

# ¿QUÉ PONGO EN EL APARTADO DE METODOLOGÍA DE MI ARTÍCULO CIENTÍFICO?

Para profesionales de la salud que hacen  
investigación clínica y epidemiológica

Laura Muñoz Ortiz



# ÍNDICE

Introducción .....	2
Diseño del estudio .....	5
Pacientes o Población de estudio .....	5
Criterios de inclusión/exclusión .....	6
Selección de la muestra .....	8
Variables de estudio .....	9
Recogida de los datos o procedimiento .....	14
Análisis estadístico o Plan de análisis .....	16
Errores habituales .....	20

## ¿Qué pongo en el apartado de metodología de mi artículo científico?

Esta es una pregunta muy común que me hacen muchos investigadores.

Y no es de extrañar porque no sé si lo sabías, pero **la metodología de un artículo científico es el apartado que los revisores de la revista evalúan con mayor detalle durante el proceso de revisión, y también es el más vulnerable para el rechazo de un manuscrito.**

Si la metodología no está correcta podemos decir que tu artículo tiene muchos números para ser rechazado.

Por eso quiero compartir contigo unas pautas para ayudarte a estructurar de manera correcta este apartado tan importante.

En este ebook **te doy indicaciones sobre qué deberías escribir en cada uno de los apartados que conforman la metodología de un artículo científico, y te muestro algunos ejemplos de artículos publicados.**

Así tendrás una buena base a partir de la cual podrás ir eliminando algún apartado que no aplique en tu estudio, o unir otros porque te sea más fácil explicarlos conjuntamente.

Un consejo importante antes de ponernos manos a la obra: recuerda siempre, antes de redactar este y cualquiera de los otros apartados de tu artículo, **consulta las instrucciones para autores y requisitos específicos de la revista científica donde quieres enviar tu artículo a publicar.**

Nadie mejor que tú sabe cómo se ha llevado a cabo tu proyecto de investigación y, por lo tanto, nadie mejor que tú sabrá explicarlo.

He redactado este ebook pensando siempre en ti, que haces investigación clínica y epidemiológica y que escribes artículos científicos.

Por eso espero, de todo corazón, que te sea muy útil.

Estoy a tu disposición para aclarar cualquier duda.

¡Empezamos!

*Nota: En este ebook te presento los apartados que conforman la metodología de un artículo original, ya que éste es el tipo de difusión más habitual de una investigación clínica y epidemiológica. Pero debes tener en cuenta que existen otras formas de hacer difusión de los resultados de una investigación (una editorial, una revisión, una nota clínica, una carta al director, una comunicación oral, póster o mesa redonda en un congreso científico, etc.) y que cada una tiene sus normas de elaboración.*

Primero vamos a situarnos.

La estructura habitual de un artículo científico original se conoce con el acrónimo **IMRD**, que representa las iniciales de sus apartados: Introducción, Material y métodos, Resultados y Discusión, además del título, el resumen y las referencias bibliográficas.

I	Introducción	¿Qué se ha estudiado?
<b>M</b>	<b>Material y métodos</b>	<b>¿Cómo se ha estudiado?</b>
R	Resultados	¿Qué se ha observado?
D	Discusión	¿Qué interpretación tienen los datos?

Para empezar, te diré que la manera más habitual para denominar a este apartado en un artículo científico es **Material y métodos**, pero también puedes denominarlo solo **Métodos**.

Algunos autores consideran que no es una expresión adecuada cuando la investigación se ha llevado a cabo en seres humanos y sugieren denominarlo **Pacientes y métodos** o **Población y métodos**. Cualquiera de estas opciones estará bien. Pero recuerda lo que te he comentado antes, consulta las instrucciones para autores de la revista científica dónde quieras enviar tu artículo a publicar.

Yo siempre digo que escribir este apartado es como contar una historia, incluso verás que el orden que te propongo ya invita a ir explicando a modo de historia **cómo has ido haciendo tu proyecto de investigación**. Y, por lo tanto, como estaremos relatando una historia pasada, es muy importante que **redactes todo este apartado en tiempo pasado**.

Tal vez llegados a este punto te estés preguntando:

### **¿Cuánta información debo incluir en este apartado y con qué detalle?**

A modo de respuesta te diré que durante todo el proceso de redacción tengas siempre presente estas dos cosas:

1. Incluye en cada apartado la información suficiente para que los lectores de tu artículo puedan entender claramente todo el proceso de la investigación.
2. Detalla todo lo suficiente como para que otros investigadores, si ellos quisieran, pudieran replicar tu estudio.

Es decir, una vez redactado, léetelo todo y responde a esta pregunta:

**Si una persona que desconoce mi proyecto leyera esto, ¿sería capaz de entender todo lo que yo he hecho durante mi investigación, hasta el punto de que sería capaz de replicarlo?**

Si tu respuesta es que sí ... ¡enhorabuena! Has redactado bien tu apartado de metodología.

Si por el contrario detectas que algunas partes están un poco flojas, no queda del todo claro lo que has hecho y cómo lo has hecho, entonces deberás intentar explicarlo mejor pensando en qué alguien va a querer replicarlo.

**¡Ojo!** Al intentar describirlo todo con detalle **no cometes el error de incluir aspectos que no sean necesarios para entender e interpretar tu estudio de investigación.** Es un error habitual.

A mí me gusta ir indicando en mi artículo cada uno de los apartados que conforman la metodología, siempre queda más ordenado y facilita la lectura, sobre todo si este apartado es muy extenso. Pero no es obligatorio, a menos que la revista te lo indique en sus instrucciones para autores.

Lo que sí es recomendable, es que redactes este apartado siguiendo la secuencia lógica que te voy a ir indicando, tanto si pones título a los apartados como si no lo haces. Y también te recomiendo que **utilices un estilo directo, con frases cortas y sencillas que describan el proceso de la investigación con claridad y concisión.**

Esta secuencia lógica incluye: **el diseño del estudio, los sujetos participantes, sus criterios de inclusión/exclusión, la selección de la muestra, las variables de estudio, intervenciones, instrumentos, fuentes de información, procedimientos y la estrategia de análisis estadístico.**

Vamos a verlos uno a uno.

## Diseño del estudio [*Study design*]

En este apartado se describe el **tipo de estudio** que se ha realizado, en una o dos frases, no hace falta más.

También se suele incluir el **ámbito de estudio** y el **período de estudio**.

Ejemplo (incluye el ámbito de estudio y el período de estudio):

Estudio observacional y retrospectivo. Se analizaron los datos de la historia clínica desde el año 2004 hasta el año 2012. Se incluyeron pacientes de 48 Centros de Atención Primaria (CAP) de la Dirección de Atención Primaria Metropolitana Nord (DAP MN) del Institut Català de la Salut (ICS), que cubre una población de 997.946 personas.

Ejemplo:

Estudio descriptivo, transversal, de base poblacional, multicéntrico, en adultos aparentemente sanos (mayores o iguales a 15 años) o sujetos no demandantes de atención médica a los que se les había realizado una ecografía hepática.

## Pacientes o Población de estudio [*Study patients or Study population*]

Aquí deberemos describir la **procedencia de los sujetos** de nuestro estudio, ya que para la interpretación de los resultados y su extrapolación no es lo mismo que procedan de un centro de atención primaria o de un hospital, por ejemplo.

Ejemplo:

Se incluyeron mujeres de 60 a 74 años que recibían tratamiento farmacológico para la osteoporosis en mayo del 2012, independientemente de si constaba o no el diagnóstico de osteoporosis en prevención primaria (ICD10: M81).

Ejemplo (también incluye criterios de exclusión, siguiente apartado):

Los pacientes se reclutaron en 33 centros de atención primaria y 5 hospitales terciarios de cinco municipios costeros catalanes (Viladecans, Gavà, Barcelona, Badalona y Mataró). Seleccionamos a todos los sujetos con un diagnóstico de EPOC de acuerdo con los criterios espirométricos de la *American Thoracic Society* y la *European Respiratory Society (ATS/ERS)* que fueron visitados en alguno de los centros de salud participantes, después de excluir a aquellos con comorbilidades graves o potencialmente mortales. Solo se incluyeron pacientes clínicamente estables (definidos como al menos 4 semanas sin antibióticos o corticosteroides orales).

Ejemplo:

Estudio multicéntrico en el que participaron equipos de Atención Primaria que abarcan una población de 360.000 habitantes de una zona urbana y semirural del norte de Barcelona y el Maresme, España.

## **Criterios de inclusión / exclusión [*Inclusion / Exclusion criteria*]**

En este apartado **se especifican y describen los criterios de inclusión y exclusión que se han seguido para seleccionar a la población de estudio.**

La definición de la enfermedad o de la exposición que se estudia es el criterio que debe definirse con mayor precisión.

Las finalidades de esta descripción son evaluar la adecuación de la población, conocer en qué tipo de individuos se ha observado la presencia o la ausencia

del efecto o de la asociación estudiados, evaluar la posibilidad de extrapolar los resultados a otras poblaciones y permitir la replicabilidad del estudio.

A veces este apartado se incluye junto con el anterior y se describe la procedencia de los sujetos a la vez que se van indicando los criterios que se han seguido para ir excluyendo a algunos pacientes.

#### Ejemplo: (criterios de inclusión y exclusión)

##### Criterios de inclusión:

Población adulta de ambos sexos entre 15 y 80 años de los equipos de Atención Primaria que aceptaron participar voluntariamente en el estudio firmando el consentimiento informado.

##### Criterios de exclusión:

Ingesta de alcohol > 30 g/día en varones y > 20 g/día en mujeres. Presencia de enfermedad hepática crónica. Presencia del antígeno de superficie del virus de la hepatitis B o presencia de anticuerpos frente a la hepatitis C. Sujetos con condiciones o enfermedades que dificulten la toma de datos y el seguimiento del estudio tales como, enfermedades incapacitantes, deterioro cognitivo, pacientes institucionalizados o sujetos sin domicilio fijo en cualquiera de las áreas básicas del estudio. Sujetos que no proporcionen su consentimiento informado por escrito para participar en el estudio.

#### Ejemplo (criterios de exclusión):

Se excluyeron las pacientes con antecedentes personales de fracturas osteoporóticas, tratamiento hormonal sustitutivo, con diagnóstico de osteoporosis con fractura (ICD 10: M80) y osteoporosis secundaria asociada a hiperparatiroidismo, trastorno de conducta alimentaria, enfermedad de Paget, malnutrición crónica y malabsorción, diabetes mellitus tipo I, artritis reumatoide, hipertiroidismo, menopausia precoz (entre 40 y 45 años) o tratamiento con glucocorticoides, tratamientos oncológicos, litio y heparina. También se excluyeron las pacientes ingresadas en residencia geriátrica y las que no habían sido visitadas en el CAP los últimos dos años.



## Selección de la muestra [*Sample selection*]

Aquí deberemos explicar **cómo se han seleccionado los sujetos de estudio**, indicando la técnica utilizada para su captación.

Se deben dar los detalles que permitan conocer las características de la muestra de sujetos incluida (si son pacientes derivados, si se captan directamente de la consulta o a partir de algún sistema de registro, si existen filtros previos a su inclusión, etc.).

En caso de haberse utilizado alguna técnica de muestreo (aleatoria o no aleatoria), debe describirse adecuadamente. Es conveniente especificar el tiempo empleado en reunir a todos los participantes.

En los estudios experimentales hay que especificar el método de asignación de los participantes a los grupos de estudio.

Si se ha realizado un cálculo del tamaño de la muestra, este apartado suele finalizar con la explicación de este cálculo, indicando qué asunciones se han utilizado para calcular el número de sujetos necesarios para la realización del estudio.

### Ejemplo (sin cálculo del tamaño de la muestra):

La secuencia de aleatorización se creó utilizando el software Stata 12.0 (StataCorp, College Station, TX, EE. UU.) y se estratificó por centro con una asignación 1:1 a los grupos de intervención o control utilizando tamaños de bloques aleatorios de 6, 8 y 10. En la segunda visita del estudio después de la inscripción, un fisioterapeuta asignó a los pacientes al grupo correspondiente mediante un archivo informático seguro, donde las asignaciones se ordenaron de acuerdo con la secuencia de aleatorización y solo estaban disponibles de una en una.

Los evaluadores de resultados y los analistas de datos permanecieron cegados a la asignación. Los fisioterapeutas que administraron la intervención, conscientes de los grupos asignados, no realizaron mediciones de resultados. Los pacientes conocían el grupo asignado, pero desconocían la existencia del grupo alternativo, según lo aprobado por los Comités de Ética.

Ejemplo (con cálculo del tamaño de la muestra):

Con un riesgo alfa de 0,05 para una precisión de  $\pm 0,03$  unidades porcentuales en un contraste bilateral para una proporción estimada de 0,20 se requirió una muestra aleatoria proporcional de 683 sujetos. Considerando que el 15% de la población presentaría consumo de alcohol o alguna enfermedad hepática previa, se incrementó el número de pacientes a estudiar en este porcentaje haciendo un total de 786 individuos necesarios para cumplir con el tamaño de muestra final.

Una muestra aleatoria de la población asignada y obtenida a través del SIAP (Sistema de Información de Atención Primaria), fue estratificada en proporción a la población asignada a cada uno de los equipos de Atención Primaria participantes en el estudio. Cabe señalar que la población asignada era prácticamente equivalente al padrón municipal actualizado.

Una vez definidos los aspectos relacionados con la población de estudio y la formación de los grupos, pasamos a describir las variables estudiadas.

## Variables de estudio [*Study variables*]

En este apartado deberemos **definir todas las variables, las técnicas de medida y las fuentes de información** utilizadas a partir de las cuales hemos obtenido los datos.

Si fuera necesario también deberemos dar estimaciones o comentarios acerca de la validez y precisión de las técnicas de medida y de las fuentes de información.

También deberemos indicar **en qué momento del estudio hemos recogido u obtenido cada una de estas variables y quien lo ha hecho.**

Si nuestro estudio cuenta con una intervención o exposición, será aquí donde deberemos explicarla con todo detalle pensando siempre en que otra persona pueda replicarla.

Cuando definamos las variables podemos diferenciar entre: **variable/s de resultado** (*outcomes*) o también llamadas variables dependientes o respuesta, **variables independientes** o explicativas, y finalmente, **covariables o variables de ajuste**. Es muy recomendable diferenciarlas, ya que facilitará muchísimo al lector entender bien los resultados de la investigación.

Importante: Muchas veces este apartado se incluye junto con el siguiente que hace referencia a la recogida de estas variables, y se va explicando que datos se han recogido y cómo se han ido recogiendo. A veces es más fácil redactarlo de manera conjunta y también es correcto. Si haces esto puedes seguir llamando a este apartado “Variables de estudio” o puedes ponerle “Variables y recogida de los datos”.

Verás que algunos de los ejemplos que te indico ya hacen esta unión de los dos apartados.

#### Ejemplo (definición de variable resultado):

El resultado primario fue el tiempo desde una prueba de PCR positiva hasta la muerte relacionada con covid-19, que se definió como una muerte confirmada por covid-19, identificada por la clasificación internacional de enfermedades (CIE-19), con código U07.1 mencionado en cualquier parte del certificado de defunción. El código U07.1 se usa cuando se ha confirmado covid-19 mediante pruebas de laboratorio, independientemente de la gravedad de los signos o síntomas clínicos, pero solo debe indicarse en un certificado de defunción cuando es la causa principal o contribuyente de la muerte.

#### Ejemplo (fuentes de información):

Utilizamos datos del Activo de datos de salud pública de la Oficina de estadísticas nacionales (ONS), un conjunto de datos vinculado que combina el censo de 2011, los registros de mortalidad, los datos del Servicio de extracción de práctica general para la planificación e investigación de pandemias, Estadísticas de episodios hospitalarios, Datos de prueba y rastreo del NHS, y datos nacionales de vacunación del Servicio Nacional de Gestión de Inmunización (NIMS). El NIMS registra todas las vacunas administradas a todas las personas que viven en Inglaterra desde que comenzó el programa de vacunación el 8 de diciembre de 2020.

*(continuación)*

Para obtener los números del NHS, el censo de 2011 se vinculó a los registros de participantes del NHS de 2011-13. De los 53.483.502 registros censales, 50.019.451 estaban vinculados de manera determinista; Se obtuvieron 555.291 coincidencias adicionales mediante el emparejamiento probabilístico (tasa de vinculación general del 94,6%). Todos los enlaces posteriores se realizaron utilizando el número del NHS. El activo de datos de salud pública de la ONS incluye datos de 35 millones de adultos, aproximadamente el 79% de la población de Inglaterra en 2020.

Ejemplo (se va indicando la validez y precisión de las técnicas de medida):

Una encuesta de 22 ítems evaluó variables sociodemográficas (edad, género, peso, altura y ocupación [académica o administrativa], niveles de actividad física, tiempo sentado, bienestar mental y pérdida de productividad laboral. Para la actividad física, la forma corta del Cuestionario Internacional de Actividad Física (IPAQ) evaluó la actividad física de marcha, moderada y vigorosa. La forma corta del IPAQ muestra una buena confiabilidad ( $\rho$  de Spearman = 0,80) y una validez de criterio moderada con acelerómetros ( $\rho$  de Spearman = 0,30) en la población general, catalana y española.

El tiempo dedicado a estas actividades se combinó para mostrar el volumen de actividad en relación con el gasto de energía (Unidades Metabólicas Equivalentes - MET), lo que arrojó una puntuación en MET-minutos semanales. Los empleados se clasificaron en categorías de AF baja ( $\geq 599$  MET-minutos/semana), moderada (al menos 600-2999 MET-minutos/semana) o alta (3000+ METminutos/semana).

Un cuestionario de sedentarismo total y específico de dominio de siete días evaluó el tiempo de sedentarismo semanal (minutos/día) en el trabajo y mientras viajaba. Estos dominios se abordaron dentro de una intervención de AF en el lugar de trabajo que tenía como objetivo reducir el tiempo de estar sentado (i) en el trabajo y (ii) mientras viajaba.

Este cuestionario tiene una alta validez y fiabilidad en la población adulta para el tiempo sentado en el trabajo entre semana ( $r = 0,69-0,74$ ), mientras que es menor para los días de fin de semana en todos los dominios ( $r = 0,23-0,74$ ). La traducción hacia adelante y hacia atrás al catalán y al castellano identificó equivalencia lingüística.

*(continuación)*

La Escala de Bienestar Mental de Warwick-Edinburg (WEMWBS) evaluó el bienestar mental positivo (funcionamiento positivo, felicidad y bienestar subjetivo) durante las dos semanas anteriores. La escala de 14 ítems tiene cinco categorías de respuesta; 1 (“Ninguno”) a 5 (“Todo el tiempo”). Las respuestas se suman para identificar la puntuación final, de 14 a 70, que indica un bienestar mental positivo de bajo a alto. WEMWBS muestra una alta fiabilidad interna (alfa de Cronbach = 0,93) y una fiabilidad test-retest ( $r = 0,97$ ) en la población española.

El Cuestionario de Limitaciones Laborales (WLQ) se utilizó para evaluar el desempeño y el grado en que los problemas de salud interferían con la capacidad de desempeñar funciones laborales. Se han desarrollado y validado versiones en español y catalán del WLQ. En el WLQ, los encuestados autoinforman los niveles de dificultad para desempeñar 25 roles laborales específicos en cuatro escalas, con puntajes expresados como un promedio de las respuestas. La "Escala de tiempo" de 5 elementos aborda la dificultad en la programación de demandas. Para la "Escala mental-interpersonal", seis ítems cubren la dificultad para realizar tareas cognitivas que involucran el procesamiento de información sensorial y la interacción con otros en el trabajo. La "Escala de producción" tiene cinco elementos que exploran las limitaciones para satisfacer las demandas de cantidad, calidad y puntualidad del trabajo completado. La "Escala física" de nueve ítems evalúa la capacidad para realizar tareas laborales que involucran fuerza corporal, movimiento, resistencia, coordinación y flexibilidad.

Las puntuaciones de las subescalas se transforman en un continuo de 0 a 100 para representar la cantidad de tiempo en las dos semanas anteriores afectado por un desempeño laboral limitado (de un índice de dificultad bajo a alto). Estas escalas estiman la pérdida de trabajo, lo que se conoce como índice WLQ, que es la suma ponderada de las puntuaciones de las escalas WLQ. En el presente estudio, el índice WLQ se calculó sumando las puntuaciones de tres escalas WLQ; la "Escala física" se excluyó de los análisis actuales porque no era relevante para estos roles de trabajo.

### Ejemplo:

El estudio consistió en cuatro visitas: una primera visita, una segunda visita una semana después, una tercera visita 12 meses después de la aleatorización y una cuarta visita una semana después. En la primera visita, todos los pacientes respondieron un cuestionario administrado por un entrevistador, que incluía datos sobre variables sociodemográficas, tabaquismo, disnea (mediante la escala modificada del Medical Research Council [mMRC]), calidad de vida relacionada con la salud mediante el Clinical COPD Questionnaire (CCQ) y EPOC Assessment Test (CAT), síntomas de ansiedad y depresión (por la Hospital Anxiety and Depression scale [HAD]), y deterioro cognitivo (por el Phototest). También medimos, siguiendo procedimientos estandarizados: capacidad de ejercicio por distancia de caminata de 6 minutos (6MWD), composición corporal (peso, talla, índice de masa corporal [IMC] e índice de masa libre de grasa [FFMI]) por examen físico e impedancia bioeléctrica, y función pulmonar (FEV1 y FVC) por espirometría antes y después del broncodilatador. Las comorbilidades, el tratamiento farmacológico y el número de ingresos hospitalarios por una exacerbación de la EPOC en los 12 meses anteriores al reclutamiento se obtuvieron de las historias clínicas. Finalmente, se indicó a los pacientes que llevaran durante una semana el acelerómetro Dynaport (McRoberts BV, La Haya, Países Bajos), previamente validado para EPOC, para medir objetivamente la actividad física. Una medición de actividad física válida se definió como un mínimo de 3 días con al menos 8 horas de tiempo de uso dentro del tiempo de vigilia; el cumplimiento con el acelerómetro se detalla en el archivo complementario en línea.

En la segunda visita, los pacientes entregaron el acelerómetro y respondieron el cuestionario Clinical-PROactive Physical Activity (C-PPAC) para medir la experiencia de actividad física. Un fisioterapeuta asignó a los pacientes al grupo correspondiente y proporcionó las intervenciones a ambos grupos como se detalla anteriormente.

En la tercera y cuarta visita (12 meses después del reclutamiento), además de los cuestionarios y pruebas utilizados al inicio del estudio, los pacientes respondieron un cuestionario sobre la satisfacción con los componentes del estudio y los posibles eventos adversos experimentados al caminar en los 12 meses anteriores. El número de exacerbaciones durante el período de seguimiento (definidas como un empeoramiento agudo de los síntomas respiratorios que dan lugar a una terapia adicional) y su gravedad (tratadas ambulatoriamente, con necesidad de urgencias o con necesidad de ingreso hospitalario) se obtuvieron de las historias clínicas. Las medidas de control de calidad se detallan en el archivo complementario en línea.

## Recogida de los datos o procedimientos [*Data collection*]

Aquí yo siempre le digo al investigador, ahora imagina que entra por la puerta un paciente candidato a participar en tu estudio, explícame, con todo detalle, que le has hecho a ese paciente para incluirlo en tu estudio, y que le has hecho desde ese momento en el que lo has incluido en tu estudio. Es decir, **explica con todo detalle desde que el paciente ha entrado por la puerta hasta que se ha ido.**

Si recopilas datos de fuentes de información, su equivalente sería: **explícame todo el proceso que has hecho para obtener esos datos.**

Podríamos decir que es como **describir la arquitectura general del estudio**, indicando la pauta de visitas y qué actividades se realizan en cada una de ellas, quiénes y cuántos son los observadores, si ha existido un entrenamiento previo, los mecanismos de captación de los pacientes, las técnicas de control de calidad de los datos recogidos, etc.

Si el estudio evalúa la eficacia de un tratamiento, debe señalarse si ha existido un período de lavado o blanqueo previo a la intervención, las técnicas de registro de la presencia de reacciones adversas, los criterios utilizados para la retirada anticipada de sujetos, las técnicas de enmascaramiento utilizadas (simple ciego, doble ciego, triple ciego, etc.).

Importante: Como te he comentado en el apartado anterior, **muchas veces este apartado se incluye junto con el anterior de las variables del estudio.**

Los ejemplos que te muestro aquí son de artículos en los que sí se ha decidido separar los dos apartados porque era más fácil explicarlos así.

### Ejemplo:

Se listaron a todos los pacientes de los centros participantes con diagnóstico de EPOC codificado en su historia clínica. Se realizó un primer cribado de estas historias clínicas de atención primaria y de los informes hospitalarios relacionados con las mismas para excluir a los pacientes con algún criterio de exclusión en su historia clínica o cuyos parámetros de espirometría registrados no fuesen congruentes (relación FEV1/FVC  $\geq$  70%) con diagnóstico de EPOC.

*(continuación)*

Los pacientes que no fueron excluidos en este primer cribado fueron seleccionados aleatoriamente hasta conseguir el número de pacientes necesarios para cada grupo de gravedad por cada centro. Los sujetos seleccionados fueron invitados a participar por carta y llamada telefónica. Si un paciente no quería participar o no se localizó tras 10 intentos de contacto en días y horarios diferentes, era sustituido por otro paciente elegido al azar del mismo centro. Los pacientes seleccionados que, en el momento de la inclusión en el estudio, tenían una exacerbación de su enfermedad respiratoria eran excluidos temporalmente y reevaluados cuatro semanas después. Si la situación de exacerbación persistía en este momento, el paciente era definitivamente excluido del estudio.

Los pacientes seleccionados acudieron a su centro de salud sin haber fumado (mínimo 6 horas antes) ni haber tomado su medicación inhalada (12 o 24 horas antes según la vida media del fármaco activo). Allí firmaron el consentimiento informado para participar en el estudio. Una enfermera específicamente capacitada para el estudio realizó una espirometría con una prueba broncodilatadora. Si el resultado de la espirometría no era compatible con un diagnóstico de EPOC, se daba por finalizado el estudio. Si el resultado de la espirometría era compatible con el diagnóstico de EPOC se realizaron las pruebas, mediciones y cuestionario descritos en el apartado de variables de estudio hasta completar los grupos de 137 pacientes leves, 102 moderados y 115 graves necesarios para el seguimiento.

La enfermera introducía la información en una base de datos durante la visita utilizando un cuestionario de recogida de datos (CRF) diseñado para el estudio con soporte electrónico y una computadora portátil.

### Ejemplo:

Se contactó con el equipo directivo de cada centro o departamento para presentar el estudio y sus objetivos. Cuando la dirección del centro autorizaba la participación en el estudio, un miembro del equipo de investigación iba a cada uno de los Centros de Atención Primaria (CAP) para informar a los profesionales, motivarlos e invitarlos a participar. Los cuestionarios se entregaron y recogieron en una sesión específica en cada CAP. Una persona en cada CAP se encargaba de la administración y recogida de los cuestionarios



*(continuación)*

de los profesionales ausentes el día de la sesión para su posterior envío al equipo investigador en un plazo máximo de dos semanas para su posterior procesamiento y análisis de los mismos. El cuestionario a los profesionales era anónimo y autoadministrado. Únicamente tenía un código indicativo del equipo al que pertenecía el profesional sanitario para obtener la tasa de respuesta de cada centro. En ningún momento los datos obtenidos podían identificar al profesional que respondía. En el procesamiento y análisis de datos tampoco era posible identificar a los centros ya que éstos estaban indexados con un código alfanumérico sin relación con la ubicación geográfica o entidad prestadora del servicio.

Para administrar los cuestionarios a las familias colaboraron los mediadores culturales de los centros asistenciales. Estos mediadores presentaron al entrevistador y lo acompañaron durante todo el proceso de cumplimentación del cuestionario para aclarar cualquier concepto o duda en relación con el mismo que pudiera surgir entre las familias encuestadas. Los mediadores participantes era expertos con formación específica en mediación intercultural. Se les instruyó para este estudio sobre las características de los cuestionarios y los objetivos y métodos del estudio. Se les formó en la aplicación del cuestionario y se explicaron los objetivos que se perseguían con cada una de las preguntas. El cuestionario se entregaba siempre bajo la supervisión de un miembro del equipo investigador. El cuestionario a las familias era anónimo. También se recopiló información mínima sobre las familias que se negaron a participar en la encuesta para después compararlas con las familias que sí aceptaron participar y poder detectar posibles diferencias.

## **Análisis estadístico o Plan de análisis [*Statistical Analysis or Analysis Plan*]**

Este es mi apartado favorito.

La mayoría de los investigadores me piden que les ayude con la redacción de este apartado. Tú también puedes pedir ayuda a tu estadístico/a de confianza, no tienes por qué saber que estrategia de análisis y que técnicas estadísticas se han utilizado para analizar tus datos si no has sido tú la persona que ha realizado el análisis estadístico.

Por lo tanto, aquí queda claro que deberá **describirse con claridad y precisión la estrategia de análisis y las técnicas estadísticas empleadas.**

Cuando se utilizan varias pruebas ha de quedar claro en qué parte del artículo se utiliza cada una de ellas. Es decir, se debe ir indicando que estrategia de análisis y que pruebas estadísticas se han utilizado para responder a cada uno de los objetivos del estudio.

Si las pruebas son de uso muy frecuente, no es necesario que se describan en el texto. Si su uso es más infrecuente, conviene incluir una breve explicación acompañada de una cita bibliográfica.

Es preferible que las citas bibliográficas de este apartado correspondan a un libro de texto o a un artículo de revisión sobre la prueba, antes que recurrir al artículo original. Los artículos originales sobre las pruebas estadísticas tienen, indudablemente, un gran valor desde el punto de vista histórico y metodológico, pero suelen ser muy complicados y ofrecen pocos ejemplos prácticos. Por lo que no es muy recomendable utilizarlos como referencia si queremos que los lectores de nuestro artículo entiendan la prueba que hemos aplicado. Por el contrario, los libros de texto o los artículos de revisión suelen ser más didácticos, lo cual facilita la comprensión del lector.

En este apartado también debe **indicarse la naturaleza de las hipótesis evaluadas (unilateral o bilateral) y el valor de significación aceptado (habitualmente 0,05).**

Finalmente se debe **indicar el programa estadístico utilizado para analizar los datos y su versión.**

Como te decía al inicio de este apartado, mi recomendación es que le dejes esta parte a tu estadística/o de confianza, no tienes por qué saber hacerlo, ¡bastantes cosas has hecho ya durante todo tu proyecto de investigación!

### Ejemplo:

Las características de la población de estudio se resumieron en general y se estratificaron por tipo de variante, utilizando medias para las variables continuas y proporciones para las variables categóricas.

Utilizamos un modelo de regresión de riesgos proporcionales de Cox de causa específica para estimar el índice de riesgo de muerte relacionada con covid-19 para personas infectadas con omicron BA.1 versus variante delta. El tiempo de seguimiento se calculó como el período desde una prueba de PCR positiva hasta la muerte por covid-19 o el final del estudio. Para las muertes no relacionadas con covid-19, las personas fueron censuradas en la fecha de la muerte si esta ocurrió antes del final de la fecha del estudio. Estimamos cuatro modelos, ajustados secuencialmente por edad, sexo, estado de vacunación e infección previa (modelo 1); más, tiempo de calendario (modelo 2); más factores socioeconómicos (modelo 3); y finalmente, más condiciones de salud preexistentes (modelo 4).

Para probar si el riesgo relativo de muerte por omicron BA.1 variaba según la edad y el sexo, incluimos interacciones entre el tipo de variante y la edad, y el tipo de variante y el sexo. Para probar si el riesgo relativo de muerte por omicron BA.1 varió según el estado de vacunación (sin vacunar, una dosis, dos dosis y refuerzo) y el número de comorbilidades (0, 1-2,  $\geq 3$ ), comparamos un modelo ajustado para las interacciones entre el tipo de variante y la edad, y la edad y el estado de vacunación (o comorbilidades) con un modelo que incluía una interacción de tres vías entre el tipo de variante, la edad y el estado de vacunación (o comorbilidades). La justificación de este enfoque fue que el estado de vacunación y el número de comorbilidades están estrechamente relacionados con la edad y, en ausencia de una interacción entre el tipo de variante y la edad, la interacción entre el estado de vacunación (o las comorbilidades) podría captar la interacción entre el tipo de variante y la edad.

Evaluamos el supuesto de riesgo proporcional probando la independencia entre los residuos de Schoenfeld escalados y el tiempo en riesgo. Utilizamos los residuos de Schoenfeld de los modelos de Cox ajustados, suavizados mediante modelos aditivos generalizados, para evaluar si las diferencias relativas en el riesgo de muerte por covid-19 entre las variantes eran constantes a lo largo del tiempo después de una prueba positiva.

Ejemplo:

La incidencia acumulada de infecciones del tracto respiratorio inferior (ITRI) durante el primer año de vida se obtuvo del cuestionario de un año. Del mismo modo, los datos de ITRI obtenidos a través del sistema de vigilancia de Barcelona se trataron como incidencia acumulada. La asociación entre ITRI y niveles de NO<sub>2</sub> se midió con el odds ratio (OR), utilizando métodos de regresión logística. Como la probabilidad es una aproximación de la incidencia acumulada, el ajuste se evaluó utilizando modelos de regresión logística multivariable. Las tasas de prescripción de antibióticos, así como la tasa de incidencia de ITRI en Barcelona se modelaron con regresión de Poisson. Las variables de confusión seleccionadas fueron aquellas que mostraban una asociación estadísticamente significativa ( $P < 0,1$ ) con ITRI en nuestro estudio (es decir, haber amamantado alguna vez, clase social, sexo, tamaño de la familia y tabaquismo materno entre un conjunto de variables que incluyen asma o atopia de los padres).

Ejemplo:

Los datos sobre las variables de resultado clave se describieron mediante frecuencias (porcentaje), medias (desviación estándar) y medianas (rango intercuartílico). Los análisis de regresión lineal bivariada evaluaron las asociaciones entre el tiempo de estar sentado (total y específico del dominio), la actividad física, el bienestar mental y la productividad laboral. El modelo se ajustó por datos demográficos y se estratificó por nivel de actividad física introduciendo un término de interacción entre el nivel de actividad física (bajo, moderado o muy activo) y el tiempo sentado en un modelo de regresión multivariable. La significación se fijó en  $p < 0,05$  y los análisis se realizaron con el software Stata, versión 12.

Pues esto es todo. Espero que te sea muy útil para redactar la metodología de tu artículo científico.

Recuerda, **intenta explicar cada apartado con el máximo detalle posible**, pero siempre **dentro de los límites del número de palabras que te indiquen las instrucciones para autores de la revista científica**.

Lo sé, es difícil resumir todo lo que has hecho durante los años que ha durado tu investigación en unas pocas páginas, pero te aseguro que **nadie mejor que tu sabrá explicarlo**. Luego **pide ayuda a un experto en metodología** para que te lo revise y te ayude si no consigues avanzar en algún apartado.

**¡Espera!**

Antes de acabar te quiero indicar cuáles son los **10 errores más habituales** que se cometen a la hora de redactar este apartado, para que tu no los cometas:

1. Describir aspectos que no son necesarios para entender e interpretar el estudio, cosas superfluas, como, por ejemplo, especificar la marca y el tipo de ordenador utilizado, ¡no es necesario! Si que es necesario indicar el software estadístico que tú o tu estadístico habéis utilizado para analizar los datos.
2. Contrariamente, omitir información necesaria para entender e interpretar el estudio o describirlo de manera incompleta, por ejemplo, no indicar todos los criterios de selección, o no definir o definir de manera incompleta aspectos tan importantes como las variables principales del estudio, o la intervención.
3. Incluir resultados en este apartado de metodología. Este es un error muy común que detecto muchas veces.
4. No indicar el diseño del estudio o indicar un diseño que no se corresponde con el realmente utilizado. Y más grave aún, indicar que se ha utilizado un diseño y que sea inapropiado para los objetivos de tu estudio, que no permita dar respuesta a tu pregunta de investigación. Por eso yo siempre recomiendo que también incluyas a un/a estadístico/a durante el diseño del protocolo de tu investigación. ¡Pon un/a estadístico/a en tu vida! o como mínimo ¡en tu investigación!

5. No describir a los pacientes o población de tu estudio y/o tampoco describir el proceso de captación/selección de estos sujetos, si han sido voluntarios, se han seleccionado aleatoriamente, etc.
6. No describir como se han asignado los sujetos a los grupos de estudio o describirlo sin el suficiente detalle como para que se puedan entender los resultados.
7. No incluir la información sobre el cálculo del tamaño de la muestra o sobre las asunciones en que se ha basado este cálculo.
8. Definir variables que posteriormente no se utilizan en el análisis. Esto es muy frecuente en artículos que informan sobre aspectos parciales de una investigación más amplia.
9. Describir la recogida de los datos o los procedimientos utilizados sin un orden lógico. Recuerda lo que te he dicho, hazlo como si explicarás una historia.
10. No indicar las pruebas estadísticas utilizadas en el análisis de los datos y redactar un plan de análisis muy general. O indicar las pruebas, pero no asociarlas a cada uno de los objetivos de tu estudio. Y más grave aún, indicar una estrategia o una prueba estadística que no sea correcta con el tipo de datos que tienes.

Ahora sí, hemos acabado.

**Si tienes dudas sobre cualquier apartado me puedes escribir a [laura@datexbio.com](mailto:laura@datexbio.com) y estaré encantada de resolver tus dudas.**

¡Hasta pronto!

Nos vemos por LinkedIn ([www.linkedin.com/in/laura-muñoz-bioestadistica](https://www.linkedin.com/in/laura-muñoz-bioestadistica))

