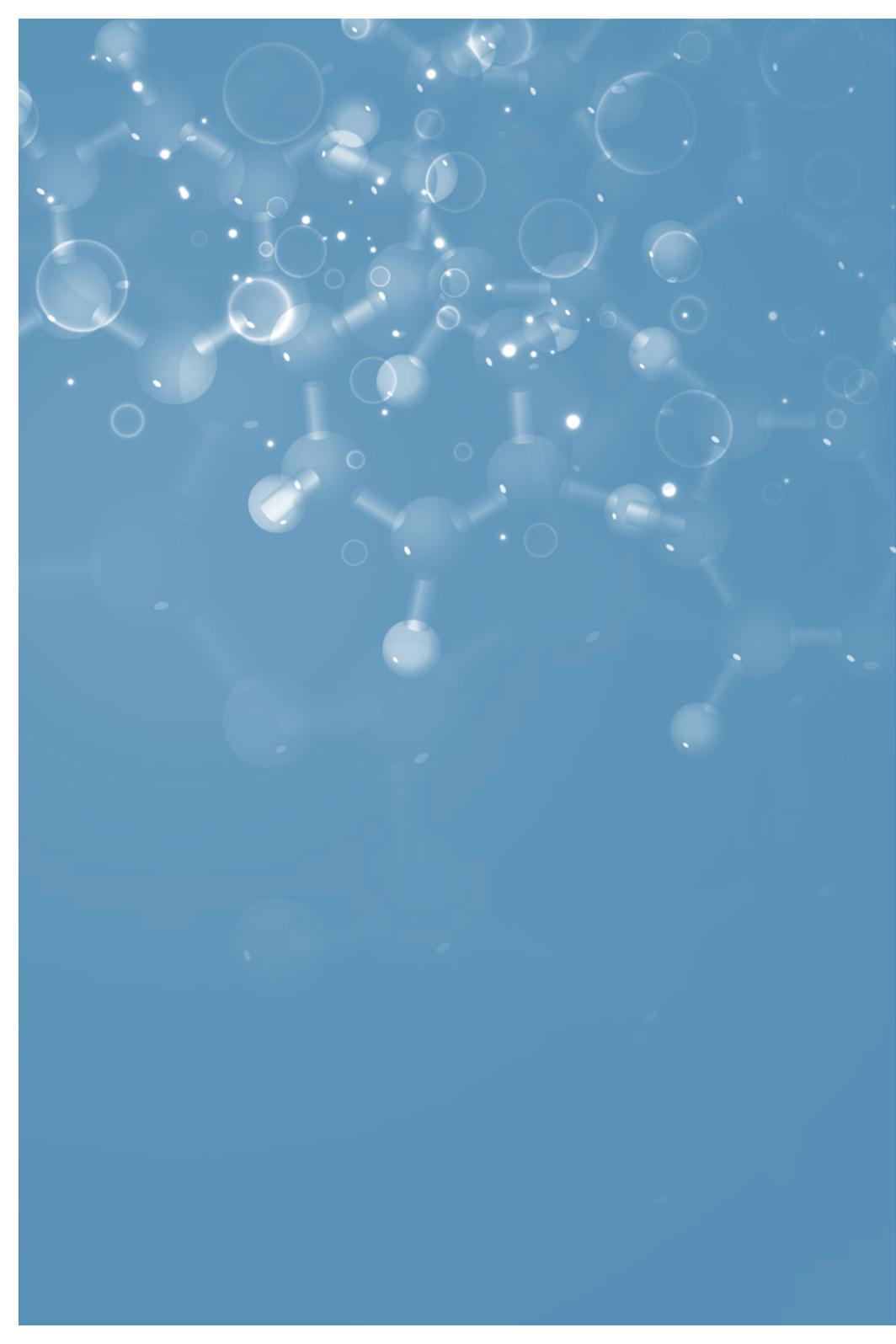
Sensibles a la realidad y a los retos a los que se enfrenta el médico internista, en Boehringer Ingelheim nos sentimos muy orgullosos de haber colaborado con SEMI en la creación de la Escuela de Investigación Clínica SEMI en 2020 y de haber puesto en marcha el primer curso para internistas dedicado íntegramente a la investigación clínica.

Afianzamos así nuestro compromiso con la especialidad de Medicina Interna para recorrer juntos nuevos caminos de colaboración en algo tan importante y estratégico para SEMI y para Boehringer Ingelheim como es la investigación de nuevos medicamentos y nuevas soluciones para mejorar la calidad de vida de las personas.

Dr. René Saito

Director Médico

Boehringer-Ingelheim, España



Es un gran honor presentar el libro ***Planteamientos de investigación clínica y sus soluciones*** de la Escuela de Investigación de la Sociedad Española de Medicina Interna (SEMI). En la actualidad, la Escuela está desarrollando un curso de investigación dirigido a internistas y que consta de cuatro módulos. Cada uno de ellos incluye un curso *online* con conocimientos teóricos, clases presenciales de ejercicios prácticos, un examen de evaluación para obtener los créditos correspondientes y un libro de lectura con ejemplos prácticos de base clínica. Este libro que presentamos, el primero de la Escuela, ha sido diseñado para enriquecer y afianzar los conocimientos adquiridos en el Módulo I de Investigación Clínica.

Con un diseño de capítulos muy cortos, el lector navegará entre problemas de diferente complejidad, muy cercanos a los que se encuentran los internistas en su trabajo diario. Para cada uno de ellos se comenta en profundidad una posible solución. El objetivo es doble. Por una parte, se pretende ayudar a comprender, desmenuzar y aplicar los resultados de los estudios publicados, por complejos que sean. Por otra, poner al alcance del lector, de forma amena y breve, herramientas de ayuda en el planteamiento y diseño de estudios de investigación clínica. Los editores creemos que el resultado es notable y quedamos en deuda con todos los autores por su extraordinario trabajo.

La SEMI no se debe únicamente a los problemas clínicos más complejos ni a los avances más recientes. Uno de sus fines primordiales es transmitir los conocimientos y valores precisos para el trabajo diario de sus asociados. Por ello, esta Escuela nació con el propósito de aproximar a todos los socios los medios e instrumentos que permiten diseñar y construir investigación clínica de alta calidad orientada a nuestros pacientes. La idea es aportar conocimientos teóricos e instruir con ejercicios prácticos muy cercanos a nuestra realidad, pensados por y para clínicos y adaptables a los distintos perfiles de nuestros socios. Entendemos que en este libro se han conseguido estos objetivos y sentimos un legítimo orgullo al prologarlo.

No podemos menos que agradecer el gran trabajo de sus autores, como se plasma en la brillantez de la obra. Además, queremos transmitir a la Junta Directiva de la Sociedad nuestro reconocimiento y profunda gratitud por su apoyo en esta aventura. Es de justicia afirmar que tenemos claro que sin el apoyo decidido de Boehringer Ingelheim hubiese sido imposible. Por ello debemos expresar nuestro mayor reconocimiento a su equipo de medicina, *alma mater* y brillante impulsor de esta obra. Por último, queremos agradecer la labor de Emilia Carretero, pues sin su ayuda estas páginas habrían encallado.

Dr. Pablo Pérez Martínez Dr. Emilio Casariego Vales

Directores y Editores
Escuela SEMI de Investigación Clínica

contenidos

CAPÍTULO 1

Resolver un problema clínico mediante un estudio de series de casos

Dr. A. Javier Trujillo Santos 10

CAPÍTULO 2

Resolver un problema clínico mediante el cálculo de una prevalencia

Dr. Jesús Díez Manglano 15

CAPÍTULO 3

Analizar el impacto de diferentes factores de riesgo para el desarrollo de una enfermedad.

Regresión logística 21

Dr. Javier Solera Santos

CAPÍTULO 4

Resolver un problema clínico mediante un estudio de casos y controles

Dr. Pedro Pablo Casado Escribano 28

CAPÍTULO 5

Resolver un problema clínico con un análisis de supervivencia mediante el estimador Kaplan-Meier

Dr. Javier Ena Muñoz 34

CAPÍTULO 6

Determinar los factores pronóstico. Análisis de Cox

Dr. Emilio Casariego Vales 41

CAPÍTULO 7

Resolver un problema clínico mediante un ensayo clínico de un fármaco frente al tratamiento habitual o placebo

Dr. Joan Carles Trullàs 49

CAPÍTULO 8

La investigación clínica sobre vacunas: la vacuna frente a la COVID-19

Dr. José López Miranda 55

CAPÍTULO 9

Resolver un problema clínico mediante un metaanálisis tradicional

Dra. Rosa Solà Alberich, Dra. Lucía Tarro,
Dra. Elisabet Llauredó y Dña. María de la Serra Besora 65

CAPÍTULO 10

Resolver un problema clínico mediante un metaanálisis en red

Dra. María Montserrat Chimeno Viñas 73

Resolver un problema clínico mediante un estudio de series de casos

Dr. A. Javier Trujillo Santos
Servicio de Medicina Interna
Hospital General Universitario Santa Lucía
Hospital Santa María del Rosell
(Cartagena, Murcia)

capítulo 1



1. Situación actual y contexto clínico

Durante la etapa inicial de la pandemia de COVID-19, los primeros artículos se centraron en las diferentes manifestaciones pulmonares de los pacientes, hasta que Klok *et al.* (véase la bibliografía) describieron que una serie de pacientes ingresados en las unidades de cuidados intensivos (UCI) con dicha infección desarrollaban con una elevada frecuencia complicaciones tromboticas venosas (tromboembolismo venoso, TEV) y arteriales, sobre todo embolia pulmonar. Esta alta incidencia, de hasta un 31%, era inesperadamente elevada y sorprendía que no se hubiera comunicado en series anteriores. Además, estos autores mostraron que las dosis de heparina de bajo peso molecular (HBPM) utilizadas para prevenir esta complicación no se habían mostrado eficaces, y lanzaban la hipótesis de utilizar una dosis superior como medida preventiva eficaz de las complicaciones relacionadas con TEV en pacientes con COVID-19 ingresados en UCI.

2. Planteamiento, pregunta de investigación y objetivos

Ante la ausencia de otros datos clínicos y la alarma que suponía una elevada incidencia de complicaciones relacionadas con TEV en los pacientes con COVID-19 ingresados en UCI, no comunicada con anterioridad, y aún más la ineficacia de la HBPM utilizada a las dosis habituales recomendadas para otras enfermedades similares en la prevención de tales complicaciones, realmente se planteaban dos problemas clínicos. El primero de ellos, la elevada incidencia de TEV en los pacientes con COVID-19 ingresados en UCI, cuestión que habría que comprobar con otras series de casos y cohortes, preferentemente de otras UCI, comparando las características de los pacientes de cada una de ellas de forma que se pudiera determinar esa inesperada incidencia. El segundo problema de investigación surge en torno a la ineficacia de la trombopprofilaxis con las dosis habituales de HBPM, circunstancia que hasta ese momento no se había planteado en enfermedades similares a la COVID-19.

El objetivo principal de los estudios futuros debería orientarse a confirmar (o refutar) la incidencia tan elevada de eventos de TEV en los pacientes con COVID-19 ingresados en UCI, por lo que sería necesario comparar con series de casos (y cohortes) de pacientes con el mismo diagnóstico de otros hospitales. Las variables fundamentales que habría que recoger para poder comparar los resultados son las consideradas como finales, es decir, el TEV, tanto embolia pulmonar como trombosis venosa profunda, diagnosticadas de una forma objetiva (mediante exploraciones complementarias estándares, es decir, angio-TAC de arterias pulmonares y ecografía venosa de compresión de los miembros inferiores, respectivamente). Es

importante que además se detallen otras variables que determinarán las características de los pacientes con el objetivo de poder establecer qué características diferenciales hacen que estos pacientes tengan una elevada incidencia de TEV. Se deberán recoger los datos demográficos de los pacientes (género, edad, peso corporal o índice de masa corporal) e información sobre las comorbilidades que pudieran influir en esa incidencia, como los antecedentes personales y familiares de TEV, la utilización de corticosteroides, la presencia de cáncer activo, las alteraciones de la coagulación previas al ingreso en UCI o durante su estancia o la utilización de fármacos anticoagulantes (y antiagregantes) y en qué dosis, y todas aquellas otras variables consideradas de interés. Por otra parte, en cuanto a un nivel superior, como es el tipo de hospital, tendría interés conocer qué características tiene, si es universitario, de carácter público o privado, si se trata de un hospital que asiste a un tipo de pacientes determinado (traumatológicos, de unas determinadas compañías de asistencia sanitaria privada con una cobertura restringida, etc.). Otro aspecto que habría que considerar son los criterios de ingreso en UCI de estos pacientes, porque, aunque existen criterios más o menos homogéneos, como la necesidad de ventilación mecánica, puede haber otros más específicos relacionados con dicho ingreso, o incluso algunos determinados por las características de la propia UCI. Solo con la comparación de las características de los pacientes incluidos en cada una de las series de casos (y cohortes) podremos estimar la incidencia real de eventos de TEV de los pacientes con COVID-19 ingresados en UCI.

En cuanto al segundo objetivo, determinar si la dosis profiláctica de HBPM es eficaz en la prevención del TEV en pacientes con COVID-19 ingresados en UCI, además de valorar los eventos finales, como se hizo en relación con el objetivo anterior, debemos recoger qué tipo de fármacos se utilizaron en la trombopprofilaxis, las dosis empleadas y la duración en días, tanto durante el ingreso del paciente en la UCI como con anterioridad. Además, si se emplearon otras medidas físicas de trombopprofilaxis, también deberán quedar explicitadas.

3. Tipo de estudio y diseño. Pruebas estadísticas y abordaje del problema

Dado que los estudios de series de casos tienen un nivel de evidencia bajo, se requiere de otro tipo de diseño de estudios que mejore esa calidad. En un escalón inmediatamente superior disponemos de los estudios de cohortes, una suerte de serie de casos mejorada que garantiza una mayor precisión a la hora de valorar la incidencia de una determinada manifestación en una entidad clínica concreta,

fundamentalmente debido a su diseño prospectivo (menos frecuentemente, retrospectivo), con un protocolo de recogida de casos perfectamente definido en cuanto a criterios de inclusión y exclusión que asegura la representatividad de los pacientes incluidos. No obstante, ambos diseños carecen de grupo de comparación con el que contrastar una eventual hipótesis. Por ello, los estudios de casos y controles podrían ser más útiles en el caso de enfermedades o manifestaciones clínicas poco frecuentes, y los ensayos clínicos quedarían como diseño ideal de comprobación de la eficacia de un determinado tratamiento frente al tratamiento establecido como estándar hasta ese momento o frente a placebo en el caso de que no existiera tratamiento disponible.

En nuestro caso, en el estudio de Klok *et al.* se planteaban dos objetivos, como ya mencionamos anteriormente. En cuanto al primero, la alta incidencia de eventos de TEV (31 %) en pacientes con COVID-19 ingresados en la UCI, resonó en la comunidad científica al no haberse comunicado en las escasas series anteriores. Rápidamente, los propios autores corroboraron sus resultados con la publicación de un nuevo artículo en el que daban más detalles de la serie de casos original, aumentaban el tiempo de seguimiento de los pacientes y realizaban un nuevo análisis de los datos. En esta nueva serie de casos, los autores no solo confirmaban sus resultados con el mismo porcentaje de eventos trombóticos (31 %), sino que realizaban un nuevo análisis en el intento de justificar tan inesperado resultado. No solo volvieron a realizar una medición de la incidencia acumulada de eventos, sino que añadieron el análisis de riesgos competitivos dada la alta mortalidad de estos pacientes, para evitar el sesgo que suponía que un paciente que muere sin haber desarrollado un evento trombótico obviamente ya no estaría en riesgo de desarrollarlo, con lo que la estimación de tal complicación aún aumentaría más, hasta un 49% en esta serie de casos.

Otras publicaciones posteriores realizaron estimaciones de la incidencia de eventos trombóticos (sobre todo, venosos) con el mismo tipo de diseño de series de casos, y en cambio no encontraron tan alta incidencia de esta complicación. Realmente, estos estudios posteriores fueron más estrictos a la hora de establecer criterios de inclusión y exclusión de los pacientes valorados en las respectivas series de casos, de forma que se pareciera más al seguimiento de una cohorte de pacientes con COVID-19 ingresados en UCI, con la intención de evitar un sesgo de selección. En una revisión sistemática en la que se analizaron todos los estudios publicados al respecto (Kunutsor y Laukkanen), se detectó que solo un estudio tenía una incidencia similar a la referida por el grupo de Klok *et al.*, con 19 estudios que detectaban una incidencia inferior, de forma que se consideró que una incidencia algo

inferior al 20 %, que era la que se obtenía en el metaanálisis de todos estos estudios, era la más cercana a la realidad.

En cuanto al segundo objetivo mencionado en el estudio original de Klok *et al.*, la recomendación que realizan los autores de valorar la utilización de dosis superiores de HBPM para la trombotoprofilaxis primaria de los pacientes con COVID-19 ingresados en UCI, a la vista de la ineficacia de las dosis habituales en su serie de casos, no deja de ser una conjetura realizada a partir de los resultados obtenidos. Tal recomendación se basa en la observación de que los pacientes desarrollaron eventos trombóticos a pesar de que todos habían recibido dosis profilácticas de HBPM, lo que indica una tasa de fallo de la profilaxis anormalmente elevada. Este diseño de serie de casos solo es válido para alertar de la situación anómala observada, a resultados de lo cual se pueden lanzar diferentes hipótesis de trabajo (aumentar la dosis de HBPM, en este caso), que se deberán comprobar en estudios con otros diseños. En este caso, el diseño ideal sería un ensayo clínico aleatorizado, en el que unos pacientes con unas características previamente definidas (similares a los de la serie de casos publicada) recibieran las dosis profilácticas estándares de HBPM, y el otro grupo de tratamiento recibiera una dosis superior de HBPM. En este último grupo habría que establecer si los pacientes deberían recibir las denominadas dosis "intermedias" de HBPM o dosis terapéuticas (en lugar de profilácticas), estableciéndose así tres grupos de comparación simultáneos o dos grupos (dosis estándar frente a dosis "intermedia" o dosis terapéutica). Este diseño de ensayo clínico (de los que hay varios en marcha, pendientes aún de resultados) establecería nuevamente la incidencia de eventos trombóticos en los pacientes que recibieran las dosis estándares de profilaxis (como el estudio de Klok *et al.*) y el eventual beneficio o riesgo de utilizar las dosis superiores recomendadas en el estudio de serie de casos original.

Las medidas cuantitativas que pueden aportar los estudios de series de casos no son *odds ratio* ni riesgos relativos, dado que no hay grupo de comparación, sino prevalencias o tasas de eventos en forma de proporción (con sus respectivas medidas de precisión). Una mayor fiabilidad tendrá si la recogida de los casos se realizó mediante la definición de un protocolo con criterios claros y, por supuesto, de forma prospectiva, lo cual es poco frecuente en este tipo de estudios.

4. Justificación y discusión

Los estudios de series de casos son los estudios de investigación más antiguos; de ellos se deriva la descripción de las distintas enfermedades. Es un diseño utilizado frecuentemente en los inicios de la actividad investigadora del médico. Se

trata de un tipo de diseño observacional, basado en la identificación y descripción de un conjunto de casos clínicos que aparecen en un intervalo de tiempo. Son estudios de fácil realización, que aportan una primera aproximación a un hecho determinado en una enfermedad o a una nueva enfermedad. Recordemos, por ejemplo, la descripción de un brote de infección comunitaria de *Pneumocystis carinii* que dio lugar a la detección de una nueva enfermedad como es el síndrome de inmunodeficiencia adquirida. En general, se puede establecer que los estudios de series de casos se caracterizan por describir un nuevo fenotipo o genotipo de enfermedad (como el mencionado del sida), reconocer una manifestación conocida o común de una enfermedad rara, reconocer una manifestación rara de una enfermedad conocida (como la serie de casos expuesta en este capítulo) o común, describir un nuevo patógeno, un efecto adverso desconocido de un fármaco o un nuevo tratamiento para una enfermedad conocida, o evidenciar mecanismos de enfermedad, con motivos didácticos y de mejora de la calidad. Son estudios de rápida difusión, en la que se llama la atención deliberadamente sobre este hecho inusual, inesperado o novedoso, basado en la recolección de datos clínicos más o menos pormenorizada.

La ventaja principal de estos estudios radica en esa rapidez de su comunicación, lo que puede ser determinante ante casos de nuevas enfermedades o epidemias, y en ellos no es frecuente que se pueda establecer una secuencia temporal entre el factor de riesgo o causal y la manifestación clínica o el evento descritos. Entre sus evidentes desventajas se incluyen la ausencia de un grupo de comparación con el que establecer una eventual significación estadística y su incapacidad para probar hipótesis de investigación de forma genérica. Son, en general, estudios de baja calidad, en los que se incluye un número limitado de personas con la característica de estudio de interés y en los que incluso el hecho de desconocer si están incluidos todos los pacientes con esa característica limita la generalización de sus hallazgos. Es decir, en el caso de que se estuviera seguro de haber incluido todos los pacientes con ese determinado hallazgo, se estaría no ya ante una serie de casos, sino ante un estudio de cohortes, con características metodológicas ciertamente solapadas, aunque este último implica la existencia de un protocolo preestablecido de recogida de los pacientes, con el consiguiente diseño prospectivo y la mayor representatividad de los casos al tener la seguridad de haber incluido todos los pacientes existentes con la enfermedad de interés (sin sesgo de selección). Por todo lo dicho, las series de casos se consideran estudios de bajo nivel de evidencia, de baja calidad, aunque en ocasiones sean el único diseño de estudio posible, pues las características intrínsecas a este tipo de estudio hacen que sea insustituible como primer paso de la investigación ante las circunstancias que definen su existencia.

5. Conclusiones

Los estudios de series de casos son estudios de investigación de baja calidad científica, dado que suelen realizarse de forma retrospectiva y carecen de un grupo de comparación, sin la seguridad de incluir todos los pacientes con la característica y la enfermedad estudiadas, con el consiguiente potencial sesgo de selección. No obstante, son útiles como primera aproximación en el estudio de determinadas situaciones, como manifestaciones inusuales de enfermedades o reacciones infrecuentes a fármacos, y en ocasiones son el único abordaje posible, como en los casos de descripción de nuevas enfermedades o, al menos, asociaciones de manifestaciones clínicas que pueden denotar una nueva entidad nosológica. Aportan una medición simple, en forma de proporción o de tasas de eventos, sin que se pueda realizar una comparación con un grupo control, pero son útiles para dar a conocer unas manifestaciones inéditas sobre las que diseñar estudios de investigación con diseños más robustos y con mayor rango de evidencia.

En nuestro ejemplo introductorio del tema se pone en evidencia la utilidad de este tipo de publicación, la serie de casos, desde el momento en que alerta de una complicación hasta ese momento poco frecuente en una enfermedad que, si bien es cierto que solo llevaba unos meses descrita, ya contaba con publicaciones suficientes para que no hubiera pasado desapercibida. En posteriores series de casos se situó una incidencia real algo inferior a la de ese primer estudio, y a partir de ellas se han iniciado ensayos clínicos para comprobar la segunda hipótesis, la dosis eficaz de tromboprolifaxis en estos pacientes. Estos son el diseño de investigación que aporta la mejor evidencia científica posible para responder a esa pregunta y actualmente hay varios en marcha.

Bibliografía

- Klok FA, Kruip MJHA, van der Meer NJM, et al. Incidence of thrombotic complication in critically ill ICU patients with COVID-19. *Thromb Res.* 2020;191:145-7.
- Klok FA, Kruip MJHA, van der Meer NJM, et al. Confirmation of the high cumulative incidence of thrombotic complication in critically ill ICU patients with COVID-19: an updated analysis. *Thromb Res.* 2020;191:148-50.
- Kunutsor SK, Laukkanen JA. Incidence of venous and arterial thromboembolic complications in COVID-19: A systematic review and meta-analysis. *Thromb Res.* 2020;196:27-30.
- Masur H, Michelis MA, Greene JB, et al. An outbreak of community-acquired *Pneumocystis carinii* pneumonia: initial manifestation of cellular immune dysfunction. *N Engl J Med.* 1981;305:1431-8.
- Murad MH, Sultan S, Haffar S, Bazerbachi F. Methodological quality and synthesis of case series and case reports. *BMJ Evid Based Med.* 2018;23:60-3.

Preguntas de autoevaluación

1. Indique cuál de las siguientes es una situación en la que los estudios de series de casos pueden ser útiles:

- a) Comparación de dosis diferentes de un tratamiento.
- b) Comprobar la causalidad de una enfermedad.
- c) Referir manifestaciones infrecuentes de enfermedades frecuentes.
- d) Dos de las afirmaciones anteriores son correctas.

La respuesta correcta es la opción c.

La comparación de dosis diferentes de tratamiento requiere un diseño de ensayo clínico. Un estudio de series de casos no puede establecer la causalidad, ya que carece de la temporalidad necesaria para su inferencia. Entre las utilidades de los estudios de series de casos destaca también la referencia de manifestaciones frecuentes de enfermedades infrecuentes y la descripción de nuevas enfermedades o de reacciones adversas inesperadas de los fármacos conocidos.

2. Indique el principal motivo por el que los estudios de series de casos se consideran de baja calidad metodológica:

- a) Porque no hacen una descripción detallada de los pacientes incluidos.
- b) Porque son retrospectivos.
- c) Porque no se puede obtener un valor de p del principal resultado.
- d) Porque no tienen grupo de comparación.

La respuesta correcta es la opción d.

Los estudios de series de casos se consideran de baja calidad metodológica en cuanto a la evidencia que demuestran debido fundamentalmente a que carecen de grupo de comparación, aunque pueden ofrecer una descripción minuciosa de los pacientes incluidos. El hecho de ser retrospectivos no limitaría en exceso su validez, siempre que se cumpla el resto de criterios metodológicos posibles, aunque existen también

estudios prospectivos. No es necesario obtener un valor de p de comparación con algún valor de referencia para que la investigación de los estudios de series de casos tenga mayor o menor validez, toda vez que siempre se puede obtener una p al comparar la tasa de incidencia o la frecuencia en la que estos estudios miden sus resultados.

3. ¿Cuál es la diferencia entre un estudio de series de casos y uno de cohortes?

- a) En el estudio de cohortes el diseño es siempre prospectivo, mientras que en el de serie de casos es retrospectivo.
- b) El estudio de cohortes posee un protocolo de inclusión y exclusión de pacientes más definido y estricto que el de los estudios de series de casos.
- c) En el estudio de cohortes se pueden obtener *odds ratio* y riesgos relativos, mientras que en el estudio de series de casos solo tasas de incidencia y frecuencias.
- d) Todas las afirmaciones anteriores son correctas.

La respuesta correcta es la opción b.

Tanto los estudios de cohortes como los de serie de casos pueden ser prospectivos y retrospectivos. En los estudios de cohortes no pueden obtenerse *odds ratio* ni riesgos relativos, pues para ello hace falta otra cohorte de comparación. En el estudio de cohortes el protocolo de inclusión y exclusión de pacientes suele ser más minucioso y detallado que en los estudios de serie de casos, aunque como criterio de calidad de estos últimos también se considera que pueden llegar a tener y deben tener ese tipo de protocolo.

Resolver un problema clínico mediante el cálculo de una prevalencia

Dr. Jesús Díez Manglano
Servicio de Medicina Interna
Hospital Royo Villanova (Zaragoza)
Presidente de la Sociedad Española de Medicina Interna (SEMI)

capítulo 2

1. Situación actual y contexto clínico

La enfermedad pulmonar obstructiva crónica (EPOC) es un trastorno de las vías aéreas que se manifiesta por tos, expectoración, disnea y malestar torácico y que está relacionado fundamentalmente con el consumo de tabaco. Se acompaña con frecuencia de manifestaciones extrapulmonares y de otras comorbilidades. En España, la EPOC afecta al 11,8% de las personas mayores de 40 años. Los pacientes con EPOC disminuyen su actividad física y permanecen largos tiempos en reposo. Es frecuente que sus síntomas empeoren por la noche y que tengan asociados trastornos del sueño.

El síndrome de piernas inquietas (SPI) consiste en una urgencia irresistible para mover las piernas acompañada de sensaciones desagradables. El síntoma aparece o empeora cuando el paciente está en reposo, sentado o tumbado, y se alivia con la actividad física, por ejemplo, al caminar o hacer estiramientos. En España, el SPI afecta al 19,7% de la población. Una de las causas más frecuentes de SPI es el déficit de hierro.

Los pacientes con EPOC y SPI tienen urgencia para moverse, pero a la vez su capacidad física está disminuida. Esto les provoca una situación de intranquilidad difícil de resolver. La relación entre EPOC y SPI está poco estudiada; por otra parte, en España desconocemos cuántos pacientes con EPOC presentan SPI. En este escenario, sería interesante hacer un estudio para conocer con qué frecuencia los pacientes con EPOC experimentan SPI.

2. Planteamiento, pregunta de investigación y objetivos

Nos planteamos hacer un estudio para conocer cuál es la prevalencia del SPI en los pacientes con EPOC y si la prevalencia se modifica con el grado de obstrucción bronquial y la gravedad de la EPOC.

La prevalencia es la proporción de la población con una enfermedad en un punto o en un periodo de tiempo. No debe confundirse con la incidencia, que es el número de casos nuevos de una enfermedad que aparecen en una población determinada y durante un periodo de tiempo definido.

La pregunta de investigación es la siguiente: ¿cuál es la proporción de pacientes con EPOC que tiene SPI en junio de 2021? O, formulada de otra manera, ¿cuál es la prevalencia de SPI en pacientes con EPOC en junio de 2021?

Para responder a la pregunta de investigación desarrollamos un estudio con los siguientes objetivos:

1. Objetivo principal: estimar la prevalencia de SPI en pacientes con EPOC.
2. Objetivos secundarios:
 - Determinar la prevalencia de déficit de hierro en los pacientes con EPOC con y sin SPI.
 - Determinar si hay diferencias significativas en la prevalencia de SPI dependiendo del grado de obstrucción bronquial.

3. Tipo de estudio y diseño. Pruebas estadísticas y abordaje del problema

Para responder a nuestra pregunta y alcanzar los objetivos formulados plantearemos un estudio con las características que se exponen a continuación.

Diseño

En nuestro estudio no realizaremos ninguna intervención sobre los pacientes (no utilizaremos fármacos, ni cambios de dieta, programas de ejercicio o intervenciones quirúrgicas). Por lo tanto, se trata de un estudio observacional.

Nuestro trabajo consistirá en seleccionar pacientes con EPOC en un momento o periodo de tiempo determinado y recoger información de ellos. Esto es lo que se denomina estudio transversal (en lengua inglesa, *cross-sectional study*).

Además, queremos realizar nuestro estudio incluyendo a pacientes que estén diagnosticados de EPOC en junio de 2021. Por lo tanto, será un estudio prospectivo. Si, por ejemplo, realizáramos el estudio revisando las historias clínicas de los pacientes con EPOC atendidos en nuestra consulta en enero de 2020, el estudio sería retrospectivo.

En resumen, haremos un estudio observacional, transversal y prospectivo.

Escenario

Se refiere al contexto en el que realizaremos el estudio. Nuestro estudio podría hacerse en diferentes escenarios, como las consultas de centros de atención primaria, las de medicina interna o las de neumología, las salas de hospitalización,

los hospitales de día e incluso entre los miembros de asociaciones de pacientes con EPOC.

Nosotros realizaremos el estudio en consultas de medicina interna de cinco hospitales españoles.

Tiempo

Hemos decidido realizar nuestro estudio en junio de 2021. Por ello, incluiremos a los pacientes con EPOC estable que acudan a consulta de medicina interna a lo largo del mes de junio de 2021.

Población del estudio. Criterios de inclusión y exclusión

Para asegurar la validez interna y externa del estudio es imprescindible definir de una forma concreta y clara qué individuos serán incluidos en el estudio. Para ello hay que establecer unos criterios de inclusión y exclusión. En nuestro estudio estableceremos los siguientes:

- Criterios de inclusión:
 - Pacientes diagnosticados de EPOC por espirometría según los criterios de la Global Initiative for Chronic Obstructive Lung Disease (GOLD), es decir cociente entre el volumen espiratorio forzado en el primer segundo (FEV_1) y la capacidad vital forzada (FVC) inferior a 0,7 en la prueba posbroncodilatadora ($FEV_1/FVC < 0,7$).
 - Pacientes con EPOC en situación estable (al menos han pasado 30 días desde la resolución de la última exacerbación).
 - Pacientes atendidos en consultas de medicina interna.
- Criterios de exclusión:
 - Pacientes que no den su consentimiento informado.
 - Pacientes con deterioro cognitivo que no puedan expresar los síntomas que experimentan.

Mediciones

Para que el estudio sea reproducible hay que especificar todas las mediciones que se van a realizar y las variables que se van a recoger, y exponer cómo se va a hacer. Además, hay que establecer con qué criterios se van a definir algunas características

de los pacientes. Por último, es necesario recabar información de algunas variables que nos permitan caracterizar a la población, por ejemplo, la edad y el sexo.

Para nuestro estudio consideraremos la edad, el sexo, el consumo de tabaco, el número de exacerbaciones en el último año, el grado de obstrucción bronquial medido con el porcentaje del valor teórico del FEV_1 en la prueba posbroncodilatadora (%p FEV_1), la presencia de SPI y la existencia de ferropenia.

Para el diagnóstico de SPI se utilizarán los criterios del International Restless Legs Syndrome Study Group.

Se considerará la existencia de ferropenia cuando los valores de ferritina en suero sean inferiores a 100 ng/mL, o entre 100 y 300 ng/mL junto con saturación de transferrina inferior al 20 %.

El grado de obstrucción bronquial se medirá con el %p FEV_1 . Se considerará obstrucción leve cuando este sea superior al 80 %, moderada si está entre el 50 % y el 80 %, grave si se sitúa entre el 30 % y el 50 % y muy grave si es inferior al 30 %.

Cálculo del tamaño muestral

Es necesario que nuestro estudio tenga poder suficiente, es decir, que incluya un número suficiente de individuos que garantice que se pueden obtener diferencias estadísticas en el caso de que existan. Para ello hay que calcular el tamaño de la muestra del estudio.

Para calcular el tamaño muestral hay que tener en cuenta algunos datos: tamaño de la población si lo conocemos, nivel de confianza, precisión y valor aproximado del parámetro que queremos medir. El nivel de confianza es la seguridad que deseamos tener en nuestra investigación, es decir, 1 menos el error α , y habitualmente es 0,95 o 95 %. La precisión es el margen de error aleatorio que estamos dispuestos a admitir. Habitualmente es menor del 20 %, y cuanto menor sea el error que admitamos mayor será el tamaño muestral necesario. El valor aproximado del parámetro que queremos calcular podemos obtenerlo de la literatura. Si no es posible, hay que optar por un valor de 50 % o 0,5. También es conveniente tener en cuenta el posible número de pérdidas de individuos durante el estudio, ya sea porque abandonan, porque faltan variables, porque retractan su consentimiento o por otros motivos.

Existen calculadoras online para determinar el tamaño muestral necesario. Una muy sencilla está disponible en Fisterra (<https://www.fisterra.com/mbe/investigacion/9muestras/9muestras2.asp>).

En un estudio italiano realizado en 2009 (Lo Coco *et al.*, véase la bibliografía), la prevalencia de SPI en pacientes con EPOC fue del 36,8%. Si queremos tener un nivel de confianza del 95%, una precisión del 5% y esperamos unas pérdidas del 10%, el tamaño de la muestra necesaria es de 393 pacientes (**tabla 2.1**).

Tabla 2.1. Cálculo del tamaño muestral

Parámetro	Valor
Total de la población	Desconocido
Nivel de confianza o seguridad ($1 - \alpha$)	95%
Precisión	5%
Valor aproximado del parámetro que queremos calcular	36%
Tamaño muestral	354
Proporción esperada de pérdidas	10%
Tamaño muestral ajustado a las pérdidas	393

Presentación de los resultados y pruebas estadísticas

Presentación de la prevalencia

La presencia o no de SPI (que se traduce en prevalencia) es una variable cualitativa y se presentará como frecuencia absoluta y porcentaje.

Veamos un ejemplo de presentación de los resultados: “Se incluyó un total de 405 pacientes con EPOC, de los cuales 126 (31%) tenían SPI”. Por lo tanto, la prevalencia de SPI fue del 31%.

Estimar el intervalo de confianza de la prevalencia

La precisión de nuestra estimación de la prevalencia viene dada por el intervalo de confianza. Este intervalo puede calcularse con un nivel de confianza del 95% o del 99%. Lo habitual es utilizar el intervalo de confianza al 95% (IC95%). Hay

diferentes métodos para calcular los intervalos de confianza, como, por ejemplo, el método exacto de Clopper-Pearson, aunque dependerá del programa estadístico que utilicemos.

Presentación de resultados: “La prevalencia fue del 31% (IC95%: 27-36%)”. Esto se interpreta como que con un 95% de probabilidad la prevalencia de SPI en los pacientes con EPOC está comprendida entre un 27% y un 36%. Podemos calcular el intervalo de confianza al 99%, que en este caso sería 31% (IC99%: 25-37%).

En general, cuanto mayor es el tamaño muestral, más estrecho será el intervalo; y cuanto mayor es el nivel de precisión deseado, más ancho es el intervalo.

Contrastar si la prevalencia observada es diferente a la del estudio italiano de Lo Coco *et al.*

Para este contraste utilizaremos la prueba z de proporciones.

En nuestro estudio, de los 405 pacientes, 129 (31%) tenían SPI; en el estudio de Lo Coco *et al.*, de los 80 pacientes, 29 (36,8%) presentaban SPI. En este caso, el valor asociado de p es 0,309. Este valor es superior a 0,05 y, por lo tanto, descartamos que haya diferencias en la prevalencia de SPI entre ambas poblaciones.

Si en nuestro estudio hubieran presentado SPI 89 pacientes (prevalencia del 22%), el valor de p sería 0,0049. En ese caso, la prevalencia de SPI sería diferente en ambas poblaciones.

Características de los pacientes con y sin SPI

Para comunicar las características de los pacientes con EPOC con y sin SPI se elaborará una tabla. Las variables cualitativas se presentarán como frecuencia absoluta y porcentaje, y las variables cuantitativas como media (desviación estándar) si siguen una distribución normal o como mediana (rango intercuartílico) si la distribución no es normal.

Veamos un ejemplo de presentación de los resultados: “Se incluyó un total de 405 pacientes con EPOC. La prevalencia de fue del 31%. Las características de los pacientes se presentan en la tabla 1” (**tabla 2.2**).

En las tablas es recomendable presentar las características de la muestra total y de cada uno de sus grupos. También pueden compararse las características de ambos grupos.

Tabla 2.2. Características de los pacientes incluidos en el estudio

Variable	Total (N=405)	Sin SPI (n=279)	Con SPI (n=126)	Comparación (p)
Edad media (desviación estándar)	62 (8)	58 (6)	67 (5)	0,02
Sexo varón n (%)	312 (77)	209 (75)	103 (82)	0,13
Fumador n (%)	182 (45)	100 (36)	82 (65)	<0,001
Grado de obstrucción n (%)				
Leve	28 (7)	22 (8)	6 (5)	0,004
Moderada	201 (50)	151 (54)	50 (40)	
Grave	144 (35)	87 (31)	57 (45)	
Muy grave	32 (8)	17 (6)	15 (12)	
Exacerbaciones en el año anterior mediana (rango intercuartílico)	3 (2)	2 (2)	4 (2)	0,009
Ferropenia n (%)	61 (15)	28 (10)	33 (26)	<0,001

Para la comparación de variables cualitativas utilizaremos la prueba de χ^2 (chi al cuadrado). Para la comparación de variables cuantitativas que sigan una distribución normal utilizaremos la prueba de *t* de Student, y para aquellas que no sigan una distribución normal la prueba de la *U* de Mann-Whitney.

Podemos decir que los pacientes con EPOC y SPI son de mayor edad, que con más frecuencia son fumadores y tienen ferropenia, y que el SPI es más frecuente en los pacientes con mayor grado de obstrucción bronquial.

4. Justificación y discusión

La prevalencia de SPI en los pacientes con EPOC en España afecta aproximadamente a uno de cada tres pacientes, de forma similar a lo observado en estudios anteriores realizados en Italia.

Este estudio tiene fortalezas. En primer lugar, es un estudio multicéntrico. En segundo lugar, se ha realizado en pacientes con EPOC en situación estable, evitando la influencia que las exacerbaciones y la hipoxia puedan tener en la aparición de SPI. Sin embargo, también tiene debilidades. Entre ellas, que no se ha comparado con un grupo de pacientes sin EPOC y, por lo tanto, no podemos asegurar que la EPOC sea un factor predisponente a padecer SPI.

5. Conclusiones

De los resultados de este estudio podemos concluir que el SPI es frecuente en los pacientes con EPOC y aparece con más frecuencia en los pacientes de más edad, fumadores, con ferropenia, con más exacerbaciones y con obstrucción bronquial más grave.

Bibliografía

Bao Vicente V, Grandas Pérez F, Bojarski JK, Lahuerta Dal-Ré J, Luquin Piudo R, por el grupo de estudio DECODE RSL-España. El síndrome de piernas inquietas: detección, diagnóstico, consecuencias sobre la salud y utilización de recursos sanitarios. Rev Clin Esp. 2009;209:371-81.

Global Initiative for Chronic Obstructive Lung Disease (GOLD). Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease. 2021 Report. GOLD; 2020. Disponible en: <https://goldcopd.org>.

International Restless Legs Syndrome Study Group [Internet]. Rochester (MN): IRLSSG. 2012 Revised IRLSSG diagnostic criteria for RLS; [consultado: 25 de enero de 2021]. Disponible en: <http://irlssg.org/Diagnostic-criteria>.

Lo Coco D, Mattaliano A, Lo Coco A, Randisi B. Increased frequency of restless legs syndrome in chronic obstructive pulmonary disease patients. Sleep Med. 2009;10:573-6.

Soriano JB, Alfageme I, Miratvilles M, et al. Prevalence and determinants of COPD in Spain: EPISCAN II. Arch Bronconeumol. 2021;57:61-9.

Ward H, Toledano MB, Shaddick G, Bavies B, Elliott P. Oxford Handbook of epidemiology for clinicians. Oxford University Press; 2012.

Preguntas de autoevaluación

1. ¿Cuál de las siguientes características de un estudio no influye en el cálculo del tamaño muestral?

- a) El nivel de confianza.
- b) El número de pruebas estadísticas que se van a utilizar en el estudio.
- c) El porcentaje de pérdidas esperado.
- d) La precisión.

La respuesta correcta es la opción *b*.

Los factores que influyen en el cálculo del tamaño muestral son el tamaño de la población, el nivel de confianza que queremos tener, la precisión que deseamos, el valor aproximado del parámetro que queremos medir y el porcentaje esperado de pérdidas de individuos en el estudio. El número de pruebas estadísticas que se van a realizar no influye en el tamaño de la muestra del estudio.

2. Si decimos que en el mes de junio de 2021 la población española es de 47 185 267 personas y, de ellas, 85 247 (0,0018 %) tienen fibromialgia, estamos dando un valor de...

- a) Incidencia de fibromialgia.
- b) Tamaño muestral de la fibromialgia.
- c) Prevalencia de fibromialgia.
- d) Concurrencia de fibromialgia.

La respuesta correcta es la opción *c*.

La prevalencia es el número de casos nuevos y antiguos, habitualmente expresados como proporción o tasa, presentes en un momento determinado de tiempo. La incidencia es el número de casos nuevos que aparecen en un periodo de tiempo. La concurrencia no es un término estadístico y hace referencia a la coexistencia de personas, sucesos o cosas en un mismo lugar o tiempo. El tamaño muestral es el número de individuos necesarios para realizar un estudio.

3. ¿Cuál de las siguientes afirmaciones es cierta?

- a) Cuanto mayor es el tamaño muestral, más ancho es el intervalo de confianza.
- b) Cuanto mayor es el nivel de precisión deseado, más estrecho es el intervalo de confianza.
- c) Los intervalos de confianza al 95 % son más estrechos que los intervalos de confianza al 99 %.
- d) Un intervalo de confianza estrecho indica poca precisión de la estimación.

La respuesta correcta es la opción *c*.

Cuanto más nivel de precisión queramos tener, mayor será el nivel de confianza que utilizaremos, por ejemplo, 99 %. En ese caso, el intervalo será más ancho. Por ejemplo, si tenemos un estimador de 1,3 con un intervalo de confianza al 99 % de 1,2-1,5, la interpretación es que la probabilidad de que el estimador real esté dentro del intervalo es del 99 %. Por ello los intervalos de confianza al 95 % son más estrechos que los intervalos de confianza al 99 % (o, viceversa, los intervalos de confianza al 99 % son más anchos que al 95 %), ya que la probabilidad de que el estimador esté dentro del intervalo es menor.

Analizar el impacto
de diferentes factores de riesgo
para el desarrollo de una enfermedad.

Regresión logística

Dr. Javier Solera Santos
Servicio de Medicina Interna
Hospital de Albacete



capítulo 3

1. Situación actual y contexto clínico

La enfermedad (COVID-19) causada por el coronavirus SARS-CoV-2 ocasiona desde una infección asintomática o leve (80 %) a una neumonía (20 %) que puede ser grave con fracaso respiratorio secundario a síndrome de dificultad respiratoria aguda (SDRA), fallo multiorgánico y muerte. La pandemia ha saturado los hospitales produciendo dificultades en la atención en salas de urgencia, unidades de cuidados intensivos (UCI) y plantas de hospitalización. La causa más frecuente de muerte es la insuficiencia respiratoria y el fallo multiorgánico. La oxigenoterapia, incluida la ventilación mecánica, junto con la administración de dexametasona y heparina y el tratamiento de las complicaciones infecciosas bacterianas o fúngicas son los pilares básicos de la atención. Se han utilizado maniobras como la colocación del paciente en decúbito prono (DP), que pueden mejorar la ventilación en algunos pacientes con insuficiencia respiratoria grave.

La colocación del paciente en DP es una técnica no invasiva que ha demostrado su eficacia en diversos estudios, previos a la aparición de la COVID-19, como medida terapéutica adicional en el tratamiento en pacientes con SDRA. El SDRA es un cuadro clínico de disnea intensa de comienzo rápido, hipoxemia e infiltrados pulmonares bilaterales. La maniobra de DP pretende mejorar la ventilación facilitando la movilidad diafragmática, la redistribución de la perfusión y la eliminación de secreciones. Sin embargo, esta técnica es difícil de realizar en algunos pacientes (obesos, embarazadas, con fijadores externos) y no está exenta de complicaciones y de un mayor coste de los cuidados. Los estudios en pacientes fuera de UCI son escasos y no concluyentes.

El objetivo de nuestro estudio sería conocer si la maniobra de DP en los pacientes con COVID-19 e insuficiencia respiratoria mejora su supervivencia.

2. Planteamiento, pregunta de investigación y objetivos

En la pandemia de SARS-CoV-2, los pacientes que precisan ingreso tienen neumonía, insuficiencia respiratoria y SDRA. No siempre es posible el ingreso en la UCI o el uso de ventilación mecánica. En este contexto, los clínicos que atienden pacientes en la sala de hospitalización convencional de un hospital comarcal se preguntan si la maniobra de DP puede ser útil en la atención de pacientes con COVID-19. El diseño ideal sería hacer un ensayo clínico o, en su defecto, un estudio de cohortes prospectivo. Estos estudios requieren tiempo y suelen hacerse en coordinación con otros hospitales. A veces la evidencia más inmediata procede de estudios observa-

cionales de casos y controles o de corte transversal. E incluso mientras preparamos un estudio de cohortes o un ensayo clínico podemos realizar un estudio de casos y controles a partir de los pacientes que han recibido o no esta maniobra en la planta de hospitalización.

3. Tipos de estudio y diseño. Pruebas estadísticas y abordaje del problema

En la literatura médica se encuentran tres tipos de estudios básicos: estudios retrospectivos de casos y controles, estudios prospectivos de cohortes y ensayos clínicos controlados y aleatorizados. Todos los estudios tienen los siguientes componentes:

- **Asignación:** Selección de los sujetos de estudio y grupo control.
- **Valoración:** Determinación del resultado de la investigación en los sujetos de estudio y en el grupo control.
- **Análisis:** Comparación de los resultados del estudio en ambos grupos.
- **Interpretación:** Conclusiones acerca del significado y de la diferencia encontrada entre el grupo de estudio y el grupo control.
- **Extrapolación:** Se extraen recomendaciones acerca del significado del estudio para individuos o situaciones no incluidos en el estudio.

Estudio de casos y controles (retrospectivo)

Nos vamos a centrar en estudios de casos y controles, que es el diseño que hemos seleccionado para obtener información sobre la utilidad de la maniobra de decúbito prono en pacientes ingresados en planta de hospitalización con COVID-19.

La característica única de este tipo de estudios es que los pacientes ya han desarrollado, o no, el resultado que se quiere investigar (variable dependiente o de desenlace). En nuestro caso, la variable de muerte del paciente durante su ingreso en el hospital. Los estudios de casos y controles son retrospectivos porque vamos hacia atrás en el tiempo buscando las características de las personas previas al desenlace. En nuestro ejemplo, buscamos si han estado expuestos a la maniobra de posición en DP. En los estudios de casos y controles, los "casos" son las personas que han presentado el desenlace y los "controles" son las personas no han presentado tal desenlace.

Para desarrollar nuestro ejemplo, procederemos siguiendo el esquema descrito anteriormente:

- **Asignación:** Seleccionamos un grupo de pacientes con infección por SARS-CoV-2 que fallecieron y un grupo lo más parecido posible de pacientes que no han muerto.
- **Valoración:** Determinar en qué pacientes del grupo de estudio y del grupo de control hubo exposición a la maniobra de DP y, si es así, cuántas horas estuvieron expuestos.
- **Análisis:** Se calcula la razón (*odds*) en el grupo de fallecidos que han estado expuestos a la maniobra de DP frente a la razón (*odds*) del grupo de pacientes que no han fallecido con exposición a dicha maniobra.
- **Interpretación:** Se obtienen conclusiones acerca del significado del uso de la exposición a la maniobra.
- **Extrapolación:** Se obtienen conclusiones acerca del significado del uso de la maniobra de DP para pacientes con COVID-19 no incluidos en el estudio.

Descripción del estudio y metodología

El objetivo del estudio es valorar si la maniobra de DP en pacientes con COVID-19 e insuficiencia respiratoria reduce la mortalidad.

- **Lugar:** Hospital comarcal con 120 camas.
- **Población:** Pacientes ingresados en marzo y abril de 2020 con diagnóstico de COVID-19.
- **Criterios de inclusión:** Todos los pacientes con diagnóstico de infección por SARS-CoV-2 mediante PCR e insuficiencia respiratoria ($\text{SatO}_2 < 90\%$, índice $\text{PaFiO}_2 < 300$) en el periodo descrito. **Criterios de exclusión:** Ausencia de datos clave, como muerte o el tratamiento de la insuficiencia respiratoria, o contraindicación para la maniobra de DP.
- **Tipo de estudio:** Estudio retrospectivo de casos y controles.
- **Variable dependiente:** Muerte o no del paciente. Se considera casos a los pacientes que fallecieron durante su estancia hospitalaria. Los controles son pacientes dados de alta del hospital por curación y mejoría.
- **Variable independiente principal:** Exposición a maniobra de DP. Covariables: edad, sexo, diabetes mellitus (DM), hipertensión arterial (HTA).

- **Organización del estudio:** Escribir el protocolo. Reunir al equipo y especificar las tareas de cada investigador y los plazos. Solicitar la aprobación del estudio por el comité ético de investigación. Extraer los datos de la historia y construir una base de datos. Realizar el análisis estadístico. Elaborar un informe con las conclusiones y recomendaciones.

Es importante hacer un control de calidad de todo el proceso, especialmente de la fiabilidad de los datos extraídos.

Pruebas estadísticas

Las variables continuas normalmente distribuidas se expresan como medias y desviación estándar y se compararon mediante la prueba *t* de Student de dos colas para datos no apareados. Las variables continuas con distribución asimétrica se muestran como medianas (rango intercuartílico) y se comparan mediante la prueba de Mann-Whitney. Las variables categóricas se presentan como números (%) y se comparan mediante la prueba de χ^2 (chi al cuadrado). Se consideró significativo un valor de *p* menor de 0,05 para todas las pruebas estadísticas. Para calcular el riesgo de mortalidad (OR) respecto a las variables independientes se utiliza regresión logística. Esta parte se describe en detalle más adelante.

Utilidad de la regresión logística en estudios observacionales para valorar el riesgo

Los métodos estadísticos se utilizan con frecuencia para evaluar la fuerza de una asociación y proporcionar evidencia causal. Un ejemplo paradigmático es en el vínculo entre el tabaquismo y el cáncer de pulmón. El análisis estadístico, junto con las ciencias biológicas, puede establecer qué factores de riesgo son causales. Sin embargo, los factores de riesgo pueden indicar correlación y no necesariamente causalidad, es decir, la correlación no prueba causalidad. Por ejemplo, no se puede decir que ser joven cause síndrome de mononucleosis infecciosa aunque los jóvenes tengan una tasa más alta de infección que los adultos. La primoinfección por virus de Epstein-Barr (VEB) causa síndrome de mononucleosis infecciosa, y los adolescentes tienen menos probabilidad de haber desarrollado inmunidad al VEB que los adultos.

En epidemiología existen distintas formas de expresar el riesgo en términos estadísticos. Esto incluye el riesgo relativo (RR) y el riesgo absoluto (RA). Cuando el riesgo se aplica a una persona concreta, es mejor utilizar el RA que el RR.

En la mayoría de las situaciones clínicas, diferentes variables interactúan entre sí para producir un efecto. Las relaciones entre estas variables pueden ser complejas. El efecto de una variable podría modificarse por la presencia de otra, y los efectos de más de una variable pueden ser mayores o menores que sus efectos individuales en conjunto. Una manera de tener en cuenta esta interacción es utilizar un análisis estratificado. Sin embargo, cuando el número de variables que hay que controlar es elevado, resulta poco eficiente.

El análisis multivariante permite tener en cuenta los efectos de muchas variables de forma simultánea. De esta forma, podemos determinar los efectos independientes de cada una. Además, podemos ordenar las variables en función de la fuerza de su contribución al resultado y seleccionar de una larga lista de variables aquellas que contribuyen de manera independiente a la variación global del resultado eliminando el resto.

Existen varios tipos de modelos para el análisis multivariante. En los estudios de cohortes y de casos y controles con variable de resultado "dicotómica" se utiliza la regresión logística. El modelo de riesgo proporcional de Cox se utiliza cuando el resultado es el tiempo esperado para que se presente la variable resultado.

Una vez realizado un estudio observacional, el análisis multivariante es el único método factible para controlar muchas variables a la vez. El apareamiento puede ser útil para controlar sesgos, pero es difícil que puedan incluir más de dos o tres variables. El análisis estratificado es más intuitivo, pero menos eficiente.

En nuestro ejemplo, la variable dependiente es la muerte o no del paciente ingresado por COVID-19. Este desenlace se asocia con diferentes características (variables independientes) de los pacientes, como la edad, la HTA o la DM. Por tanto, si queremos conocer si la maniobra de DP está asociada con la mortalidad (aquí en sentido negativo), tendremos que "controlar" por otras variables que influyen en la mortalidad.

Tamaño muestral

No existen fórmulas para determinar el tamaño de muestra necesario en regresión logística ni en cualquier otra técnica de análisis multivariante. No obstante, hay recomendaciones que pueden ser útiles. Una regla práctica es usar diez sujetos por cada variable dependiente que se incluya en el estudio. Los investigadores nunca deben hacer un análisis de regresión a menos que el número de sujetos sea el doble del número de variables.

Presentación de resultados y discusión

En la **tabla 3.1** se refleja el análisis bivalente entre la mortalidad (variable dependiente) y la maniobra de DP (variable independiente). Se incluyen otras covariables relacionadas con la mortalidad en la COVID-19, como el sexo, la edad, la HTA y la DM.

En la primera columna de la **tabla 3.1** se reflejan los pacientes en los que se realizó la maniobra de DP y los que tenían HTA y DM junto con características demográficas básicas como la edad y el sexo. En las siguientes dos columnas se presentan estos datos separados por aquellos que murieron o no durante su ingreso. En la tercera columna se refleja la significación estadística mediante el valor de p . Este valor se calcula con un test de χ^2 para las variables dicotómicas (maniobra de DP, sexo, HTA y DM) y mediante una prueba de t de Student para la variable continua edad. La edad podría haberse introducido como variable dicotómica si elegimos un punto de corte, por ejemplo, 60 años, que dejaría a los pacientes repartidos en dos grupos: mayores de 60 años frente a menores de 60 años. Incluso podemos transformar la variable continua en más de una categoría, por ejemplo, menores de 40, entre 40 y 59 años y mayores de 60 años. En general, se considera que al transformar una variable continua en categórica se pierde información. Aun así, en la práctica clínica es frecuente categorizar variables continuas en cálculos de riesgo. Como ejemplo, el índice CALL (*CALL score*) utiliza la edad mayor o menor de 60 años o un valor de linfocitos mayor o menor de 1000/ μL para calcular el riesgo de progresión en pacientes con neumonía por COVID-19.

Tabla 3.1. Relación de variables con exitus. Análisis univariante

Total de pacientes = 165	Exitus (n=41)	No exitus (n=124)	p
Sexo varón (n=113; 68%)	28 (68,3%)	85 (68,5%)	0,976
Hipertensión arterial (n=103; 62%)	29 (70,7%)	74 (59,7%)	0,205
Diabetes mellitus (n=50; 30%)	14 (35%)	36 (29,3%)	0,495
Maniobra de decúbito prono (n=93; 56%)	23 (56,1%)	70 (56,5%)	0,968
Edad en años	71,51 ± 7,916	63,61 ± 9,349	<0,001

Valor de p calculado mediante χ^2 para variables dicotómicas (sexo, hipertensión arterial, diabetes mellitus y maniobra de decúbito prono) y t de Student para la variable continua (edad).

Análisis multivariante con regresión logística del estudio

En estudios observacionales como el que exponemos de casos y controles es importante añadir al análisis univariante un análisis multivariante. El método más utilizado cuando la variable es dicotómica es la regresión logística. La regresión logística permite calcular el riesgo de muerte para cada una de las variables independientes ajustando por el resto de las variables. El modelo completo se muestra en la **tabla 3.2**. La significación estadística para cada variable se puede valorar considerando $Z = b/DE(B)$ como una desviación normal estándar y calcular el valor de p . Sin embargo, el programa SPSS utiliza la prueba de Wald, que es la que aparece en la tabla. En la prueba estadística de Wald, la estimación de máxima verosimilitud \hat{O} del parámetro O se compara con el valor propuesto O_0 bajo la suposición de que la diferencia tipificada entre ambos seguirá aproximadamente una distribución normal. Típicamente, el cuadrado de la diferencia se compara con una distribución de χ^2 o con una distribución normal. Se puede usar para variables categóricas o continuas y permite el cálculo de p . Al utilizar χ^2 , para ver la p necesitamos conocer los grados de libertad, en este caso 1, como se muestra en la **tabla 3.2**.

En nuestro ejemplo, la variable de estudio (maniobra de DP) no está asociada a la mortalidad. Tampoco se ve asociación con las variables independientes de HTA, ser varón o DM. La única variable significativa es la edad, con $p < 0,001$.

En la **tabla 3.2**, la columna Exp(B) es el riesgo (OR) de muerte si está presente o no el factor (por ejemplo, DM). Si el valor es 1 es que no hay diferencia de riesgo. Si es mayor de 1, el riesgo de muerte es mayor, como ocurre en el caso de la edad, la DM y la maniobra de DP. Si es menor de 1, es un factor protector, ya que el riesgo de morir es menor en los que presentan esa característica, como sucede aquí con la HTA. No obstante, en la siguiente columna nos aparece el intervalo de confianza al 95% para ese estimador de riesgo, y sabemos que si el intervalo incluye el 1 no es significativo, lo que concuerda con un valor de p mayor de 0,05.

En este ejemplo, los pacientes en los que se intentó una maniobra de posición en DP no tuvieron menor mortalidad, tanto en el análisis crudo (univariante) como en el análisis ajustado (multivariante). La única variable que se correlacionó con la mortalidad fue la edad.

Tabla 3.2. Regresión logística binaria para la variable de exitus

	B	ET	Wald	GL	p	Exp(B)	IC95% para Exp(B)	
							Inferior	Superior
Varón	-0,051	0,428	0,014	1	0,905	0,950	0,410	2,199
Hipertensión arterial	-0,198	0,481	0,170	1	0,680	0,820	0,320	2,105
Diabetes mellitus	0,027	0,433	0,004	1	0,950	1,028	0,440	2,402
Maniobra de decúbito prono	0,178	0,402	0,195	1	0,659	1,194	0,543	2,627
Edad	0,113	0,027	18,173	1	0,000	1,120	1,063	1,179
Constante	-8,743	1,832	22,767	1	0,000	0,000		

B: coeficiente de regresión; ET: error estándar; GL: grados de libertad; p: valor de p (significación si es $< 0,05$); Exp(B): estimación puntual del riesgo para cada variable respecto a la muerte (en este caso, *odds ratio*); IC95% para Exp(B): intervalo de confianza al 95% del *odds ratio* para cada variable.

4. Conclusiones

En medicina, la idea de riesgo se refiere a la probabilidad de que suceda un acontecimiento desfavorable. Los médicos, con frecuencia, hacen una evaluación de riesgo en cada encuentro con el paciente. La mayoría de las consultas incluyen una investigación sobre los factores de riesgo que pueden producir problemas de salud en el futuro, especialmente aquellos modificables. También los pacientes están interesados sobre el riesgo de enfermedades como el Alzheimer, el cáncer, la osteoporosis o una enfermedad del corazón.

Cuando el clínico solicita pruebas diagnósticas o propone un tratamiento, los pacientes quieren conocer la precisión del diagnóstico y el beneficio potencial del tratamiento. La precisión de las respuestas a estas preguntas depende del cálculo de riesgo. La investigación clínica sobre el riesgo (o beneficio) es la base de los avances en la mejora de la salud de los últimos cien años. El conocimiento sobre factores de riesgo cardiovasculares o del cáncer han ayudado a disminuir considerablemente la mortalidad de las enfermedades en los países desarrollados.

El riesgo y la predicción son fundamentales en la medicina moderna y se apoyan en la epidemiología clínica. La epidemiología nos permite hacer predicciones numéricas sobre un problema médico de nuestro paciente utilizando resultados de los estudios basados en grupos de pacientes independientes. Estas predicciones

deben expresarse de forma precisa incluyendo su probabilidad y el intervalo de confianza de la estimación.

El análisis multivariante nos permite analizar un grupo de observaciones que incluyen más de una variable independiente. Al analizar varias variables independientes a la vez, el análisis multivariante ofrece tres ventajas: 1) se puede comparar la capacidad de las variables independientes para estimar la variable dependiente; 2) el efecto de las variables de confusión puede controlarse; y 3) el problema de las comparaciones múltiples con frecuencia puede evitarse.

El análisis multivariante con variable dependiente categórica puede realizarse con tres métodos: regresión logística, análisis discriminante y análisis de logaritmo lineal. El primero solo puede utilizarse cuando la variable dependiente es dicotómica. El análisis discriminante permite analizar variables con más de dos valores. La regresión logística se utiliza más en biología y medicina, y los otros dos más en ciencias sociales. Los paquetes estadísticos son muy útiles y permiten analizar grandes bases de datos para explorar variables de interés (factores de riesgo o factores protectores) controlando por otras variables de confusión o correlacionadas. Los estudios observacionales que queramos publicar en revistas de prestigio deben incluir un análisis multivariante.

Bibliografía

- Altman DG. Practical statistics for medical research. London: Chapman and Hall; 1993.
- Dawson-Saunders B, Trapp RG. Bioestadística. Mexico DF: Manual moderno; 1993.
- Fletcher Grant S. Epidemiología clínica. 6.ª ed. Philadelphia (PA): Wolters Kluwer; 2020.

Preguntas de autoevaluación

1. Respecto a la regresión logística, señale cuál de las siguientes afirmaciones es correcta:

- El análisis discriminante se emplea más en biomedicina que la regresión logística.
- La regresión logística, al controlar por factores de confusión, confirma la causalidad entre la exposición y el resultado.
- Exp(B) es una expresión del riesgo.
- El intervalo de confianza al 95% de Exp(B) no está relacionado con el valor de p .

La respuesta correcta es la opción c.

El análisis estadístico, como la regresión logística, puede indicar asociación, pero no prueba causalidad. Para establecer pruebas epidemiológicas de una relación causal se utilizan los criterios de Hill, que son los siguientes: fuerza y tamaño del efecto, consistencia o reproducibilidad, especificidad, temporalidad, gradiente biológico, plausibilidad, coherencia, pruebas experimentales, analogía y reversibilidad del efecto al eliminar la causa.

2. En un estudio sobre la hipertensión arterial, en un análisis de regresión logística se obtuvieron los siguientes valores para el cálculo de riesgo Exp(B): ser fumador, -0,068; obesidad, 0,695; y roncar por la noche, 0,872. ¿Cuál de las siguientes afirmaciones cree que es correcta?

- Ser fumador es un factor de riesgo.
- Ser obeso es un factor de riesgo.
- Ser obeso y roncar son factores de riesgo.
- Ninguna de las afirmaciones anteriores es correcta.

La respuesta correcta es la opción c.

Para calcular el riesgo, aquí *odds ratio* (OR), esto es, Exp(B), elevamos e (número de Euler) al coeficiente B , que son los números que aparecen

en la pregunta. El resultado para fumador es 0,939; para la obesidad, 1,62; y para roncar, 2,34. Los valores mayores de 1 indican una asociación positiva o de riesgo (obesidad y roncar), mientras el valor menor de 1 (aquí, ser fumador) es un factor protector.

3. Respecto a la tabla 3.2 (véase el texto del capítulo), donde se expresan los resultados del análisis multivariante, una de las siguientes afirmaciones es incorrecta:

- B es el coeficiente de regresión.
- El intervalo de confianza al 95% hace referencia a dónde se encuentra el valor de Exp(B) en 5 veces que realicemos el experimento.
- Si el valor de p es $<0,05$ el intervalo de confianza no debe incluir el 1.
- Las siglas ET de la tabla hacen referencia al error típico o estándar.

La respuesta correcta es la opción b.

El intervalo de confianza al 95% (IC95%) hace referencia al intervalo en el que podría encontrarse el valor. Si realizáramos un experimento igual 100 veces, en 95 veces los resultados estarían entre los valores que marca el IC95%. En la opción señalada como incorrecta, en vez de 100 veces se dice 5 veces.

Resolver un problema clínico mediante un estudio de casos y controles

Dr. Pedro Pablo Casado Escribano
Servicio de Medicina Interna
Hospital Universitario de La Princesa (Madrid)



capítulo 4

1. Situación actual y contexto clínico

Nos encontramos pasando la consulta de medicina interna y atendemos a una paciente conocida nuestra desde hace al menos 6 años. La seguimos por alto riesgo cardiovascular y es diabética, hipertensa, dislipémica y obesa. Desde la última revisión (hace 8 meses) nos refiere que ha tenido un ingreso a causa de un tromboembolismo pulmonar (TEP). A esta paciente le habíamos añadido a su tratamiento antidiabético un inhibidor del cotransportador de sodio y glucosa 2 (iSGLT-2) hace algo más de un año, la única modificación terapéutica destacable. A nuestra memoria viene el caso de otro paciente en la consulta que hace un año aproximadamente tuvo un TEP. También era diabético y creemos recordar que también contaba con un iSGLT-2 dentro de su tratamiento.

En el momento actual, dentro de los potenciales eventos adversos recogidos en la ficha técnica para los diferentes iSGLT-2 comercializados en nuestro territorio no está incluida la enfermedad tromboembólica venosa (ETE) en sus diferentes formas de manifestación (TEP y trombosis venosa profunda [TVP]).

2. Planteamiento, pregunta de investigación y objetivos

El hecho de no estar recogido el riesgo de ETE de los iSGLT-2 en su ficha técnica no asegura que no exista esta asociación. De forma relativamente reciente, y precisamente con este grupo farmacológico, se notificó posteriormente a la comercialización la presencia de casos de cetoacidosis euglicémica con su empleo en pacientes con diabetes mellitus tipo 2 (DM2), posible evento que acabó incorporándose a su ficha técnica.

Nuestra experiencia reciente con un caso y la sospecha de un segundo caso hace que nos planteemos la idoneidad de estudiar esta posible asociación. Así, nos planteamos la siguiente pregunta de investigación: ¿está el uso de iSGLT-2 en pacientes con DM2 asociado con una mayor presentación de ETE?

El objetivo principal es determinar la magnitud de la asociación de presentar ETE en los pacientes DM2 en dependencia de la toma de iSGLT-2. Las variables son las siguientes:

- Variable dependiente: presencia de ETE.
- Variable principal: tratamiento con iSGLT-2.
- Variables secundarias: aquellas que puedan suponer un factor de riesgo para la ETE (y, como consecuencia, puedan ser un factor de confusión), como la

edad, una cirugía reciente, el hábito tabáquico, la obesidad, la coagulopatía, el tratamiento anticoagulante, un cáncer activo, el embarazo, la toma de anticonceptivos hormonales o la inmovilidad reciente.

3. Tipo de estudio y diseño. Pruebas estadísticas y abordaje del problema

Para dar respuesta a nuestra pregunta podríamos seleccionar pacientes diabéticos tratados o no con iSGLT-2 y seguirlos en el tiempo evaluando las incidencias de ETE en los dos grupos (estudio observacional de cohortes), pero eso está fuera de nuestras posibilidades por el tiempo que deberíamos esperar para detectar un número mínimo de eventos (a menos que tuviéramos una muestra muy amplia de pacientes, hecho que tampoco está a nuestro alcance).

Por todo ello, parece que nuestro caso se amoldaría mejor al diseño de casos y controles. Los casos serían pacientes con DM2 que han padecido una ETE, mientras que los controles estarán exentos de esta circunstancia (variable dependiente). Analizaremos de forma retrospectiva el empleo de iSGLT-2 en cada uno de los grupos (variable principal), recogiendo también otras variables que pudieran influir en el resultado de que el paciente desarrollara una ETE (factores de riesgo para ETE que pudieran convertirse en factores de confusión para nuestro caso).

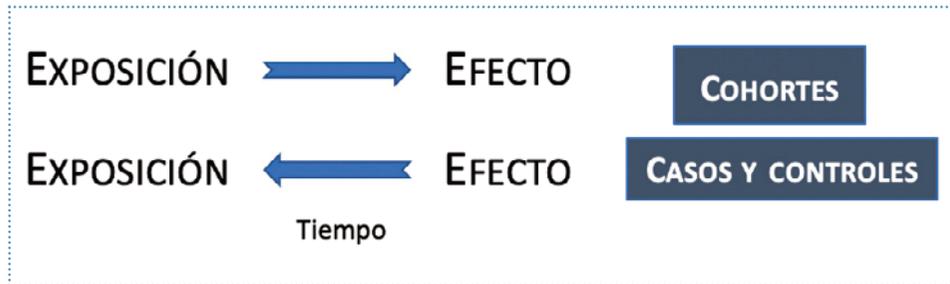
Se trataría, pues, de un estudio observacional, al no realizar por nuestra parte intervención sobre los pacientes; solo “observamos” la posible relación entre dos variables (la principal y la dependiente).

En los estudios de casos y controles, como hemos mencionado, partimos de la presencia de un evento o efecto y se busca de forma retrospectiva la potencial asociación con una exposición, lo cual es inverso a los estudios de cohortes, también observacionales, pero que van desde la exposición prospectivamente en búsqueda del efecto (**figura 4.1**). De estas diferencias se derivan algunas de las ventajas y desventajas de los estudios de casos y controles.

Dado que se parte directamente del efecto, es un tipo de diseño útil cuando estudiamos eventos (efectos) de baja incidencia o de larga latencia de aparición desde que sucede la exposición; no obstante, como desventaja, no podrán calcularse prevalencias por el mismo motivo.

El primer paso será seleccionar nuestros casos y controles, punto clave y crucial para la validez del estudio. Tanto casos como controles deberían reunir tres propiedades: representatividad (los casos deben ser todos los existentes en un lugar

Figura 4.1. Relación entre la dirección de la temporalidad en los estudios observacionales y la elección de cohortes o de casos y controles



y tiempo determinados y los controles deben representar a sujetos de esa misma población que se puedan convertir en casos), simultaneidad (los controles se obtienen en el mismo tiempo del que surgieron los casos) y homogeneidad (los controles se obtendrán de la misma cohorte de donde surgieron los casos).

Para nuestro estudio, los casos se definen como los pacientes con DM2 que han sufrido una ETEV (TEP o TVP), por lo que inicialmente pensamos en recoger aquellos informes de alta de hospitalización en los que exista el diagnóstico de ETEV y DM2. No obstante, con ello podríamos incurrir en un sesgo de selección, pues puede que haya pacientes con diagnóstico en urgencias y que no precisen ingresar (eventos de menor gravedad). Por ello pensamos que tal vez nos sirvan los informes de alta solo del servicio de urgencias (así tendríamos tanto los ingresados como los dados de alta en dicho servicio), pero de esta forma no estaríamos recogiendo los ETEV que sucedieran durante el transcurso del ingreso. Por ello decidimos, finalmente, recoger las historias tanto de urgencias como de hospitalización (del año 2020 de nuestro centro), evaluar aquellas con diagnóstico de ETEV y seleccionar solo las que se refieran a pacientes con DM2.

En cuanto a los controles, decidimos coger los mismos informes (altas de planta y urgencias) del año 2020 (así aseguramos la representatividad, la simultaneidad y la homogeneidad) con pacientes diabéticos sin diagnóstico de ETEV.

Sobre el tamaño muestral, existen fórmulas específicas para su cálculo, con una importante limitación, y es que en dichas fórmulas se incluye como elemento la proporción esperada del evento en casos y controles, que es justo lo que se está tratando de averiguar en este caso. Por este motivo, con frecuencia no se realiza un cálculo del tamaño muestral en los trabajos que emplean un diseño de casos y

controles (no obstante, en la bibliografía aportada puede encontrarse la fórmula para realizar el cálculo muestral en un estudio de casos y controles).

Para nuestro trabajo decidimos coger el conjunto de casos existentes con los criterios descritos para todo el año 2020. Revisando las historias de urgencias y planta de hospitalización, encontramos un total de 527 casos de ETEV, de los cuales 230 son en pacientes diabéticos (esa es la cifra de nuestros casos). Se puede incrementar la eficiencia estadística aumentando el número de controles con respecto al de los casos, pero este incremento de la eficiencia estadística deja de producirse cuando superamos los cuatro controles por cada caso. Para nuestro estudio decidimos coger dos controles por cada caso (460).

El siguiente paso sería esclarecer cuántos de los casos y cuántos de los controles han estado sometidos a la exposición a estudio (tratamiento con iSGLT-2). En este punto, en los estudios de casos y controles hay que tener precaución con el sesgo de memoria (sesgo de información), dado que los sujetos con el efecto pueden estar más predispuestos a recordar dicha exposición (o la insistencia del investigador puede ser distinta al conocer qué pacientes han sufrido el evento y cuáles no).

En nuestro estudio vamos a utilizar una base de datos digitalizada que nos permite conocer el tratamiento del paciente durante el último año, de tal manera que evitamos este sesgo, y consideramos expuesto a aquel paciente que estuviera recibiendo tratamiento con un iSGLT-2 al menos un mes antes del momento de la inclusión en el estudio. Adviértase aquí otra de las posibles debilidades de un estudio de casos y controles, pues no es un diseño adecuado para exposiciones poco frecuentes (si, por ejemplo, en vez de iSGLT-2, grupo farmacológico aceptablemente extendido dentro del tratamiento de la DM2, hubiéramos buscado la exposición a una molécula como la pioglitazona, empleada actualmente en menos del 5% de los pacientes diabéticos, esta reducida exposición nos limitaría de forma notable la potencia estadística de nuestro estudio).

Una vez definidos los casos y controles y conocida la exposición de cada grupo, podríamos construir una tabla de 2x2 como la siguiente:

	Casos (DM2 + ETEV)	Controles (DM2 sin ETEV)	
Expuestos (tratamiento con iSGLT-2)	58 (A)	91 (B)	149
No expuestos (sin tratamiento con iSGLT-2)	172 (C)	369 (D)	541
	230	460	690

La forma que tenemos en este tipo de estudios de ver la magnitud de la asociación (para nuestro caso, ver si los iSGLT-2 están ligados a la ETEV) es mediante el cálculo del *odds ratio* (OR, o razón de momios), que sería la probabilidad de que el caso haya estado expuesto al factor de estudio con respecto a la probabilidad de que un control lo haya estado. En nuestra tabla de 2x2, esto se representaría como se muestra en la **figura 4.2**.

Figura 4.2. Odds ratio de exposición en casos y controles

$$\frac{\frac{\text{Casos expuestos}}{\text{Casos no expuestos}}}{\frac{\text{Controles expuestos}}{\text{Controles no expuestos}}} = \frac{A/C}{B/D} = \frac{A \times D}{B \times C}$$

La interpretación de el OR es la siguiente:

- Si OR=1, la exposición es igual en casos que en controles (no hay asociación).
- Si OR<1, la exposición es un factor protector para el evento o efecto estudiado.
- Si OR>1, la exposición se comporta como un factor de riesgo para el evento o efecto estudiado.

Si introducimos nuestros datos en la fórmula de la **figura 4.2**, obtendríamos un OR de 1,36.

Con este valor numérico, y atendiendo a las reglas de interpretación que se expusieron anteriormente, tendríamos la inclinación a indicar que la toma de iSGLT-2 en el paciente diabético supone un factor de riesgo para el desarrollo de ETEV. Pero antes de poder hacer esta afirmación tenemos que evaluar si el valor obtenido está determinado por el azar. Para ello se puede o bien aplicar una prueba de hipótesis (la prueba de χ^2 de Mantel y Haenszel suele ser la utilizada para tablas de 2x2) o, lo más habitual en los estudios de casos y controles, calcular un intervalo de confianza del 95% (IC95%) para ese OR. Con ello, debemos reescribir la interpretación del OR teniendo en cuenta su IC95%:

- Si OR=1 o el IC95% contiene el 1, la exposición es igual IC en casos que en controles (no hay asociación).

- Si OR<1 y el límite superior del IC95% no incluye el 1, la exposición es un factor protector para el evento o efecto estudiado.
- Si OR>1 y el límite inferior del IC95% no incluye el 1, la exposición se comporta como un factor de riesgo para el evento o efecto a estudio.

Volviendo a nuestro caso, el OR de 1,36 se acompaña de un IC95% de 1,08-1,56.

Pero aún nos falta un último paso antes de poder concluir que la exposición es un factor de riesgo, y es corregir los posibles sesgos de confusión. Al tratarse de un tipo de diseño no aleatorizado, no podemos asegurar el reparto proporcionado de diferentes variables que pudieran tener un peso relevante en la aparición del evento de estudio. A esto se lo conoce como sesgo de confusión, y para entenderlo en nuestro caso pondremos un ejemplo. La obesidad es un factor de riesgo para la ETEV, y tal vez haya una mayor proporción de obesos entre los sujetos a los que se administra un iSGLT-2 (expuestos), dado que es un fármaco que logra la reducción del peso corporal. Si estuviéramos en un ensayo clínico aleatorizado, el reparto del iSGLT-2 estaría aleatorizado y deberíamos encontrar la misma proporción de obesos en ambos grupos de exposición, pero en un estudio de casos y controles partimos desde el evento y analizamos de forma retrospectiva, por lo que el reparto de variables no es aleatorio, y de ahí que puedan existir variables confusoras en el resultado obtenido.

Existen varias formas de minimizar el impacto de posibles variables confusoras. Algunas de ellas están en la fase de diseño; por ejemplo, limitando la entrada de esas variables mediante los criterios de inclusión y exclusión (por ejemplo, en nuestro estudio podríamos haber limitado la entrada solo a pacientes con DM2 y un índice de masa corporal por debajo de cierta cifra) o emparejando los casos y los controles (cogeríamos por cada caso un control con un peso similar –estas dos opciones nos imposibilitarían evaluar el impacto de esa variable en el efecto analizado–).

Pero el sistema más utilizado para controlar estas variables confusoras es mediante el análisis multivariante, y es importante a la hora de analizar e interpretar los resultados de un estudio de casos y controles si estamos ante el OR crudo (análisis sin ajuste, como sería nuestro OR de 1,36 ya mencionado anteriormente) o el OR ajustado (con análisis multivariante).

De esta manera, concluimos que, tras aplicar el análisis multivariante a nuestros datos, obtenemos un OR de 1,08 que se acompaña de un IC95% de 0,90-1,32. Con ello podríamos concluir en que no encontramos asociación entre la toma de iSGLT-2 y la aparición de ETEV en pacientes con DM2.

4. Justificación y discusión

Los estudios de casos y controles son útiles para evaluar la posible relación entre exposiciones y efectos, partiendo desde la presencia de un evento determinado y analizando retrospectivamente la asociación temporal con una o más exposiciones. Son estudios relativamente sencillos y baratos, con una ejecución rápida, lo que ha hecho que sean de los más frecuentemente utilizados en investigación sanitaria.

Son estudios muy útiles para evaluar una o varias exposiciones ante situaciones de efectos infrecuentes o con periodos de latencia prolongados.

Sin embargo, debemos tener cuidado con los posibles sesgos que se nos presentarán: de selección, de información (memoria) y de confusión. Y tener en cuenta sus limitaciones, como la de no poder calcular prevalencias o que pierde utilidad cuando estudiamos exposiciones infrecuentes.

El valor del OR, con la interpretación de su intervalo de confianza, será la medida de la magnitud de la asociación, y nos indicará si estamos ante un factor de riesgo ($OR > 1$), un factor protector ($OR < 1$) o neutro ($OR = 0$).

5. Conclusiones

Los estudios de casos y controles son una herramienta accesible para dar respuesta a cuestiones que pueden surgir en la práctica clínica diaria y que no requieren de una logística y tiempo de dedicación desproporcionados, por lo que todo clínico debería estar familiarizado con sus conceptos para poder desarrollar un estudio de estas características o poder interpretar correctamente los datos provenientes de un estudio de casos y controles.

Bibliografía

Rothman KJ, Greenland S, Lash TL. Modern epidemiology. 3.^a ed. Philadelphia (PA): Lippincott Williams & Wilkins; 2008.

Ruano-Ravira A, Perez-Rios M, Barros-Dios JM. Population-based Vs Hospital-based Controls; are they comparable?. Gac Sanit. 2008;22(6):609-13.

Sierra López A, Sáenz González MC, Fernández-Crehuet Navajas J, et al. (dirs). Piédrola Gil. Medicina preventiva y salud pública. 11.^a ed. Elsevier Masson; 2008.

Walcholder S, McLaughlin JK, Silverman DT, Mandel JS. Selection of controls in case-control studies. Principles. Am J Epidemiol. 1992 May 1;135(9):1019-28.

Walcholder S, McLaughlin JK, Silverman DT, Mandel JS. Selection of controls in case-control studies. Types of controls. Am J Epidemiol. 1992 May 1;135(9):1029-41.

Walcholder S, McLaughlin JK, Silverman DT, Mandel JS. Selection of controls in case-control studies. Design options. Am J Epidemiol. 1992 May 1;135(9):1042-50.

Preguntas de autoevaluación

1. Señale la afirmación incorrecta respecto a los estudios de casos y controles:

- No permiten calcular prevalencias.
- Se parte de la exposición y se busca el efecto prospectivamente.
- Son estudios observacionales.
- Son estudios adecuados ante enfermedades de baja incidencia.

La respuesta correcta es la opción *b*.

Los estudios de casos y controles con observacionales (no hay intervención), y van desde el evento o efecto hacia la exposición de forma retrospectiva (la diferencia de los estudios de cohortes, que van de exposición a efecto, sentido inverso), por lo que no podemos calcular prevalencias. No obstante, son útiles ante entidades de baja incidencia o con largo periodo de latencia entre exposición y factor.

2. ¿Cuál de los siguientes es un sesgo muy característico y propio de los estudios de casos y controles?

- Sesgo de pertenencia.
- Sesgo de memoria.
- Sesgo de detección.
- Ninguno de los anteriores es propio de los estudios de casos y controles.

La respuesta correcta es la opción *b*.

Los estudios de casos y controles son un tipo de estudios muy vulnerables a los sesgos al tener un sentido retrospectivo. Algunos de sus sesgos más típicos son los de selección, de memoria y de confusión. Los otros dos sesgos enunciados son propios o más característicos de otro tipo de diseños.

3. Si estoy estudiando la posible asociación del consumo de tabaco con el desarrollo de esclerosis múltiple (EM) mediante un estudio de casos y controles, y obtengo un *odds ratio* (OR) de 1,7 con un intervalo de confianza al 95% (IC95%) que no incluye el 1, ¿qué interpretación debo asignarle?

- Que el consumo de tabaco es un factor protector para el desarrollo de EM.
- Que el consumo de tabaco es un factor de riesgo para el desarrollo de EM.
- Que el consumo de tabaco parece comportarse como un factor protector para el desarrollo de la EM, pero que al no estar englobado el 1 por el intervalo de confianza no puede demostrarse estadísticamente.
- Que el consumo de tabaco parece comportarse como un factor de riesgo para el desarrollo de la EM, pero que al no estar englobado el 1 por el intervalo de confianza no puede demostrarse estadísticamente.

La respuesta correcta es la opción *b*.

La medida de asociación en los estudios de casos y controles es el *odds ratio* (OR) o razón de momios, cuya interpretación es la siguiente:

- Si $OR = 1$ o el IC95% contiene el 1, la exposición es igual en casos que en controles (no hay asociación).
- Si $OR < 1$ y el límite superior del IC95% no incluye el 1, la exposición es un factor protector para el evento o efecto estudiado.
- Si $OR > 1$ y el límite inferior del IC95% no incluye el 1, la exposición se comporta como un factor de riesgo para el evento o efecto a estudio.

Resolver un problema clínico
con un análisis de supervivencia
mediante el estimador Kaplan-Meier

Dr. Javier Ena Muñoz
Servicio de Medicina Interna
Hospital Marina Baixa (Villajoyosa, Alicante)



capítulo 5

1. Situación actual y contexto clínico

Las infecciones causadas por el coronavirus tipo 2 causante del síndrome respiratorio agudo grave (SARS-CoV-2) se manifiestan tras un corto periodo de incubación con síntomas similares a los de la gripe. En algunas personas, la infección progresa hasta causar un fracaso respiratorio hipoxémico. La base fisiopatológica del deterioro clínico parece estar en una intensa respuesta inflamatoria que simula un síndrome de liberación de citocinas. En esta fase, los pacientes muestran una marcada elevación de mediadores inflamatorios, entre los que se incluyen la interleucina 6 (IL-6), la ferritina y la proteína C-reactiva. Se ha demostrado que existe una asociación directa de los niveles de IL-6 en suero y los valores de viremia de SARS-CoV-2, la progresión del fracaso respiratorio a la necesidad de ventilación mecánica y, finalmente, la muerte.

Estos hallazgos indican que el uso de un fármaco que bloquee el receptor de IL-6 podría interrumpir la cascada inflamatoria y evitar la progresión de la enfermedad.

2. Planteamiento, pregunta de investigación y objetivos

El tocilizumab, un anticuerpo monoclonal humanizado que se une al receptor de la IL-6, podría resultar útil en el tratamiento de la infección grave por SARS-CoV-2. La administración del fármaco debería realizarse durante una fase relativamente precoz de la infección para evitar la progresión al fracaso respiratorio hipoxémico, reducir la necesidad de ventilación mecánica y mejorar la mortalidad.

3. Tipo de estudio y diseño. Pruebas estadísticas y abordaje del problema

El diseño de estudio que da una mejor respuesta a la valoración de la eficacia de una intervención es un ensayo clínico aleatorizado, controlado con placebo y preferiblemente doble ciego. Para ilustrar este capítulo vamos a recurrir a los resultados del estudio llevado a cabo por Stone *et al.* (véase la bibliografía), en el que se evaluó la eficacia del tocilizumab en pacientes hospitalizados con COVID-19.

En este estudio se incluyó a pacientes mayores de 18 años y menores de 80 años con infección confirmada por SARS-CoV-2 mediante prueba de PCR que fueron ingresados en plantas de pacientes no críticos y que presentaban infección grave por SARS-CoV-2, definida por al menos dos de los siguientes criterios clínicos: fiebre superior a 38 °C, infiltrados pulmonares en la radiografía de tórax y necesidad de oxígeno suplementario para mantener una saturación de oxígeno supe-

rior al 92%; junto a al menos un criterio de laboratorio: ferritina, >500 ng/mL; proteína C-reactiva, >50 mg/L; lactato-deshidrogenasa (LDH), >250 UI/L; o dímero D, >1000 ng/mL. Se excluyó a pacientes con infección no controlada por bacterias, virus u otras infecciones virales, tuberculosis activa o tuberculosis latente con menos de un mes de tratamiento, cualquier inmunoterapia o inmunosupresión grave, un recuento de neutrófilos inferior a 500/mm³ o de plaquetas inferior a 50 000/mm³, elevación de las transaminasas (GOT/GPT) superior a cinco veces el límite de la normalidad o tratamiento con plasma procedente de sujetos convalecientes de COVID-19. Los pacientes firmaron un consentimiento informado.

Los pacientes fueron asignados 2:1 a recibir cuidados estándar más una dosis de tocilizumab (8 mg por kilogramo de peso por vía intravenosa, sin exceder de 800 mg) o placebo. La aleatorización se realizó en bloques de 3 y 6 pacientes. Los pacientes podían recibir además remdesivir, hidroxicloroquina o glucocorticoides. Ningún paciente recibió dexametasona.

El desenlace primario evaluado fue intubación (o muerte en quienes murieron antes de la intubación) tras la administración de tocilizumab o placebo, teniendo en cuenta el tiempo hasta un evento (*time-to-event analysis*).

Resultado principal

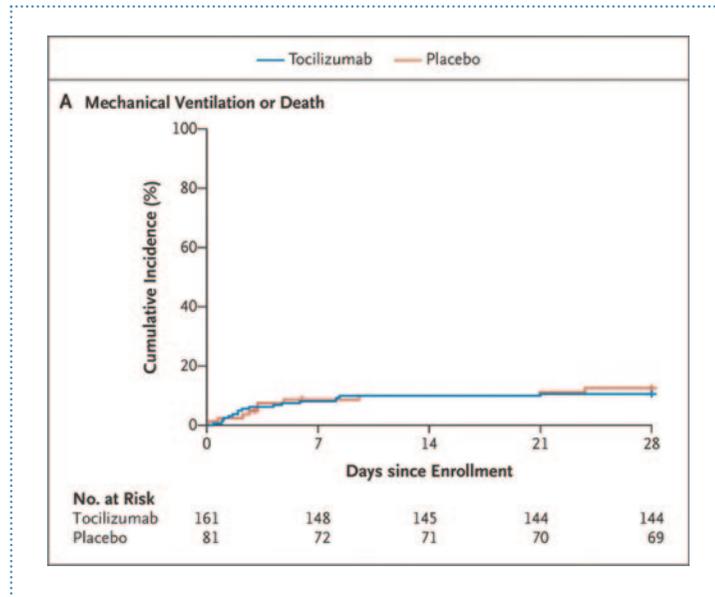
En este estudio se incluyó a 243 pacientes; 141 (58%) eran hombres, y 102 (42%) mujeres. La edad mediana fue de 59,8 años (rango, 21,7 a 85,4), y el 45% de los pacientes eran de origen hispano. El cociente de riesgos instantáneo (*hazard ratio*) para intubación o muerte en el grupo de tocilizumab en comparación con placebo fue 0,83 (IC95%: 0,38-1,81; $p=0,64$).

La siguiente figura, procedente del artículo de Stone *et al.*, muestra el análisis mediante curvas de Kaplan-Meier (tiempo hasta el evento) del desenlace principal (**figura 5.1**). En este caso, en el eje de ordenadas se muestra la incidencia acumulada (porcentaje de intubación o muerte), y en el eje de abscisas el tiempo dividido en fragmentos de 7 días hasta completar 28 días de observación.

Podemos resumir los valores del estudio como se muestra en la siguiente tabla (**tabla 5.1**).

Basándonos en los resultados de este estudio, vamos a reproducir los cálculos de las curvas de Kaplan-Meier.

Figura 5.1. Análisis mediante curvas de Kaplan-Meier del desenlace principal (intubación o muerte)



Fuente: Stone JH, Frigault MJ, Serling-Boyd NJ, et al. Efficacy of tocilizumab in patients hospitalized with Covid-19. N Engl J Med. 2020 Dec 10;383(24):2333-44.

Tabla 5.1. Principales resultados del estudio de Stone *et al.*

Tiempo hasta el evento	Tocilizumab		Placebo	
	UCI o muerte	Ningún evento	UCI o muerte	Ningún evento
0-7 días	13	148	9	72
8-14 días	3	145	1	71
15-21 días	1	144	1	70
22-28 días	0	144	1	69

Fuente: Stone JH, Frigault MJ, Serling-Boyd NJ, et al. Efficacy of tocilizumab in patients hospitalized with Covid-19. N Engl J Med. 2020 Dec 10;383(24):2333-44.

Las curvas de Kaplan-Meier dibujan la fracción de pacientes supervivientes (los que aún no han presentado un evento) con respecto al tiempo en cada grupo de tratamiento. La altura de la curva de Kaplan-Meier al final de cada intervalo de tiempo se calcula tomando la proporción de pacientes libres de evento al final del intervalo previo y multiplicándola por la fracción de pacientes que sobreviven el presente intervalo de tiempo sin presentar el evento. Los intervalos de tiempo vienen definidos por el momento en que cada paciente alcanza el desenlace o desaparece del seguimiento si no lo completa. En general, los cálculos son muy laboriosos a poco que crezca el tamaño de la muestra, por lo que se hace necesario acudir a procesos automatizados de cálculo. Con objeto de simplificar los cálculos para el presente análisis vamos a suponer que los eventos suceden cada 7 días desde el ingreso, en vez del tiempo real hasta el evento. Esto simplificará el número de tablas que haya que realizar.

Vamos a calcular a continuación las tasas de supervivencia de cada intervalo de 7 días (**tabla 5.2**) mediante la siguiente fórmula:

$$\text{Supervivencia del intervalo } i = \frac{\text{Expuestos al comienzo del intervalo } i - \text{casos de intubación o muerte en el intervalo } i}{\text{Expuestos al comienzo del intervalo } i}$$

Tabla 5.2. Supervivencia para cada intervalo

Intervalo	Tocilizumab	Placebo
1 (0-7 días)	$[161 - 13]/161 = 0,919$	$[81 - 9]/81 = 0,889$
2 (8-14 días)	$[148 - 3]/148 = 0,987$	$[72 - 1]/72 = 0,986$
3 (15-21 días)	$[145 - 1]/145 = 0,993$	$[71 - 1]/71 = 0,986$
4 (22-28 días)	$[144 - 0]/144 = 1$	$[70 - 1]/70 = 0,986$

Por último, calculamos las tasas de supervivencia acumulada en un intervalo como producto de la anterior y la supervivencia propia del intervalo (**tabla 5.3**). Precisamente, la representación gráfica de estas tasas constituye las denominadas curvas de supervivencia.

Si quisiéramos estimar el riesgo acumulado en cada intervalo, realizaríamos los cálculos que se describen en la **tabla 5.4**.

Tabla 5.3. Supervivencia acumulada

Intervalo	Tocilizumab	Placebo
1 (0-7 días)	0,919	0,889
2 (8-14 días)	$0,919 \cdot 0,987 = 0,907$	$0,889 \cdot 0,986 = 0,877$
3 (15-21 días)	$0,907 \cdot 0,993 = 0,901$	$0,877 \cdot 0,986 = 0,865$
4 (22-28 días)	$0,901 \cdot 1 = 0,901$	$0,865 \cdot 0,986 = 0,853$

Tabla 5.4. Incidencia acumulada (en probabilidad)

Intervalo	Tocilizumab	Placebo
1 (0-7 días)	$13/161 = 0,081$	$9/81 = 0,111$
2 (8-14 días)	$0,081 + (3/148) = 0,101$	$0,111 + (1/72) = 0,124$
3 (15-21 días)	$0,101 + (1/145) = 0,102$	$0,124 + (1/71) = 0,138$
4 (22-28 días)	$0,102 + (0/144) = 0,102$	$0,138 + (1/70) = 0,152$

Para introducir los datos en SPSS, generaríamos una base de datos con las siguientes cuatro variables:

- ID del paciente.
- Tratamiento asignado (0: Placebo; 1: Tocilizumab).
- Desenlace (0: Sin ingreso en UCI ni muerte; 1: Ingreso en UCI o muerte).
- Seguimiento en días hasta el desenlace.

La gráfica de la función de supervivencia que aparece en el SPSS se muestra en la **figura 5.2**. Desafortunadamente, SPSS ajusta los intervalos de días de 5 en 5 en el seguimiento.

La gráfica de la función de riesgo acumulado aparece en el SPSS como se muestra en la **figura 5.3**.

Figura 5.2. Función de supervivencia libre de evento en pacientes con COVID-19 en tratamiento con tocilizumab o placebo

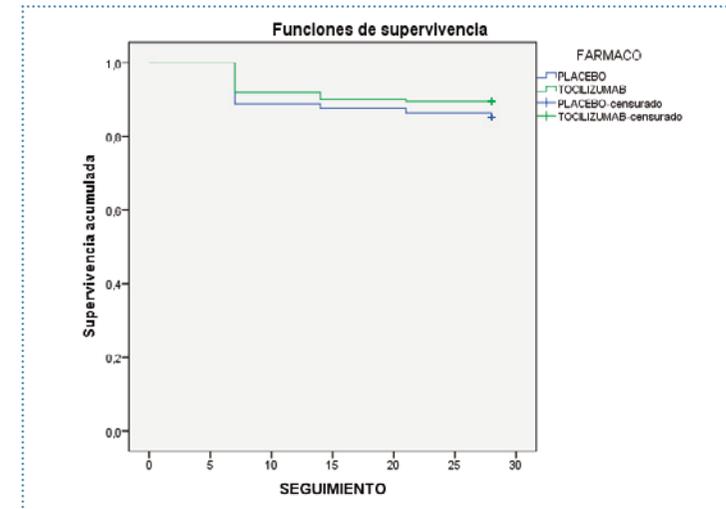
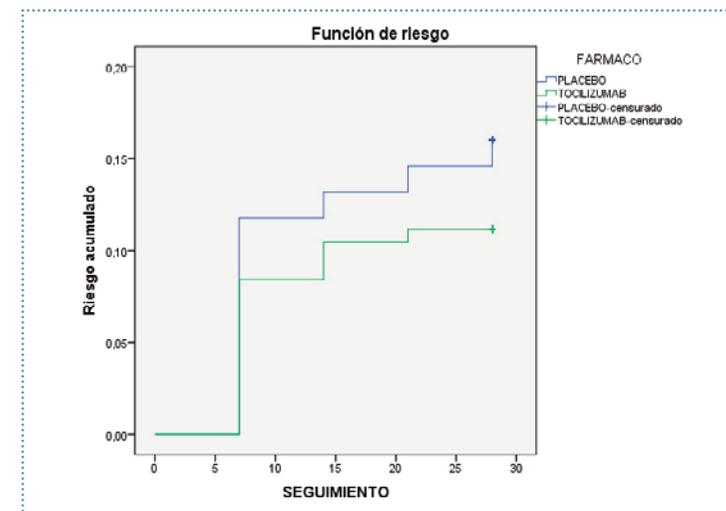


Figura 5.3. Función de incidencia acumulada de ingreso en UCI o muerte en pacientes con COVID-19 en tratamiento con tocilizumab o placebo



Las curvas de Kaplan-Meier representan de forma gráfica la supervivencia de cada grupo estudiado a lo largo del tiempo, pero no ofrecen información sobre la significación estadística. La significación estadística se establece con el test de *log-rank*. Esta prueba tiene la limitación de no tener en consideración el efecto de diversas variables de confusión, como podrían ser en un estudio de cohortes las diferencias en las características demográficas de la población.

Para realizar el test de *log-rank* distribuiremos los datos del estudio de tocilizumab frente a placebo en forma de incidencia de evento (ingreso en UCI o muerte) para cada uno de los cuatro intervalos de tiempo (0-7 días, 8-14 días, 15-21 días, 22-28 días) (tablas 5.5 a 5.8).

Tabla 5.5. Incidencia según el grupo de tratamiento para los días 0-7

Grupo de tratamiento	UCI o muerte	Ni UCI ni muerte	Total
Tocilizumab	13	148	161
Placebo	9	72	81
Total	22	220	242

Tabla 5.6. Incidencia según el grupo de tratamiento para los días 8-14

Grupo de tratamiento	UCI o muerte	Ni UCI ni muerte	Total
Tocilizumab	3	145	148
Placebo	1	71	72
Total	4	216	220

Tabla 5.7. Incidencia según el grupo de tratamiento para los días 15-21

Grupo de tratamiento	UCI o muerte	Ni UCI ni muerte	Total
Tocilizumab	1	144	145
Placebo	1	70	71
Total	2	214	216

Tabla 5.8. Incidencia según el grupo de tratamiento para los días 22-28

Grupo de tratamiento	UCI o muerte	Ni UCI ni muerte	Total
Tocilizumab	0	144	144
Placebo	1	69	70
Total	1	213	214

Para determinar la significación estadística mediante el test de *log-rank* hay que llevar a cabo un test de Mantel-Haenzel en cada una de las tablas de 2x2 y calcular la siguiente fórmula:

$$\chi^2 LR = \frac{([O - E] - 0,5)^2}{V}$$

donde $\chi^2 LR$ es una chi al cuadrado de *log-rank*, O es el número de casos observados y E es el número de casos esperados:

$$E = \sum_{i=1}^k E_i = \sum_{i=1}^k \frac{(a_i + b_i)(a_i + c_i)}{N_i}$$

$$V = \sum_{i=1}^k E_i = \sum_{i=1}^k \frac{(a_i + b_i)(c_i d_i)(a_i + c_i)(b_i d_i)}{N_i^2 (N_i - 1)}$$

teniendo en cuenta la distribución entre la exposición y los estratos de tiempo en las tablas de 2x2:

Grupo de tratamiento	UCI o muerte	Ni UCI ni muerte	Total
Tocilizumab	a_i	b_i	n_{i1}
Placebo	c_i	d_i	n_{i2}
Total	$a_i + c_i$	$b_i + d_i$	N_i

La hipótesis nula sigue una distribución de χ^2 con un grado de libertad.

Para un test de dos colas con un nivel de significación α , si $\chi^2 LR > \chi^2_{1, 1-\alpha}$, entonces se rechaza la hipótesis nula; si $\chi^2 LR \leq \chi^2_{1, 1-\alpha}$, entonces se acepta la hipótesis nula.

El test solo se puede usar si el valor de V es ≥ 5 .

El número de eventos observados (O) para tocilizumab es el siguiente:

$$13 + 3 + 1 + 0 = 17$$

El número de eventos esperados (E) para tocilizumab es el siguiente:

$$\frac{22 \cdot 161}{242} + \frac{4 \cdot 148}{220} + \frac{2 \cdot 145}{216} + \frac{1 \cdot 144}{214} = 14,63 + 2,69 + 1,34 + 0,68 = 19,34$$

El valor V es el siguiente:

$$\frac{161 \cdot 81 \cdot 22 \cdot 220}{242^2 \cdot 241} + \frac{148 \cdot 72 \cdot 4 \cdot 216}{220^2 \cdot 219} + \frac{145 \cdot 71 \cdot 2 \cdot 214}{216^2 \cdot 215} + \frac{144 \cdot 70 \cdot 1 \cdot 213}{216^2 \cdot 215} = 4,47 + 0,89 + 0,44 + 0,22 = 6,02$$

Y, finalmente, para $\chi^2 LR$ se obtiene el siguiente valor:

$$\frac{[(17 - 19,34) - 0,5]^2}{6,02} = \frac{[-2,84]^2}{6,02} = 1,33$$

Como el valor de $\chi^2_{1, 0,95} = 3,84 > 1,33$, tenemos que $p > 0,5$, por lo que no hay asociación entre el uso de tocilizumab y la disminución del ingreso en la UCI o los casos de muerte.

4. Justificación y discusión

Los datos del presente ensayo clínico no apoyan el concepto de que una intervención precoz para bloquear la IL-6 sea una estrategia efectiva en pacientes con COVID-19 de gravedad moderada. La hipótesis del estudio es que en los pacientes con COVID-19 que no han requerido hasta el momento intubación se interrumpiría la tormenta de citocinas y se prevendrían las consecuencias más graves de la enfermedad. Debido al amplio intervalo de confianza en la comparación de la efi-

cia, no se puede concluir la posibilidad de que tocilizumab pueda tener algún beneficio o perjuicio para los pacientes.

En el artículo original los autores realizan un análisis multivariante de Cox teniendo en cuenta la edad, el sexo, la raza, la presencia de diabetes y las concentraciones iniciales de IL-6. El *hazard ratio* ajustado fue 0,66 (IC95%: 0,28-1,52). La diferencia entre el *hazard ratio* crudo y el ajustado se debía al superior porcentaje de pacientes con mayor edad en el grupo de tocilizumab.

5. Conclusiones

En los estudios epidemiológicos que tienen un seguimiento prospectivo es útil el uso del análisis de supervivencia (llamado también *time-to-event analysis*). El método de análisis univariante se denomina curvas de Kaplan-Meier y el análisis estadístico utilizado para comparar las curvas se llama test de *log-rank*. El análisis de Kaplan-Meier tiene como limitación el no incorporar el efecto de variables de confusión. Si los estudios de supervivencia se aplican a ensayos clínicos aleatorizados de muestra amplia, es posible que las variables de confusión estén distribuidas de forma equitativa en ambos grupos (grupo de intervención y grupo control). Los ensayos clínicos que no presentan una muestra suficientemente grande, y especialmente en los estudios de cohortes, en los que existe el sesgo de asignación, la diferente distribución de variables pronósticas hace necesario que el análisis de supervivencia mediante las curvas de Kaplan-Meier deba complementarse en el modelo de riesgos proporcionales de Cox para ajustar por variables de confusión.

Bibliografía

- Carrasco JL. El método estadístico en la investigación médica. Madrid: Editorial Ciencia 3; 1986.
- Lee ET. Statistical methods for survival analysis. 2.^a ed. New York: John Wiley & Sons; 1992.
- Martínez-González MA, Alonso A, Lopez-Fidalgo J. ¿Qué es un hazard ratio?. Nociones de análisis de supervivencia. Med Clin (Barc). 2008;131(2):65-72.
- Rosner B. Fundamentals of biostatistics. 3.^a ed. Boston, Massachusetts: PWS-Kent Publishing Company; 1986.
- Stone JH, Frigault MJ, Serling-Boyd NJ, et al. Efficacy of tocilizumab in patients hospitalized with Covid-19. N Engl J Med. 2020 Dec 10;383(24):2333-44. doi: 10.1056/NEJMoa2028836.

Preguntas de autoevaluación

1. El análisis tiempo-hasta-evento se denomina...

- a) Análisis χ^2 (chi al cuadrado).
- b) Prueba t de Student.
- c) Análisis de supervivencia.
- d) Test de ANOVA.

La respuesta correcta es la opción *c*.

El análisis tiempo-hasta-evento se denomina análisis de supervivencia. El análisis de χ^2 va dirigido a comparar variables categóricas en dos o más grupos. La prueba t de Student compara medias en dos grupos independientes con distribución normal o paramétrica. El test ANOVA compara medias en tres o más grupos independientes con distribución normal o paramétrica.

2. La comparación de las curvas de supervivencia entre dos grupos se denomina...

- a) Análisis de Kaplan-Meier.
- b) Test de *log-rank*.
- c) Test de McNemar.
- d) Test de Wilcoxon.

La respuesta correcta es la opción *b*.

La comparación de las curvas de supervivencia entre dos grupos se denomina test de *log-rank*. El análisis de Kaplan-Meier es la representación gráfica de las curvas de supervivencia. El test de McNemar compara variables categóricas de observaciones repetidas en un mismo individuo (tanto χ^2 , con o sin corrección de Yates, y eventualmente la prueba exacta de Fisher no pueden aplicarse cuando se tienen que evaluar muestras pareadas). El test de Wilcoxon es una prueba no paramétrica que compara el rango medio de dos muestras relacionadas. Se utiliza como alternativa a la prueba t de Student cuando no se puede suponer la normalidad de dichas muestras.

3. ¿Cuál es la limitación del análisis de Kaplan-Meier?

- a) No tiene en cuenta el ajuste por variables de confusión.
- b) Solo es aplicable a variables dicotómicas.
- c) No indica la significación estadísticas entre dos curvas de Kaplan-Meier.
- d) Todas las anteriores son limitaciones del análisis de Kaplan-Meier.

La respuesta correcta es la opción *d*.

Todas son limitaciones del análisis de Kaplan-Meier. Efectivamente, entre las limitaciones de las curvas de Kaplan-Meier están el no tener en cuenta el ajuste por variables de confusión, que solo es aplicable a variables dicotómicas y que no indica la significación estadística (es necesario realizar un test de *log-rank*).

Determinar los factores pronóstico.

Análisis de Cox

Dr. Emilio Casariego Vales
Servicio de Medicina Interna
Hospital Universitario Lucus Augusti (Lugo)

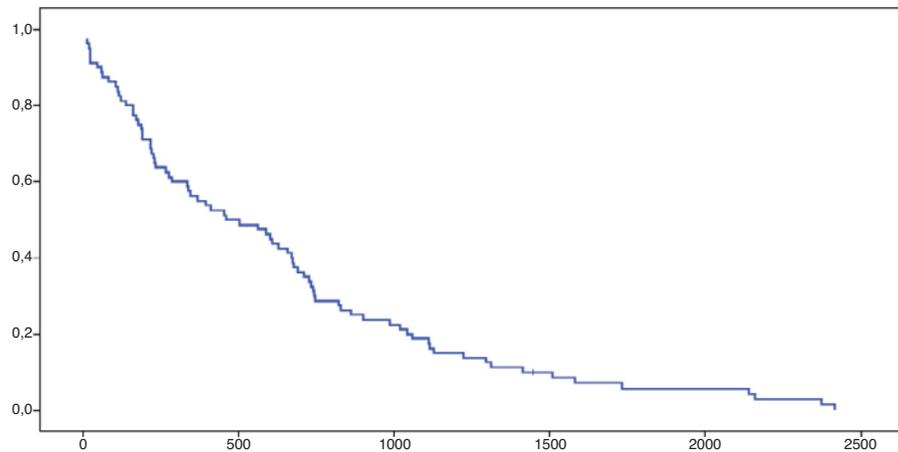


capítulo 6

1. Situación actual y contexto clínico

Uno de los problemas clínicos más habituales es intentar estimar el tiempo hasta que ocurre un suceso o evento. Problemas como “Si me contagio con SARS-CoV-2, ¿cuánto tiempo transcurre hasta que aparece la fase más grave de la enfermedad?”, o “Si iniciamos el tratamiento por brucelosis, ¿cuál es el tiempo hasta la remisión de la fiebre?”, o bien “¿Cuál es el tiempo de supervivencia para un paciente intervenido por cáncer gástrico en estadio IIIa?, ¿es distinto al de los pacientes en estadio II?”. Puesto que la respuesta a las preguntas previas no es única, ¿podemos determinar los factores que “explican” las diferencias? El conjunto de técnicas que permiten estudiar esos tiempos se denominan *análisis de supervivencia*. Sin embargo, puesto que lo que realmente analizan son los tiempos hasta que ocurre “algo”, quizá fuese más preciso una denominación más general, como *análisis de tiempos de espera*.

Figura 6.1. Probabilidad de supervivencia (expresada en días) de personas que han cumplido 100 años

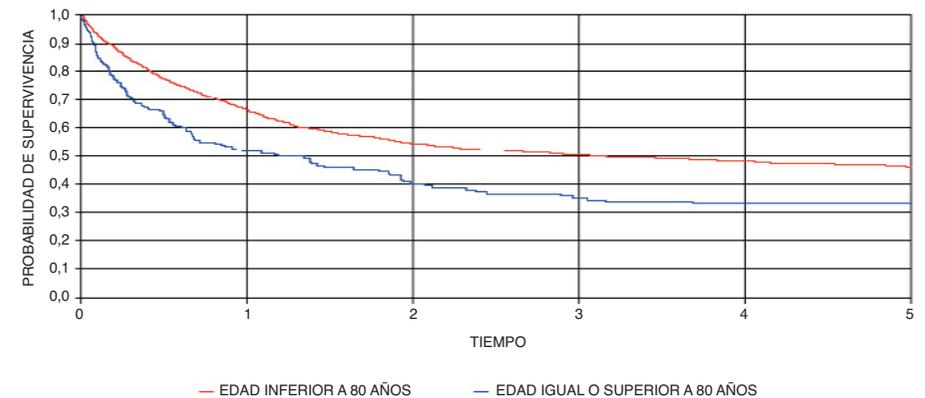


Puede observarse que la probabilidad al cabo de 500 días es, aproximadamente, del 50%.

Básicamente, los clínicos tienen como objetivo conocer la respuesta en tres escenarios distintos:

- Estimar los tiempos hasta el suceso de interés (por ejemplo, hasta el fallecimiento), describirlos y resumirlos. De esta forma, podemos usarlos para predecir el comportamiento futuro de pacientes con características similares (**figura 6.1**). Este aspecto se estudia en profundidad en el capítulo 5.
- Comparar el patrón de supervivencia de diferentes grupos de pacientes. El método más utilizado es el test *log-rank* (**figura 6.2**). Nuevamente, este aspecto se estudia en profundidad en el capítulo 5.
- Determinar qué factores modifican la supervivencia. Esto es, entre toda la constelación de variables que pueden modificar el pronóstico (edad, sexo, gravedad del proceso, etc.), se pretende conocer el impacto de cada una de ellas. Así pues, en este caso no se pretende conocer el tiempo, sino identificar los hechos que pueden modificarlo. Para este propósito el modelo de Cox es el más habitual y es el protagonista de este capítulo.

Figura 6.2. Probabilidad de supervivencia hasta el quinto año tras el diagnóstico de cáncer gástrico según la edad al diagnóstico



El método *log-rank* permite comparar ambas curvas.

2. Planteamiento, pregunta de investigación y propósito

La fibrilación auricular (FA) es la arritmia cardiaca más frecuente en la práctica clínica diaria. Afecta a un 1-2% de la población general y hasta el 10% por encima de los 80 años. En el ámbito de la atención primaria, presentan FA hasta el 17,7% de los pacientes mayores de 80 años asistidos por cualquier motivo. En el hospital alcanza hasta el 32,3% de los ingresados en servicios de medicina interna. Estos pacientes se caracterizan por una elevada complejidad: pluripatología con índices elevados de fragilidad, mortalidad, deterioro funcional, polifarmacia y frecuente situación de dependencia funcional. Como consecuencia, existe una frecuente demanda de atención, en distintos niveles y de manera difícilmente programable, con agudizaciones de las patologías conocidas y aparición de otras interrelacionadas que agravan la situación del paciente. Como es lógico suponer, la mortalidad entre estos pacientes ingresados en el hospital es elevada. De hecho, son las distintas combinaciones, duración y severidad de las enfermedades crónicas y sus interacciones con las complicaciones agudas las que marcan el pronóstico en un proceso muy heterogéneo y complejo.

La pregunta de investigación, en consonancia con lo anterior, sería la siguiente: ¿Cuáles son los factores que determinan el pronóstico de los pacientes con FA ingresados en servicios médicos de un hospital general?

Los objetivos del estudio se resumen así:

1. Describir las características generales de los pacientes con FA que han ingresado en alguno de los servicios médicos de un hospital general.
2. Determinar los factores independientes que influyen en su mortalidad, incluyendo la propia FA.

3. Tipo de estudio y diseño. Pruebas estadísticas y abordaje del problema

A continuación detallamos los pasos necesarios para contestar a nuestra pregunta de investigación.

Diseño

El propósito de nuestro análisis es estudiar el pronóstico de pacientes con FA que ingresan en el área médica de un hospital y establecer cuáles son sus determinantes.

El primer paso es plantear un diseño del estudio que permita obtener datos válidos y no sesgados. El estudio de cohortes es el "esqueleto" más adecuado para este fin. En este tipo de estudio se seleccionan los pacientes a partir de su enfermedad (FA en nuestro caso). Por ello se diseñó un estudio de cohortes retrospectivo de todos los pacientes ingresados en todos los servicios y secciones del área médica de un hospital general entre los años 2000 y 2018. De ellos se seleccionaron aquellos con FA y se estudiaron tanto sus características clínicas como la mortalidad durante y tras la estancia hospitalaria. Para ello se revisaron las características clínicas, la fecha y el motivo principal de ingreso hospitalario y los diagnósticos secundarios y, al alta, su fecha y su causa. En todos los casos se realizó seguimiento hasta el fallecimiento o la fecha de censura a 31 de diciembre de 2020. Este esquema permite disponer de "buenos" datos: el origen y el evento que se va a medir son claros y la escala de tiempo, días. Además, si los datos provienen de nuestros registros hospitalarios, o son verificables, no tendremos pérdidas.

Con todo ello dispondremos de una base de datos, una matriz, donde cada columna será una variable en estudio y cada fila un paciente. Debemos verificar que disponemos de todas las fechas de ingreso, de todas las fechas de alta y el estado al alta (vivo/muerto), así como de todas las características clínicas que consideremos de interés para explicar el fallecimiento.

Tamaño muestral

En nuestro ejemplo, los pacientes necesarios para estimar los eventos de interés reflejados en los objetivos del estudio con un nivel de confianza del 95% y una precisión de $\pm 1\%$ son 9604 pacientes. Un estudio piloto previo en nuestro archivo mostró que, en el periodo señalado, 14.538 pacientes (18,2%) presentaban FA. Por ello, consideramos que podemos analizar los parámetros en investigación con total precisión al contar con un tamaño muestral muy superior al preciso.

Variables y definiciones principales

Una vez identificados los pacientes, el siguiente aspecto es identificar la constelación de características que pueden definir o modificar el riesgo de fallecer en una persona. En cada análisis serán distintos y múltiples, aunque, a efectos prácticos, pueden agruparse en fijos (por ejemplo, el sexo) o variables (tratamientos, valores de un parámetro analítico, etc.). Lo importante es recoger todos aquellos que, *a priori*, pensamos que podrían influir en el fallecimiento de los pacientes. Lógicamente,

esto se puede extender a cualquier evento. Por ejemplo, tiempo y condicionantes precisos para aprobar el carnet de conducir.

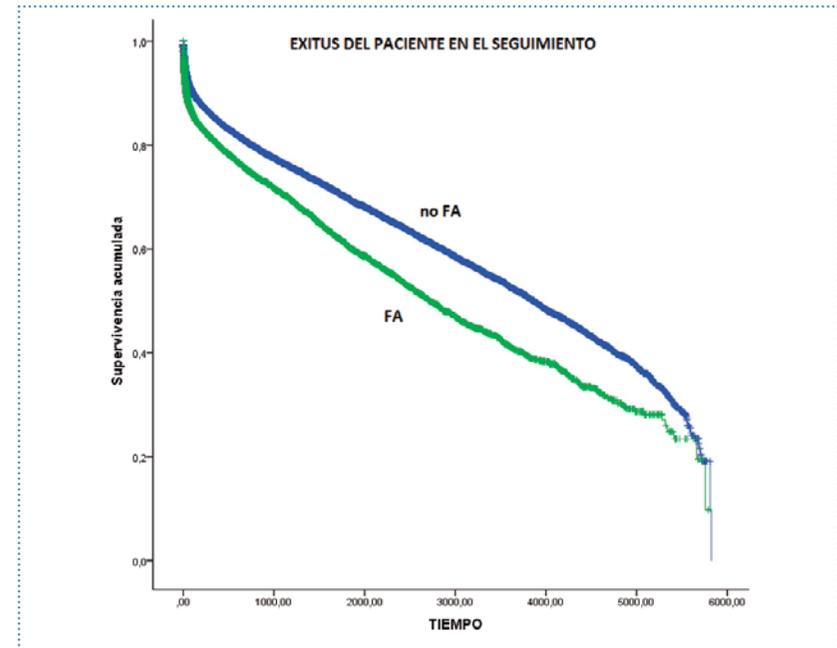
Una vez identificados, el siguiente paso es establecer criterios operativos capaces de definir cada variable de manera precisa. A efectos prácticos esto incluye los criterios de inclusión y exclusión de los casos y también la definición inequívoca de cada una de las variables. En cualquier caso, además de ser precisa, ha de tener sentido clínico. Por ejemplo, podemos definir insuficiencia cardiaca como la suma de hallazgos clínicos, radiológicos y analíticos definidos de manera muy precisa; pero también podemos decir que un paciente la ha sufrido si este diagnóstico consta en el informe de alta. Lo importante es tener un criterio único, con sentido clínico, y aplicarlo en todos los casos.

Pruebas estadísticas y presentación de resultados

Tras confeccionar una base de datos completa y libre de errores, el siguiente paso es realizar un análisis descriptivo de las variables. Para conocerlas en profundidad utilizaremos tablas de frecuencias en las categóricas y medidas de tendencia central y dispersión en las numéricas. Esto nos permite no solo conocer bien nuestros datos (y empezar a plantearnos posibles análisis), sino también descubrir errores, identificar valores extremos, etc. En la comparación de porcentajes podemos utilizar el test de χ^2 (o chi al cuadrado) –y el test exacto de Fisher en caso de que sea necesario–. Para comprobar la igualdad de dos medias aplicamos la prueba *t* de Student, y en el caso de comparaciones múltiples el test de análisis de la varianza (ANOVA). Previamente es preciso verificar la igualdad de varianzas mediante el test *F* de Snedecor o test de Levene. En los pocos casos de distribución no normal se utilizaron métodos no paramétricos.

Llegados a este punto, con un conocimiento adecuado de nuestros datos, es posible realizar el análisis univariado de supervivencia. Como hemos visto, la probabilidad de supervivencia la podemos estimar, tanto de forma global como para cada una de las características en estudio, con el método no paramétrico de Kaplan-Meier. Se estimará la función de riesgo y riesgo acumulado para detectar la probabilidad de los eventos de las situaciones de interés clínico. Luego hemos de comparar las curvas de supervivencia. Para ello utilizaremos la prueba *log-rank*, presentando los gráficos de las curvas de supervivencia estimadas. En la **figura 6.3** se observa la probabilidad de supervivencia, tras el primer ingreso, de los pacientes con FA frente a los que no la presentaban. En la comparación entre ambas se ha utilizado la prueba *log-rank*.

Figura 6.3. Probabilidad de supervivencia tras el primer ingreso de los pacientes con fibrilación auricular frente a los pacientes sin fibrilación auricular



FA: fibrilación auricular. $p < 0,0001$.

Análisis multivariante

Hasta aquí hemos visto los métodos que permiten estimar las funciones de riesgo y supervivencia para una muestra aleatoria y comparar dos estimaciones. El objetivo ahora es plantear un modelo que nos permita seleccionar, de entre toda una constelación de factores de riesgo, aquellos que contribuyen a predecir el tiempo de supervivencia teniendo en cuenta las posibles influencias de los otros. En nuestro caso, a partir de la supervivencia de pacientes con FA y de variables como la edad, el sexo, la presencia de insuficiencia cardiaca, la anemia, etc., se desea conocer cuáles de todos estos factores ayudan a predecir la supervivencia. Por lo tanto, el objetivo es plantear un modelo de regresión que permita estimar la supervivencia en función de diferentes variables “explicatorias”. Por ejemplo, la supervivencia a dos tratamientos alternativos para una enfermedad puede depender no solo del tra-

tamiento, sino también de las características del paciente, del retraso diagnóstico, etc. Conocer esas variables que “explican” el resultado puede ser de gran importancia para decidir qué tratamiento utilizar.

Si bien se han propuesto varios modelos de regresión para afrontar este problema, el más utilizado, por su sencillez y facilidad de interpretación, es el modelo de riesgos proporcionales o modelo de Cox. En este modelo, el primer paso es buscar la relación entre los riesgos de muerte de dos individuos expuestos a factores de riesgo diferentes. Para ello, el modelo parte de una hipótesis fundamental, la de que los riesgos son proporcionales.

Para comprender esta noción definiremos previamente la denominada *razón o cociente de riesgos* (*hazard ratio*, HR; riesgo relativo, RR). Típicamente, se evalúa en el numerador el grupo de mayor riesgo y en el denominador el de menor. En tal caso, se espera que el HR sea mayor que 1. Es más fácil la interpretación si excede del valor base de la unidad (que indica que tienen el mismo riesgo) que si disminuye de la unidad. En el modelo de Cox se parte de la premisa de que los riesgos son proporcionales. La hipótesis de riesgos proporcionales significa explícitamente que la razón de riesgos tiene el mismo valor a lo largo del periodo de seguimiento, ya que se suponen variables no dependientes del tiempo. Para utilizar la regresión de Cox para unos datos determinados será necesario verificar previamente que se cumple esta hipótesis de proporcionalidad de riesgos. Existen varios métodos. Por un lado, puede utilizarse un método gráfico: si una variable, por ejemplo, toma únicamente los valores 0 y 1 (por ejemplo, la presencia o ausencia de *shock* séptico), pueden representarse las curvas de supervivencia para los dos grupos. Puesto que la presencia de *shock* supone tener mayor riesgo, analizamos cómo las dos curvas se separan de manera progresiva, sin cruzarse. También pueden utilizarse técnicas más rigurosas, como el método de los residuos de Schoenfeld.

En la construcción del modelo el primer paso es seleccionar las variables candidato. En nuestro caso seleccionamos aquellas que en el análisis univariado tuvieron un nivel de significación inferior al 5%. Sin embargo, en nuestro caso, se decidió añadir al modelo otras variables de carácter sociodemográfico que, aun no siendo significativas, sí las considerábamos clínicamente relevantes y que podrían modificar el resultado final. De esta manera, seleccionamos un total de 26 variables (edad, sexo, nivel socioeconómico, año de ingreso, patologías acompañantes, etc.), y para todas ellas la proporcionalidad de los riesgos se evaluó por métodos gráficos. Hemos de estar seguros de que todas las variables seleccionadas tienen en común que son relevantes clínicamente, que son fáciles de medir y, en el caso de las categóricas,

que tienen una estructura ya conocida de estudios previos. Sin embargo, debemos tener cuidado de no incluir un número excesivo de variables, ya que se corre el riesgo de obtener resultados finales erróneos. Es habitual que el clínico le solicite al estadístico que incluya todas las variables posibles. Esto puede ser un error, y para evitarlo se sugiere valorar el tamaño de la muestra en relación con el número de eventos (por ejemplo, la muerte) por variable. Se considera adecuado un valor superior a 20. Con todo ello ya disponemos del modelo máximo; esto es, un conjunto de variables dispuestas para que sean valoradas como contribuyentes, o no, al pronóstico.

En la construcción del modelo máximo los conocimientos y las decisiones clínicas son relevantes, y clínico y estadístico han de trabajar codo con codo. A partir de aquí el peso del análisis es más estadístico. En nuestro caso, decidimos una estrategia “hacia atrás”, esto es, partir del modelo máximo y retirar paulatinamente las variables que no son significativas, iniciando el proceso con los factores de interacción entre variables.

Interpretación del análisis de Cox

En la **tabla 6.1** se muestra el resultado obtenido tras concluir el análisis de Cox en nuestro estudio. En la primera columna se detallan las variables finales seleccio-

Tabla 6.1. Modelo final del análisis de Cox: probabilidad de supervivencia de pacientes con fibrilación auricular ingresados en servicios médicos

Variable	p	Exp. (B)	IC95% para Exp. (B)	
			Inferior	Superior
Sexo (referencia: mujer)	0,000	1,204	1,162	1,243
Edad	0,000	1,043	1,042	1,049
Fibrilación auricular	0,273	1,029	0,971	1,067
Ictus	0,000	1,161	1,0884	1,227
Insuficiencia cardíaca congestiva	0,000	1,501	1,423	1,544

Exp. (B): riesgo relativo; IC95%: intervalo de confianza al 95%.

nadas como pronósticas en esta situación clínica. En las variables insuficiencia cardiaca, ictus y FA, lo que se ofrece es el riesgo relativo (columna "Exp. (B)") de sufrir cada una de estas patologías frente a no tenerla. En el caso del sexo, se especifica cuál es el de referencia. En el caso de la edad, se expresa como, por cada año que aumenta la edad del paciente, el riesgo de fallecer se incrementa 1,048 veces con respecto al de menor edad.

En las cuatro siguientes columnas se especifican sucesivamente el nivel de significación estadística, el RR y sus límites inferior y superior. El RR se interpreta en relación con el valor 1 (neutro). En nuestro caso, la insuficiencia cardiaca muestra un valor de 1,501; esto es, su presencia incrementa en un 50% el riesgo de fallecer. Puesto que sus límites no incluyen el 1, hemos de interpretar que este dato es significativo. Por el contrario, el RR asociado a la presencia de FA, si bien supera levemente el 1, este está incluido en los límites. En este caso, interpretamos que la presencia de FA, por sí misma, no es un factor que influya en el pronóstico de este grupo de pacientes. En nuestro ejemplo no ocurre, pero en otros podemos encontrar una variable con un valor inferior a 1 y cuyos límites no lo incluyen el 1. La interpretación en este caso sería que se trata de un factor protector.

4. Justificación y discusión

Establecer los factores que determinan el pronóstico en una situación clínica concreta es un proceso complejo. Es evidente que se requiere mucho conocimiento clínico para señalar un objetivo, diseñar un estudio adecuado o seleccionar las variables precisas. Pero, además, para usar de manera correcta el modelo de Cox es preciso tener experiencia previa con esta técnica. Por ello, por lo general, se recomienda que el clínico trabaje conjuntamente con un estadístico acostumbrado a las peculiaridades del mundo sanitario. Debemos entender que no es preciso que los clínicos seamos unos expertos en el manejo de estas herramientas; lo necesario es que seamos capaces de entenderlas o interpretarlas, o de plantear un estudio cuya parte estadística las incluya. Si las conocemos y las entendemos, podremos plantear estudios de investigación clínica de elevada complejidad.

Identificar estos factores que modifican el pronóstico en diferentes situaciones, como, por ejemplo, con el uso de un nuevo tratamiento, es muy importante tanto

para tomar decisiones como para informar a nuestros pacientes. Cuando realizamos un análisis de supervivencia es habitual que encontremos dos tipos de factores pronósticos. Unos factores fijos, que no podemos modificar, como el sexo u otros hechos inamovibles con los que hemos de contar. Pero en muchos casos es posible que identifiquemos otros factores que esté en nuestra mano cambiar o impedir. En nuestro caso, tanto el ictus como la insuficiencia cardiaca son situaciones clínicas que se pueden prevenir o modificar.

Por otra parte, el conocimiento de estas técnicas permite una lectura más adecuada y profunda de muchos artículos de la literatura médica actual. Si bien se trata de una técnica relativamente reciente, su uso es muy habitual en artículos que modifican la práctica y el conocimiento clínico. Entender los conceptos, el desarrollo o interpretar sus resultados son claves en la formación del médico de hoy en día.

Bibliografía

- Abraira V, Pérez de Vargas A. Métodos multivariantes en bioestadística. Madrid: Editorial Centro de Estudios Ramón Areces; 1996.
- Casariego E, Pita S, Rigueiro MT. Supervivencia del cáncer gástrico y factores que modifican el pronóstico. Análisis en 2334 pacientes. Med Clin (Barc). 2001;117:361-5.
- Iñiguez Vázquez I, Matesanz Fernández M, et al. Predictor of enhanced mortality in patients with multimorbidity and atrial fibrillation in an acute hospital setting. QJM. 2020 May 1;113(5):330-35. doi: 10.1093/qjmed/hcz302.
- Lee ET, Wang WU. Statistical methods for survival data analysis. Wiley Series in probability and statistics. 4.^a ed. New York: John Wiley and Sons; 2013.
- Rabuñal R, Pita S, Rigueiro MT, et al. Presentation and prognosis of cancer gastric in patients aged 80 years and older. World J Surg. 2004;28:155-9.
- Santabárbara J, Rubio E, Martínez T. Manual de análisis de supervivencia: curvas de supervivencia y regresión de Cox. Zaragoza: Pressas de la Universidad de Zaragoza; 2017.

Preguntas de autoevaluación

1. Evaluamos un artículo que describe la probabilidad de supervivencia de pacientes de 100 años o más, evaluados en su domicilio en situación de estabilidad clínica. Se evaluaron 181 personas de ambos sexos y el examen incluyó una entrevista, una exploración, un electrocardiograma y una analítica básica. Todos los pacientes fueron seguidos hasta su fallecimiento. Con los datos de esa única entrevista, los autores confeccionaron este análisis de Cox:

	B	p	HR	IC95% para HR
Edad	-0,17	0,154	0,84	0,66-1,07
Sexo	0,16	0,561	1,17	0,69-1,98
Glucemia > 126	0,80	0,032	2,22	1,07-4,59
Fibrilación auricular o flutter auricular	0,71	0,011	2,03	1,18-3,50
Albúmina < 3,5	1,24	0,000	3,47	1,86-6,46
Índice de Barthel	0,57	0,024	1,77	1,08-2,92

HR: hazard ratio; IC95%: intervalo de confianza al 95%.

A la vista de estos resultados, ¿qué conclusión podemos extraer?

- La edad durante la entrevista es un factor que tiene poca influencia sobre el pronóstico.
- El índice de Barthel es el factor con mayor influencia.
- La presencia de fibrilación auricular en el momento de la entrevista multiplica por 3,50 el riesgo de fallecer.
- Todas las afirmaciones anteriores son incorrectas.

La respuesta correcta es la opción a.

Puede observarse que el intervalo de confianza en el caso de la edad incluye el 1; por lo tanto, su valor pronóstico no es significativo. Por lo tanto, la edad del paciente en el momento de la entrevista no es un factor pronóstico. Sí lo son los cuatro últimos factores, y el de mayor peso

es la presencia de un valor bajo de albúmina. El índice de Barthel es el de menor peso. Por último, el valor señalado para la fibrilación auricular es el del límite superior del intervalo de confianza, no el del HR.

2. En un estudio pronóstico tras el diagnóstico de cáncer gástrico observamos el siguiente modelo final de Cox:

	Coficiente	ES	RR	IC95% del RR	p
Edad^a					
50-64	0,2652	0,1556	1,3037	0,9611-1,7684	0,08
65-79	0,2736	0,1471	1,3147	0,9854-1,754	0,06
>80	0,4477	0,1967	1,5647	1,0642-2,3006	0,02
Localización^b					
Tercio medio	-0,4182	0,3032	0,6583	0,3634-1,1925	0,16
Muñón	0,0023	0,1304	1,0023	0,7762-1,2942	0,98
Tercio superior	0,2065	0,1124	1,2294	0,9863-1,5324	0,06
Generalizado	0,4429	0,1125	1,5573	1,2492-1,9414	<0,001
Clasificación TNM^c					
Ib	0,5987	0,2479	1,8197	1,1193-2,9583	0,01
II	0,8383	0,2429	2,3125	1,4364-3,7228	<0,001
IIIa	0,8152	0,2452	2,2597	1,3973-3,6542	<0,001
IIIb	0,8415	0,2819	2,3199	1,2251-4,0312	0,002
IV	1,2924	0,2438	3,6414	2,258-5,8724	<0,001
Intención^d					
Paliativa	0,8830	0,1098	2,4181	1,9499-2,9987	<0,001
No cirugía	1,3918	0,1132	4,0221	3,2221-5,0208	<0,001
Tipo histológico de Lauren^e					
Difuso	0,1481	0,0850	1,1592	0,9816-1,3699	0,08
Mixto	0,5607	0,1959	1,7519	1,1932-2,5722	0,004

Deviance: 7361,4644; $p < 0,001$. ES: error estándar; RR: riesgo relativo; IC95%: intervalo de confianza al 95%. ^a Referencia: Edad < 50. ^b Referencia: tercio inferior. ^c Referencia: Ia. ^d Referencia: intención curativa. ^e Referencia: tipo intestinal.

Tras analizar los datos, ¿cuál de las siguientes afirmaciones debería aparecer entre nuestras conclusiones?

- a) Los tumores del tercio inferior de la cámara gástrica son los de peor pronóstico.
- b) En este estudio el incremento de la edad no se asocia a un peor pronóstico.
- c) El factor que tiene un mayor peso es la imposibilidad de intervenir al paciente.
- d) Todas las afirmaciones son correctas.

La respuesta correcta es la opción c.

Observando los HR, podemos afirmar que el correspondiente a la no cirugía es el de mayor valor, incluso superior al del grado de extensión según la clasificación TNM. Por el contrario, la localización del tumor en el tercio inferior es la que se utiliza de referencia y, por lo tanto, con mejor pronóstico. Con respecto a la edad, se puede apreciar cómo con el paso de los años se incrementa de forma consistente el HR.

3. A la hora de seleccionar las variables que vamos a utilizar para realizar una regresión de Cox, debemos tener en cuenta que...

- a) Es imprescindible seleccionar exclusivamente aquellas que han sido significativas en un análisis univariado previo.
- b) Podemos seleccionar las significativas en el análisis univariado e incorporar alguna variable que no haya resultado significativa si la consideramos de especial relevancia clínica.
- c) Lo habitual es incluir todas las variables disponibles, ya que eso permite explorar mejor la situación clínica y podemos sacarle un mayor rendimiento.
- d) Todas las afirmaciones anteriores son correctas.

La respuesta correcta es la opción b.

Lo más adecuado es incluir las variables con significación en el estudio univariado. Además, es posible añadir alguna variable sin significación en el estudio previo, siempre que consideremos que es muy relevante y que sin ella el análisis podría ser incompleto. Sin embargo, no es conveniente introducir todas las disponibles, ya que no se produce un mayor beneficio y podemos superar los límites recomendados de evento por variable.



Resolver un problema clínico mediante un ensayo clínico de un fármaco frente al tratamiento habitual o placebo

Dr. Joan Carles Trullàs
Servicio de Medicina Interna
Hospital de Olot i Comarcal de la Garrotxa (Olot, Gerona)

capítulo 7

1. Situación actual y contexto clínico

Este problema clínico se va a centrar en el tratamiento de la insuficiencia cardiaca aguda (ICA) con sobrecarga de volumen. La ICA es una situación clínica muy prevalente, pero existe poca evidencia científica sobre su manejo terapéutico. El tratamiento de elección son los diuréticos del asa (DA), que en más de la mitad de los casos son el único tratamiento administrado a estos pacientes.

En ocasiones el tratamiento con DA no será suficiente para conseguir la descongestión, y hablaremos de resistencia al tratamiento. Esto puede suceder por muchos motivos, como, por ejemplo, la malabsorción del fármaco, la mala distribución y llegada del fármaco a la luz tubular, la hipertrofia e hiperfunción de los segmentos distales de la nefrona tras el uso continuado de DA, etc. Independientemente del motivo, cuando el tratamiento con DA no es suficiente, algunos pacientes persistirán con sobrecarga de volumen o congestión residual.

Cuando eso sucede, la opción terapéutica recomendada es combinar los DA con otras clases de diuréticos (diuréticos tiacídicos, antagonistas de la aldosterona, inhibidores de la anhidrasa carbónica, etc.), pero hasta el momento ningún estudio ha demostrado que la adición de uno de ellos sea superior (y, por tanto, de elección) frente a otro.

2. Planteamiento, pregunta de investigación y objetivos

En este escenario, puede ser interesante desarrollar un estudio para explorar si añadir un segundo diurético al DA es una estrategia que permite vencer la resistencia al tratamiento, eliminar la congestión residual y conseguir la completa descongestión. Vamos a plantear la resolución de esta situación clínica mediante un ensayo clínico que proponga resolver el problema añadiendo acetazolamida (un inhibidor de la anhidrasa carbónica). Para ello vamos a hacer una aproximación similar a la que hizo un grupo de investigadores belgas al diseñar el ensayo clínico ADVOR, un ensayo clínico real que actualmente está en fase de reclutamiento y cuyo diseño se publicó en 2018 (véase la bibliografía).

La hipótesis de trabajo o pregunta de investigación es que la acetazolamida (añadida al DA) podría mejorar la eficacia del DA al facilitar una mayor natriuresis y diuresis y, por tanto, una mejor y más rápida descongestión.

El objetivo (o *endpoint*) principal de este estudio sería determinar la tasa de éxito a la hora de lograr una descongestión completa en pacientes con ICA y sobrecarga de volumen. Hay que definir también las variables que se van a utilizar para

medir el objetivo principal de forma objetiva y homogénea entre los distintos investigadores. En este estudio la variable para medir el objetivo principal sería valorar la proporción de pacientes que han alcanzado una descongestión completa después de tres días de tratamiento aleatorizado, es decir, en la mañana del día 4 del estudio.

Al margen del objetivo principal, normalmente se plantean también otros objetivos, llamados secundarios, y en este estudio se podrían plantear los siguientes: mortalidad a los tres meses, reingresos hospitalarios a los tres meses, estancia hospitalaria, calidad de vida, etc. También habrá que definir muy bien las variables para determinar estos objetivos secundarios.

3. Tipo de estudio y diseño. Pruebas estadísticas y abordaje del problema

Tipo y diseño de estudio. Criterios de inclusión/exclusión. Aleatorización y enmascaramiento

Como se ha comentado previamente, las distintas estrategias diuréticas para el manejo de la ICA tienen poca evidencia científica y la mayoría de las recomendaciones se basan en opiniones de expertos. Cuando esto sucede es deseable diseñar estudios que aporten la mayor evidencia científica posible. Es por ello que, para alcanzar el objetivo principal de este problema clínico, plantearíamos un estudio de intervención o ensayo clínico multicéntrico, aleatorizado, doble ciego y controlado con placebo.

Lo primero que hay que definir son los criterios de inclusión/exclusión de los pacientes. Los criterios de inclusión de este estudio podrían ser los siguientes:

- Obtener el consentimiento informado antes de empezar cualquier procedimiento del estudio.
- Edad igual o superior a 18 años.
- Ingreso hospitalario por ICA y con un signo como mínimo de sobrecarga de volumen: edemas, ascitis o derrame pleural (confirmado por pruebas de imagen).
- Tratamiento previo con DA como mínimo durante un mes antes del ingreso (dosis mínimas de 40 mg de furosemida o equivalentes de torasemida/bumetadina).

- Niveles elevados de péptidos natriuréticos en el momento de la selección (NT-proBNP > 1000 pg/mL o BNP > 250 ng/mL).
- Tener documentada la fracción de eyección del ventrículo izquierdo en los doce meses previos a la inclusión.

También hay que definir los criterios de exclusión, normalmente extensos y que suelen incluir algunos de los siguientes: hipotensión arterial, inestabilidad hemodinámica, alteraciones graves de la función renal (normalmente se define un punto de corte de filtrado glomerular) o necesidad de terapia renal sustitutiva, uso previo de los fármacos que se investigan en el estudio, malas condiciones (médicas, psíquicas o sociales) que comprometan el buen cumplimiento y desaconsejen la inclusión del paciente en el ensayo, etc.

Una vez establecidos estos criterios, se debe definir cuál va a ser la intervención. En este estudio, la intervención sería la administración del fármaco en investigación (acetazolamida) o placebo de la siguiente manera: una vez que el paciente ingresa, se cambia la vía de administración del DA de vía oral a vía intravenosa a una dosis protocolizada (por ejemplo, el doble de la dosis que recibía en casa por vía oral). Al mismo tiempo, todos los pacientes recibirán de forma aleatorizada tratamiento con acetazolamida o placebo durante tres días consecutivos. El ajuste de dosis del DA (y su posterior paso a vía oral) se realizará cada día durante los tres días de tratamiento aleatorizado y de acuerdo a un *score* de congestión predefinido (que incluye la presencia o ausencia de edema, derrame pleural y ascitis) y a criterio clínico del médico o investigador que trata al paciente.

La aleatorización se define como que todos los individuos tengan la misma probabilidad de entrar en el grupo de tratamiento y la misma probabilidad de pertenecer al grupo placebo. Las técnicas de enmascaramiento intentan evitar la influencia sobre los resultados que pueda tener el que los sujetos participantes o los investigadores conozcan el tratamiento administrado. El estudio será "simple ciego" si el sujeto participante está enmascarado; "doble ciego" cuando además el observador está enmascarado; y "triple ciego" cuando también lo está el analista de los resultados.

Cálculo del tamaño muestral

Antes de empezar un ensayo clínico los investigadores deben asegurar que va a tener suficiente potencia para evaluar el objetivo principal, habitualmente mediante el cálculo del tamaño muestral mínimo necesario. El primer paso para calcular el tamaño muestral es estimar en qué proporción de pacientes se va alcanzar el obje-

tivo principal (en este caso, la descongestión completa) en cada rama de tratamiento. Para ello es necesario revisar los resultados de otros estudios ya finalizados, que incluyeran una población de estudio similar a la nuestra y que analizaran el mismo objetivo principal.

En este caso, podemos recurrir al ensayo clínico DOSE (véase la bibliografía), que comparó dos estrategias de tratamiento con DA intravenoso en pacientes que ingresaron por ICA: dosis bajas (la misma que tomaba en casa vía oral) frente a dosis altas (la dosis habitual vía oral multiplicada por 2,5). En este estudio llegaron a la descongestión completa el 11% y el 18% de los pacientes del grupo de dosis bajas y altas, respectivamente. Tomando el valor promedio de estos porcentajes, podríamos estimar que un 15% de los pacientes de la rama de tratamiento con furosemida y placebo de nuestro estudio alcanzarán la completa descongestión. Respecto a la rama de pacientes a los que se les asigne tratamiento con furosemida y acetazolamida, no es posible hacer esta misma aproximación, pues no existen estudios de este tipo hasta la fecha. Cuando esto sucede, hay que hacer una aproximación arbitraria y suponer que el grupo de tratamiento va a conseguir la completa descongestión en un 25% de los pacientes. Una vez que tenemos fijadas estas dos estimaciones (15% y 25%), debemos fijar también el error que estamos dispuestos a asumir y las posibles pérdidas de pacientes estimadas.

Cuando disponemos de toda esta información, ya podemos proceder a calcular el tamaño muestral, algo que se puede hacer de forma manual (si se tienen suficientes conocimientos estadísticos) o mediante la ayuda de calculadoras del tamaño muestral (muchas de ellas son gratuitas y están disponibles en línea). Así, en nuestro estudio, si asumimos un error de tipo I o α de 0,05 y un error de tipo II o β de 0,2 (lo que significa una potencia estadística del 80%), el tamaño de la muestra calculado sería de 494 pacientes. Si además anticipamos que vamos a perder un 5% de los pacientes, el número final de pacientes que hay que incluir sería 519.

Análisis estadístico

Antes de empezar el estudio hay que plantear cómo se va a realizar el análisis estadístico. En primer lugar, hay que definir cómo se van a expresar las variables y qué test estadísticos vamos a aplicar para su comparación en función de si son variables cualitativas o cuantitativas, si siguen una distribución normal o no y cuando comparemos dos grupos o más.

Lo habitual es comparar los grupos de tratamiento utilizando modelos lineales (si el *endpoint* es una variable cuantitativa), modelos de regresión logística (si el *endpoint*

es una variable binaria) y modelos de Cox o curvas de Kaplan-Meier cuando la variable tiene en cuenta el tiempo (tiempo hasta el evento). En nuestro estudio, la variable principal “alcanzar o no la descongestión completa” es una variable dicotómica (sí o no), por lo que deberíamos utilizar un modelo de regresión logística. Si vamos a comparar la proporción de pacientes que alcanzan este objetivo en las dos ramas de tratamiento, debemos aplicar test de comparación de proporciones.

Una vez aplicados los test estadísticos para comparar los dos grupos, obtendremos los resultados expresados en unos valores o medidas de asociación. Estas medidas establecen si la frecuencia de una variable (en este caso, el porcentaje de pacientes que consiguen descongestión completa) es diferente entre los expuestos a una determinada intervención (en este caso, al fármaco acetazolamida). La medida de asociación en la que se suelen expresar los resultados de los ensayos clínicos es el *hazard ratio* (HR), que es un cociente o razón entre dos tasas (la de expuestos y la de no expuestos). Su interpretación es cuántas veces se incrementa (o reduce) el riesgo de presentar un evento dependiendo de la exposición: si es >1 , implica incremento del riesgo; si es <1 , reducción de riesgo; y si es 1 indica efecto nulo.

Las medidas de asociación se acompañan de un intervalo de confianza al 95% (IC95%). El IC95% ofrece una manera de estimar, con alta probabilidad, un rango de valores en el que se encuentra el verdadero valor poblacional de la variable en estudio. Esta probabilidad ha sido fijada por consenso en un 95% basándose en supuestos de normalidad, pero rangos entre el 90% y el 99% se utilizan también en la literatura científica. Para analizar si existe o no diferencia estadísticamente significativa con el IC95% debemos observar los extremos del intervalo. Independientemente de si el estimador puntual muestra beneficio, debemos verificar si alguno de los extremos del intervalo pasa sobre la línea del efecto nulo (en el HR, valor igual a 1). Si es así, existe la posibilidad de que el valor poblacional real corresponda al no efecto o incluso tenga un efecto opuesto al esperado. En este caso, no existiría diferencia estadísticamente significativa entre aplicar o no la intervención.

También se debe preespecificar el valor de p del contraste de hipótesis a partir del que consideraremos que las diferencias encontradas son significativas. La habitual obsesión de los investigadores por obtener una $p < 0,05$ se basa en un convenio general que plantea que encontrar un resultado por azar no más de una vez de cada 20 ensayos podría considerarse significativo. Sin embargo, la búsqueda p depende de la magnitud de la diferencia, pero también del tamaño de la muestra, y basta incrementarlo para obtener resultados “estadísticamente significativos”. Nunca hay que olvidar que el valor de la p no mide la magnitud del efecto. Al inves-

tigador corresponde interpretar si un resultado estadísticamente significativo es clínicamente relevante.

¿Entonces, cómo evaluamos la significación estadística, con el intervalo de confianza al 95% o con el valor de p ?

Hacer pruebas de hipótesis (valores de p) o hacer estimaciones (intervalos de confianza) son técnicas validadas que contribuyen a la precisión de los efectos clínicos de cualquier investigación original. Ambas están estrechamente relacionadas, y es raro que un IC95% ofrezca un resultado significativo y un valor de p no. No obstante, la información ofrecida por un intervalo de confianza es más detallada que la contenida en los valores de p , y permite al clínico además estimar si el rango de una diferencia entre tratamientos es lo suficientemente grande como para justificar su uso en una patología relevante.

En la **tabla 7.1** se presentan ejemplos de posibles resultados de nuestro estudio. En las dos primeras columnas se expone el porcentaje de pacientes que han alcanzado el objetivo principal en cada grupo de tratamiento, y en las siguientes columnas el HR, el IC95% y el valor de p .

En el primer ejemplo, el “riesgo” de presentar el objetivo primario (en este caso, beneficioso, pues significa la completa descongestión) es 1,7 veces mayor en el grupo de pacientes que reciben el fármaco en comparación con el grupo placebo. Además, el IC95% es estrecho (lo que indica que la estimación del valor poblacional

Tabla 7.1. Posibles resultados del estudio sobre la acetazolamida frente a placebo en el tratamiento de la insuficiencia cardiaca aguda

	Placebo	Acetazolamida	<i>Hazard ratio</i>	Intervalo de confianza al 95%	p
1	0,15	0,25	1,7	1,4-1,9	0,002
2	0,15	0,17	1,1	1,1-1,5	0,03
3	0,15	0,17	1,1	0,9-1,4	0,12
4	0,10	0,20	2,0	1,2-5,1	0,04
5	0,16	0,10	0,6	0,5-0,8	0,02

real es bastante precisa) y no incluye el valor 1 (o de nulo efecto de la intervención). Estos resultados son clínicamente relevantes y además estadísticamente significativos (tanto por el IC95% como por el valor de p).

En el segundo ejemplo, la diferencia a favor de la intervención sigue siendo estadísticamente significativa, pero la magnitud o relevancia clínica es menor, pues el HR es solo de 1,13 y con un IC95% cercano al valor 1.

En el tercer ejemplo, el IC95% incluye el 1, por lo que deja de ser estadísticamente significativo (igual que el valor de p).

El cuarto ejemplo muestra también un efecto beneficioso y significativo de la intervención, pero en esta ocasión hay que fijarse bien en el IC95%, que es más amplio (o menos preciso) y con el límite inferior cercano al 1, por lo que la estimación no es tan precisa como en el ejemplo 1 (a pesar de ser estadísticamente significativa y clínicamente relevante, al ser el HR de 2).

En el último ejemplo, estaríamos ante una situación en la que la intervención tendría menos probabilidad (HR inferior a 1) de conseguir el objetivo principal, algo que en este caso sería desfavorable (los pacientes tratados con acetazolamida presentarían menos descongestión completa). Además, este resultado sería estadísticamente significativo, al no incluir el 1 y tener un valor de p inferior a 0,05.

Análisis por intención de tratar

Es habitual predefinir también si se va a realizar un análisis por intención de tratar. El diseño ideal de un ensayo clínico, mediante la aleatorización, procura que los dos grupos sean iguales en todo, excepto en el tratamiento recibido, para que así, si al final hay diferencias entre ellos, estas puedan atribuirse al efecto del tratamiento (evitar el sesgo de selección). Sin embargo, en la práctica es dudoso que este ideal se alcance, pues casi siempre hay algún paciente que no satisface los criterios de entrada, no completa el tratamiento al que fue asignado inicialmente o no permanece en el estudio hasta el final, lo que impide recoger todos sus datos. El modo generalmente recomendado de evitar los sesgos que pueden introducir estas desviaciones del protocolo es el denominado análisis por intención de tratar, que, según su definición más ampliamente aceptada, consiste en incluir (en el análisis) a todos los pacientes en los grupos que se le asignaron aleatoriamente, independientemente de que cumplieran o no los criterios de entrada, del tratamiento que realmente recibieran y de que abandonaran el tratamiento o se desviaran del protocolo.

4. Justificación y discusión

Un estudio de intervención o ensayo clínico es el que siempre nos aportará un mayor nivel de evidencia científica. Su fortaleza es la distribución aleatoria de la intervención, que procurará que los dos grupos de tratamiento estén bien distribuidos y balanceados en sus características basales. Ello no ocurre en los estudios observacionales, en los que es necesario recurrir a análisis estadísticos multivariados ajustados para superar esta limitación de no aleatorización. A pesar de estos ajustes estadísticos, el nivel de evidencia científica de los estudios observacionales es inferior.

Los ensayos clínicos también tienen limitaciones. A veces son demasiado restrictivos en los criterios de inclusión, lo que dificulta el reclutamiento de pacientes y hace que sean poco representativos de toda la población, sino más bien de solo una parte de ella. Por otro lado, su complejidad organizativa y administrativa y su elevado coste económico dificultan la posibilidad de llevarlos a cabo.

5. Conclusiones

El mejor estudio para responder a una pregunta científica sin respuesta como la de nuestro problema es un ensayo clínico y, a pesar de las limitaciones antes mencionadas, los investigadores clínicos deberían aspirar siempre a diseñar y llevar a cabo este tipo de estudios.

Bibliografía

- Abraira V. ¿Qué es el análisis por intención de tratar? SEMERGEN. 2000;26:393-4.
- Doménech JM. Fundamentos de diseño y estadística. Barcelona: Signo; 2008. 9.ª ed. UD 7, Investigación científica: diseño de estudios.
- Felker GM, Ellison DH, Mullens W, et al. Diuretic therapy for patients with heart failure: JACC State-of-the-Art Review. J Am Coll Cardiol. 2020;75:1178-95.
- Felker GM, Lee KL, Bull DA, et al. Diuretic strategies in patients with acute decompensated heart failure. N Engl J Med. 2011;364:797-805.
- García Garmendia JL, Maroto Monserrat F. Interpretation of statistical results. Med Intensiva. 2018;42:370-9.
- Martínez-González MA, Alonso A, López Fidalgo J. ¿Qué es una hazard ratio? Nociones de análisis de supervivencia. Med Clin (Barc). 2008;131:65-72.
- Mullens W, Verbrugge FH, Nijst P, et al. Rationale and design of the ADVOR (Acetazolamide in Decompensated Heart Failure with Volume Overload) trial. Eur J Heart Fail. 2018;20:1591-1600.

Preguntas de autoevaluación

1. En relación con las técnicas de enmascaramiento de los ensayos clínicos, seleccione la afirmación correcta:

- a) Intentan evitar la influencia sobre los resultados que pueda tener el que los sujetos participantes o los investigadores conozcan las variables del objetivo principal.
- b) El estudio será "simple ciego" si el observador está enmascarado.
- c) El estudio será "doble ciego" si el sujeto participante y el observador están enmascarados.
- d) El estudio será "doble ciego" si el sujeto participante y el analista están enmascarados.

La respuesta correcta es la opción c.

Las técnicas de enmascaramiento intentan evitar la influencia sobre los resultados que pueda tener el que los sujetos participantes o los investigadores conozcan el tratamiento administrado. El estudio será "simple ciego" si el sujeto participante está enmascarado; "doble ciego" cuando además el observador está enmascarado; y "triple ciego" cuando también lo está el analista de los resultados.

2. En relación con el cálculo del tamaño muestral de un ensayo clínico, ¿cuál de las siguientes afirmaciones es falsa?

- a) Necesitamos fijar un error de tipo I.
- b) Necesitamos fijar un error de tipo II.
- c) No es necesario ajustar el cálculo a la pérdida de pacientes estimada.
- d) Es necesario conocer resultados de estudios previos ya finalizados.

La respuesta correcta es la opción c.

Para calcular el tamaño muestral es necesario revisar los resultados de otros estudios ya finalizados, que incluyeran una población de estudio similar a la nuestra y que analizaran el mismo objetivo principal. También es necesario fijar el error que estamos dispuestos a asumir (de tipo I y de tipo II) y las posibles pérdidas de pacientes estimadas.

3. En un ensayo clínico se analizó si el tratamiento con el fármaco X (en comparación con el placebo) mejoraba el pronóstico (*endpoint* principal, supervivencia a los 12 meses) de la enfermedad Y, para la que hasta la fecha no hay ningún tratamiento. El fármaco X, en comparación con el placebo, obtuvo un *hazard ratio* de 1,8, con un intervalo de confianza del 95% de 1,5-4,3 y un valor de *p* de 0,002. ¿Cuál de las siguientes interpretaciones de este resultado es más acertada?

- a) El fármaco X multiplica por 1,8 la probabilidad de supervivencia en comparación con el placebo. Estos resultados son clínicamente relevantes y estadísticamente significativos.
- b) El fármaco X multiplica por 1,8 el riesgo de muerte en comparación con el placebo. Estos resultados son clínicamente relevantes y estadísticamente significativos.
- c) El fármaco X multiplica por 1,8 la probabilidad de supervivencia en comparación con el placebo. Estos resultados no son clínicamente relevantes pero sí estadísticamente significativos.
- d) El fármaco X multiplica por 1,8 la probabilidad de supervivencia en comparación con el placebo. Estos resultados no son clínicamente relevantes ni estadísticamente significativos.

La respuesta correcta es la opción a.

Se incrementa 1,8 veces (pues el valor del *hazard ratio* es 1,8) la probabilidad de supervivencia y no de muerte (pues el *endpoint* primario es la supervivencia). Los resultados son estadísticamente significativos, pues el IC95% no contiene el valor 1 y el valor de *p* es inferior a 0,05). En este caso, el incremento de la supervivencia 1,8 veces es algo relevante desde el punto de vista clínico (especialmente ante una enfermedad para la que no hay tratamiento).

La investigación clínica sobre vacunas: la vacuna frente a la COVID-19

Dr. José López Miranda
Servicio de Medicina Interna
Hospital Universitario Reina Sofía (Córdoba)

capítulo 8



1. Situación actual y contexto clínico

El desarrollo de una vacuna es un proceso largo y complejo que a menudo tarda de 10 a 15 años, e implica la participación combinada de organizaciones públicas y privadas. El sistema actual para crear, probar y regular vacunas se creó durante el siglo XX, cuando los grupos implicados estandarizaron sus procedimientos y reglamentos.

Historia, regulación y supervisión del gobierno

Estados Unidos

A finales del siglo XIX ya se habían creado varias vacunas para humanos, entre ellas las de la viruela, la rabia, la plaga, el cólera y la fiebre tifoidea. Sin embargo, no existía ninguna regulación sobre la producción de las vacunas. El 1 de julio de 1902, el Congreso de Estados Unidos aprobó una ley para regular “la venta de virus, sueros, toxinas y productos análogos”, que posteriormente se conoció como la Ley de Control de Productos Biológicos (aunque la expresión “productos biológicos” no aparece en el texto de la ley). Esta fue la primera legislación federal moderna para controlar la calidad de los medicamentos, y surgió, en parte, como una respuesta a los sucesos de contaminación de 1901 en San Luis y Camden, que afectaban a la vacuna contra la viruela y la antitoxina de la difteria. La ley creó el Laboratorio Higiénico del Servicio de Salud Pública de Estados Unidos para supervisar la fabricación de medicamentos biológicos, que posteriormente se convertiría en los Institutos Nacionales de Salud. La ley estableció el derecho por parte del gobierno para controlar los establecimientos donde se producían las vacunas.

La Ley de Servicio Público de Estados Unidos de 1944 ordenaba que el gobierno federal emitiera autorizaciones oficiales para los productos biológicos, incluidas las vacunas. Después de un accidente en 1954 con una vacuna contra el poliovirus (conocido como el incidente Cutter), se formó la División de Normas para Productos Biológicos (DBS, por sus siglas en inglés) con el fin de supervisar la seguridad y la regulación de las vacunas. Posteriormente, la DBS cambió de nombre a Oficina de Productos Biológicos, y se convirtió en parte de la Administración de Medicamentos y Alimentos; ahora es el Centro para la Evaluación e Investigación de Productos Biológicos.

Resto del mundo

En la Unión Europea, la Agencia Europea de Medicamentos supervisa la regulación de vacunas y otras medicinas, y un comité de la Organización Mundial de la Salud

(OMS) hace recomendaciones para los productos biológicos utilizados a nivel internacional; muchos países han adoptado las normas de la OMS.

2. Planteamiento, pregunta de investigación y objetivos

La enfermedad por coronavirus de 2019 (COVID-19) ha afectado a decenas de millones de personas en todo el mundo desde que la OMS la declaró una pandemia el 11 de marzo de 2020. Los adultos mayores, las personas con ciertas enfermedades crónicas coexistentes y los trabajadores de primera línea están en mayor riesgo en relación con esta enfermedad y sus complicaciones. Los datos recientes nos muestran además unas tasas crecientes de infección por SARS-CoV-2, el virus causante de la COVID-19, y del síndrome respiratorio agudo grave en otras poblaciones, incluidos los adultos más jóvenes. Por consiguiente, se necesitan con urgencia vacunas profilácticas seguras y eficaces para contener la pandemia, que ha tenido consecuencias médicas, económicas y sociales devastadoras.

En este capítulo mostraremos las fases del proceso de investigación para desarrollar una vacuna y nos centraremos en el estado actual del desarrollo de las vacunas contra la COVID-19.

3. Etapas en el desarrollo de vacunas

El desarrollo de las vacunas sigue un conjunto estandarizado de pasos. Las primeras etapas son de naturaleza exploratoria. La regulación y la supervisión aumentan conforme la vacuna candidata avanza en el proceso.

Primeros pasos: estudios de laboratorio y en animales

Etapas de exploración

Esta etapa incluye la investigación básica de laboratorio, y a menudo dura de 2 a 4 años. Los científicos identifican antígenos naturales o sintéticos que podrían ayudar a prevenir o tratar una enfermedad. Estos antígenos pueden ser partículas parecidas a virus, virus o bacterias debilitados, toxinas bacterianas debilitadas u otras sustancias derivadas de patógenos.

Etapas preclínicas

Los estudios preclínicos usan sistemas de cultivos de tejidos o cultivos de células y pruebas en animales, que pueden ser ratones o monos, para evaluar la seguridad de la vacuna candidata y su capacidad inmunógena, o su capacidad de provocar una

respuesta inmunológica. Estos estudios les dan una idea a los investigadores sobre las respuestas celulares que podrían esperar en los humanos, y también podrían sugerir una dosis inicial segura para la siguiente fase de la investigación, así como un método seguro para aplicar la vacuna.

Los investigadores pueden adaptar la vacuna candidata durante la fase preclínica para tratar de hacerla más eficaz. También pueden realizar estudios de exposición con animales, lo cual significa que se vacuna a los animales y luego se trata de infectarlos con el patógeno objetivo. Este tipo de estudios nunca se realizan en humanos.

Muchas vacunas candidatas no van más allá de esta etapa, ya que no pueden desencadenar la respuesta inmunológica deseada. A menudo, las etapas preclínicas duran de 1 a 2 años.

Siguientes pasos: estudios clínicos con humanos

Fase I de los ensayos con la vacuna

Este primer intento por evaluar la vacuna candidata en humanos involucra a un pequeño grupo de adultos, entre 20 y 80 por lo general. Si la vacuna está dirigida a los niños, los investigadores harán pruebas primero en adultos, y poco a poco reducirán la edad de los sujetos de prueba hasta llegar al objetivo. Los ensayos de fase I pueden ser no ciegos (conocidos también como estudios abiertos, pues los investigadores, y tal vez los sujetos incluidos en el ensayo clínico, saben si se usa una vacuna o un placebo).

Las metas de los estudios de fase I son evaluar la seguridad de la vacuna candidata y determinar el tipo y el alcance de la respuesta inmunitaria que provoca. Un ensayo prometedor de fase I avanzará a la siguiente etapa.

Fase II de los ensayos con la vacuna

Un grupo más grande de varios cientos de personas participa en los ensayos de fase II. Algunas de las personas que se incluyen en esta fase pueden pertenecer a grupos en riesgo de contraer la enfermedad; los ensayos son aleatorizados e incluyen a un grupo de placebo.

Las metas de los estudios de fase II son estudiar la vacuna candidata en cuanto a su seguridad, capacidad inmunogénica, dosis propuestas, programa de vacunación y método de aplicación.

Fase III de los ensayos con la vacuna

Las vacunas candidatas que tienen éxito en la fase II avanzan a ensayos con un mayor tamaño muestral, en los que participan de miles a decenas de miles de personas. Los estudios de fase III son aleatorizados y doble ciego, e incluyen el uso de la vacuna experimental, que se prueba contra un placebo (el placebo puede ser una solución salina, una vacuna para otra enfermedad o alguna otra sustancia).

Una meta de la fase III es evaluar la seguridad de la vacuna en un grupo grande de personas. Algunos efectos secundarios poco usuales podrían no ser evidentes en grupos más pequeños de personas que formaron parte de las fases anteriores. Por ejemplo, para que un efecto adverso de baja frecuencia relacionado con una vacuna candidata pudiera ocurrir en 1 de cada 10 000 personas y se pudiera detectar una diferencia significativa en el ensayo, este tendría que haber incluido 60 000 sujetos, la mitad de ellos en el grupo de control o en el grupo sin vacuna.

La eficacia de la vacuna también se prueba, lo cual puede incluir los siguientes factores:

1. ¿La vacuna candidata previene la enfermedad?
2. ¿Previene la infección por el patógeno?
3. ¿Conduce a la producción de anticuerpos u otros tipos de respuestas inmunitarias relacionadas con el patógeno?
4. ¿Previene las complicaciones graves derivadas de la infección?

Vigilancia posterior una vez emitida la autorización oficial

Diversos sistemas vigilan las vacunas después de que se hayan aprobado. Entre ellos se incluyen los ensayos de la fase IV, el Sistema de Información sobre Eventos Adversos a una Vacuna (Vaccine Adverse Event Reporting System) y el Enlace de Datos sobre la Seguridad de las Vacunas (Vaccine Safety Datalink).

Fase IV de los ensayos

Los ensayos de la fase IV son estudios opcionales que pueden realizar las compañías de medicamentos después de que se lanza una vacuna. El fabricante puede seguir realizando pruebas con la vacuna en cuanto a seguridad, eficacia y otros posibles usos o pautas de administración diferentes.

4. Estrategias en el desarrollo de vacunas frente al SARS-CoV-2

Se han dirigido muchos esfuerzos hacia el desarrollo de las vacunas contra el SARS-CoV-2 para acabar con la pandemia, y la mayoría de los candidatos a vacunas en desarrollo han utilizado la proteína S del virus. En la actualidad, el panorama mundial de la vacuna contra el SARS-CoV-2 incluye al menos 198 candidatos a vacunas, de los que 145 se encuentran en la etapa preclínica o exploratoria de su desarrollo. Las vacunas que se encuentran en investigación clínica y desarrollo se basan en vacunas inactivadas o virus vivos atenuados, subunidades de proteínas, partículas similares a virus, vector viral (replicante y no replicante), ADN, ARN, nanopartículas, etc., cada una con ventajas e inconvenientes (**tabla 8.1**). La distribución porcentual de los diferentes tipos de vacunas para la COVID-19 se representa en la **figura 8.1**.

Tabla 8.1. Ventajas e inconvenientes de los diferentes tipos de vacunas en frente a la COVID-19

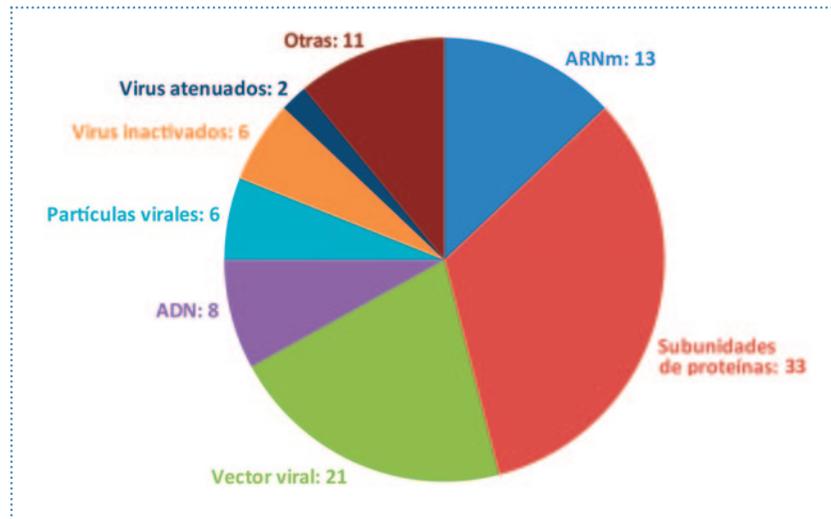
Tipo de vacuna	Ventajas	Inconvenientes
Virus atenuados	<ol style="list-style-type: none"> 1. Tiene la capacidad intrínseca de estimular el sistema inmunitario mediante la inducción de receptores tipo toll (TLR), a saber: TLR 3, TLR 7/8 y TLR 9 del sistema inmunitario innato, que involucra a las células B y T CD4 y CD8. 2. Puede derivarse de cepas de virus "adaptadas al frío", reordenamientos y genética inversa. 	<ol style="list-style-type: none"> 1. Requiere una amplia y extensa investigación previa para establecer la seguridad y la eficacia. 2. Existe una probabilidad de sustitución de nucleótidos durante la replicación viral, lo que resulta en la creación de recombinantes después de la vacunación.
Virus inactivados	<ol style="list-style-type: none"> 1. Estable y más segura en comparación con las vacunas de virus atenuados. 2. Cuenta con la tecnología e infraestructura preexistente requerida para su desarrollo. 3. Ya ha sido probado para la infección por SARS-CoV y otras enfermedades. 4. Se puede utilizar junto con adyuvantes para aumentar su inmunogenicidad. 	<ol style="list-style-type: none"> 1. Requiere inyecciones de refuerzo para mantener la inmunidad. 2. Además, se deben manipular grandes cantidades de virus y se debe mantener la integridad de las partículas inmunogénicas.

Tabla 8.1 continuación. Ventajas e inconvenientes de los diferentes tipos de vacunas en frente a la COVID-19

Tipo de vacuna	Ventajas	Inconvenientes
Subunidades de proteínas	<ol style="list-style-type: none"> 1. No tiene ningún componente vivo de la partícula viral. 2. Es segura y tiene menos efectos secundarios. 	<ol style="list-style-type: none"> 1. Induce una respuesta inmunitaria más débil. 2. La memoria para respuestas inmunitarias futuras es dudosa.
Basada en vectores virales	<ol style="list-style-type: none"> 1. Presenta una exposición de genes altamente específica en la célula huésped y una respuesta inmunitaria vigorosa. 2. Evita la manipulación de cualquier partícula infecciosa y se ha utilizado ampliamente para la infección por MERS-CoV con resultados positivos en los ensayos. 	<ol style="list-style-type: none"> 1. El huésped puede tener inmunidad contra el vector debido a una exposición previa, lo que reduce la eficacia. 2. Puede causar otros problemas de salud debido a la integración del genoma viral en el genoma del hospedador.
Vacunas de ADN	<ol style="list-style-type: none"> 1. El ADN sintético es estable a la temperatura ambiente y sin cadenas de frío. 2. Puede desarrollarse a un ritmo acelerado. 3. No requiere la manipulación de la partícula viral infecciosa. 	<ol style="list-style-type: none"> 1. Aunque produce inmunidad tanto citotóxica como humoral, los títulos permanecen bajos. 2. La inserción de ADN extraño en el genoma del huésped puede causar anomalías en la célula. 3. Puede inducir la producción de anticuerpos contra el huésped.
Vacunas de ARN	<ol style="list-style-type: none"> 1. La traducción del ARNm se produce en el citosol de la célula huésped y se evita el riesgo de cualquier tipo de integración en el genoma del huésped. 2. La proteína sintetizada mantiene la estructura tridimensional. 	<ol style="list-style-type: none"> 1. Se ha informado de problemas de seguridad con reactogenicidad para varias vacunas basadas en ARN. 2. También muestra inestabilidad. 3. Requiere condiciones de almacenamiento rigurosas.

Fuente: Kaur SP, Gupta V. COVID-19 Vaccine: A comprehensive status report. Virus Res. 2020;288:198114.

Figura 8.1. Tipos de vacunas en desarrollo para la COVID-19 (porcentajes sobre el total)



Fuente: Jeyanathan M, Afkhami S, Smaill F, Miller MS, Lichty BD, Xing Z. Immunological considerations for COVID-19 vaccine strategies. *Nat Rev Immunol.* 2020;20(10):615-32.

Para mejorar la inmunogenicidad se han utilizado diversas tecnologías adyuvantes, como AS03 (GSK), MF-59 (Novartis), CpG 1018 (Dynavax), etc. El análisis inmunoinformático también se utiliza para la identificación de epítopos candidatos a la vacuna. De esta forma, se pueden identificar epítopos citotóxicos significativos de células T y células B en las proteínas virales para seleccionar la mejor estrategia en su desarrollo.

Vacunas basadas en subunidades de proteínas virales

Una vacuna de subunidad es aquella que se basa en péptidos sintéticos o proteínas antigénicas recombinantes, que son necesarias para inducir la respuesta inmunitaria protectora o terapéutica de larga duración. Las vacunas de una de las subunidades, sin embargo, exhiben baja inmunogenicidad y requieren un soporte auxiliar de un adyuvante para potenciar las respuestas inmunitarias inducidas por la vacuna. La adición de un adyuvante puede mejorar la vida media biológica del material antigénico o puede mejorar la respuesta de las citocinas inmunomoduladoras.

La adición de un adyuvante, por lo tanto, ayuda a superar las deficiencias de las vacunas de subunidades de proteínas.

La proteína S del SARS-CoV-2 es el antígeno más adecuado para inducir los anticuerpos neutralizantes contra el patógeno. Esta proteína consta de dos subunidades. La subunidad S1 tiene los dominios NTD, RBD y RBM, mientras que la subunidad S2 comprende FP, HR 1 y 2. El virus entra en la célula a través de la endocitosis utilizando la unión mediada por la proteína S al receptor ACE2. Por consiguiente, la proteína S y sus fragmentos antigénicos son los principales objetivos para la producción de la vacuna de subunidad. La glicoproteína S es una proteína dinámica, que posee dos estados conformacionales, es decir, estado de prefusión y posfusión. Por lo tanto, el antígeno debe mantener su estructura química de superficie y el perfil de la proteína previa a la fusión original para preservar los epítopos y desencadenar respuestas de anticuerpos de buena calidad.

Vacunas basadas en vectores virales

Una vacuna basada en vectores virales es una solución preventiva prometedora contra un patógeno. Estas vacunas son altamente específicas a la hora de proporcionar genes a las células diana, y muy eficientes en la transducción de genes y la inducción de la respuesta inmunitaria. Ofrecen un alto nivel de expresión de proteínas antigénicas a largo plazo y, por lo tanto, tienen un gran potencial para uso preventivo, ya que activan e inducen las células T citotóxicas, lo que finalmente conduce a la eliminación de las células infectadas por el virus.

Vector de adenovirus recombinante tipo 5

Es un vector de adenovirus tipo 5 recombinante, defectuoso en la replicación (Ad5), que expresa la proteína S recombinante del SARS-CoV-2. Se preparó clonando un gen optimizado de longitud completa de la proteína S junto con el gen del péptido señal del activador del plasminógeno en el vector Ad5 desprovisto de los genes E1 y E3. Esta vacuna se construyó utilizando el sistema Admax de Microbix Biosystem. Los ensayos clínicos de fase I han establecido una respuesta de anticuerpos positiva o seroconversión. Se observó un aumento de cuatro veces en los anticuerpos neutralizantes específicos de la proteína RBD y S dentro de los 14 días posteriores a la inmunización y alcanzó su punto máximo el día 28 después de la vacunación. Además, la respuesta de las células T CD4+ y CD8+ alcanzó su punto máximo el día 14 después de la vacunación. Sin embargo, la inmunidad anti-Ad5 preexistente limitaba en parte tanto las respuestas de anticuerpos como de células T.

Coroflu

M2SR es una versión autolimitante del virus de la gripe, que se modifica mediante la inserción de la secuencia del gen SARS-CoV-2 de la proteína S. Además, la vacuna expresa la proteína hemaglutinina del virus de la gripe, lo que induce una respuesta inmunitaria contra ambos virus. El M2SR es autolimitante y no se replica porque carece del gen M2. Puede ingresar a la célula, induciendo así la inmunidad contra el virus. Se administrará por vía intranasal, imitando la ruta natural de la infección viral. Esta ruta activa varios modos del sistema inmunitario y tiene una mayor inmunogenicidad en comparación con las inyecciones intramusculares.

Vacuna LV-SMENP-DC

La vacuna LV-SMENP-DC se prepara modificando las células dendríticas con el vector lentiviral que expresa los dominios conservados de las proteínas estructurales del SARS-CoV-2 y la proteasa utilizando los minígenos SMENP. La inoculación subcutánea de la vacuna presenta los antígenos en las células presentadoras de antígeno, que finalmente activan las células T citotóxicas e inician la respuesta inmunitaria.

Vacuna de adenovirus recombinante ChAdOx1

La vacuna de adenovirus recombinante ChAdOx1 se desarrolló utilizando la glicoproteína S de codón optimizado y se sintetizó con la secuencia del activador de plasminógeno tisular (tPA) en el extremo 5'. El genoma del vector de adenovirus se construye en el cromosoma artificial bacteriano insertando el gen SARS-CoV-2 S en el locus E1 del genoma del adenovirus ChAdOx1. La ausencia de cualquier ARN subgenómico en los animales vacunados por vía intramuscular de los ensayos preclínicos es indicativa de la inmunidad inducida contra el virus. Los estudios previos han sugerido que una sola dosis debería dirigir la respuesta inmunitaria, aunque recientemente se ha demostrado que una segunda dosis administrada 56 días después de la primera mejora y aumenta dicha respuesta.

Recientemente se han publicado los resultados preliminares de cuatro ensayos clínicos controlados, aleatorizados y ciegos en fase 3 sobre la eficacia de esta vacuna que se están realizando en el Reino Unido, Brasil y Sudáfrica. Los participantes de 18 años o más fueron asignados al azar (1:1) a la vacuna ChAdOx1 nCoV-19 o al grupo de control (vacuna antimeningocócica conjugada de los grupos A, C, W e Y o solución salina). Los participantes en el grupo ChAdOx1 nCoV-19 recibieron dos dosis que contenían 5×10^{10} partículas virales (dosis estándar, cohorte SD/SD); y un

subgrupo del ensayo del Reino Unido recibió media dosis como primera dosis (dosis baja) y una dosis estándar como segunda dosis (cohorte LD/SD). El análisis de eficacia principal incluyó la presentación de una forma de COVID-19 sintomática en participantes seronegativos con un test positivo en la prueba de PCR 14 días después de una segunda dosis de vacuna. Los participantes fueron analizados de acuerdo con el tratamiento recibido, con un corte de datos el 4 de noviembre de 2020. Entre el 23 de abril y el 4 de noviembre de 2020, se incluyeron 23 848 participantes en total, y 11 636 participantes (7548 en el Reino Unido, 4088 en Brasil) en el análisis de eficacia primario provisional. En los participantes que recibieron dos dosis estándar, la eficacia de la vacuna fue del 62,1 %, y en los participantes que recibieron una dosis baja seguida de una dosis estándar, la eficacia fue del 90 %. La eficacia global de la vacuna en ambos grupos fue del 70,4 %. A partir de los 21 días tras la administración de la primera dosis, hubo diez casos hospitalizados por COVID-19, todos en el grupo de control; dos fueron clasificados como COVID-19 grave, uno de los cuales murió. En el periodo de seguimiento (74 341 meses-persona de seguimiento de seguridad; mediana de 3,4 meses por persona) ocurrieron 175 eventos adversos graves en 168 participantes, 84 eventos en el grupo ChAdOx1 nCoV-19 y 91 en el grupo de control. Tres eventos se clasificaron como posiblemente relacionados con una vacuna: uno en el grupo ChAdOx1 nCoV-19, uno en el grupo de control y uno en un participante que permanece oculto a la asignación del grupo.

Vacunas ARNm

Las vacunas basadas en ARNm constituyen una plataforma emergente, no infecciosa y no integradora con casi ningún riesgo potencial de mutagénesis insercional. Actualmente, se están estudiando el ARN no replicante y los ARN autorreplicantes derivados de virus. La inmunogenicidad del ARNm se puede minimizar y se pueden realizar alteraciones para aumentar la estabilidad de estas vacunas. Además, la inmunidad antivector también se evita, ya que el ARNm es el vector genético mínimamente inmunogénico, lo que permite la administración repetida de la vacuna. Esta plataforma ha potenciado el programa de desarrollo rápido de vacunas debido a su flexibilidad y capacidad para imitar la estructura y expresión del antígeno, como se ve en el curso de una infección natural.

Vacuna ARNm-1273 (Moderna TX)

Es una vacuna compuesta de ARNm sintético encapsulado en nanopartículas lipídicas que codifica la proteína espiga (S) estabilizada previa a la fusión de longitud completa del SARS-CoV-2. Tiene el potencial de provocar una respuesta antiviral

altamente específica de la proteína S. Además, se considera relativamente segura, ya que no está compuesta por el patógeno inactivado ni por las subunidades del patógeno vivo. La vacuna obtuvo una aprobación rápida de la FDA para realizar los ensayos de fase II. La compañía ha publicado los datos de anticuerpos de fase I provisionales de ocho participantes que recibieron varios niveles de dosis. Los participantes del grupo de dosis de 25 µg dieron resultados comparables a los del suero de convalecientes, mientras que en los participantes que recibieron la dosis de 100 µg las concentraciones de anticuerpos neutralizantes esencialmente superaron las encontradas en el suero de convalecientes. Se encontró que la vacuna era predominantemente segura y bien tolerada en las cohortes de dosis de 25 µg y 100 µg, mientras que tres participantes experimentaron síntomas sistémicos de grado 3 después de la administración de la segunda dosis en los niveles de dosis de 250 µg.

Recientemente se han publicado los resultados de un ensayo clínico con esta vacuna de fase 3 aleatorizado y controlado con placebo. Se incluyó a voluntarios con alto riesgo de infección por SARS-CoV-2 o sus complicaciones, que se asignaron al azar en una proporción de 1:1 para recibir dos inyecciones intramusculares de ARNm-1273 (100 µg) o placebo con 28 días de diferencia. El criterio de valoración principal fue la prevención de la enfermedad por SARS-CoV-2 con inicio al menos 14 días después de la segunda inyección en participantes que no habían sido infectados previamente por el virus. En este estudio se incluyó a 30 420 voluntarios que fueron asignados al azar en una proporción de 1:1 para recibir la vacuna o el placebo (15 210 participantes en cada grupo). La enfermedad sintomática por SARS-CoV-2 se confirmó en 185 participantes del grupo placebo (56,5 por 1000 personas-año) y en 11 participantes del grupo que recibió la vacuna de ARNm-1273 (3,3 por 1000 personas-año); la eficacia de la vacuna fue del 94,1%. La presencia de COVID-19 grave ocurrió en 30 participantes, con una muerte; los 30 estaban en el grupo placebo. Los eventos adversos graves fueron raros y la incidencia fue similar en los dos grupos.

BNT162b1 (BioNTech-FosunPharma-Pfizer)

BNT162b1 es una vacuna de ARNm de codón optimizado que codifica el dominio de unión al receptor trimerizado del SARS-CoV-2. La vacuna presenta un aumento de la inmunogenicidad debido a la adición del dominio de trimerización del pliegue derivado de la fibritina T4 al antígeno del dominio de unión al receptor. El ARNm está encapsulado en nanopartículas de lípidos catiónicos ionizables de 80 nm, lo que asegura una administración eficiente. Los ensayos clínicos de fase 1/2 han

revelado valores elevados de anticuerpos IgG específicos, con una concentración media geométrica de hasta entre 8 y 46,3 veces el título del suero de pacientes convalecientes.

Se han publicado recientemente los resultados del gran ensayo clínico de eficacia en fase 3 de esta vacuna. En este ensayo de eficacia multinacional y en curso, comparado con placebo y ciego al observador, se distribuyó aleatoriamente a personas de 16 años de edad o mayores en una proporción de 1:1 para recibir dos dosis, con 21 días de diferencia, de placebo o de la vacuna candidata BNT162b2 (30 µg por dosis). BNT162b2 es una vacuna de ARN modificada con nucleósidos, formulada con nanopartículas lipídicas, que codifica la proteína S de longitud completa del SARS-CoV-2. Los criterios de valoración principales fueron la eficacia de la vacuna contra la COVID-19 confirmada por laboratorio y la seguridad. Un total de 43 548 participantes fueron aleatorizados, de los cuales 43 448 recibieron inyecciones: 21 720 con BNT162b2 y 21 728 con placebo. Hubo 8 casos de COVID-19 con inicio al menos 7 días después de la segunda dosis entre los participantes asignados a la administración de BNT162b2 y 162 casos entre los asignados al placebo; la vacuna tuvo una eficacia del 95% en la prevención de la COVID-19 (intervalo de confianza del 95%: 90,3-97,6). Se observó una eficacia similar de la vacuna (generalmente, del 90% al 100%) en los subgrupos definidos por la edad, el sexo, la raza, la etnia, el índice de masa corporal inicial y la presencia de enfermedades coexistentes. Entre los 10 casos de COVID-19 grave que comenzaron después de la primera dosis, 9 ocurrieron en receptores de placebo y 1 en receptores de BNT162b2. El perfil de seguridad de BNT162b2 se caracterizó por la presencia de dolor a corto plazo de leve a moderado en el lugar de la inyección, fatiga y dolor de cabeza. La incidencia de eventos adversos graves fue baja y similar en los grupos de vacuna y placebo.

Vacunas de ADN

El enfoque más revolucionario de la vacunación es la introducción de la vacuna de ADN, que codifica el antígeno y un adyuvante que induce la respuesta inmunitaria adaptativa. Las células transfectadas expresan el transgén, que proporciona un suministro constante de proteínas específicas del transgén, que es bastante similar al virus vivo. Además, el material antigénico es endocitado por las células dendríticas inmaduras que, en última instancia, presentan el antígeno a las células T CD4+ y CD8+ en asociación con los antígenos MHC 2 y MHC 1 en la superficie celular, estimulando así respuestas inmunitarias humorales eficaces y mediadas por células.

Vacunas de virus vivos atenuados

Esta vacuna (DeJNS1-SARS-CoV2-RBD, Universidad de Hong Kong) consiste en una cepa de vacuna basada en la de la gripe con una delección en el gen NS1. Se reorganiza para expresar el dominio de unión al receptor de la proteína S del SARS-CoV-2 en su superficie y se cultiva en el embrión del pollo o células de riñón canino de Madin Darby. Es potencialmente más inmunogénico que el virus de la gripe de tipo salvaje y puede administrarse como un aerosol nasal.

Otros tipos de vacunas

El descubrimiento de la estructura y el genoma del SARS-CoV-2 ha conducido al rápido desarrollo de varias vacunas candidatas con potencial inmunogenicidad, pero también reacciones adversas. El proceso del desarrollo de vacunas es largo y complicado, lo que requiere su evaluación en ensayos clínicos de larga duración. Varias empresas biotecnológicas están utilizando diferentes tecnologías para el desarrollo de sus vacunas candidatas; British and American Tobacco Company (BAT) desplegó recientemente la vacuna COVID-19 utilizando su nueva tecnología de plantas de tabaco de rápido crecimiento, mientras que la Universidad de Tianjin ha desarrollado una vacuna oral que ha empleado con éxito *Saccharomyces cerevisiae* para producir la proteína S. El estado GRAS (generalmente considerado como seguro) de la levadura proporciona alta escalabilidad, solidez y producción rentable de las dosis necesarias para combatir esta pandemia. Además, estudios *in silico*, utilizando varias bases de datos como VaxiJen, han revelado que las secuencias de epítotos WTAGAAAYY y YDPLQPEL pueden emplearse para la formulación de vacunas peptídicas basadas en otros epítotos más inmunogénicos.

5. Discusión y conclusiones

La epidemia de COVID-19 continúa creciendo, especialmente en países que no han podido o no han querido instituir fuertes medidas de salud pública. El regreso a la normalidad depende cada vez más del éxito de las vacunas para prevenir la enfermedad y limitar la propagación de la infección. Sin embargo, esta esperanza se había visto atenuada hasta hace pocas semanas por varias incógnitas. No se había demostrado que las vacunas existentes fueran eficaces contra la infección por ningún betacoronavirus, la familia que incluye el SARS-CoV-2. La epidemia de síndrome respiratorio agudo grave (SARS) causada por otro betacoronavirus en 2003 terminó por sí sola antes de que se emprendieran serios esfuerzos en el desarrollo de vacunas, y el número bastante pequeño de casos de síndrome respiratorio de

Oriente Medio (MERS) aún no ha justificado el esfuerzo y la inversión a gran escala necesarios para determinar si los candidatos a vacunas preclínicas son eficaces. Además, las estrategias para aumentar la velocidad del desarrollo de vacunas solo han tenido pruebas limitadas. Por otra parte, un número relativamente pequeño de personas habían recibido vacunas con vectores de adenovirus y aún no se habían probado en la población general los efectos a largo plazo de vacunas basadas en tecnologías de ARNm. ¿Serán estos nuevos productos eficaces y seguros? Esta es la pregunta que toda la comunidad científica se plantea.

En la actualidad, y gracias a un proceso de investigación clínica sin precedentes en un corto espacio de tiempo (menos de un año), comenzamos a tener parte de la respuesta y esta es muy alentadora y prometedora. Las dos vacunas basadas en un ARN modificado que codifica una versión de la proteína S del SARS-CoV-2 y que contiene mutaciones que bloquean la proteína en una conformación que puede inducir respuestas de anticuerpos neutralizantes han demostrado que inducen inmunidad humoral y celular y que estas respuestas protegen frente al desarrollo de la enfermedad sintomática con una eficacia superior al 95%, hecho de extraordinaria relevancia e importancia desde un punto de vista clínico y epidemiológico. Los resultados de los dos grandes ensayos clínicos en fase III publicados hasta la fecha son realmente impresionantes, con prevención extraordinaria no solo del desarrollo de infección y enfermedad sintomática, sino del desarrollo de formas graves de la enfermedad.

Esto supone un verdadero triunfo de la ciencia y de la investigación clínica. La mayoría de las vacunas han tardado décadas en desarrollarse, pero es probable que estas vacunas contra la COVID-19 pasen de la concepción a la implementación a gran escala en menos de un año. La secuencia del virus que condujo al desarrollo de la secuencia de ARN viral específica requerida para diseñar la vacuna no se conoció hasta que fue determinada y ampliamente difundida por el Centro Chino para el Control y la Prevención de Enfermedades en enero de 2020. Y todo este éxito es gracias a todos: a los científicos que compartieron datos y que desarrollaron los métodos utilizados para implementar y crear las vacunas, a los investigadores clínicos que realizaron un trabajo de alta calidad en el contexto de una emergencia de salud, a los miles de participantes que se ofrecieron como voluntarios para los ensayos clínicos, y a los gobiernos que ayudaron a crear estándares de desempeño para acelerar todo el proceso de creación de la vacuna. Y todo esto sirve como modelo para las muchas otras vacunas contra la COVID-19 actualmente en desarrollo, algunas de las cuales ya han completado sus ensayos de fase I-II. Sin embargo, aún quedan preguntas importantes por responder. Solo menos de 40 000 personas han

recibido estas vacunas en ensayos clínicos a corto plazo. ¿Surgirán problemas de seguridad inesperados cuando el número aumente a millones y posiblemente a miles de millones de personas a más largo plazo? ¿Aparecerán efectos secundarios con un seguimiento más prolongado? Implementar una vacuna que requiera dos dosis es un desafío. ¿Cuánto tiempo seguirá siendo eficaz la vacuna? ¿La vacuna previene la enfermedad asintomática y limita la transmisión? ¿Y qué pasa con los grupos de personas que no estuvieron representados en estos ensayos, como niños, mujeres embarazadas y pacientes inmunodeprimidos de varios tipos? Todas estas preguntas necesitan respuesta y para ello contamos con todas las herramientas que la ciencia moderna nos ofrece en la actualidad y, sobre todo, con la investigación clínica de calidad.

No obstante, el importante y relevante nivel de seguridad y eficacia que la vacuna ha demostrado hasta ahora hace que este sea un problema que deberíamos acoger con satisfacción. Lo que parece ser un éxito espectacular para la vacunación es cómo se puede traducir en salvar incontables vidas humanas y darnos un camino para salir de lo que ha sido un desastre a escala mundial.

Bibliografía

Baden LR, El Sahly HM, Essink B, et al. Efficacy and safety of the mRNA-1273 SARS-CoV-2 vaccine. *N Engl J Med*. 2020 Dec 30;NEJMoa2035389.

Corey L, Mascola JR, Fauci AS, Collins FS. A strategic approach to COVID-19 vaccine R&D. *Science*. 2020;368(6494):948-50.

Graham BS. Rapid COVID-19 vaccine development. *Science*. 2020;368(6494):945-6.

Jackson LA, Anderson EJ, Roupael NG, et al. An mRNA vaccine against SARS-CoV-2 - preliminary report. *N Engl J Med*. 2020;383(20):1920-31.

Jeyanathan M, Afkhami S, Smaill F, Miller MS, Lichty BD, Xing Z. Immunological considerations for COVID-19 vaccine strategies. *Nat Rev Immunol*. 2020;20(10):615-32.

Kaur SP, Gupta V. COVID-19 Vaccine: A comprehensive status report. *Virus Res*. 2020;288:198114.

Lipsitch M, Dean NE. Understanding COVID-19 vaccine efficacy. *Science*. 2020;370(6518):763-5.

Polack FP, Thomas SJ, Kitchin N, et al. Safety and efficacy of the BNT162b2 mRNA Covid-19 vaccine. *N Engl J Med*. 2020;383(27):2603-15.

Ramasamy MN, Minassian AM, Ewer KJ, et al. Safety and immunogenicity of ChAdOx1 nCoV-19 vaccine administered in a prime-boost regimen in young and old adults (COV002): a single-blind, randomised, controlled, phase 2/3 trial. *Lancet*. 2021;396(10267):1979-93.

Thanh Le T, Andreadakis Z, Kumar A, et al. The COVID-19 vaccine development landscape. *Nat Rev Drug Discov*. 2020;19(5):305-6.

Voysey M, Clemens SAC, Madhi SA, et al. Safety and efficacy of the ChAdOx1 nCoV-19 vaccine (AZD1222) against SARS-CoV-2: an interim analysis of four randomised controlled trials in Brazil, South Africa, and the UK. *Lancet*. 2021;397:99-111.

Wiersinga WJ, Rhodes A, Cheng AC, Peacock SJ, Prescott HC. Pathophysiology, transmission, diagnosis, and treatment of coronavirus disease 2019 (COVID-19): A Review. *JAMA*. 2020;324(8):782-93.

Zhu FC, Li YH, Guan XH, et al. Safety, tolerability, and immunogenicity of a recombinant adenovirus type-5 vectored COVID-19 vaccine: a dose-escalation, open-label, non-randomised, first-in-human trial. *Lancet*. 2020;395(10240):1845-54.

Preguntas de autoevaluación

1. Señale la afirmación incorrecta en relación con los estudios en fase III con vacunas:

- a) Se realizan en todas las vacunas independientemente de los resultados en fase II.
- b) Incluyen a miles de decenas de pacientes candidatos a la vacuna.
- c) Son estudios aleatorizados y doble ciego.
- d) Su objetivo principal suele ser determinar la eficacia y la seguridad de la vacuna.

La respuesta correcta es la opción a.

De las cuatro opciones de respuesta que se presentan, solo la a es incorrecta. Esto se basa en que solo las vacunas candidatas que tienen éxito en la fase II avanzan a ensayos con un mayor tamaño muestral, en los que participan de miles a decenas de miles de personas. Los estudios de fase III son aleatorizados y doble ciego, e incluyen el uso de la vacuna experimental, que se prueba contra un placebo (el placebo puede ser una solución salina, una vacuna para otra enfermedad o alguna otra sustancia). Una meta de la fase III es evaluar la seguridad de la vacuna en un grupo grande de personas.

2. Señale la afirmación incorrecta en relación con la vacuna de adenovirus recombinante ChAdOx1:

- a) La vacuna de adenovirus recombinante ChAdOx1 se desarrolló utilizando la glicoproteína S.
- b) La eficacia global de la vacuna en los grupos de tratamiento del ensayo clínico en fase III fue del 70,4%.
- c) La eficacia de los voluntarios que recibieron como primera dosis una dosis baja y como segunda una dosis estándar fue inferior a la eficacia que se observó en los voluntarios que recibieron ambas dosis estándar.
- d) En el ensayo clínico en fase III de esta vacuna, a partir de los 21 días tras la administración de la primera dosis hubo diez casos hospitalizados por COVID-19, todos en el grupo de control.

La respuesta correcta es la opción c.

Recientemente se han publicado los resultados preliminares de cuatro ensayos clínicos controlados, aleatorizados y ciegos en fase 3 sobre la eficacia de esta vacuna que se están realizando en el Reino Unido, Brasil y Sudáfrica. En los participantes que recibieron dos dosis estándar, la efi-

ca de la vacuna fue del 62,1%; en los que recibieron una dosis baja seguida de una dosis estándar, del 90%.

Por otra parte, la eficacia global de la vacuna en ambos grupos fue del 70,4%. A partir de los 21 días tras la administración de la primera dosis, hubo diez casos hospitalizados por COVID-19, todos en el grupo de control; dos fueron clasificados como COVID-19 grave, uno de los cuales murió.

Esta vacuna se desarrolló utilizando la glicoproteína S de codón optimizado y se sintetizó con la secuencia del activador de plasminógeno tisular (tPA) en el extremo 5'.

3. Señale la afirmación incorrecta en relación con la vacuna BNT162b1 de Pfizer-BioNTech:

- a) Es una vacuna basada en ARNm.
- b) Su administración debe ser intramuscular y está basada en una formulación de nanopartículas lipídicas.
- c) La incidencia de eventos adversos graves con esta vacuna es baja y similar a la del grupo de placebo.
- d) La eficacia global de la vacuna es del 85%.

La respuesta correcta es la opción d.

BNT162b1 es una vacuna de ARNm de codón optimizado que codifica el dominio de unión al receptor trimerizado del SARS-CoV-2. El ARNm está encapsulado en nanopartículas de lípidos catiónicos ionizables de 80 nm, lo que asegura una administración eficiente.

En el ensayo clínico de eficacia en fase 3 multinacional, comparado con placebo y ciego al observador, se distribuyó aleatoriamente a personas de 16 años de edad o mayores en una proporción de 1:1 para recibir dos dosis, con 21 días de diferencia, de placebo o de la vacuna candidata BNT162b2 (30 µg por dosis). Hubo 8 casos de COVID-19 con inicio al menos 7 días después de la segunda dosis entre los participantes asignados a la administración de BNT162b2 y 162 casos entre los asignados al placebo; la vacuna tuvo una eficacia del 95% en la prevención de la COVID-19.

El perfil de seguridad de BNT162b2 se caracterizó por la presencia de dolor a corto plazo de leve a moderado en el lugar de la inyección, fatiga y dolor de cabeza. La incidencia de eventos adversos graves fue baja y similar en los grupos de vacuna y placebo.

Resolver un problema clínico mediante un metaanálisis tradicional

Dra. Rosa Solà Alberich

*Servicio de Medicina Interna
Hospital Universitari Sant Joan de Reus (Reus, Tarragona)

*Dra. Lucia Tarro, Dra. Elisabet Llauredó y Dña. Maria de la Serra Besora



capítulo 9

1. Situación actual y contexto clínico

La sarcopenia es una patología multifactorial definida por una pérdida de masa muscular, fuerza muscular y función física. Dicha pérdida de masa muscular en las extremidades inferiores y de fuerza muscular se estima del 1-5% por año en mayores de 50 años. La prevalencia general de sarcopenia en Europa en personas mayores se situaba, aunque los criterios de definición aún no están claros, en un 11,1% en 2016 y se espera un incremento progresivo. Además, cabe destacar que la sarcopenia está infradiagnosticada. Así, diagnosticar y tratar a tiempo la sarcopenia permitiría reducir sus consecuencias, como el riesgo de caídas, fracturas y mortalidad en personas mayores.

En el abordaje de la sarcopenia las intervenciones basadas en la nutrición y la actividad física tienen un papel importante. En concreto, del componente dietético, la vitamina E, por su poder antioxidante y antiinflamatorio, puede producir una mejora de la disfunción y participar en la regeneración muscular, con un efecto positivo sobre la sarcopenia; pero se considera que se necesitan más investigaciones en esta línea. En este escenario, será interesante recopilar de forma sistematizada la información científica sobre el efecto de la vitamina E sobre la sarcopenia y, si es posible, cuantificar el tamaño del efecto mediante un metaanálisis.

2. Planteamiento, pregunta de investigación y objetivos

Dado que existe una carencia de conocimiento sobre el tratamiento de la sarcopenia, nos planteamos estudiar si la vitamina E puede ser un punto clave. La hipótesis del estudio es que la suplementación de vitamina E se asocia con una reducción de la sarcopenia en ancianos.

El objetivo principal del estudio es determinar si la suplementación con vitamina E reduce la sarcopenia en personas mayores. Los objetivos secundarios son evaluar la relación entre la suplementación de vitamina E y los parámetros de sarcopenia (masa, fuerza o función muscular).

3. Tipo de estudio y diseño. Pruebas estadísticas y abordaje del problema

Para alcanzar el objetivo sobre los efectos se plantea una revisión sistemática y posteriormente, si se obtienen datos sobre dichos efectos, se realizará un metaanálisis de estudios controlados aleatorizados (ECA) como mínimo con dos estudios sobre una de las variables. A continuación, se detalla de forma estructurada la metodología del metaanálisis.

Registro del protocolo

El primer paso es registrar el protocolo en PROSPERO (Centre for Reviews and Dissemination del National Institute for Health Research del Reino Unido), donde se genera un número de registro. El registro quiere evitar que haya revisiones duplicadas y reduce la posibilidad de sesgos, y se debe realizar antes del inicio del metaanálisis. El número de registro en PROSPERO se debe incluir en la descripción de un metaanálisis.

Criterios PRISMA

En la elaboración de revisiones sistemáticas y metaanálisis se recomienda la utilización de los criterios PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-analysis), que son 27 criterios para facilitar la redacción de este tipo de estudios.

Características de los estudios

Seguidamente, se definen los criterios de inclusión y exclusión de los estudios a partir de los criterios PICOS (*population, intervention, comparison, outcomes, study design*). En nuestro caso, pueden ser los que se muestran a continuación (tabla 9.1):

- **Criterios de inclusión:** ECA sobre suplementación de vitamina E y sarcopenia (masa, fuerza o función muscular), con una población diana de 65 años o más con sarcopenia, escritos en inglés y publicados entre 2005 y 2020. En este caso, los ECA son el tipo de estudio elegido, ya que son los estudios más indi-

Tabla 9.1. Criterios de inclusión y exclusión de los estudios según los criterios PICOS

Criterios PICOS	Traslación al ámbito de la investigación
Población (<i>population</i>)	Personas mayores (>65 años) con sarcopenia
Intervención (<i>intervention</i>)	Intervención nutricional basada en la suplementación de vitamina E
Comparación (<i>comparison</i>)	Grupo de control (con placebo o sin intervención)
Resultados (<i>outcomes</i>)	Efecto de la suplementación de vitamina E sobre la sarcopenia
Diseño del estudio (<i>study design</i>)	Estudio controlado aleatorizado

cados para evaluar la efectividad de un tratamiento o una intervención en general. La elección del inicio del periodo de publicación de los artículos, 2005, se corresponde con el año a partir del cual se identifican la mayor parte de los artículos sobre sarcopenia.

- **Criterios de exclusión:** ECA que incluyan otras patologías aparte de la sarcopenia, así como los estudios que no cumplan con los criterios de inclusión.

Bases de datos científicas y estrategia de búsqueda

Hay que elegir las bases de datos de publicaciones científicas y definir la estrategia de búsqueda. En el ámbito clínico las bases de datos más utilizadas son Medline/PubMed®, Scopus, Cochrane Library y Web of Science, entre otras. Es aconsejable utilizar más de una base de datos para la búsqueda de artículos científicos.

Posteriormente, se debe conformar la estrategia de búsqueda eligiendo correctamente las palabras clave que se utilizarán (a partir de los criterios PICOS identificados previamente) basándose en los términos MeSH (*medical subject headings*). Por ejemplo, se podrían utilizar los siguientes: *sarcopenia, muscle mass, muscle strength, physical performance, vitamin E, vitamin E supplementation y elderly*.

Además, para la elección de los artículos se pueden aplicar filtros, como pueden ser el tipo de estudio, la edad de la población, la fecha de publicación del estudio, el idioma o si se trata de estudios en humanos o en animales. El hecho de aplicar estos filtros reduce el número de resultados y a la vez focaliza la búsqueda. Asimismo, es importante no excederse en la aplicación de filtros, ya que podría conllevar la pérdida de artículos científicos que serían de utilidad.

A modo de ejemplo, en el caso de la sarcopenia, escogeríamos las bases de datos Medline/PubMed® y Cochrane Library. La estrategia de búsqueda podría ser la que se muestra a continuación:

Estrategias de búsqueda

```
((((sarcopenia[MeSH Terms]) OR (muscle strength[MeSH Terms])) OR (muscle mass[Title/Abstract]) OR (physical performance[Title/Abstract])) AND ((vitamin e[Title/Abstract]) OR (vitamin e supplementation[Title/Abstract])) AND (elderly[MeSH Terms]) Filters: Randomized Controlled Trial, From 2005 to 2020, Humans
```

Selección de los estudios

A partir de los estudios obtenidos en las distintas bases de datos, se deben seleccionar los que deberán incluirse en el metaanálisis.

El proceso de selección de los estudios puede ser extenso, y para facilitar dicho proceso hay diferentes programas informáticos que ayudan en la gestión de los artículos de un metaanálisis, como, por ejemplo, Covidence o Rayyan QCRI.

La selección de estudios se presenta con el diagrama de flujo PRISMA (**figura 9.1**), que permite esquematizar todas las fases del metaanálisis. Este diagrama engloba varios aspectos:

- a) El total de artículos extraídos a partir de las bases de datos.
- b) El proceso de selección con el total de artículos incluidos y excluidos.
- c) Las causas de exclusión de los artículos.
- d) El número total de artículos que constan tanto en la revisión sistemática (evaluación cualitativa) como en el metaanálisis (evaluación cuantitativa).

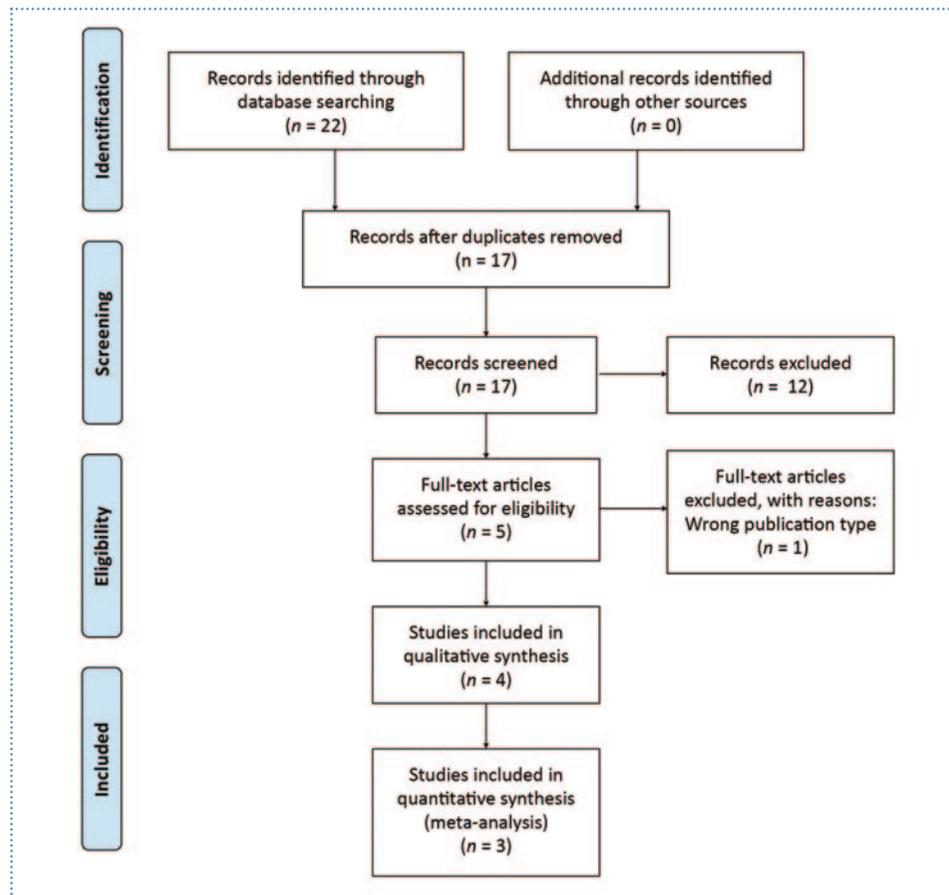
Además, el diagrama se compone de varias fases:

1. **Eliminación de artículos duplicados:** Hay que eliminar las referencias duplicadas en las diferentes bases de datos. Algunos programas tienen un sistema para eliminar los duplicados de forma automática, pero, como otros programas no disponen de este sistema, se requiere del uso de un gestor de referencias bibliográficas como Mendeley, Reference Manager, EndNote o Refworks.
2. **Selección por título:** Se debe realizar una primera selección basándose en el título y las palabras clave que se hayan definido previamente. Los estudios que en el título no contengan las palabras clave o similares deberán eliminarse.
3. **Selección por resumen:** En los artículos seleccionados por título se analizarán los resúmenes a partir de las palabras clave para comprobar si se cumplen los criterios de inclusión y exclusión. Los estudios que no cumplan estas características deben eliminarse.
4. **Revisión a texto completo:** Por último, se revisará el texto completo de los artículos seleccionados por su resumen, para determinar si cumplen con los criterios de inclusión y exclusión y con los criterios PICOS. Los estudios que

no cumplan con los criterios de inclusión y con los criterios PICOS se eliminarán y se deberá justificar el motivo de exclusión: diseño del estudio, tipo de intervención, resultados no comparables o población de estudio, por ejemplo.

Los estudios que hayan pasado todas las fases y cumplan con todos los criterios se incluirán en el estudio para la elaboración del metaanálisis.

Figura 9.1. Diagrama de flujo PRISMA



Fuente: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA Statement. PLoS Med. 2009;6(7):e1000097. doi:10.1371/journal.pmed1000097.

Extracción de datos de los estudios incluidos

A continuación, a partir de los estudios incluidos se realizará la extracción de datos. Se aconseja construir una base de datos en formato tabla con toda la información relevante de cada artículo incluido.

En la tabla, siguiendo con el ejemplo de la sarcopenia y la suplementación de vitamina E, se recogerá la información de las siguientes variables:

- Primer autor.
- Año de publicación.
- País de desarrollo del estudio.
- Diseño del estudio.
- Duración del estudio.
- Número de participantes.
- Edad y género de los participantes.
- Método de valoración de la sarcopenia.
- Variables de la sarcopenia evaluadas (masa, fuerza y función muscular).
- Tipo de intervención (dosis de vitamina E suplementada y duración de la intervención).
- Tipo de grupo de control (con o sin placebo).
- Resultados (diferencias pre- y postintervención y diferencias entre el grupo de intervención y el grupo de control).

Calidad de los artículos

De los artículos incluidos, se evaluará su calidad para reducir el riesgo de sesgo. De acuerdo con el ejemplo basado en los ECA, una de las herramientas recomendadas para evaluar la calidad de este tipo de estudios es la de Cochrane-RoB 2. Esta herramienta permite evaluar la calidad de los artículos a partir de la identificación de los riesgos de sesgo derivados del proceso en cinco dominios: aleatorización, intervención prevista, falta de datos en los resultados, medición de los resultados y selección del resultado.

Análisis estadístico, metaanálisis

El metaanálisis se realizará después de la extracción de datos y el proceso es el siguiente:

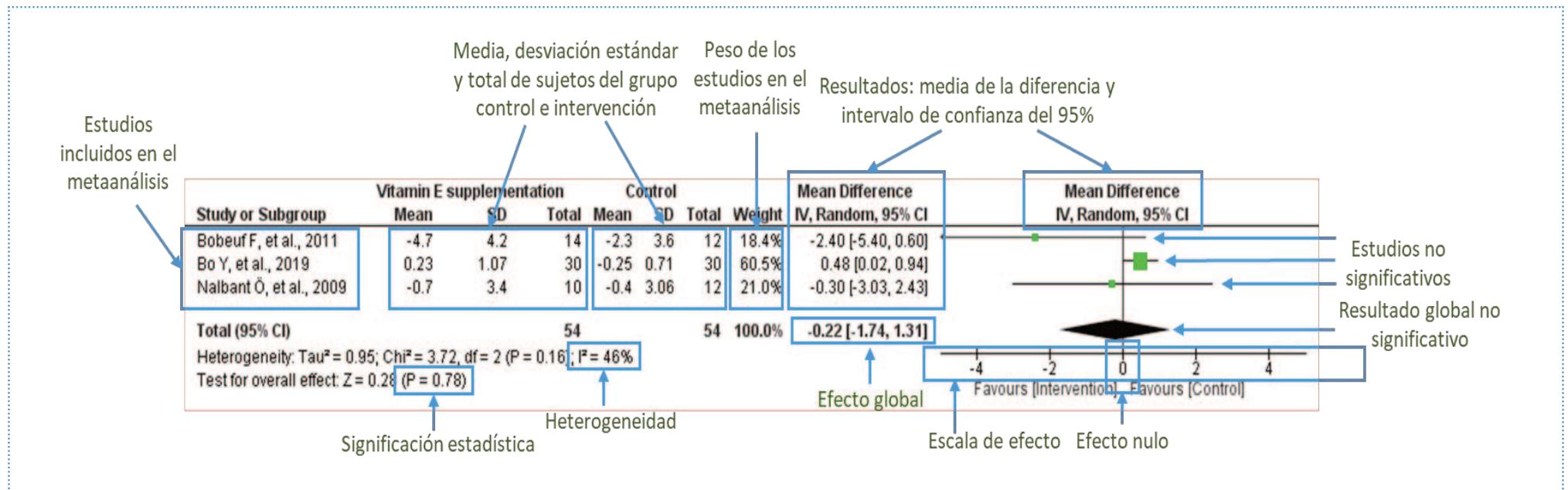
1. De todos los estudios incluidos, se deben seleccionar los que evalúen la variable principal con el mismo método o herramienta.
2. Es necesario un mínimo de dos estudios para la realización de un metaanálisis, aunque cuantos más estudios se puedan incluir, mejor. Volviendo a nuestro ejemplo, de todos los estudios incluidos sobre sarcopenia y suplementación de vitamina E, solo tres podrán incluirse en el metaanálisis, ya que son los únicos que evalúan la sarcopenia mediante la masa muscular.
3. Para el análisis estadístico se recomienda el uso del programa RevMan (Cochrane). En él se insertarán los datos numéricos de todos los estudios incluidos en el metaanálisis para que se genere un diagrama de efectos o *forest plot*.

4. Para realizar el *forest plot* (figura 9.2) se plantea el siguiente proceso:

- a) En la primera columna, incluir el apellido del primer autor, la letra inicial de su nombre y el año de publicación, lo que facilita la identificación del estudio. Ejemplo: Bobeuf F *et al.* 2011, Nalbant Ö *et al.* 2009, Bo Y *et al.* 2019.
- b) Identificar si la variable principal es una variable continua o categórica. En el ejemplo de la sarcopenia, la variable de la masa muscular es una variable continua. Así, se necesitará para hacer el cálculo mediante el *forest plot* la diferencia de la media de masa muscular y la desviación estándar entre el final y el inicio del estudio.

En algunos estudios estas características de la variable principal se detallan en las tablas de resultados, como es el caso de Bo Y *et al.* 2019. En cambio, en otros estudios no se muestra el valor de la diferencia y se debe calcular.

Figura 9.2. *Forest plot* de la relación entre la masa muscular y la suplementación de vitamina E en ancianos con sarcopenia



- c) Para calcular las diferencias cuando no aparecen directamente en el artículo publicado, el programa RevMan dispone de una calculadora que determina de forma automática la media de la diferencia y la respectiva desviación estándar a partir de otros valores, como el valor de p de la diferencia, los intervalos de confianza de la variable principal, etc.
- d) Otro aspecto importante es la población total en el grupo de intervención y en el grupo de control. La población incluida es uno de los factores que determinará el peso de los resultados de un estudio sobre el resultado final del metaanálisis. En la **figura 9.2** se muestran unos cuadrados verdes que reflejan cada estudio, que son más o menos grandes en relación con el peso del estudio sobre el resultado final del metaanálisis.
- e) Si las desviaciones estándar de los estudios (que se indican con las líneas que salen de los cuadrados) no tocan la línea vertical que refleja el 0 (sinónimo de efecto nulo), informan de que el estudio tiene un efecto significativo. Además, dependiendo de si el cuadrado está en la derecha o en la izquierda en relación con la línea vertical, se puede determinar si el tratamiento favorece al grupo de intervención o al grupo de control.

Para interpretar los resultados del *forest plot* se consideran distintos aspectos:

- a) La heterogeneidad muestra las diferencias metodológicas que hay entre estudios y, por lo tanto, si estos son comparables entre sí. Es importante identificar si hay diferencias en los tratamientos, los pacientes incluidos y cómo se definieron las variables de resultados.
 - Por ejemplo, Bo Y *et al.* 2019 basa la intervención en un suplemento rico en proteína, vitamina D y vitamina E.
 - En cambio, Bobeuf F *et al.* 2011 combina la suplementación de vitamina C y E a la vez que el ejercicio de resistencia.
 - Por otra parte, Nalbant Ö *et al.* 2009 se centra únicamente en la vitamina E como tratamiento juntamente con el ejercicio aeróbico.
 - Por lo tanto, podemos ver que estos estudios son heterogéneos en cuanto al tratamiento.
- b) Es muy importante que, si los estudios tienen poca heterogeneidad entre sí, utilicemos los efectos fijos (una opción dentro del programa RevMan) para el

cálculo del *forest plot*. En cambio, si los estudios tienen heterogeneidad, utilizaremos los efectos aleatorios.

- c) La heterogeneidad se valora mediante el estadístico I^2 que aparece en el *forest plot*, y se clasifica en baja (25%), moderada (50%) o alta (75%). No obstante, lo más importante es la valoración que hace el investigador. En el ejemplo con el que estamos trabajando (**figura 9.1**), la heterogeneidad es del 46% (moderada), pero lo más importante es que ya se ha identificado que los tres estudios tienen diferencias importantes en el tratamiento.
- d) La valoración de la sensibilidad determina si el metaanálisis es fiable; se puede verificar con alguno de los siguientes métodos:
 - La eliminación en el *forest plot* de los estudios uno a uno para corroborar que los resultados no cambian drásticamente.
 - La eliminación de los estudios de poca calidad para comprobar si los resultados del *forest plot* siguen manteniéndose.
- e) Finalmente, un valor de p inferior a 0,05 nos muestra que el metaanálisis es estadísticamente significativo. En el *forest plot* de ejemplo, el valor de p es 0,78; por lo tanto, las intervenciones basadas en la suplementación de vitamina E no presentan un efecto significativo en la mejora de la masa muscular.

4. Justificación y discusión

El metaanálisis, que se identifica como el estudio con un grado de evidencia mayor, tiene diferentes niveles de evidencia científica según la clasificación de la Scottish Intercollegiate Guidelines Network (SIGN) (1-, 1+ o 1++). Los metaanálisis que se clasifican como evidencia 1++ son los que incluyen ensayos clínicos de alta calidad y con pocos sesgos. Son uno de los requisitos de la recomendación A, que están en la base de la traslación a las recomendaciones clínicas, es decir, solo los metaanálisis clasificados como 1++ podrán derivar al establecimiento de recomendaciones clínicas. Es importante tener en cuenta que se pueden realizar metaanálisis de distintos tipos de estudios, pero los que están basados en ECA son los de mayor evidencia para estimar la efectividad de una intervención.

Como se observa en el ejemplo del metaanálisis sobre efectos de intervenciones basadas en la suplementación de vitamina E para la mejora de la sarcopenia, se puede concluir que la vitamina E, ya sea sola o con otros tratamientos añadidos, no muestra un efecto significativo sobre la mejora de la masa muscular.

En las limitaciones de los metaanálisis destaca que si no hay un número elevado de artículos incluidos los resultados pueden ser confusos y de utilidad dudosa. Otra de las limitaciones es que la mayoría de los artículos con resultados negativos o no significativos no se publican (sesgo de publicación) y limitan la información de los metaanálisis. También es una limitación abarcar estudios de baja calidad, ya que no permiten extrapolar los resultados a una recomendación clínica.

5. Conclusiones

Los metaanálisis facilitan la cuantificación del efecto de un tratamiento sobre una variable concreta que deriva en demostrar o no su efectividad a partir de la síntesis de la evidencia publicada. Como resultado, en especial de los metaanálisis de alta calidad, la evaluación de la efectividad puede ser de gran ayuda para la toma de decisiones de los profesionales clínicos para una mejor atención a los pacientes.

Bibliografía

Bloom I, Shand C, Cooper C, Robinson S, Baird J. Diet quality and sarcopenia in older adults: A systematic review. *Nutrients*. 2018;10(3).

Chung E, Mo H, Wang S, et al. Potential roles of vitamin E in age-related changes in skeletal muscle health. *Nutr Res*. 2018;49:23-36.

Covidence [Internet]. Melbourne (Australia): Veritas Health Innovation; [consultado: 23 de diciembre de 2020]. Disponible en: <https://www.covidence.org>.

Cruz-Jentoft AJ, Bahat G, Bauer J, et al. Sarcopenia: revised European consensus on definition and diagnosis. *Age Ageing*. 2019;48(1):16-31.

Ethgen O, Beaudart C, Buckinx F, Bruyère O, Reginster JY. The future prevalence of sarcopenia in Europe: a claim for public health action. *Calcif Tissue Int*. 2017;100(3):229-34.

Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *Br Med J*. 2003;327(7414):557-60.

Hutton B, Salanti G, Caldwell DM, et al. The PRISMA extension statement for reporting of systematic reviews incorporating network meta-analyses of health care interventions: Checklist and explanations. *Ann Intern Med*. 2015;162(11):777-84.

Keller K, Engelhardt M. Strength and muscle mass loss with aging process. *Age and strength loss. Muscles Ligaments Tendons J*. 2013;3(4):346-50.

Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. *Syst Rev*. 2016;5(1):210.

Scottish Intercollegiate Guidelines Network. SIGN 50: A guideline developers' handbook. Edimburgo: Scottish Intercollegiate Guidelines Network; 2011 [SIGN publication 50]. Disponible en: https://www.sign.ac.uk/assets/sign50_2011.pdf.

Sterne JAC, Savović J, Page MJ, et al. RoB 2: A revised tool for assessing risk of bias in randomised trials. *BMJ*. 2019;366:l4898.

Preguntas de autoevaluación

1. En relación con la realización de un metaanálisis, ¿cuál de las siguientes afirmaciones es correcta?

- a) Es importante definir los criterios de inclusión y exclusión de los artículos y las características antes de empezar la búsqueda.
- b) Se tiene que hacer la búsqueda antes de definir las características de los artículos.
- c) Para poder realizar un metaanálisis hay que utilizar las guías internacionales CONSORT.
- d) Es necesario leer el texto completo de todos los artículos encontrados en la búsqueda para saber si cumplen o no los criterios de inclusión.

La respuesta correcta es la opción a.

Los criterios de inclusión y exclusión de los artículos se tienen que diseñar antes de realizar la búsqueda, siguiendo los criterios PICOS.

2. Con respecto a la heterogeneidad de los estudios incluidos en un metaanálisis, señale la afirmación correcta:

- a) Solo se tiene que considerar la variable I^2 .
- b) Se tiene que considerar la variable I^2 , pero el investigador tiene que tener en cuenta también si hay diferencias en los tratamientos, los pacientes incluidos y la definición de las variables.
- c) Solo se tienen que considerar las diferencias en los tratamientos, los pacientes incluidos y la definición de las variables.
- d) Un metaanálisis con un 80% de heterogeneidad muestra que hay mucha diferencia de resultados entre los estudios incluidos.

La respuesta correcta es la opción b.

Es importante considerar la I^2 , ya que nos da una aproximación de la heterogeneidad entre estudios, pero a la vez es fundamental también que el investigador considere si hay diferencias metodológicas entre los estudios incluidos, ya que esto no se puede observar mediante la I^2 .

3. Para poder generar recomendaciones clínicas a partir de estudios científicos es necesario...

- a) Un metaanálisis de nivel de evidencia A con ensayos clínicos de alta calidad.
- b) Un metaanálisis de nivel de evidencia 1+ con ensayos clínicos de alta calidad y pocos sesgos.
- c) Un metaanálisis de nivel de evidencia 1++ con ensayos clínicos de alta calidad y pocos sesgos.
- d) Un metaanálisis es el tipo de estudio con mayor grado de evidencia, por lo tanto, siempre tiene una fuerza de recomendación A y se pueden derivar de él recomendaciones clínicas.

La respuesta correcta es la opción c.

El nivel de evidencia es 1++ y no grado A, porque grado A es el nivel de recomendación.

Resolver un problema clínico mediante un metaanálisis en red

Dra. María Montserrat Chimeno Viñas

Servicio de Medicina Interna

Hospital Universitario Virgen de la Concha (Zamora)

Vicepresidenta segunda de la Sociedad Española de Medicina Interna (SEMI)

capítulo 10



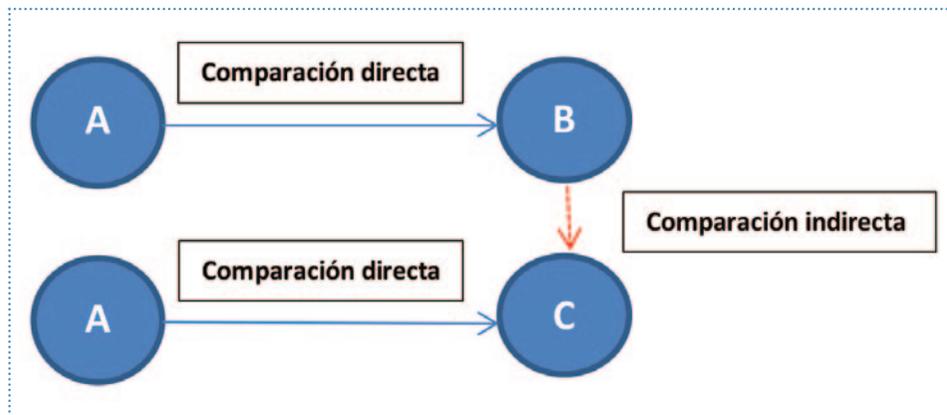
1. Estado del arte y contexto clínico

En el momento actual, disponemos de múltiples herramientas para la toma de decisiones, evidentemente, siempre basadas en la mejor evidencia disponible. Las revisiones sistemáticas y los metaanálisis proporcionan una ayuda fundamental para la práctica asistencial y la salud pública, pero presentan un inconveniente importante: no permiten comparar múltiples alternativas al mismo tiempo. Además, mucha evidencia proviene de comparaciones realizadas contra placebo, lo que resulta menos útil cuando se debe recomendar el tratamiento más adecuado.

Para solventar estos inconvenientes se han desarrollado técnicas de metaanálisis en red (MAR), también conocidos como metaanálisis con comparaciones múltiples o metaanálisis con comparaciones mixtas. Permiten obtener estimaciones de los efectos relativos de las diferentes alternativas a partir de comparaciones directas (si las hubiera) o indirectas, es decir, a partir de comparaciones mixtas.

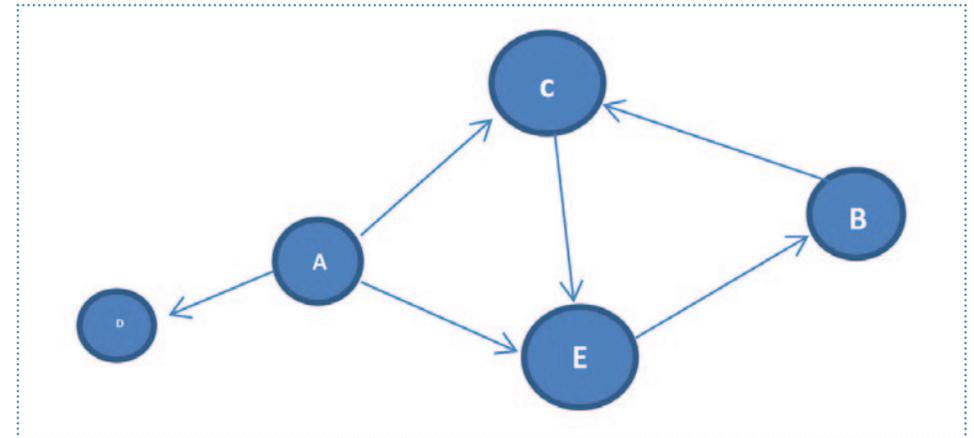
Se habla de comparación directa cuando dos tratamientos se han evaluado el uno frente al otro (A frente a B o A frente a C), y de comparación indirecta si dos tratamientos nunca se han comparado uno frente al otro directamente (B frente a C), pero estos dos tratamientos se han comparado frente a un comparador común (A, placebo). Se puede realizar así una comparación indirecta entre los tratamientos utilizando los efectos relativos de los dos tratamientos frente al comparador común (figura 10.1).

Figura 10.1. Comparación directa frente a comparación indirecta



La realización de múltiples comparaciones directas e indirectas es lo que da lugar al MAR. Este representa de forma gráfica la red de las diferentes comparaciones entre las distintas intervenciones estudiadas, con la obtención de información de las múltiples comparaciones directas, indirectas y mixtas. Para representar cada tratamiento se utilizan nodos y se dibuja una línea entre los nodos para los cuales existen comparaciones directas (figura 10.2).

Figura 10.2. Esquema de metaanálisis en red



2. Planteamiento, pregunta de investigación y objetivos

Responder a preguntas como “¿Cuál es el tratamiento que mejor controla el dolor entre los que habitualmente se pautan para la neuralgia del trigémino?”, “¿Cuál es el grupo farmacológico de antihipertensivos más eficaz en la insuficiencia cardiaca?” o “¿Son más eficaces los inhibidores del cotransportador de sodio y glucosa 2 (iSGLT-2) que la metformina en el tratamiento complementario con la insulina en la diabetes mellitus tipo 1?” no es posible con los metaanálisis convencionales que evalúan la eficacia o seguridad de un tratamiento (o grupo) en comparación con un único comparador.

Si nos planteamos “¿Cuál sería la estatina, sola o en combinación con ezetimiba, y la dosis de esta más eficaz en el enfermo con insuficiencia renal crónica?”, inicialmente debemos definir que se entenderá como más eficaz aquel tratamiento

que demuestre ser capaz de asociarse a una menor probabilidad de sufrir un nuevo evento cardiovascular y que consideraremos insuficiencia renal la existencia de un filtrado glomerular inferior a 60 mL/min/1,73 m². Nuestro objetivo será el análisis de todas las pruebas respecto al impacto del tratamiento con diferentes estatinas solas o con ezetimiba en el riesgo cardiovascular del enfermo renal.

3. Tipo y diseño de estudio. Pruebas estadísticas a aplicar y analizar

Evidentemente, no es posible responder nuestra pregunta con un metaanálisis tradicional, dado que, además de necesitar la realización de una estimación numérica de diversas comparaciones indirectas o mixtas, precisamos proponer una jerarquía de intervenciones. Para ello, debemos plantearnos un MAR. Este nos permite evaluar el efecto de cada estatina, sola o en combinación con ezetimiba, en los individuos con enfermedad renal crónica.

El concepto de MAR fue propuesto por Lumley en el año 2002. Proporcionó una herramienta estadística para la comparación de múltiples tratamientos, pero hasta el año 2005 la cantidad de MAR fue escasa.

El MAR, también denominado metaanálisis con comparaciones múltiples o metaanálisis con comparaciones mixtas, es una técnica estadística que combina estudios experimentales con similares características y que no han sido comparados directamente mediante la aplicación de una red de análisis.

Los pasos que se deben seguir en la elaboración de un metaanálisis en red son los siguientes:

- Selección de los estudios.
- Extracción de la información relevante.
- Aplicación de los métodos estadísticos para obtener las comparaciones indirectas y para combinar los resultados de los estudios seleccionados.

Para mejorar la calidad y la consistencia de las revisiones sistemáticas y los metaanálisis tradicionales se diseñó la guía PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analysis), y tras el desarrollo de los MAR se amplió. La declaración PRISMA-NMA (Network Meta-Analysis) incorporó nuevos conceptos y terminologías.

La mayoría de los MAR utilizan un enfoque estadístico único; el más frecuente es el modelo bayesiano (se evalúa la relación de causa y efecto). Existen diferentes programas para el análisis de los diferentes modelos y su representación gráfica (WinBUGS, SAS, STATA...). El manejo de estos programas requiere amplios conocimientos estadísticos.

Los MAR basan sus estimaciones en dos asunciones o supuestos específicos: la transitividad (similitud) y la coherencia (consistencia). El término *transitividad* se refiere a que asumimos que podemos comparar dos tratamientos a través de un tercero en común. La otra gran asunción es la coherencia o "consistencia", que expresa el grado de acuerdo entre la evidencia directa e indirecta, lo cual es fundamental para decidir combinar o no esa evidencia mixta dentro de un MAR. La coherencia solo puede evaluarse para todas las comparaciones que cuentan con ambas estimaciones de efecto (directa e indirecta).

La metodología del MAR permite la estimación numérica de diversas comparaciones indirectas o mixtas, explorar sesgos y proponer una jerarquía o clasificación de intervenciones de acuerdo con la probabilidad de ser el mejor tratamiento. Sin embargo, hay que utilizarlos con cautela porque pueden ser engañosos. Por ejemplo, la intervención que encabeza esa clasificación de eficacia no necesariamente es la que se basa en la mejor calidad de la evidencia.

Para responder a nuestra pregunta se realizó una revisión bibliográfica en la que se seleccionaron los ensayos clínicos aleatorizados que estudiaron el tratamiento con estatinas con o sin ezetimiba en los pacientes con enfermedad renal crónica. El desenlace estudiado fueron los eventos cardiovasculares adversos mayores, entre los que figuraban todos los eventos coronarios fatales y no mortales, los procedimientos de revascularización de todo tipo y todos los eventos cerebrovasculares, incluidos los accidentes isquémicos transitorios.

Previamente al análisis matemático de la información de los ensayos incluidos, se evaluaron los sesgos en ellos con la herramienta de la colaboración Cochrane diseñada para este fin.

El efecto combinado correspondiente a los esquemas de tratamiento con estatinas solas o con ezetimiba se presenta en *odds ratio* (OR), con sus correspondientes intervalos de credibilidad al 95 % (ICr95%) para los cálculos a nivel bayesiano en red.

Se estimaron la heterogeneidad (variabilidad entre estudios) y el sesgo de publicación y se clasificaron las pautas de tratamiento con estatinas solas o con ezetimiba

mediante el cálculo de la superficie bajo la clasificación acumulativa (*surface under the cumulative ranking*, SUCRA).

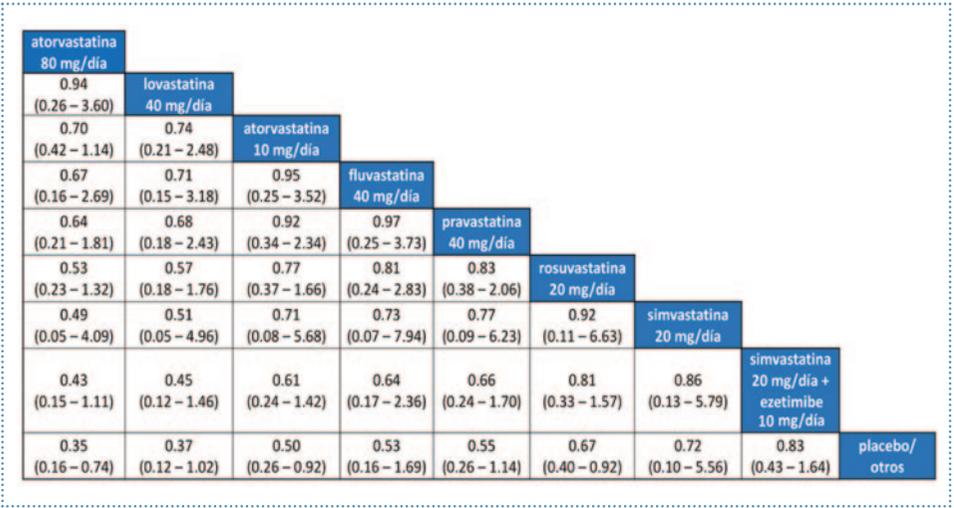
Se seleccionaron, finalmente, 22 ensayos controlados aleatorizados. Las exclusiones se produjeron, en su mayoría, al identificar estudios en los que no se evaluaban resultados o desenlaces relacionados con la enfermedad cardiovascular, o estudios que incluyeron sujetos sin enfermedad renal crónica afectados por otras patologías.

Para el análisis de la información se utilizó WinBUGS. Este programa procesa los cálculos, que NetMetaXL presenta luego en la interfaz de Windows.

NetMetaXL proporciona el siguiente gráfico (figura 10.3) con los valores de OR y los intervalos de confianza al 95% correspondientes a las pautas de tratamiento con estos fármacos presentes en los ensayos incluidos en el análisis.

En la figura 10.3 se muestra que solo tres estatinas son realmente eficaces en relación con la disminución del número de eventos cardiovasculares adversos mayores en esta población: atorvastatina 80 mg/día, atorvastatina 10 mg/día y rosuvastatina 20 mg/día.

Figura 10.3. Tabla de ligas con el efecto combinado de cada estatina con o sin ezetimiba



Las intervenciones y exposiciones estudiadas en el MAR se clasificaron de acuerdo al método de la SUCRA después de verificar la convergencia y la inconsistencia entre las comparaciones generadas.

Tras calcular la SUCRA para cada una de las pautas de tratamiento con estatinas solas o con ezetimiba, atorvastatina 80 mg/día, atorvastatina 10 mg/día y rosuvastatina 20 mg/día se clasifican como las mejores respecto al resto (tabla 10.1). No obstante, se observa una marcada diferencia entre tratar y no tratar, incluso considerando separadamente del placebo los otros fármacos utilizados para tratar la dislipemia en el enfermo renal descritos en los estudios incluidos.

Es necesario señalar que el efecto combinado de una pauta (dosis) determinada de tratamiento con estatinas solas o con ezetimiba se refleja mejor en la clasificac-

Tabla 10.1. Clasificación de las pautas de estatinas con y sin ezetimiba en pacientes con insuficiencia renal según la SUCRA

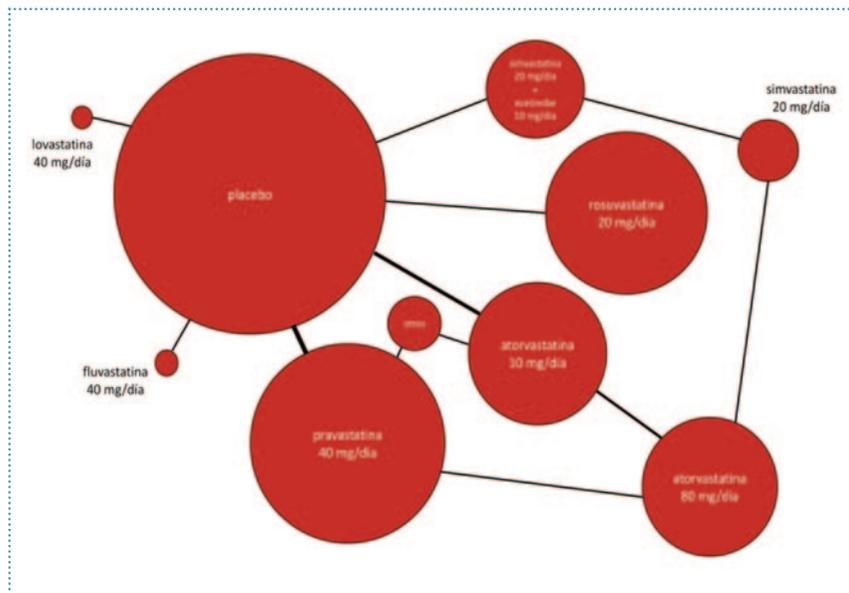
Pautas con estatinas con o sin ezetimiba	SUCRA
Atorvastatina 80 mg/día	0,8318
Atorvastatina 10 mg/día	0,8182
Rosuvastatina 20 mg/día	0,7799
Simvastatina 20 mg/día	0,6205
Simvastatina 20 mg/día con ezetimiba 10 mg/día	0,6110
Lovastatina 40 mg/día	0,4786
Fluvastatina 40 mg/día	0,3832
Pravastatina 40 mg/día	0,2597
Otros tratamientos	0,1984
Placebo	0,0187

SUCRA: superficie bajo la clasificación acumulativa.

ción por el valor de la SUCRA cuanto mayor era el número de participantes que recibieron esa pauta. Este hecho lleva a una interpretación cautelosa de la información generada (en el caso de llevarse a cabo más estudios, los resultados de este trabajo podrían variar).

A continuación se muestra la geometría de la red bayesiana generada para este análisis con las combinaciones directas de los estudios que se consideraron y la densidad de pruebas por combinación: los nexos de unión son más gruesos cuando la cantidad de estudios en una combinación dada es mayor, y los círculos de mayor tamaño cuando la cantidad de datos numéricos disponibles por cada pauta analizada es mayor (**figura 10.4**).

Figura 10.4. Red bayesiana para el efecto del tratamiento con estatinas con o sin ezetimiba en la insuficiencia renal



Evidentemente, una de las limitaciones de este estudio (y de otros muchos trabajos) es el sesgo de publicación. La heterogeneidad en los cálculos es, también, un punto débil, aunque inherente a este tipo de trabajos.

4. Justificación y discusión

Evidentemente, el tipo de pregunta que planteamos no puede contestarse con un metaanálisis pareado. Pretendíamos comparar el efecto de las distintas estatinas, solas o en combinación con ezetimiba, y las diferentes dosis de estas en los pacientes con insuficiencia renal crónica (filtrado glomerular < 60 mL/min/1,73 m²) para poder determinar una jerarquía en ellas.

Las técnicas de MAR permiten evaluar el efecto de cada fármaco en la población de interés (en este caso, individuos con enfermedad renal crónica). Los análisis se llevan a cabo en contextos bayesianos (es decir, se evalúa la relación causa y efecto) incluso si alguno de los fármacos no se ha estudiado de manera directa en los trabajos seleccionados tras una exhaustiva búsqueda bibliográfica.

Los estudios realizados con estatinas son muy numerosos, la mayoría usando como comparador el placebo. Son menos numerosos los trabajos publicados en pacientes con enfermedad renal crónica, una patología cada vez más prevalente y que precisa recomendaciones específicas. En el MAR expuesto de manera sucinta se analizan las pruebas de mejor calidad disponibles para el tratamiento de la hipercolesterolemia en el enfermo renal.

Es necesario subrayar algunos aspectos. La validez de los resultados y las conclusiones de los MAR, al igual que sucede en los metaanálisis convencionales, dependerán de la calidad de los estudios seleccionados, de la variabilidad entre estudios y de que los sesgos de información sean mínimos. Respecto a este último punto, el sesgo de información cobra una importancia especial en los MAR, ya que la falta de información para un tratamiento concreto podría afectar al cálculo de la probabilidad de que los distintos tratamientos sean mejores o peores.

Los resultados deben interpretarse siempre con cautela en los MAR, dado que la intervención que encabeza la clasificación de eficacia no necesariamente se basa en la mejor calidad de la evidencia.

5. Conclusiones

El MAR permite obtener estimaciones de los efectos de diferentes intervenciones a partir de comparaciones directas (si las hubiera) e indirectas y proponer una jerarquía de intervenciones. Se deben realizar en el marco de búsquedas bibliográficas rigurosas que garanticen la calidad metodológica de los estudios seleccionados, y precisan el uso de programas complejos para el análisis de la información cuyo manejo requiere amplios conocimientos estadísticos.

Tras una búsqueda exhaustiva bibliográfica, un análisis estadístico riguroso, una valoración de la heterogeneidad, la transitividad y los sesgos posibles, la conclusión de nuestro MAR establece como mejores pautas de tratamiento con estatinas, solas o en combinación con ezetimiba, en el enfermo renal crónico las siguientes: atorvastatina 80 mg/día, atorvastatina 10 mg/día y rosuvastatina 20 mg/día. No obstante, los resultados de los MAR deben interpretarse con prudencia; en ocasiones, la red de evidencia es débil (existen pocos estudios para algunas comparaciones o son de escasa calidad o de pequeño tamaño).

Bibliografía

Catalá-López F, Tobías A, Roqué M. Conceptos básicos del metaanálisis en red. *Aten Primaria*. 2014;46:573-81.

Fau C, Nabzo S, Nasabun V. Metaanálisis en red. *Rev Mex Oftalmol*. 2018; 92(3):153-9.

Herrera-Gomez F, Chimeno MM, Martin-Garcia D, et al. Cholesterol-lowering treatment in chronic kidney disease: multistage pairwise and network meta-analyses. *Sci Rep*. 2019;9:8951.

Hutton B, Salanti G, Caldwell DM, et al. The PRISMA extension statement for reporting of systematic reviews incorporating network meta-analyses of health care interventions: checklist and explanations. *Ann Intern Med*. 2015;162:777-84.

Lumley T. Network meta-analysis for indirect treatment comparisons. *Stat Med*. 2002;21(16):2313-24.

Mills EJ, Thorlund K, Ioannidis JP. Demystifying trial networks and network meta-analysis. *BMJ*. 2013;346:f2914.

Molina Arias M. El metaanálisis en red. *Rev Pediatr Aten Primaria*. 2016;18:279-82.

Preguntas de autoevaluación

1. Señale la afirmación correcta entre las siguientes:

- a) El concepto de metaanálisis en red fue propuesto en el año 1992.
- b) Aunque a menudo se argumenta que las comparaciones indirectas solo son necesarias cuando las comparaciones directas no existen, es importante tener en cuenta que los resultados de pruebas indirectas en combinación con las directas pueden reforzar la evaluación entre tratamientos.
- c) La comparación indirecta está sujeta a menos sesgos potenciales que la comparación directa.
- d) El metaanálisis pareado permite comparar múltiples alternativas terapéuticas al mismo tiempo.

La respuesta correcta es la opción *b*.

El concepto de metaanálisis en red fue propuesto por Lumley en el año 2002. Evidentemente, las comparaciones indirectas son complementarias a las evaluaciones directas cuando estas existen, aunque no las sustituyan en ningún caso. Obviamente, la comparación indirecta está sujeta a más sesgos potenciales que la comparación directa. Es el metaanálisis en red el que permite comparar múltiples alternativas terapéuticas al mismo tiempo.

2. ¿Cuál de los siguientes términos no se utiliza en los metaanálisis en red?

- a) Heterogeneidad.
- b) Análisis bayesiano.
- c) Intervalo de confianza.
- d) Comparación directa.

La respuesta correcta es la opción *c*.

La terminología utilizada en los metaanálisis en red es muy específica. Heterogeneidad (presencia de variabilidad entre los efectos de los tra-

tamientos que se combinan en un metaanálisis), homogeneidad (ausencia de heterogeneidad), comparación directa (evaluación entre dos tratamientos que se han comparado directamente el uno frente al otro), análisis bayesiano (emplea el conocimiento previo en el análisis de los datos de modo que, si se conoce la probabilidad de que ocurra un suceso, su valor será modificado cuando se disponga de esa información externa), etc., son términos que habitualmente se utilizan en los metaanálisis en red.

El intervalo de confianza en los análisis bayesianos se denomina intervalo de credibilidad.

3. Ante un paciente con un filtrado glomerular de 48 mL/min/1,73 m², ¿cuál de las siguientes opciones de tratamiento sería mejor según las evidencias del metaanálisis expuesto en el capítulo?

- a) Lovastatina 40 mg/día.
- b) Atorvastatina 10 mg/día con ezetimiba 10 mg/día.
- c) Atorvastatina 10 mg/día.
- d) Rosuvastatina 10 mg/día.

La respuesta correcta es la opción *c*.

Teniendo en cuenta los resultados del metaanálisis en red expuesto, atorvastatina 80 mg/día, atorvastatina 10 mg/día y rosuvastatina 20 mg/día se clasifican como las mejores pautas de tratamiento respecto al resto.



Patrocinado por:

