

Capítulo 1:

Introducción

Presentación

Este capítulo introduce las bases de la comunicación científica. En la primera parte se sientan los principios para interpretar los valores que toman las variables y así desarrollar la capacidad de razonamiento numérico o cuantitativo; a partir de la definición de medida, se razona sobre el significado de un valor observado; por ejemplo, 78 kg al hablar de peso.

En la segunda parte se presentan las guías de publicación consensuadas internacionalmente.

Objetivos

Al terminar este capítulo, un lector que haya realizado los ejercicios:

- Delante de una variable, distinguirá entre escala de medida nominal, ordinal y de intervalo.
- Delante de una variable, dirá que está en escala ordinal si puede ordenar sus categorías.
- Delante de una variable, dirá que está en escala de intervalo si dispone de unidad de medida.
- Delante de una variable, distinguirá entre la variable observada y su objetivo de medida.
- Delante de una variable, se preguntará si es válida (en el sentido de medir el objeto de su medida).
- Delante de una variable, se preguntará si es fiable (en el sentido de obtener resultados similares en determinaciones repetidas).
- Identificará Buenas Prácticas Clínicas con el respeto de los derechos de los participantes y con la fiabilidad de los datos.
- Conocerá la existencia de recomendaciones internacionales para el desarrollo y el informe de estudios clínicos.
- Conocerá la recomendación CONSORT para la publicación de ensayos clínicos en revistas biomédicas.
- Conocerá las normas ICH como referente para el desarrollo de intervenciones terapéuticas.

El proceso de cuantificar la observación. Escalas de medida

En el lenguaje literario es excitante percibir, para una misma frase, diferentes perspectivas, matices o significados. De la misma forma, un amigo nos hará reír con un «doble sentido».

Historieta



El amigo Carlos está obsesionado por una calvicie que sólo él percibe. Al llegar al grupo de amigos, Ramón le dice: «¡Cuánto tiempo sin verte el pelo!». Y mientras se apagan las risas, Miguel Ángel le aconseja en tono confidente: «Carlos, esto del pelo deberías quitártelo de la cabeza».

Por el contrario, una característica fundamental del proceso científico es establecer un lenguaje común, con idéntico significado para cualquier observador. Es decir, «llamar a las cosas por su nombre».

Nota técnica



En el tema de escalas de medida va a seguirse la línea de Stevens (1,2).

Escala nominal

Clasificar consiste en agrupar los objetos estudiados en categorías: dos objetos de la misma categoría deben ser equivalentes entre sí y diferentes de los objetos de otra categoría.

Definición



La escala **nominal** clasifica las unidades de acuerdo con ciertas características, de forma que dos unidades de la misma categoría son iguales entre sí y diferentes de las restantes categorías.

Ejemplo 1.1



Una burda y primera clasificación de las personas las dividiría en enfermos y sanos. Dos enfermos compartirán ciertas características comunes que los hacen diferentes de los sanos.

Ejercicio 1.1



Proponga ejemplos de otras clasificaciones posibles.

Nótese que si todas las unidades fueran iguales, todas pertenecerían a la misma categoría y el proceso de clasificación no tendría ninguna utilidad. Es decir, si no existiera variabilidad, no nos plantearíamos ni clasificar ni medir.

Ejemplo 1.2

Los códigos de identificación personal pertenecen a la escala nominal. Incluso aunque estén formados por números, como lo hace el DNI. Nótese que los números del DNI o del teléfono no tienen auténtico significado de número: no significa nada tener un número de teléfono mayor; el DNI sólo «clasifica» y, por tanto, está en escala nominal.

Escala ordinal

En el ejemplo anterior de enfermo y sano, la inmediata ambición del clínico y del científico es poder matizar más y, por ejemplo, establecer grados de intensidad: leve, moderado y grave. Igual que antes, dos unidades de la misma categoría serán iguales entre sí y diferentes de las restantes categorías. Pero ahora, además, se puede establecer una **relación de orden** y decir que grave es más que moderado. Y como moderado es más que leve, entonces grave también es más que leve.

Definición

La escala **ordinal** cumple las propiedades de la escala nominal y, además, permite ordenar las categorías.

Ejemplo 1.3

La reacción a la prueba de la tuberculina puede valorarse en una escala ordinal que va de 0 a 3 cruces (0 / + / ++ / +++).

Ejercicio 1.2

Proponga algún otro ejemplo de variable en escala ordinal.

Escala de intervalo

Una vez establecido un orden, la siguiente ambición del científico será comparar las diferencias entre categorías sucesivas. En el ejemplo de la tuberculina, ¿existe el mismo salto de + a ++, que de ++ a +++? Si todos los saltos tuvieran el mismo significado, se podría hablar de una misma **unidad de medida**, lo que permitiría comparar diferentes intervalos y decir, por ejemplo, que la diferencia entre + y +++ es mayor que la diferencia entre 0 y +. Mientras no se disponga de esta unidad de medida, los «saltos» tendrán diferente significado: no será lo mismo, por ejemplo, pasar de + a ++, que de ++ a +++.

Definición

La escala de **intervalo** cumple las propiedades de la escala ordinal y, además, dispone de unidad de medida.

En la escala de intervalo, las categorías se han convertido en números que disfrutan de una unidad que se aplica por igual a todos ellos: ya tienen significado de número. Como todos contienen las mismas unidades, se pueden restar entre ellos, lo que permite, por ejemplo, comparar la amplitud de varios intervalos, lo que da nombre a la escala.

Ejemplo 1.4



Se puede decir que entre dos cuerpos, uno a 19 °C y otro a 20 °C, hay la misma diferencia que entre uno a 29 °C y otro a 30 °C. O incluso, que el intervalo entre 10 °C y 20 °C es 5 veces mayor que entre 30 °C y 32 °C.

Ejercicio 1.3



Proponga algún otro ejemplo de variable en escala de intervalo.

Escala de razón

Historieta



Dice: «¿Qué tiempo hace?». Y contesta: «Según el termómetro, ni frío, ni calor: 0 °C».

Cuando hay unidad de medida es interesante preguntarse si existe un **cero absoluto**. Es decir, si el valor 0 de la escala tiene significado de «absoluta ausencia de ...». En el ejemplo de la temperatura en grados centígrados, ¿significa 0 °C ausencia absoluta de temperatura? ¡Cuánto tuvo que avanzar la física para poder contestar a esta pregunta! Si se puede afirmar que existe cero absoluto, entonces se dice que se está en escala de razón y es posible también comparar proporciones.

Ejemplo 1.5



Se puede decir que un cuerpo que está a 400° Kelvin tiene el doble de temperatura que un cuerpo a 200 °K.

Para los propósitos clínicos habituales, la distinción entre escala de intervalo y de razón es irrelevante.

Ejercicio 1.4



La variable «¿tiene cefalea?» admite las categorías «nunca», «a veces», «muchas veces» y «siempre». ¿En qué escala de medida se encuentra?


Ejercicio 1.5

¿En qué escala de medida se encuentra la variable «número de cigarrillos que fuma al día»?

Escalas de medida y tipos de variables

La tabla 1-1 resume las propiedades que corresponden a las escalas de medida. Estas propiedades son acumulativas, ya que para tener una propiedad superior es necesario cumplir la anterior. Así, una variable en escala de intervalo, además de poseer unidad constante, tiene sus valores ordenados.

Nota




Más adelante se verá que, en el momento de escoger un tipo de análisis estadístico, se puede optar por renunciar a las propiedades superiores y utilizar uno que corresponda a las inferiores. Por ejemplo, se verá que la edad tiene unidad de medida y permite calcular la media, pero también se pueden hacer categorías (joven, adulto...) y calcular frecuencias.

Escala	Propiedades
Nominal	Equivalencia
Ordinal	Orden
Intervalo	Unidad
Razón	Cero absoluto


Tabla 1-1 Tipos de escala y propiedades acumulativas

Nota técnica



Otra clasificación divide las variables en **cualitativas** y **cuantitativas**. Las primeras hacen referencia a la escala nominal, y las últimas a la de intervalo con una unidad que permite comparar cantidades. La escala ordinal puede tener un papel más ambiguo y corresponderse con ambas, ya que las propiedades de orden podrían aplicarse a categorías o cualidades (como la clase social) o a expresiones cuantitativas (como el número de puntos obtenidos en un score como el índice de Apgar).


Nota técnica



Otra división muy habitual consiste en clasificar las variables en **discretas** o **continuas**. Un recuento (el número de hermanos, por ejemplo) es una variable discreta ya que sólo puede tomar un número limitado de valores. En cambio, la altura, que en teoría puede tomar cualquier valor, es una variable continua. Suele considerarse a las variables en escala nominal como discretas, mientras que las otras dos escalas pueden ser tanto discretas como continuas.

Teoría representativa de la medida

Ejemplo 1.6



¿Podemos utilizar la edad como aproximación del grado de madurez de una persona? Veamos, para empezar, en qué escala de medida podemos considerar que se encuentra la variable «edad». Si se habla de la edad que figura en el DNI,

Ejemplo 1.6 (Cont.)

si se pretende utilizar la edad para aproximarse al grado de madurez del sujeto, se podría discutir si representa el mismo incremento en madurez pasar de 2 a 3 años que de 42 a 43, con lo que se estaría poniendo en entredicho la escala de intervalo. Pero, además, se podrían encontrar ejemplos de personas con menos años pero más maduras, con lo que se pondría en entredicho la propiedad de orden. Finalmente, incluso se podría argumentar que no tienen la misma madurez dos individuos de la misma edad, con lo que no se tendría la propiedad de equivalencia y no se podría considerar que la edad es una medida de la madurez. Pero, por otro lado, puede ser útil observar la edad de una persona para considerar qué comportamiento esperamos de ella. Así pues, si no se quieren perder estas posibilidades que ofrece la edad, conviene redefinir el proceso de medida.

Las escalas nominal, ordinal y de intervalo corresponden a una visión «operativa» de la medida: se define una variable por la forma de medirla. Esta visión permitiría definir «el cociente de inteligencia (CI)», como la variable con la que se cuantifica la inteligencia, pero nunca permitiría definir el concepto de inteligencia en sí mismo.

Definición

Medida es el proceso que conecta un concepto con una variable latente y ésta, con variables observables.

Es decir, existe por un lado un atributo latente que no es directamente observable (por ejemplo, la inteligencia) y por otro lado, una o varias variables que pretenden cuantificar dicho atributo (por ejemplo, el CI). El CI será tanto mejor medida de inteligencia cuanto más intensa sea su relación con la misma y menor dependencia tenga con otros factores.

La teoría representativa de la medida es estadística en el sentido de que acepta cierta variabilidad en los resultados. Dos individuos que obtengan exactamente la misma puntuación en una prueba de inteligencia no tienen por qué ser exactamente idénticos, pero cabe esperar que sean más similares que dos individuos con valores alejados.

Dos propiedades son especialmente deseables para poder decir que la variable observable (el CI en el ejemplo) es una buena medida del concepto o variable latente (la inteligencia en sí misma): la **validez** y la **fiabilidad**. Si una variable es válida y fiable, la variabilidad de sus valores depende exclusivamente de la variabilidad de su objeto de medida (el **concepto latente**): por no depender de otros conceptos, no tendrá error sistemático y se dirá que es válida; y por no tener error aleatorio de medida, se dirá que es fiable o repetible.

Definición

Se dice que una variable mide de manera **válida** un concepto representado por una variable latente si está relacionada con esta variable latente y sólo con ella.

Se dice que una variable mide de manera **fiable** un concepto si las variaciones de esta variable sólo están relacionadas con variaciones en el concepto.

La validez hace referencia a la ausencia de error sistemático y la fiabilidad a un error aleatorio razonablemente pequeño. Así, validez implica que se esté valorando el concepto y nada más: que la medida es «pertinente», y que variaciones en la variable comporten variaciones en la medida. Por su parte, la fiabilidad requiere obtener valores próximos en mediciones repetidas en el mismo individuo en condiciones similares.

Ejemplo 1.7

Los logros sanitarios en la cantidad de vida están permitiendo que cada vez sea más relevante la calidad de vida. Desde el punto de vista del paciente, la calidad de vida es su objetivo. Es pertinente y relevante. En una palabra, es una medida válida. Pero para un clínico es una variable algo incómoda, ya que un mismo paciente puede tener oscilaciones en su puntuación de la calidad de vida que no son explicables por variaciones en sus parámetros clínicos. Dicho de otra manera, determinaciones repetidas de la calidad de vida de un paciente, que teóricamente está en las mismas condiciones, no resultan en la misma puntuación. Es decir, no es una medida fiable en su sentido técnico.

Lectura

Si quiere saber más sobre las posibilidades de la conceptualización estadística del proceso de medida puede consultar, por ejemplo, Bollen (3) o Hand (4).

Nota técnica

Estas escalas de medida no se deben tomar como un procedimiento automático para decidir el análisis estadístico, vease Velleman (5).

Ejercicio 1.6

La variable «recuento de linfocitos CD4» suele emplearse en el seguimiento de pacientes con sida. ¿En qué escala de medida se encuentra? Si se decidiera utilizarla como indicador de la evolución, ¿qué opina de su validez y de su fiabilidad?

Ejercicio 1.7

El proceso de aprendizaje universitario, como unos estudios de Medicina, pretende que aquellos que lo finalicen sean capaces de ejercer como profesionales. ¿En relación a la validez y fiabilidad, qué le parece el examen MIR comparado con, por ejemplo, la observación de su trabajo delante de un paciente real?

Ejercicio 1.8

Cierto fármaco tiene un efecto directo sobre la presión arterial. Se ha decidido hacer un ensayo clínico para demostrar su efecto y se discute si medirlo en dicha presión (al cabo de cierto plazo) o en la aparición de complicaciones vasculares (en un plazo muchísimo más largo). El responsable de Sanidad argumenta que el objetivo auténtico es prevenir el riesgo vascular, pero el investigador argumenta que la aparición de accidentes vasculares es muy caprichosa y variable. ¿Quién se preocupa por la validez y quién por la fiabilidad?

Guías internacionales consensuadas

Para mejorar la calidad de la investigación y de los informes que la documentan, la administración, la industria y las revistas científicas han desarrollado varias iniciativas.

Normas de Buena Práctica en Estudios Clínicos

Para garantizar la autenticidad de los datos obtenidos, de su análisis y de su comunicación, se deben seguir una serie de principios y procedimientos conocidos como normas de Buena Práctica en Estudios Clínicos (BPC).

Definición

La **BPC** es la denominación que ha recibido un conjunto internacional de normas éticas y de calidad científica que deben regir el diseño, realización, registro de datos y comunicación de ensayos clínicos (EC).

Lectura

Real Decreto 223/2004, de 6 de febrero (6), por el que se regulan los ensayos clínicos con medicamentos. «Especial mención merece la obligación de aplicar las normas de buena práctica clínica a la planificación, realización, registro y comunicación de todos los ensayos clínicos que se realicen en España, como conjunto de requisitos éticos y científicos de calidad reconocidos a escala internacional y como garantía de la protección de los derechos, la seguridad y el bienestar de los sujetos del ensayo, así como la fiabilidad de sus resultados.»

Nota técnica

Otros términos que también resultan descriptivos del verdadero sentido de BPC son: Buena Práctica de Investigación Clínica, utilizado por la Asociación Británica de la Industria Farmacéutica, o Buena Práctica de EC, utilizado por el Consejo Nórdico de Medicina. Todos ellos son preferibles a normas de Buena Práctica Clínica, que no contextualiza bien su entorno de aplicación.

El cumplimiento de dichas normas garantiza que se han protegido los derechos, bienestar y confidencialidad de los sujetos participantes en el estudio.

En todos los países, los medicamentos deben ser autorizados por las Autoridades Sanitarias antes de su comercialización. Si se observa la definición de BPC, éstas adquieren mayor importancia cuando la información obtenida de un EC se va a utilizar como documentación para el registro de un fármaco y su posterior comercialización.

Comentario

El EC es la herramienta metodológica de investigación clínica requerida por las Autoridades Reguladoras para demostrar la eficacia de un tratamiento y autorizar su uso comercial. Por eso la BPC está referida a los EC. No obstante, cualquier otro estudio de investigación clínica debería también realizarse bajo esta filosofía de trabajo. Prueba de ello es la regulación de los estudios postautorización, que no deben ser necesariamente EC.

Ejercicio 1.9

¿Cuáles son los objetivos principales que pretenden garantizar las BPC?

Directrices ICH (International Conference of Harmonisation)

Para evitar la heterogeneidad de requisitos de las Autoridades Reguladoras de diferentes países, las asociaciones de la industria farmacéutica y autoridades de Estados Unidos, la Unión Europea y Japón fundaron en 1989 la Conferencia Internacional de Armonización de los Requisitos Técnicos para el Registro de Productos Farmacéuticos para Uso Humano.

Ejercicio de Navegación

Entre en la página que se indica a continuación y vea las diferentes directrices expuestas, con especial atención en la ICH E6.

<http://www.ich.org>

Lectura



La Conferencia Internacional de Armonización organiza reuniones con grupos técnicos para preparar directrices (guidelines) consensuadas para la investigación y desarrollo de fármacos (7). Estas actividades se desarrollan en interés de los consumidores y de la salud pública con el fin de desarrollar, de la manera más eficiente posible, medicamentos seguros y eficaces. Con ello se evitan duplicaciones innecesarias de los EC en humanos y se minimiza la experimentación animal, sin comprometer los requisitos normativos necesarios para demostrar eficacia y seguridad de los nuevos medicamentos. Además, también persigue que estas directrices sean aceptadas por las autoridades reguladoras y sean seguidas por la industria farmacéutica como método común, facilitando así la evaluación y el registro de nuevos medicamentos en el denominado territorio ICH.

La tabla 1-2 lista 10 directrices relevantes para realizar y documentar EC.

Código	Tema
E1	Tiempo de exposición y número de caso para valorar la seguridad clínica de fármacos
E2	Gestión de datos sobre seguridad y tolerabilidad
E3	Informe de estudios clínicos
E4	Información sobre relación dosis-respuesta
E5	Factores étnicos en estudios clínicos foráneos
E6	Buenas prácticas en estudios clínicos
E7	Población geriátrica
E8	Consideraciones generales para EC
E9	Aspectos estadísticos en el diseño de EC
E10	Elección del grupo control

Tabla 1-2 Directrices ICH sobre estudios clínicos.

Los documentos que se aceptan en la Conferencia Internacional de Armonización deben ser implementados en la normativa de cada país.

Lectura



Desde 1996 se dispone del documento ICH de BPC (ICH E6 Good Clinical Practice, tabla 1-3) que proporciona un estándar para la realización de EC, válido para la Unión Europea, Japón y Estados Unidos, que pretende facilitar la aceptación mutua de los resultados de los EC por parte de las autoridades sanitarias responsables del registro de medicamentos en dichos países (8).

Capítulo 1.	Glosario de términos referentes a investigación clínica con medicamentos
Capítulo 2.	Principios de Buena Práctica Clínica
Capítulo 3.	Responsabilidades, funciones y procedimientos de actuación de los Comités Éticos
Capítulo 4.	Responsabilidades del investigador
Capítulo 5.	Responsabilidades del promotor
Capítulo 6.	Protocolo
Capítulo 7.	Manual del investigador
Capítulo 8.	Documentos esenciales para la realización de un EC

Tabla 1-3 Índice del contenido de la directriz ICH E6 sobre BPC

Recuerde



Las guías ICH ofrecen un marco de referencia para el desarrollo y la investigación de nuevas intervenciones terapéuticas.

Guías de las revistas biomédicas

Las revistas más prestigiosas del campo biomédico han coordinado sus esfuerzos para ofrecer recomendaciones a sus autores que mejoren la calidad de sus presentaciones y permitan al lector evaluar apropiadamente la relevancia de los resultados presentados.

Se verá a lo largo de este curso que la estadística ofrece diferentes herramientas para obtener pruebas empíricas sobre el rendimiento de alternativas terapéuticas, diagnósticas y pronósticas.

Recuerde



Diferentes objetivos médicos requieren diferentes técnicas estadísticas.

Así, el método por excelencia para demostrar las mejoras producidas por una intervención es el ensayo clínico; pero para establecer el pronóstico de una determinada enfermedad se recurre a la regresión; y para evaluar la calidad de un procedimiento diagnóstico, a estimar las probabilidades de acierto.

Lectura



The Revised CONSORT (9) Statement for Reporting Randomized Trials: Explanation and Elaboration. Hay en la actualidad abundantes datos que demuestran que la calidad de la publicación de los ensayos clínicos controlados y realizados con asignación aleatoria (ECA) está por debajo de los niveles considerados óptimos. En estudios de carácter metodológico publicados recientemente se señala que su deficiente publicación y diseño se asocian a la aparición de sesgos en la estimación de los efectos de los tratamientos evaluados. Este tipo de error sistemático está deteriorando gravemente los EC. Por lo que su eliminación se ha convertido en un objetivo prioritario. El error sistemático en los EC refleja ciencia de bajo nivel que amenaza los principios éticos. Un grupo de científicos y editores elaboró la declaración CONSORT (Consolidated Standards of Reporting Trials) con el objetivo de incrementar la calidad en la redacción y publicación de los EC. En esta declaración se incluye una lista de ítems y diagrama de flujo que pueden utilizar los autores para la publicación de un EC. Muchas de las principales revistas médicas y de los grupos editoriales internacionales más importantes han adoptado la declaración CONSORT, que facilita el estudio e interpretación críticos de los EC, ofreciendo a los autores una guía para mejorar la publicación de los resultados obtenidos en sus estudios.

Ejercicio 1.10



Los 22 puntos de la guía CONSORT (tabla 1-4) requieren combinar habilidades y conocimientos tanto clínicos como estadísticos. Diga 3 que sean eminentemente clínicos y 3 estadísticos.

Ejercicio 1.11

¿Cuál es la diferencia más relevante entre las ICH y la CONSORT?

La tabla 1-4 muestra los puntos de la guía CONSORT que han de incluirse en el informe de un ensayo clínico aleatorizado, mientras que la figura 1-1 presenta el diagrama de flujo del progreso de los individuos a través de las fases del mismo.

Lectura



En diciembre de 2005, Medicina Clínica publicó un número extraordinario con las principales listas de comprobación para autores, revisores y editores de revistas médicas. (10)

	Punto n.º	Descriptor	Informado en pág. n.º
Título y resumen	1	Cómo se asignaron los pacientes a las intervenciones (p. ej. «asignación aleatoria» o «aleatorizado»).	
Introducción			
Antecedentes	2	Antecedentes científicos y razón de ser del estudio	
Métodos			
Participantes	3	Criterios de selección de pacientes y ámbito y lugares en que se recogieron los datos	
Intervenciones	4	Detalles precisos de las intervenciones pretendidas para cada grupo y cómo y cuándo efectivamente se administraron	
Objetivos	5	Objetivos e hipótesis específicos	
Resultados	6	Definición clara de las medidas de los resultados principal y secundarios y, cuando proceda, de cualquier método utilizado para mejorar la calidad de las medidas (p. ej., observaciones múltiples, entrenamiento de evaluadores)	
Tamaño muestral	7	Cómo se determinó el tamaño muestral y, cuando proceda, explicación de cualquier análisis intermedio y reglas de la interrupción del ensayo	
Aleatorización	8	Método utilizado para generar la secuencia de asignación aleatoria, incluyendo detalles sobre cualquier restricción (p. ej., bloques)	
Asignación oculta	9	Método utilizado para implementar la secuencia de asignación aleatoria (p. ej., contenedores de medicación numerados, aleatorización centralizada por teléfono), y clarificar si la secuencia se mantuvo oculta hasta el momento de la asignación	
Implementación	10	Quién generó la secuencia de asignación, quién incluyó a los participantes y quién los asignó a los grupos	
Cegado (enmascaramiento)	11	Si los participantes, quienes administraron las intervenciones o quienes evaluaron los resultados, conocían o no la intervención asignada. Si procede, cómo se evaluó el éxito del enmascaramiento	
Métodos estadísticos	12	Métodos estadísticos empleados para comparar los grupos en el resultado principal; métodos utilizados en análisis adicionales, como análisis de subgrupos o análisis ajustados	
Resultados			
Flujo de participantes	13	Flujo de participantes en cada fase (se recomienda un diagrama). Específicamente, para cada grupo, documentar los números de participantes asignados aleatoriamente, que recibieron el tratamiento pretendido, que completaron el protocolo del estudio y a los que se incluyó en el análisis del resultado principal. Describir las desviaciones del protocolo planificado y los motivos	
Reclutamiento	14	Fechas que limitan los períodos de reclutamiento y de seguimiento	
Datos basales	15	Características demográficas y clínicas basales en cada grupo	

Tabla 1-4 Lista de comprobación para publicar un ensayo clínico aleatorizado (11)

	Punto n.º	Descriptor	Informado en pág. n.º
Números analizados	16	Número de participantes (denominador) de cada grupo incluidos en cada análisis, y si el análisis se realizó «por intención de tratar». Expresar los resultados mediante números absolutos cuando sea factible (p. ej., 10/20 en lugar del 50%)	
Resultados y estimación	17	Para cada resultado principal y secundario, un resumen de resultados por grupo y el efecto estimado y su precisión (p. ej., intervalo de confianza del 95%)	
Análisis complementarios	18	Considerar la multiplicidad, e informar sobre cualquier otro análisis realizado, incluidos análisis de subgrupos y análisis ajustados, indicando los preespecificados y los exploratorios	
Eventos adversos	19	Todos los eventos adversos o efectos colaterales importantes en cada grupo de intervención	
Discusión Interpretación	20	Interpretación de los resultados, teniendo en cuenta las hipótesis del estudio, las fuentes de sesgo o imprecisión potenciales y los peligros asociados a la multiplicidad de análisis y de variables	
Generalización	21	Generalización (validez externa) de los hallazgos del ensayo	
Evidencia actual	22	Interpretación general de los resultados en el contexto de la evidencia	

Tabla 1-4 Lista de comprobación para publicar un ensayo clínico aleatorizado (Cont.)

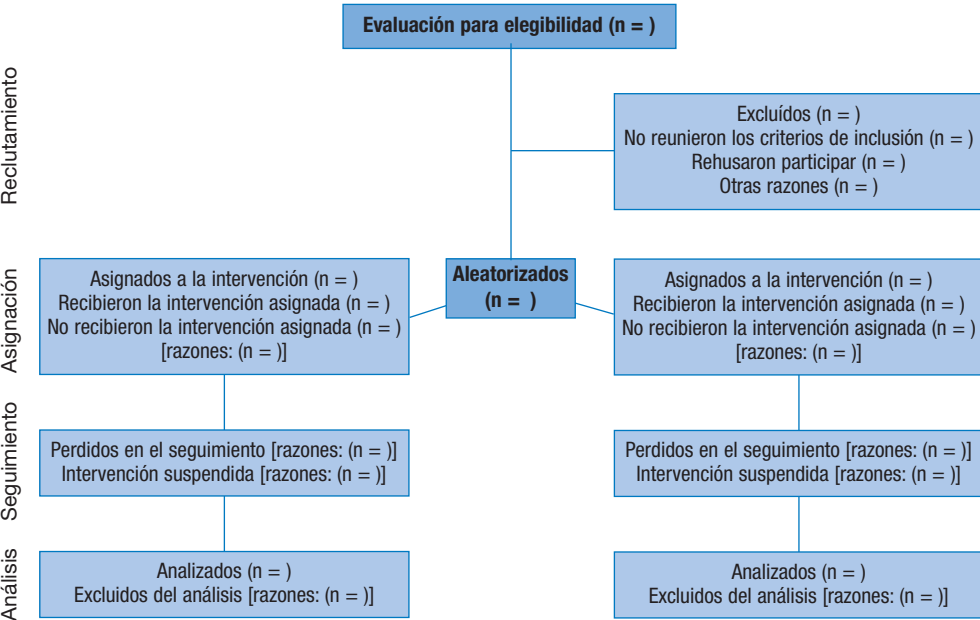


Figura 1-1 CONSORT: diagrama de flujo del progreso de los individuos a través de las fases de un estudio aleatorizado.

Ejercicio 1.12

Algunas de las «n» que figuran en el diagrama de flujo (fig. 1-1) de la CONSORT reflejan pérdidas que podrían comprometer la credibilidad de los resultados y que se estudian con detalle más adelante. Ahora, para empezar diga, para un estudio ideal, ¿cuáles de estas pérdidas le gustaría que fueran cero?

Soluciones a los ejercicios

1.1 La clasificación más habitual de los seres humanos es en masculino o femenino. Otro ejemplo posible de clasificación serían las especialidades médicas, siempre que no exista solapamiento entre las categorías. Para que sea una auténtica clasificación, las unidades deben pertenecer siempre a una categoría y solamente a una. (La teoría de los conjuntos borrosos –*fuzzy sets*– relaja este requisito y permite que una unidad pertenezca a un conjunto con cierta probabilidad.)

1.2 La clase social, en alta, media y baja es otro ejemplo de escala ordinal. Un *score*, tipo test de Apgar, entre 0 y 10, también cumple las propiedades de orden: un recién nacido con valor de 10 está mejor que otro con 9, y así sucesivamente.

1.3 El peso de un paciente está en escala de intervalo: la diferencia entre un sujeto con 65 y otro con 70 kg es la misma que la que existe entre otros dos, de 85 y 90 kg.

1.4 «¿Tiene cefalea?» está claramente en escala ordinal.

1.5 El número de cigarrillos que se fuma al día es más problemático. Para un comercial de una empresa tabacalera, tiene claramente unidad de medida y está en escala de intervalo, ya que su empresa obtiene el mismo beneficio cuando un fumador pasa de 20 a 21 cigarrillos al día, que cuando pasa de 0 a 1 cigarrillo al día. En cambio, para un neumólogo de Sanidad es muy diferente el incremento de riesgo que supone pasar de 0 a 5 cigarrillos/día que de 20 a 25. El profesional sanitario estará más cómodo con representaciones del tipo «no fumador», «fumador moderado» y «fumador severo». Ambos profesionales (el comercial y el neumólogo) están utilizando la misma variable observada (el número de cigarrillos por día) para «acercarse» a dos variables subyacentes (o latentes) muy diferentes: el beneficio comercial de la empresa o el riesgo potencial del paciente.

1.6 Desde el punto de vista de escala de medida, el recuento de CD4 posiblemente estaría en una escala cuantitativa de intervalo, con un mismo significado del incremento al pasar de 150 a 200 que de 550 a 600. Una primera dificultad aparece si el aparato de medida precisa un valor mínimo, pongamos 20, para poder detectar los linfocitos. Si fuera así, se trataría de una variable «censurada», en la que todos los valores inferiores a 20 han sido reconvertidos en un único valor «no detectado». De esta forma, se dispondría de una variable parcialmente de intervalo y parcialmente nominal u ordinal. Otra dificultad es si esta variable se pretende utilizar como indicadora de una variable subyacente, no directamente observable, como podría ser la evolución de ese paciente ante su enfermedad. ¿Aún significa lo mismo un incremento de 150 a 200 que de 550 a 600? ¿O de 375 a 425? Posiblemente no. Todo apunta a que debemos movernos desde la visión operativa de la medida hacia la visión representativa y preguntarnos, no por la escala, sino por la validez y la fiabilidad. Esta última será posiblemente alta en el sentido de que, si se repite la determinación de CD4, se obtienen valores similares. Pero esta fiabilidad será no tan alta si lo que se pretende que sea similar es la evolución, por lo que deberá matizarse también cómo se define la fiabilidad. En cuanto a la validez, se trata de estudiar cómo ayudan los

valores de CD4 a predecir esta evolución, lo que puede estudiarse, por ejemplo, con la ayuda de los términos sensibilidad y especificidad estudiados más adelante.

1.7 El examen MIR es menos válido, ya que mide la capacidad de contestar unas preguntas, no la de actuar profesionalmente. En cambio, es mucho más fiable, en el sentido de que si se repite la evaluación de un mismo individuo (con otras preguntas) se obtendrán puntuaciones mucho más similares, sea quien sea el evaluador, que si se cambia el paciente-caso o el examinador.

1.8 El responsable de Sanidad, que quiere medir el impacto en el auténtico objetivo sanitario, se preocupa por la validez. El investigador desea una variable más fiable, que tenga menor variabilidad aleatoria y permita, de esta forma, una muestra de menor tamaño.

1.9 Los derechos de los participantes y la fiabilidad de los datos.

1.10 Por ejemplo, las más clínicas son: 2, 4, 21 y 22. Y las más estadísticas son: 1, 5, 7, 8, 12, 15 y 18. El resto combinan habilidades de ambas disciplinas.

1.11 Las ICH son recomendaciones de las agencias reguladoras para facilitar la obtención de la autorización de uso comercial de una nueva intervención, usualmente farmacológica. La CONSORT, en cambio, recoge recomendaciones de las revistas científicas para la publicación de EC.

1.12 Todas. Como se verá más adelante, las pérdidas después de la asignación de la intervención (excluidos del análisis pérdidas de seguimiento y no recibieron la intervención) atentan a la validez interna y comparabilidad de los resultados, mientras que las pérdidas previas (rechazaron participar) atentan a la validez externa y generalidad.