
METODOLOGÍA DE LA INVESTIGACIÓN

¿Investigar en Osteopatía?

Ciencia: estudio de las relaciones funcionales entre objetos.

La ocurrencia de enfermedades (prevalencia, incidencia) y las características diagnósticas pronósticas o terapéuticas relacionadas no son constantes en la naturaleza.

La relación entre dichas características y la ocurrencia de variaciones en el estado salud/enfermedad.

La investigación es un proceso fundamentalmente empírico que se realiza a través de la recolección sistemática de observaciones.

I - ÁMBITO (SETTING, DOMAINE):

La base del estudio (experiencia concreta realizada en el estudio)
Hay que dar sentido a los datos y nos va a llevar datos empíricos.

$$Y = f(x)$$

- Diagnóstico= probabilidad
- Pronóstico= incidencia de muerte
- Etiológico= asociación a exposición
- Terapéutico= asociación a intervención

Componentes de la Función (relación a estudio).

Variable de resultado (out-come).

Cuya ocurrencia se estudia (mortalidad) y cómo se mide días- horas.

Parámetro de resultado: (eventos, tasa de mortalidad) verificación en la población.

Ámbito (setting, domaine): el tipo de situación que se estudia.

BASES DEL ESTUDIO (experiencia concreta realizada en el estudio)

- PLANOS CONCEPTUAL Y OPERATIVO
 - Entidad conceptual—operativa
 - Escala conceptual –operativa
- VALIDEZ y eficiencia de los métodos.
- GENERALIZACIÓN de la relación empírica observada.

1) Sujeto:

- Criterios de inclusión.
- Distribución de las variables.
- Tamaño de la muestra.

2) Duración del estudio: hasta la recogida de informaciones del resultado.

3) Obtención de la información (muestra, direccionalidad, temporalidad, calidad).

METODOLOGÍA E INVESTIGACIÓN



Diseño = Validez + Fiabilidad = Control
Metodología

a) Medición de variables:

Asignación de valores a las variables de interés en cada sujeto del estudio de acuerdo con reglas previamente explicadas.

b) Estimación de los parámetros poblacionales (no necesariamente humanos).

Confección matemática de un indicador sumario (estimador) que cuantifique la magnitud de las variables de interés o de la fuerza de su asignación (de su asociación).

b) Contraste estadístico de hipótesis:

Verificar hasta que punto el azar ha influido en los resultados de nuestro estudio (estimados).



A - PROCESO DE LA INVESTIGACIÓN MÉDICA

Empírico: instalado en la realidad temporo-espacial, necesita cuantificación o tratamiento numérico de los factores o variables involucrados en el estudio, a través de 3 procedimientos relacionados:

1) Características de los métodos de Investigación Biomédica:

La certeza: en la medición viene dada por 2 condiciones de una medida: la validez y la fiabilidad.

Validez: grado en que una medida mide realmente lo que quiere medir, aquellos para lo que está destinada, o en el plano más cualitativo, adecuación entre las variables seleccionadas y el concepto a medir.

Fiabilidad: grado en que una medida proporciona resultados similares cuando se repite en las mismas condiciones.

2) Preguntas de Investigación:

Fenómeno de incertidumbre en el proceso de explicar la realidad:

- Nace de la Experiencia
- Esfuerzo y búsqueda activa
- Conocimiento previo
- Provocado por la necesidad
- La clave consiste en no contestar todo a la vez
- No es un problema de escasez de ideas, sino de saber conformar una idea en un proyecto factible y válido.

Proyecto: esfuerzo intelectual para hacer operativa la hipótesis. Demostrar en pacientes...

Protocolo: descripción exhaustiva de los componentes. Se deriva del proyecto de investigación.

Requisitos, características metodológicas de fases y actividades.

Propuesta: documentos escritos investigados.

3) Estructura de un proyecto de Investigación:

- Presentación
- ¿qué es?
- ¿De qué se trata?
- Razonamiento del Proyecto
- Bibliografía

METODOLOGÍA E INVESTIGACIÓN

- Justificación: Pertenencia y factibilidad (¿disponemos de la capacitación necesaria?)
- No viable
- Muy amplia
- No hay sujetos suficientes
- Metodología muy compleja
- Muy caro
- Irrelevante
- No ética

B - ERROR ALEATORIO:

- Baja precisión: ninguno de los observadores coinciden en el resultado. Discoincidencia entre observadores en las mediciones.

- No error sistemático:

El valor medio de las observaciones se acerca al valor verdadero.

- Tamaño de la muestra:

Al aumentar el tamaño de la muestra disminuye el error aleatorio.

NOTA:

Una diferencia estadística de 1.5 es relevante.

El estudio piloto indica el porcentaje de mejoría que se va a estudiar.

C - ERROR SISTEMÁTICO:

- Alta precisión: los 5 observadores coinciden en el resultado. Hay coincidencias de los resultados pero son falsos.

- Error sistemático:

El valor medio se aleja mucho del valor verdadero.

- Tamaño de la muestra: No modifica el error.

El error aleatorio es fuente de variabilidad:

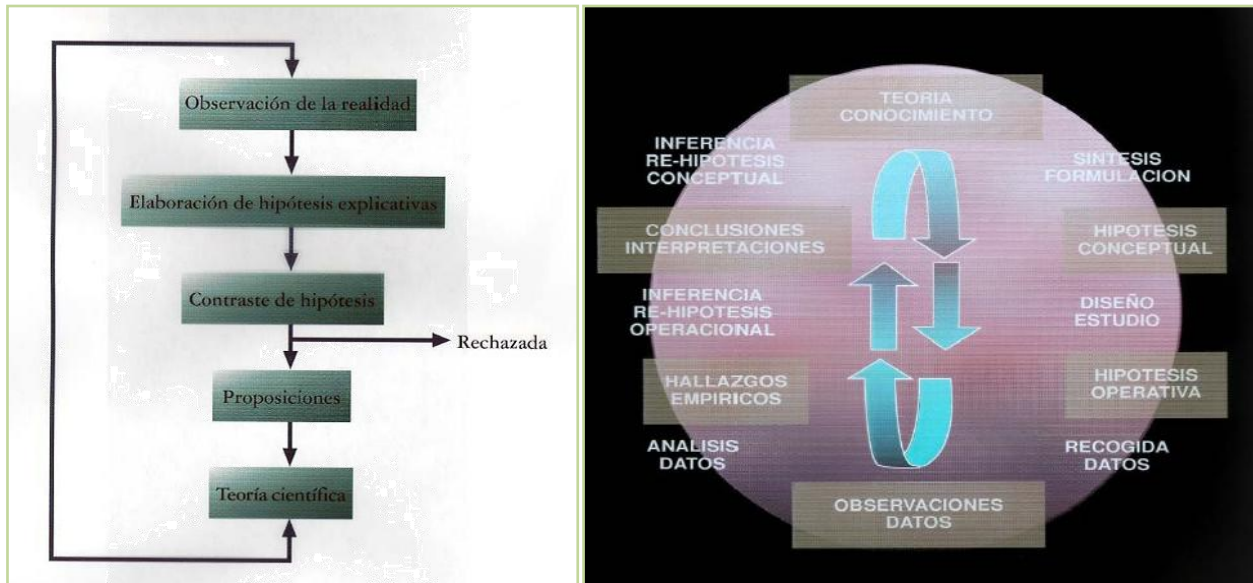
- Variación de la biología individual.
- Variación del observador.
- Variación del instrumento utilizado.

D - INVESTIGAR = MEDIR



METODOLOGÍA E INVESTIGACIÓN

- Exactitud (fiabilidad): compara con paciente de referencia.
- Precisión (Validez): Compara mediciones repetidas y ver su grado de acuerdo.



Conceptos básicos en investigación

II - PROYECTO- FORMA

1. Identificación del problema.
2. Razonamiento del problema.
3. Hipótesis y objetivos.
4. Estrategia.
5. Sujeto estudiados.
6. Variables.
7. Manejo de la información.
8. Plan de análisis.
9. Plan de acción.
10. Generalización.

VARIABLES:

**Es aquella característica que se mide en el proyecto de investigación.
Es lo que se mide o cuenta.**

A - VARIABLES INDEPENDIENTES

Características del paciente.

1. Edad.
2. Sexo.
3. Presencia de lesión: variable de categoría (Presente/Ausente).
4. Grupo de tratamiento.

B - VARIABLES DEPENDIENTES DE RESULTADO

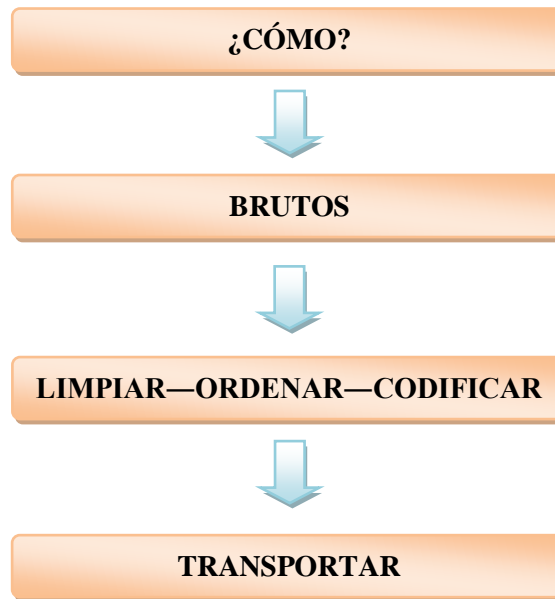
Al paciente se le aplica un tratamiento y se consigue un efecto, constatado en lo que medimos:

1. Dolor (Escala analógica).
2. Número de comprimidos de analgésico.
3. Ausentismo laboral.
4. Necesidad de reposo.
5. Grado de discapacidad.
6. Nivel de salud global.
7. Grado de satisfacción del paciente.
8. ...

Puede medirse múltiples días tras su inclusión.

Listado de variables:

METODOLOGÍA E INVESTIGACIÓN



¿Cuántas? Tantas como sean necesarias y tan pocas como sean posible.

- conceptual
- operativo
- pertinencia
- factibilidad
- viabilidad
- Datos: unidad de información que extraemos en el proyecto.

C - MEDICIÓN DE VARIABLES

Un estudio de investigación trata de obtener una respuesta válida y fiable a la pregunta planteada. Un estudio bien diseñado, ejecutado y analizado, fracasará si la información que se obtiene es inexacta o poco fiable.

Medir es asignar valores a las variables. Algunas son sencillas de medir. La sensibilidad de un microorganismo a un antibiótico, o el peso de una persona no presentan grandes dificultades de conceptualización ni de medición. Estas variables se denominan objetivas, ya que su definición es clara, deja poco margen a la interpretación por parte del sujeto o del investigador y se dispone de un buen instrumento para medirlas. Sin embargo, muchos de los problemas que se investigan en ciencias de la salud son difíciles de medir. La gravedad de la enfermedad, la intensidad del dolor, o el propio concepto de salud o de calidad de vida, son variables subjetivas, para las que no existe una definición ni un proceso de medición universalmente aceptados.

En cualquier estudio, el equipo investigador debe asegurar la validez y la fiabilidad de las medidas que utilice. La validez expresa el grado en que el valor que se obtiene con el proceso de medición se corresponde con el verdadero valor de la variable en un sujeto, es decir, que no existe un sesgo (error sistemático) en el proceso de medición.

La fiabilidad expresa el grado en que se obtendrían valores similares al aplicar el proceso de medición en más de una ocasión, concepto muy relacionado con el de variabilidad aleatoria.

1) FIABILIDAD

El término fiabilidad es sinónimo de precisión. Si, por ejemplo, el resultado de la medición de los valores de colesterol en un sujeto es de 180 mg/dl y acto seguido, una segunda medición efectuada con la misma técnica y en el mismo individuo, proporciona el resultado de 240 mg/dl, de inmediato se creará que el proceso de medición es poco fiable. Cuanta menor variabilidad se obtenga al repetir una medición, más fiable será el proceso.

El que una medida sea muy precisa no significa que tenga que ser válida. Un ejemplo sencillo sería pesar dos veces consecutivas a un mismo individuo

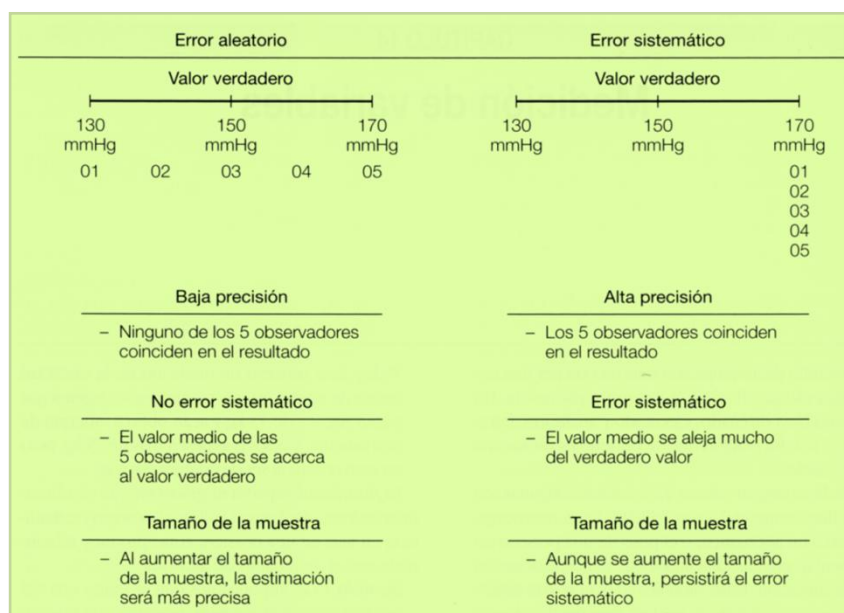


Figura Relación entre error aleatorio y error sistemático. Medición de la presión arterial sistólica de un paciente por cinco observadores independientes

2) FUENTES DE VARIABILIDAD

La fiabilidad de una medida no es una propiedad intrínseca de ella, sino más bien de un instrumento cuando se aplica en una muestra de individuos concreta, en unas condiciones dadas. La medición de un fenómeno clínico está sujeta a la variación biológica individual, del observador o del instrumento utilizado. Así, si dos médicos toman la presión arterial de un paciente y no obtienen el mismo resultado, puede ser debido a la propia variación de la tensión arterial entre una medición y otra, a que el esfigmomanómetro utilizado proporciona resultados variables, a que los clínicos difieran en la forma de medir y registrar la presión arterial, o a una combinación de todas ellas.

Aunque se describirá cada una de las fuentes de variación por separado, en la práctica todas ellas actúan de forma simultánea.

3) VARIACIÓN INDIVIDUAL

La mayoría de fenómenos biológicos varían de un momento a otro. Una medición realizada en un momento en el tiempo puede considerarse una muestra de todas las posibles mediciones de este fenómeno durante un período determinado de tiempo y no tiene por qué representar su verdadero valor.

Para minimizar la variación biológica intraindividual, una estrategia consiste en repetir la medición varias veces y utilizar alguna medida promedio de los resultados.

4) VARIACIÓN DEBIDA AL INSTRUMENTO

Por instrumento se entiende cualquier vehículo útil para recoger datos de una forma organizada, por ejemplo, un esfigmomanómetro, un autoanalizador o un cuestionario. Siempre pueden presentarse variaciones al aplicar un instrumento de medida y en la técnica empleada.

Por ejemplo, al tomar la presión arterial, la presión del brazal o la posición del estetoscopio pueden cambiar, lo que originará inconsistencias en las mediciones. Por ello, es importante estandarizar los procedimientos y seleccionar el instrumento que presente menor variabilidad.

5) VARIACIÓN DEBIDA AL OBSERVADOR

La tercera fuente de variabilidad es el observador. Las personas que llevan a cabo la medición de la presión arterial pueden variar, por ejemplo, el ángulo de observación, las condiciones en que realizan las mediciones, o aumentar su experiencia a medida que progresa el estudio.

Para reducir la variabilidad debida a los observadores es necesario entrenarlos debidamente antes de iniciar el estudio y elegir medidas lo más objetivas posible, ya que, cuanto más subjetivas sean, mayor es la probabilidad de que exista variabilidad de una medición a otra, y entre los distintos observadores. Por último, siempre que sea factible es conveniente evaluar la variabilidad inter e intraobservadores.

6) EVALUACIÓN DE LA FIABILIDAD

La fiabilidad se evalúa repitiendo el proceso de medición con el fin de analizar la concordancia de las distintas medidas. El índice más utilizado para ello, en caso de variables cualitativas, es el coeficiente kappa, que corrige la concordancia observada por la que se esperaría simplemente por azar. Si las variables son cuantitativas, se utiliza el coeficiente de correlación intraclass, que combina un análisis de la correlación de las medidas con una prueba de la diferencia de las medias entre ellas, o el método gráfico de Bland y Altman.

Deben estudiarse diferentes aspectos de la fiabilidad: la repetibilidad (fiabilidad test-retest), la concordancia intraobservador y la concordancia interobservador.

7) REPETIBILIDAD

La evaluación de la repetibilidad tiene por objetivo determinar si una prueba da los mismos resultados, o similares, cuando se aplica a una misma persona en más de una ocasión. Requiere que las condiciones de aplicación sean iguales en ambas ocasiones.

El hecho de aplicar la misma prueba en más de una ocasión en la misma muestra de individuos conlleva algunos problemas, ya que algunas características pueden variar con el tiempo. Es el caso de una medida sobre el estado físico o de salud de una persona; cuanto más tiempo transcurra entre ambas pruebas mayor es la probabilidad de que hayan ocurrido cambios reales en su estado de salud, o de que su opinión o actitud hacia un problema de salud haya cambiado. En esta situación, la fiabilidad de la prueba será infravalorada. Esta dificultad se puede subsanar, al menos parcialmente, efectuando la segunda medición poco tiempo después de la primera. Sin embargo, si el intervalo de tiempo entre la aplicación de las dos pruebas es corto, es posible que el resultado obtenido en la segunda ocasión esté influido por el aprendizaje adquirido la primera vez. En otras palabras, ambas mediciones no son independientes, y el coeficiente de fiabilidad estará artificialmente elevado.

Cabe la posibilidad de que los participantes no accedan a que se les efectúe una prueba en más de una ocasión en poco tiempo, en especial cuando sea dolorosa o molesta, o implique mucho tiempo de realización o desplazamiento. Ello motiva que se pierda mucha información, ya que si un individuo no ha completado las dos pruebas no será útil para el análisis.

8) CONCORDANCIA INTRA E INTEROBSERVADOR

La concordancia intraobservador se refiere al grado de consistencia al leer o interpretar un resultado de un observador consigo mismo. Al igual que con la repetibilidad, las dos medidas que realiza cada observador deben ser independientes una de otra, lo que puede ser muy difícil de conseguir cuando la observación de interés es un hallazgo físico que requiere la presencia directa del paciente, ya que el recuerdo de la primera exploración puede condicionar el resultado de la segunda. Esta dificultad queda soslayada si la observación es, por ejemplo, una radiografía, o un electrocardiograma.

Por concordancia interobservador se entiende la consistencia entre dos observadores independientes sobre una medida practicada en el mismo individuo. Para asegurar la independencia de las medidas entre ambos observadores, uno no debe conocer el resultado proporcionado por el otro (observadores ciegos).

Con frecuencia, la concordancia intra e interobservador se evalúa en un mismo estudio. En este caso, hay que asegurar la independencia de todas las medidas, para lo que puede ser muy útil aplicar técnicas como la aleatoriedad en la secuencia de aplicación de las medidas y las técnicas de enmascaramiento.

El análisis de la concordancia inter e intraobservador es, en muchas ocasiones, un paso previo a la validación de una nueva prueba diagnóstica, ya que, si los distintos observadores discrepan en los resultados, la prueba, aunque teóricamente pueda ser válida, tendrá poca utilidad clínica.

9) CONSECUENCIAS DE UNA MEDIDA POCO FIABLE

La fiabilidad de las medidas no se ha de analizar en cada estudio. Sólo es necesario cuando se desarrolle una nueva medida o como control de la calidad de las mediciones. Sin embargo, al planificar un trabajo de investigación es imprescindible considerar las posibles medidas alternativas de las distintas variables y escoger las que hayan demostrado ser más fiables en estudios anteriores.

Si la medición de la variable de respuesta tiene poca fiabilidad, la consecuencia directa es que la estimación del efecto que se obtenga será poco precisa, es decir, su intervalo de confianza será muy amplio. Este problema puede soslayarse aumentando el tamaño de la muestra, ya que el número de sujetos necesarios depende, entre otros factores, de la variabilidad de las medidas, aunque ello supone un incremento de la complejidad y coste del estudio.

10) VALIDEZ

El término validez se refiere al grado en que una variable mide realmente aquello para lo que está destinada. Es un concepto ligado al de error sistemático. Cuanto menos válida sea una medida más probabilidades han de cometer un sesgo.

Unas variables son más válidas que otras. Por ejemplo, la hemoglobina glucosilada representa mejor el grado de control de un diabético que una medición aislada de la glucemia. Los responsables del estudio deben procurar escoger siempre las medidas más válidas, en especial cuando se trata de las variables importantes del estudio.

11) FUENTES DE ERROR SISTEMÁTICO

Las posibles fuentes de error sistemático son las mismas que las enumeradas en el apartado destinado a la fiabilidad.

12) ERROR DEBIDO AL INDIVIDUO

Ocurre cuando el error en la medición de un fenómeno se introduce a causa de los participantes en el estudio. El ejemplo más sencillo es el sesgo de memoria. Los pacientes que sospechen que el problema de salud que padecen está relacionado con alguna variable en estudio, es posible que recuerden con mucha más exactitud su historia pasada de exposición que aquellos participantes libres de la enfermedad. Este diferente recuerdo se traduce en un error en la medición de la variable en estudio.

En otras ocasiones, los participantes pueden falsear de forma deliberada las respuestas simplemente para complacer al entrevistador o porque piensen que una determinada respuesta puede no estar bien vista socialmente, por ejemplo, cuando se interroga sobre los hábitos tóxicos de una persona.

13) ERROR DEBIDO AL OBSERVADOR

Ocurre cuando el observador mide una variable de forma errónea, por ejemplo, al utilizar una técnica inadecuada para la toma de la presión arterial (mala posición del brazal, tomarla en el brazo inadecuado, etc.) o al administrar un cuestionario de forma incorrecta.

En los estudios analíticos, puede ocurrir también cuando el observador mide una variable de forma sistemáticamente distinta en un individuo u otro, según el grupo de estudio al que pertenece. Si un investigador supone que los pacientes que reciben un tratamiento tienen con más frecuencia un determinado efecto secundario es posible que, incluso de forma inconsciente, lo busque con más insistencia (con preguntas con más frecuencia o con más pruebas complementarias) en aquellos que lo reciben que en los participantes del otro grupo.

14) ERROR DEBIDO AL INSTRUMENTO

Ocurre cuando la medida empleada no es la adecuada para el fenómeno que se desea analizar, o en el caso de que el instrumento utilizado no esté bien calibrado.

Ejemplo. En un estudio se revisaron los esfigmomanómetros de 9 centros de salud, lo que suponía una muestra de 44 aparatos de mercurio y 58 aneroides, comparándose sus lecturas con las de un aparato testigo debidamente certificado (Borrell et al, 1984). Se consideró que los aparatos funcionaban incorrectamente cuando tenían fugas de aire que producían un descenso de la columna de mercurio superior a 5 mmHg, aun con la válvula cerrada, registraban diferencias superiores o inferiores a 4 mmHg en los valores de 106 o 96, o mostraban este mismo error (4 mmHg) en dos de las cifras siguientes: 180, 160 o 146. El 13,6% de los aparatos de mercurio y el 67,2% de los aneroides funcionaban incorrectamente. Estos resultados ponen de manifiesto la necesidad de asegurar la validez de los instrumentos que se van a utilizar en una investigación.

15) EVALUACIÓN DE LA VALIDEZ

Para evaluar la validez de una medida se comparan sus valores con los obtenidos con una técnica de referencia objetiva, fiable y ampliamente aceptada como una buena medida del fenómeno de interés (validez de criterio).

Cuando la variable es dicotómica, los índices de validez de criterio más usados son la sensibilidad y la especificidad. La sensibilidad se define como la probabilidad de que una persona que presente la característica sea clasificada correctamente por la medida empleada en el estudio. La especificidad se define como la probabilidad de que una persona que no tenga la característica sea correctamente clasificada por la medida

utilizada. Para que una medida sea útil deberá ser lo más sensible y específica posible. Si la variable es cuantitativa, una forma de evaluar la validez es el análisis de las diferencias individuales entre los valores que proporciona la medida que se está utilizando y la obtenida con el método de referencia, sintetizadas en su media y desviación estándar.

A veces es difícil disponer de un criterio de referencia adecuado, debido a que no existe o porque no está al alcance del investigador. En estos casos, el procedimiento más empleado es evaluar la validez de constructo o de concepto, que analiza la correlación de la medida con otras variables que se cree relacionadas con ella (validez convergente) y la correlación con otras que se sospecha que no tienen relación alguna (validez divergente). Si la correlación es alta en el primer caso y baja en el segundo, puede considerarse que la medida tiene validez de constructo.

Otro aspecto a tener en cuenta es la evaluación de la validez de contenido de la medida, es decir, si contempla o abarca todas las dimensiones del fenómeno que se quiere medir. Por ejemplo, la concentración de triglicéridos puede ser una medida fácil de definir, fiable y con cifras muy exactas por parte del laboratorio. Sin embargo, puede no ser una medida válida de arteriosclerosis. Es conveniente abordar estos problemas en la fase de diseño de un estudio, ya que después ayudará a matizar los resultados.

16) CONSECUENCIAS DE UNA MEDIDA POCO VÁLIDA

La utilización de medidas no válidas conduce a una respuesta errónea a la pregunta de investigación. Si se trata de un estudio descriptivo que estime, por ejemplo, la frecuencia de una enfermedad, se obtendrá una prevalencia errónea. Si se trata de estudios analíticos, los sesgos de información conducirán a una estimación errónea del efecto del factor de estudio. En este tipo de estudios, conviene distinguir entre los errores diferenciales y los no diferenciales.

17) ERROR NO DIFERENCIAL

Es el error que ocurre al medir el factor de estudio y/o la variable de respuesta y que se produce por igual en todos los participantes, con independencia del grupo al que pertenecen o del valor de cualquier otra variable. En otras palabras, significa que cuando la variable es cualitativa, la sensibilidad y la especificidad para medirla permanecen constantes en todos los participantes, al margen de otros factores. Si la variable es continua lo que permanece constante es la diferencia entre la media observada y la verdadera media.

Ejemplo. Consideremos un estudio en el que un 60% de los individuos del grupo A desarrollan la enfermedad, mientras que en el grupo B, la incidencia es sólo del 20%. La verdadera razón de incidencias es igual a 3. Supongamos que la prueba que se aplica para detectar la enfermedad tiene una sensibilidad del 85% y una especificidad del 90% en ambos grupos.

Por ejemplo, en el grupo A se esperaría que de las 60 personas con la enfermedad, la prueba identificara correctamente a 51 ($60 \times 0,85$), y que 36 ($40 \times 0,90$) se clasificaran correctamente como que no la tienen.

De esta forma, la incidencia acumulada en el grupo A ha disminuido de un 60 hasta un 55%, mientras que en el grupo B ha aumentado de un 20 hasta un 25%. Esto es así porque cuando la incidencia o la prevalencia no son muy altas, como sucede en el grupo B, incluso una pequeña proporción de falsos positivos puede compensar una mayor proporción de falsos negativos.

Grupo A Resultados de la prueba	Enfermedad		Total	
	Presente	Ausente		
Positivos	51	4	55	Incidencia observada
Negativos	9	36	45	
Total	60	40	100	55%
<small>Incidencia acumulada verdadera: 60%. Sensibilidad de la prueba: 85%. Especificidad de la prueba: 90%.</small>				
Grupo B Resultados de la prueba	Enfermedad		Total	
	Presente	Ausente		
Positivos	17	8	25	Incidencia observada
Negativos	3	72	75	
Total	20	80	100	25%
<small>Incidencia acumulada verdadera: 20%. Sensibilidad de las pruebas: 85%. Especificidad de las pruebas: 90%.</small>			<small>Razón de incidencias verdadera: 3. Razón de incidencias observada: 2.2.</small>	

TABLA Estudio hipotético que ilustra un error sistemático no diferencial.

18) MEDICIÓN DE VARIABLES

Si se calcula la razón de incidencias, se observa que ha disminuido de 3 hasta 2,2 como resultado de una mala clasificación no diferencial.

Cuando se comparan dos grupos y se usa una medida poco válida que afecta a todos los participantes por igual, se introduce un error no diferencial que conduce a una infraestimación del verdadero efecto o asociación.

Cuando el efecto real es de gran magnitud, esta infraestimación no tiene una gran trascendencia. Sin embargo, en aquellos estudios que no encuentran un efecto, el investigador debe valorar la posibilidad de que la prueba empleada sea poco sensible o específica y que ésta sea la verdadera causa por la que no se ha encontrado una asociación.

19) ERROR DIFERENCIAL

Ocurre cuando la medición del factor de estudio y/o la variable de respuesta depende del valor de otra variable, y afecta de modo diferente a los participantes en el estudio según al grupo al que pertenecen.

Cuando la medida es poco válida, y la sensibilidad y la especificidad de la prueba son diferentes en cada grupo, la mala clasificación que resulta puede sesgar los resultados en

cualquier dirección. Una verdadera asociación puede quedar enmascarada, disminuida o aumentada, o bien, puede encontrarse un efecto que en realidad no existe.

Ejemplo. En un estudio similar al del ejemplo 14.7 se utiliza una prueba de detección de la enfermedad que tiene la misma especificidad en ambos grupos (98%), mientras que la sensibilidad es del 98% en el grupo A y sólo del 70% en el B. El resultado es que se sobreestima la verdadera razón de incidencias, que de 3 ha pasado a 3,75.

Grupo A Resultados de la prueba	Enfermedad		Total	
	Presente	Ausente		
Positivos	59	1	60	Incidencia observada
Negativos	1	39	40	
Total	60	40	100	60%
Incidencia acumulada verdadera: 60%. Sensibilidad de la prueba: 99%. Especificidad de la prueba: 98%.				
Grupo B Resultados de la prueba	Enfermedad		Total	
	Presente	Ausente		
Positivos	14	2	16	Incidencia observada
Negativos	6	78	84	
Total	20	80	100	16%
Incidencia acumulada verdadera: 20%. Sensibilidad de las pruebas: 70%. Especificidad de las pruebas: 98%.			Razón de incidencias verdadera: 3. Razón de incidencias observada: 3,75.	

TABLA Estudio hipotético que ilustra un error sistemático diferencial

Un ejemplo clásico de error diferencial es el sesgo de memoria en los estudios de casos y controles, en el que el conocimiento de la enfermedad influye sobre el recuerdo de la exposición. Otro ejemplo se encuentra en los estudios prospectivos en los que no se aplican o no es posible aplicar técnicas de ciego: el conocimiento de la exposición (o la intervención) puede condicionar la búsqueda, consciente o inconsciente, de un efecto en este individuo de forma diferente que en otro no expuesto. En ambos casos, la medición no se realiza de forma independiente y puede sobrevalorarse o infraestimarse la verdadera magnitud del efecto o asociación.

20) ESTRATEGIAS PARA AUMENTAR LA FIABILIDAD Y LA VALIDEZ

Las cinco primeras estrategias tienen efecto tanto sobre la validez como sobre la precisión, la sexta sólo mejora la fiabilidad, mientras que las restantes están destinadas a aumentar la validez de la medición.

1. Seleccionar las medidas más objetivas posibles. Si la variable de respuesta es, por ejemplo, la mortalidad total no existirán variaciones en su determinación por parte de los observadores. Si, por el contrario, es la aparición o el alivio de un síntoma, es más probable que la evaluación se lleve a cabo de forma dispar entre los observadores.
2. Estandarizar la definición de las variables. Es imprescindible que las variables estén definidas de forma operativa en el protocolo o en el manual de instrucciones,

de forma que todos los investigadores utilicen los mismos criterios incluso ante situaciones dudosas.

3. Formar a los observadores. Su entrenamiento no permitirá controlar totalmente la variación entre un observador y otro, pero sí reducirla.

4. Utilizar la mejor técnica posible. Por ejemplo, si se tiene la posibilidad de escoger entre un aparato anerode y otro de mercurio para medir la presión arterial, se elegirá este último ya que es más fiable.

5. Utilizar instrumentos automáticos. Si es factible, y el instrumento automatizado es más preciso, permitirá reducir la variabilidad interobservador.

6. Obtener varias mediciones de una variable. El promedio de varias medidas tiende a ser más preciso que una medida aislada. El ejemplo más clásico es la presión arterial: dado que es una medida muy variable, la práctica más habitual es tomarla en más de una ocasión y obtener la media de las distintas mediciones. De la misma forma, se obtendrá una estimación más precisa de un fenómeno como la calidad de vida si se realizan varias preguntas relacionadas con el tema en un cuestionario. Esta estrategia sólo mejora la precisión, no la validez.

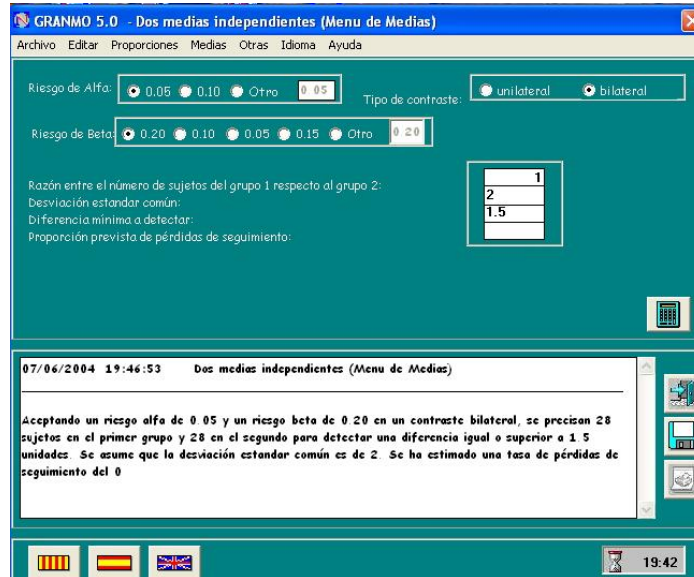
7. Emplear técnicas de enmascaramiento (ciego). En los ensayos clínicos controlados con placebo donde es posible utilizar las técnicas de enmascaramiento, y más concretamente las de doble ciego, es muy difícil incurrir en un error diferencial, ya que ni los observadores ni los participantes conocen la intervención que están recibiendo. En algunos estudios observacionales es factible mantener ciegos a los observadores pero no así a los participantes. En esta situación se elimina el error potencial que se deriva de las expectativas de los observadores, pero no el que puede provenir de los participantes. De todos modos, las técnicas de ciego no evitan los errores no diferenciales.

8. Calibrar los instrumentos. El ejemplo correspondiente a un estudio sobre la calibración de los esfigmomanómetros, ilustra lo que puede ocurrir si no se toma esta precaución. Hay que calibrar los instrumentos con el fin de evitar medidas sesgadas.

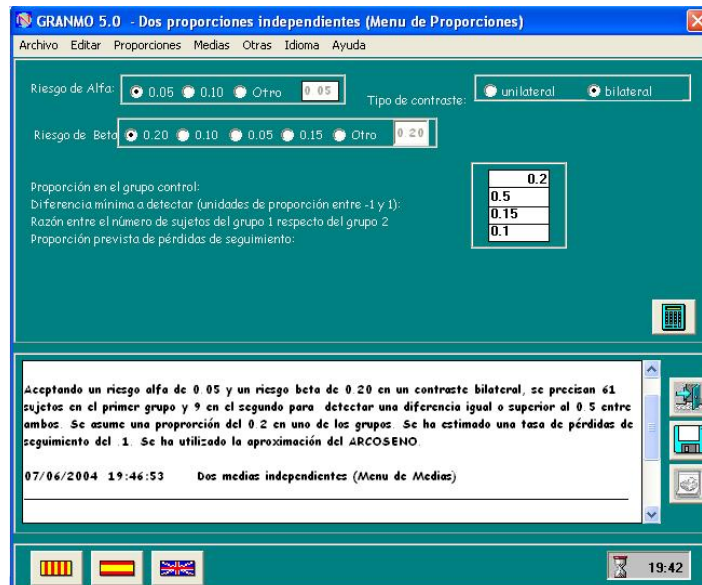
La puesta en marcha de todas las estrategias simultáneamente es muy difícil. No obstante, la estandarización de las variables y la formación de los observadores deben realizarse siempre.

La aplicación del resto de estrategias dependerá de las características de las variables que se están estudiando, el diseño empleado y la disponibilidad de recursos.

III - CÁLCULO DEL TAMAÑO DE LA MUESTRA CON GRAMMO 5.0



Permiten calcular las medias: la desviación estándar se basa en el estudio piloto.



IV- LOS DATOS DE LA INVESTIGACIÓN:

A - OBTENCIÓN: ya explicado

B - ANÁLISIS:

1) DATOS PRIMARIOS:

- Observación directa.
- Encuestas.

2) DATOS SECUNDARIOS:

Fuentes de información:

- Historial clínico
- Registros
- Certificados

3) SELECCIÓN DEL MÉTODO:

Aquél que con más certeza responde a nuestro objetivo.

Certeza + validez + fiabilidad

a) Síntesis de datos:

- Dar sentido a los datos crudos.
- No perder información.
- hacerlos manejables.
 - Observación de tablas y gráficos
 - Estadística descriptiva: Según el resultado primordial esperado determina el procedimiento estadístico: cualitativas- categóricas.
- Descripción univariante
- Variable cuantitativa
- Medidas de tendencia central
 - Media-mediana-moda.
 - Medidas de dispersión.
 - Varianza, desviación, estándar.

- Percentiles.

b) Descripción numérica

Descripción de un fenómeno en el tiempo:

- Tasa.
- Serie cronológica.

c) Estimación parámetros población error estándar: intervalo de confianza

Margen de fluctuación, todo procede de una muestra que no es representativa de la población

4) DESCRIPCIÓN DE LA RELACIÓN ENTRE DOS VARIABLES:

a) Dos variables cuantitativas:

- Correlación, regresión.
- Categorización.

a) Dos variables cualitativas:

- Tablas 2 por N.
- Relacionar frecuencias relativas y absolutas entre categorías.

c) 1 cualitativa y otra cuantitativa:

Medidas de tendencia central, dispersión

d) Análisis Bivariante:

- Test de hipótesis.
- Verificar si existe asociación estadística.
- Identificación de asociaciones.

e) Análisis estadístico:

- Test de hipótesis y fuerza de asociación.
- Influencia de factores extraños.
- Estratificación.
- Consistencia.
- Valoración de sesgos.

- Según el tipo de variables

C - VALIDEZ INTERNA.

Es el grado en que los cambios observados se pueden atribuir a la manipulación experimental. Estudia hasta que punto una causa puede ser atribuida a un efecto. Ej.: Ensayo clínico: tiene el máximo grado de validez interna.

Teniendo en cuenta la validez interna de mayor a menor grado los diseños los podemos clasificar en los siguientes grupos:

1. Experimentales auténticos: Verdaderos, puros, pues no tienen problemas de validez interna (True Desing).
2. Cuasiexperimentales: No se pueden descartar la presencia de variables de confusión, pues no es posible eliminarlas todas. El investigador sabe que A es causa de B, pero no está seguro que A también pueda ser causa de otros factores como C ó D.
3. No experimentales: Están cerca de los anteriores en cuanto a validez interna, aunque presentan más variables de confusión, pueden ser:
 - 3.1. Longitudinales: (Prospectivo / Retrospectivo)
 - 3.2. Transversales.

Cuántas más variables entran en un diseño van restando validez interna.

Las variables de confusión afectan al diseño, forman parte de las amenazas a la validez interna.

AMENAZAS A LA VALIDEZ INTERNA.

1. **HISTORIA.** Hay amenaza de historia, cuando hay acontecimientos externos que ocurren simultáneamente con éste y que pueden alterar o influir. Ej. : Programa educativo ---- Problemas de T.A. ---- Disminución de la T.A. Pero al margen de este programa los individuos practican gimnasia y se sigue una dieta sana, estos factores influyen en los resultados.
2. **SELECCIÓN.** Cuando los grupos de estudio son diferentes. Ej. Raza, sexo, ...
3. **MADURACION.** Son los cambios producidos por evolución natural. Tiene relevancia en salud y confunde el efecto del cambio de la variable con el de la causa. Ej. : Una herida mejora hagamos o no hagamos nada, pero ¿cuánto depende la mejoría de la herida de lo que hemos hecho sobre ella?.
4. **EFFECTOS RELATIVOS DEL PRE-TEST.** Es la influencia que produce el pre-test. Ej.: Si medimos la T.A. en un barrio y posteriormente aplicamos el programa y volvemos a medir. ¿Cuánto del cambio es por el programa y cuánto por ser la primer a vez que se le mide la T.A.?.
5. **MORTALIDAD.** El que desaparezcan sujetos de los grupos de comparación. No sabemos qué sujetos se pierden, unos se mueren y otros se van.

METODOLOGÍA E INVESTIGACIÓN

6. **INSTRUMENTACION.** Uso de instrumentos no fiables ni validos.
7. **REGRESION ESTADISTICA.** Los sujetos seleccionados representan situaciones o puntuaciones en alguna variable. Cuando se usan sujetos extremos. Sucede cuando para probar los efectos algo se escogen a los sujetos más extremos. Ej.: Para probar los efectos de una dieta seleccionamos a los más gordos.

El tamaño muestral afecta a la validez interna.

En la validez interna se pueden observar dos tipos de errores: El error aleatorio y el error sistemático.

- **ERROR SISTEMATICO O SESGO:** Pérdida de validez o de exactitud. El error sistemático aparece cuando existen diferencias entre lo que el estimador está realmente estimando (O°) y el parámetro verdadero, medida verdadera que pretendemos conocer (O). Este error es atribuible a la forma en que se seleccionan los sujetos, la calidad de la información obtenida o a la presencia de otras variables diferentes al factor de estudio y la enfermedad o efecto de interés.
- **ERROR ALEATORIO:** Pérdida de Fiabilidad o de Precisión. El error aleatorio aparece cuando existen diferencias entre el estimador (\hat{O}) obtenido con nuestros datos y el valor del parámetro que está siendo realmente estimado (O). Dicho error es, en esencia, atribuible a variación en el muestreo, (tamaño muestral, varianza).

O, O°, \hat{O}

Error aleatorio	Error sistemático
<p style="text-align: center;">x</p> <p>130 mmHg 150 mmHg 170 mmHg</p> <p>01 02 03 04 05</p>	<p style="text-align: center;">x</p> <p>130 mmHg 150 mmHg 170 mmHg</p> <p style="text-align: right;">01 02 03 04 05</p>
<p>Baja precisión</p> <p>Ninguno de los 5 observadores coinciden en el resultado.</p>	<p>Alta precisión</p> <p>Los 5 observadores coinciden en el resultado.</p>
<p>No error sistemático</p> <p>El valor medio de las 5 observaciones se acerca al valor verdadero.</p>	<p>Error sistemático</p> <p>El valor medio se aleja mucho del verdadero valor</p>
<p>Tamaño de la muestra</p> <p>Al aumentar el tamaño de la muestra disminuirá el error aleatorio.</p>	<p>Tamaño de la muestra</p> <p>Aunque se aumente el tamaño de la muestra, persistirá el error sistemático.</p>

Bases Metodológicas de la Investigación en Ciencias de la Salud 2003

D - VALIDEZ EXTERNA.

Es el grado en que los resultados de un estudio pueden ser generalizados a muestras o condiciones espacio-temporales diferentes. Ej. "A" causa "B", pero seguiría causando "B" con otros:



- Sujetos.
- Contexto ---- validez ecológica.
- Momentos.

Los estudios descriptivos (encuestas) son los que más se preocupan por la validez externa.

La validez externa está afectada por los siguientes aspectos:

- Por la variable independiente. Es el nivel de operacionalización del estudio.
- “Efecto Rosenthal”: es el efecto derivado de las expectativas, es decir, el efecto derivado de que se presupone o se espera que ocurra, cuando algo se espera un efecto favorece que se produzca. Afecta tanto a la variable interna como a la validez Externa.
- “Efecto Hawthorne”: son las expectativas que el sujeto tiene sobre si mismo, es el efecto de la autoexpectativa.

En el Efecto Rosenthal las expectativas se reflejan en el otro sujeto, mientras que el Efecto Hawthorne es el producido por las expectativas del sujeto sobre si mismo.

E - CLASIFICACIÓN DE LOS TIPOS DE SESGOS GENÉRICOS

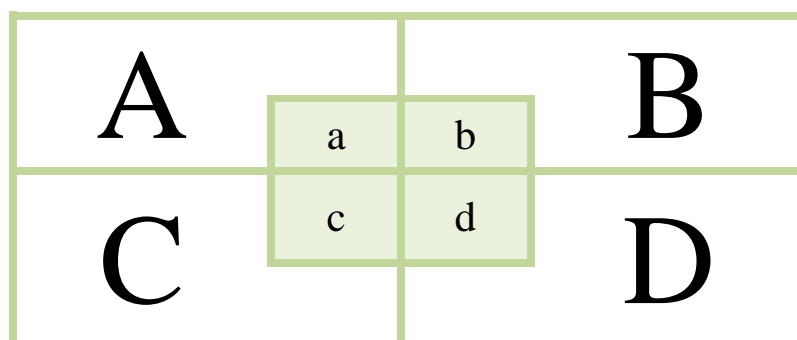
1) SESGO DE SELECCIÓN:

Es una distorsión en la estimación de un efecto causado por la forma en que se han seleccionado los sujetos de estudio.

Como ejemplos:

- Pérdidas en el seguimiento
- Controles inadecuados
- Supervivencia selectiva

Sesgo de Selección:

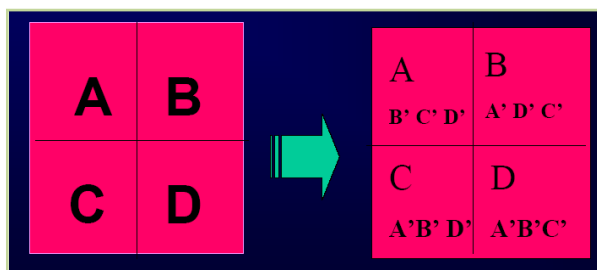


2) SESGO DE INFORMACIÓN (MALA CLASIFICACIÓN):

Es una distorsión en la estimación de un efecto debido a una medición errónea o mala clasificación de los sujetos de estudio con respecto a una o varias variables.

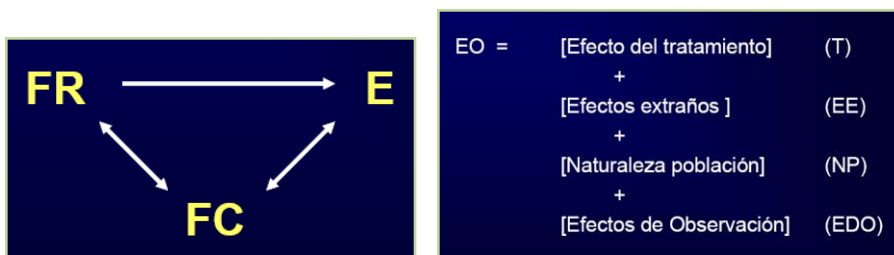
Ejemplos:

- Aparatos o instrumentos defectuosos
- Cuestionarios o registros incorrectos
- Criterios diagnósticos incorrectos
- Cualquier fuente de datos errónea.



3) SESGO DE CONFUSIÓN:

- Distorsión debida a que el efecto del factor de estudio está mezclado en los datos con los efectos de otros factores extraños al estudio.
- Asociación del factor de confusión, al factor de estudio y al efecto o enfermedad.



4) EFECTOS EXTRAÑOS:(EE)

- Alteración de la identidad de la intervención:
 - Interferencia de eventos.
 - Cambio endógeno.

- Tendencia secular.
- Comparabilidad de efecto extraño.
- Medios para intentar corregirlo:
 - Específico actividades primordiales.
 - Idéntico de recurso.
 - Asigno responsabilidad.
 - Cronograma.
 - Ejecución.

5) NATURALEZA DE LA POBLACIÓN: (NP)

- Pérdida de idoneidad
- Selección incontrolada:
 - autoselección.
 - pérdidas.

6) EFECTOS DE OBSERVACIÓN: (EDO)

- Pérdida de objetividad:
 - Efectos estocásticos.
 - Placebo.
 - How torne (sentirse observado).
 - Falta de validez.
 - Falta de fiabilidad.

No hay forma de aislar los efectos de la intervención sin comparación.

- Control aleatorio.
- Control reflexivo.
- Control caso-testigo.
- Control histórico.
- Otros...

A) COMPARABILIDAD DEL EFECTO EXTRAÑO:

Medios para intentar corregirlo:

- Distribución aleatoria.
- Comparación con tratamiento de placebo.
- Utilización de técnicas de enmascaramiento.

B) COMPARABILIDAD DE LAS POBLACIONES:

Medios para intentar corregirlo:

- Restricción simple.
- Apareamiento (Matching).
- Distribución aleatoria (Randomización).
- Medición y control analítico de los factores pronósticos.

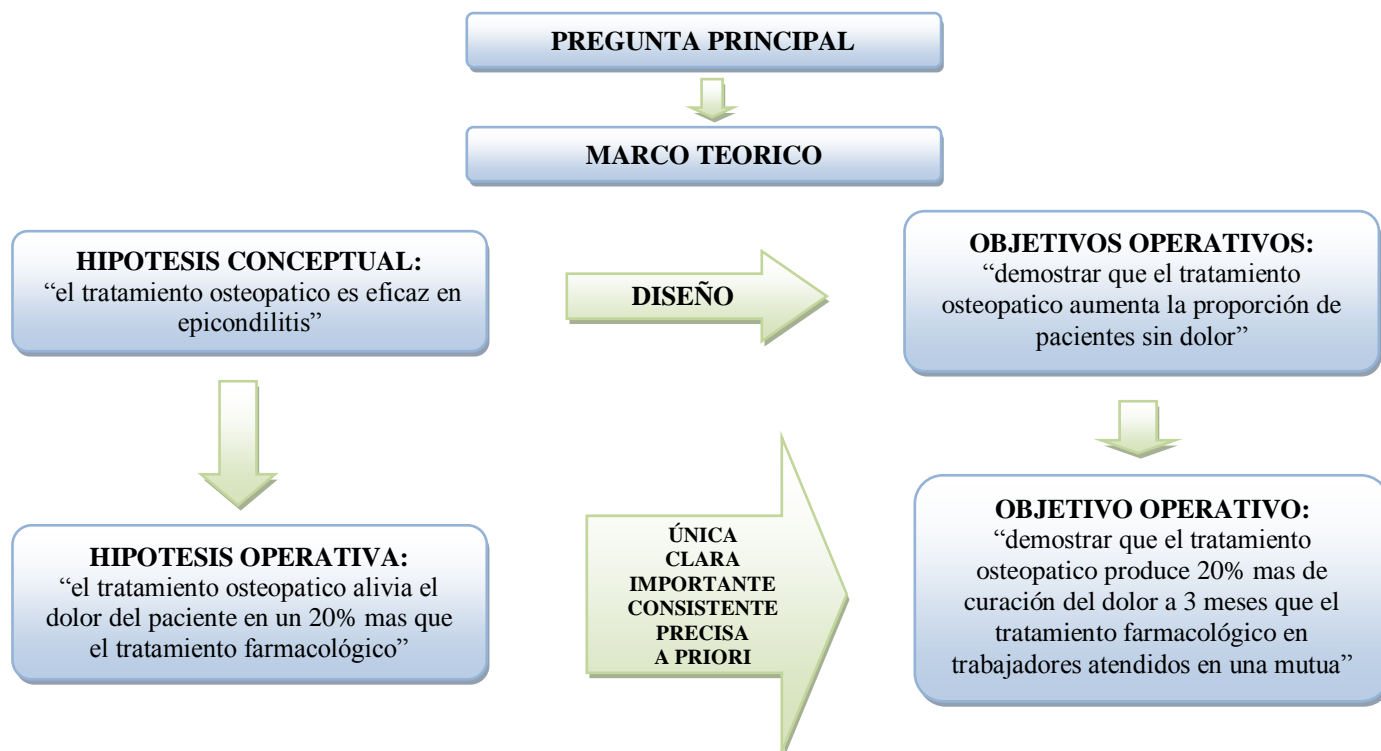
c) COMPARABILIDAD DE LAS OBSERVACIONES:

- OBJETIVIDAD; INDEPENDENCIA
- Técnicas de enmascaramiento (ciego)
- Ensayo pragmático: No posible aumentar.
- Ensayo explicativo: Paciente y observador ciego
- Hard vs. Sofá endpoints.
- Estandarización, protocolización.
- Sistematización, ENTRENAMIENTO

F - HIPÓTESIS

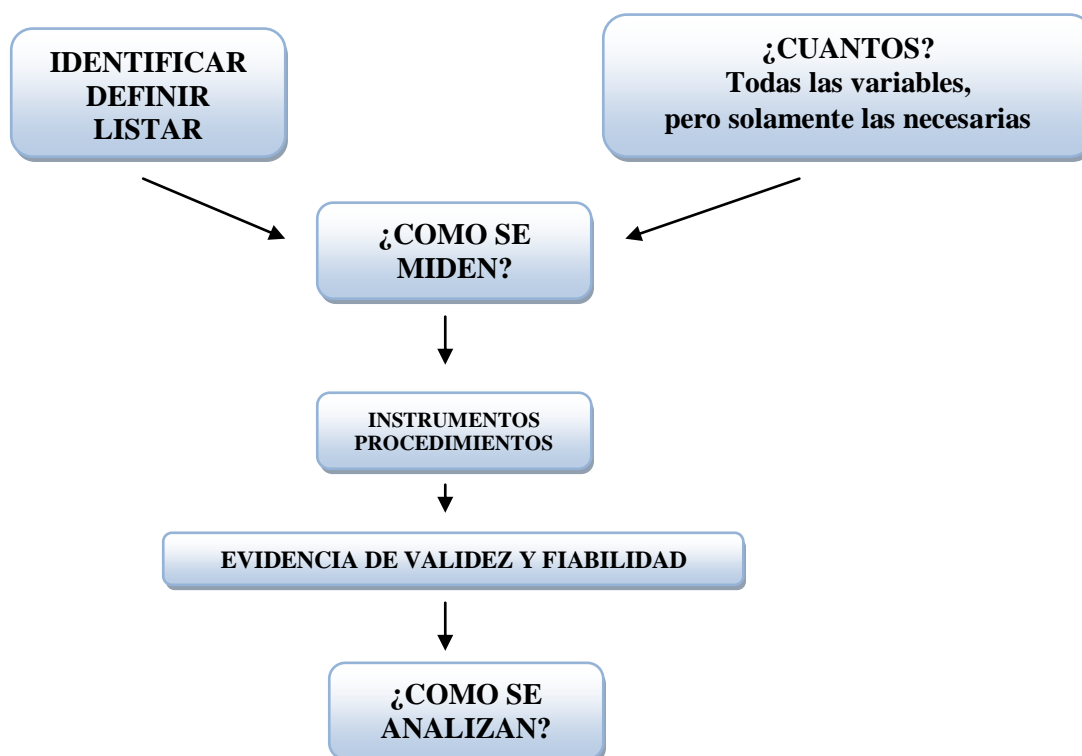
Es el enunciado formal de las relaciones entre al menos una variable dependiente X y otra independiente.

Preguntas secundarias e hipótesis de subgrupos.



G - VARIABLES DEL ESTUDIO

Es lo que se va a medir en el proceso de investigación.



Son las características de los sujetos del estudio que se deben medir o contar en la investigación. Cada individuo se diferencia de otro por una serie de características cualitativas o cuantitativas: a esas características las denominamos variables. Las variables son todos los conceptos operacionales, y la mayor parte de las investigaciones clínicas son conducentes a comprender por qué varían las funciones orgánicas en una determinada situación clínica y por qué aparece ésta.

1. Las variables que se refieren a fenómenos no apropiados para ser cuantificables a menudo pueden ser medidas clasificándolas en categorías; son las variables categóricas, que pueden ser de tres tipos:

a) Variables dicótomas: son las variables categóricas en las cuales sólo hay dos valores posibles: Hombre/Mujer, Sano/Enfermo, etc.

b) Variables nominales: son las variables categóricas que no implican un orden en la clasificación; por ejemplo, provincia de origen: Lérida, Granada, o los diversos fenotipos del grupo sanguíneo ABO (A, B, AB, O). Las variables nominales tienen un carácter cualitativo y absoluto que los convierte en fáciles de medir, pero las opciones de análisis estadístico son limitadas.

c) Variables ordinales: son variables categóricas en las cuales la clasificación implica un orden de intensidad definido, con intervalos que no son cuantificables.

Por ejemplo: dolor leve, moderado y grave, o bien, estadio evolutivo de una enfermedad: estadio I, estadio II, etc.

2. Variables numéricas, en las cuales hay intervalos perfectamente cuantificables. Hay varios tipos de variables numéricas:

a) Variables continuas: son fenómenos que se pueden contar o medir con una escala numérica continua, como la talla, la glucemia, etc., y se les denomina variables discretas si tienen un número limitado de valores; por ejemplo, el número de cigarrillos consumidos por día.

b) Variables continuas de proporciones o razones, tal como un porcentaje, las cuales son consideradas óptimas por los estadísticos.

Las variables pueden ser atributos del individuo estudiado (peso, talla, glucemia, etc.) o ser variables activas creadas artificialmente por el investigador (fármaco A/fármaco B).

En un estudio analítico, si se investiga la relación entre dos o más variables para predecir unos resultados o inferir una causalidad, la variable que precede a la otra se llama variable predictora o independiente, y a la variable resultante se le llama dependiente. Esta condición de independencia o dependencia es relativa y propia de cada estudio concreto. En un estudio observacional puede haber varias variables independientes y varias dependientes. Las variables independientes o predictoras, cuando son múltiples, pueden influir de un modo perturbador sobre los resultados, y se les denomina perturbadoras o intervinientes. Para eliminar su efecto hay que corregir los resultados para las mismas o emplear análisis multivariable.

La variable predictora o independiente en un estudio experimental (terapéutico) es el tratamiento efectuado (intervención), y en el estudio de un método diagnóstico es el resultado del test (+/-). La variable dependiente en el estudio experimental (terapéutico) es el resultado del mismo (curación, alivio de los síntomas, etc.); en este caso también la denominan algunos autores, variable de criterio o variable de respuesta, y en la investigación de un medio diagnóstico es la presencia o ausencia de enfermedad.

La medida de las variables consiste en traducir a números los fenómenos observados para poder someterlos a un tratamiento estadístico. La validez externa de un estudio depende en una buena parte del grado en que las variables diseñadas para el estudio representan los fenómenos de interés, y la validez interna depende de cómo las medidas efectuadas representan a estas variables. En cualquiera de estos aspectos puede haber errores: así, puede haber errores de muestreo y de medida. El acto material de la investigación clínica es medida de variables. Si no hay una objetivación de los fenómenos clínicos para su medida o cuantificación, no es posible la investigación clínica. Si no existe un procedimiento o instrumento objetivo de medida o cuantificación debemos crearlo y validarlo antes de iniciar la investigación.

Tipos de variables:

a) Variables numéricas, cuantificables en una escala aritmética: por ejemplo, la talla, el peso. Se las denomina variables discretas si tienen un número entero de valores, por ejemplo, el número de hijos, el número de plaquetas, etc.

b) Variables continuas de proporciones o razones, tal como un porcentaje.

c) Variables categóricas: Se refieren a fenómenos que no se pueden cuantificar ni medir y pueden ser dicótomas, nominales u ordinales.

- Variables dicótomas cuando tienen sólo dos posibilidades: respuesta terapéutica/no respuesta.
- Variables nominales: cuando no se pueden ordenar: por ejemplo, grupos sanguíneos A, B, AB, O.
- Variables ordinales: como el dolor grave, moderado o leve, el estadio evolutivo de una enfermedad, etc.

H - EL CUADERNO DE RECOGIDA DE DATOS:

Los valores del cuadro de datos se tratan con programas de estadísticas (SPSS).

I - CUANTIFICACIÓN DEL PROYECTO DE INVESTIGACIÓN

¿Que quiero hacer?

- ¿Describir?
- ¿Medir asociaciones?
- ¿Medir impactos? Ejemplo: 60% de los pacientes que reciben manipulaciones mejoran.

RAZÓN:

Es el:

- Resultado de dividir una cantidad por otra cantidad. Son cantidades distintas (nº de mujeres /de hombres, incidencia E/incidencia E-). Es además un concepto adinámico.
- La incidencia de lumbalgia en trabajadores de fuerza.
- Concepto de magnitud relativa.

$$\text{Rango} = 0 \dots \dots \dots \alpha.$$

PROPORCIÓN:

- Es la Razón o coeficiente en el cual el numerador está incluido en el denominador.

$$a / a + b$$

TASA:

- Es la medida del cambio de una variable (y) cuando cambia otra variable (x).

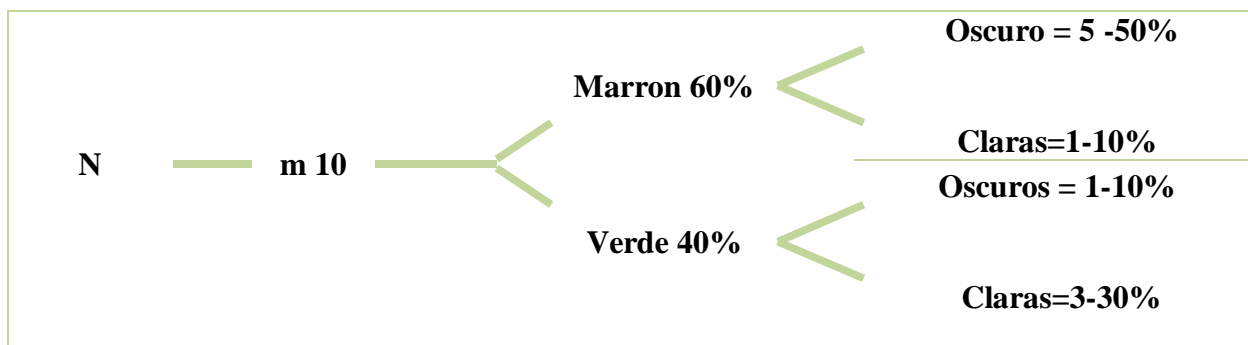
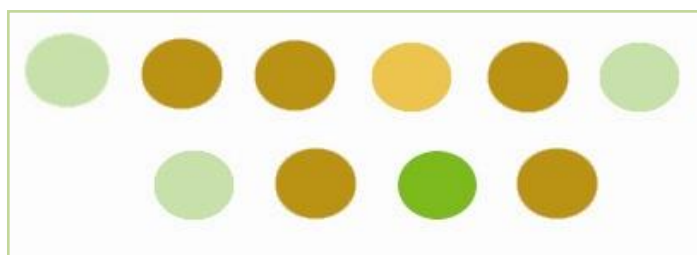
Tasa de promedio de cambio:

$$T = \frac{\Delta y}{\Delta x}$$

El cambio puede ser lineal o exponencial: es la relación entre 2 variables.



PREVALENCIA: (%).



	Oscuro	Clara	
Verde	1a 10%	3b 30%	4 n ₁
Marron	5c 50%	1d 10%	6 n ₀
	6 m ₁	4 m ₀	

Proporción de prevalencia:

- $P_1 = a / n_1 = 25\%$
- $P_0 = c / n_0 = 80\%$
- $P = m_1 / n = 60\%$

Razón de prevalencia:

$$PR = P_1/P_0 = 0,25/0,8 = 0,3$$

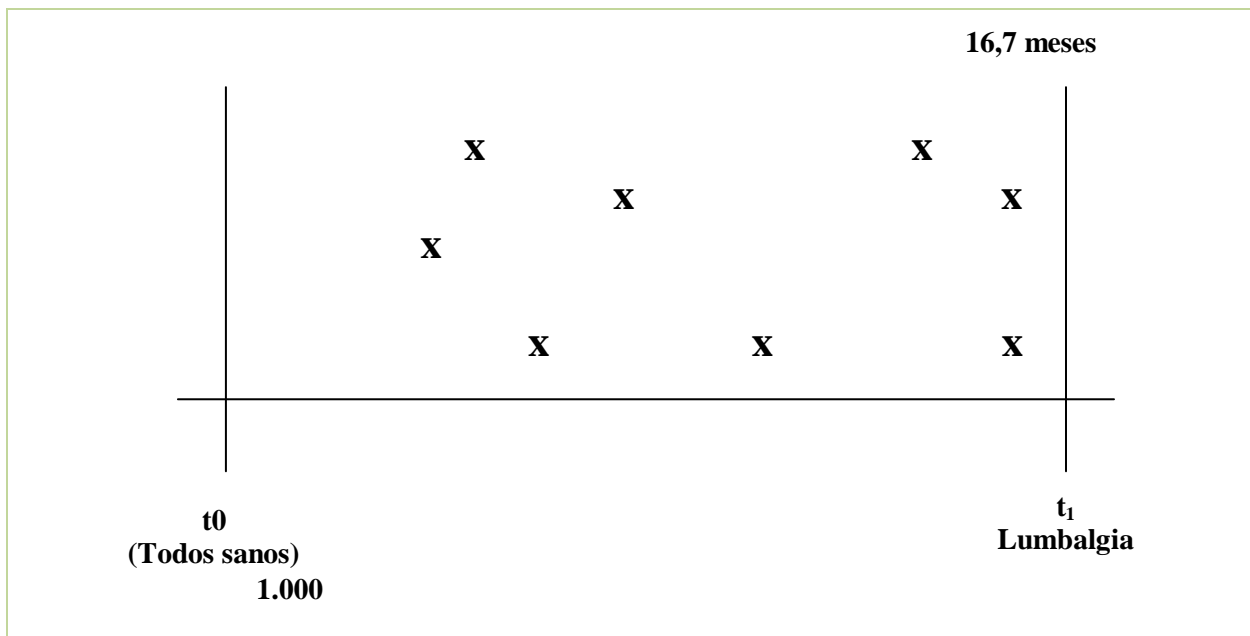
Es la prevalencia puntual, la probabilidad de que un individuo presente lumbalgia.

C_t = Casos existentes en el momento t .

$$P_t = \frac{C_t}{N_t}$$

N_t (Número de personas en la población en el momento t).

CAMBIO EN LA SITUACIÓN:



$9/1.000/16,7$ meses.

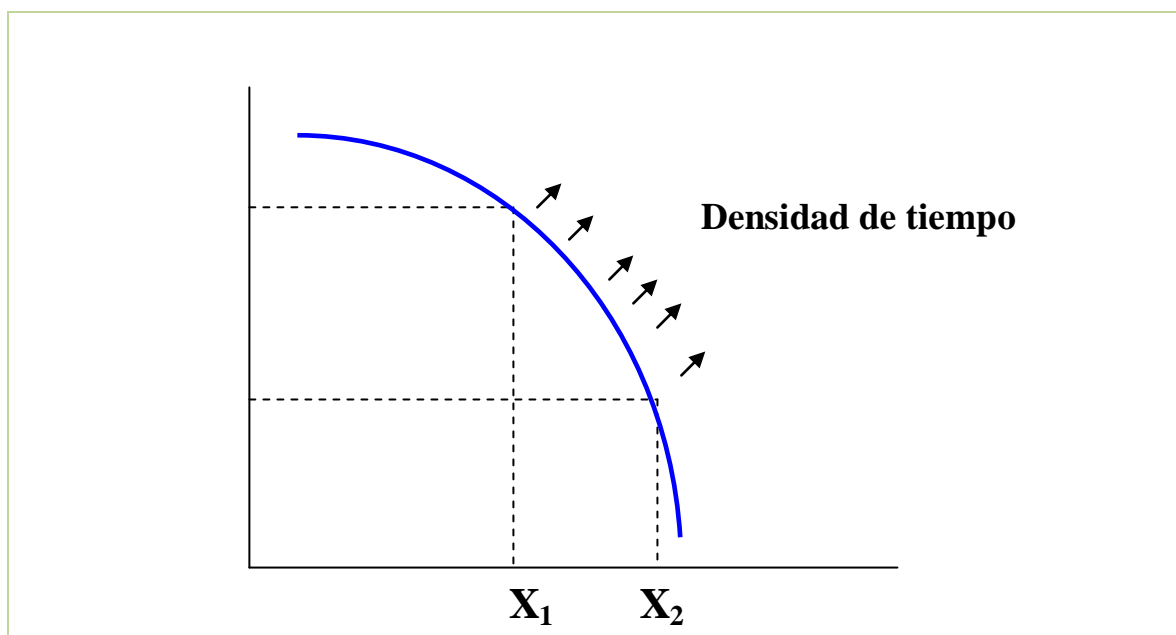
$\frac{9}{1.000}$ = Incidencia acumulada: noción de cambio en el tiempo.

CÁLCULO DEL RIESGO: MÉTODO ACUMULATIVO SIMPLE.

Es la incidencia acumulada (IC), la proporción de sujetos que desarrollan la enfermedad durante el periodo (t_0, t) .

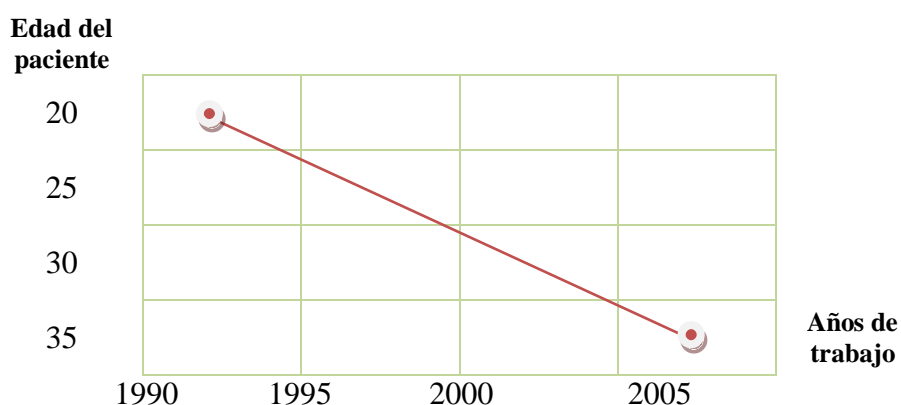
$$IC(t_0, t) = \frac{I(\text{casos aparecidos durante el periodo } t_0, t)}{N(\text{Tamaño de la población libre de enfermedad en el momento } t_0)}$$

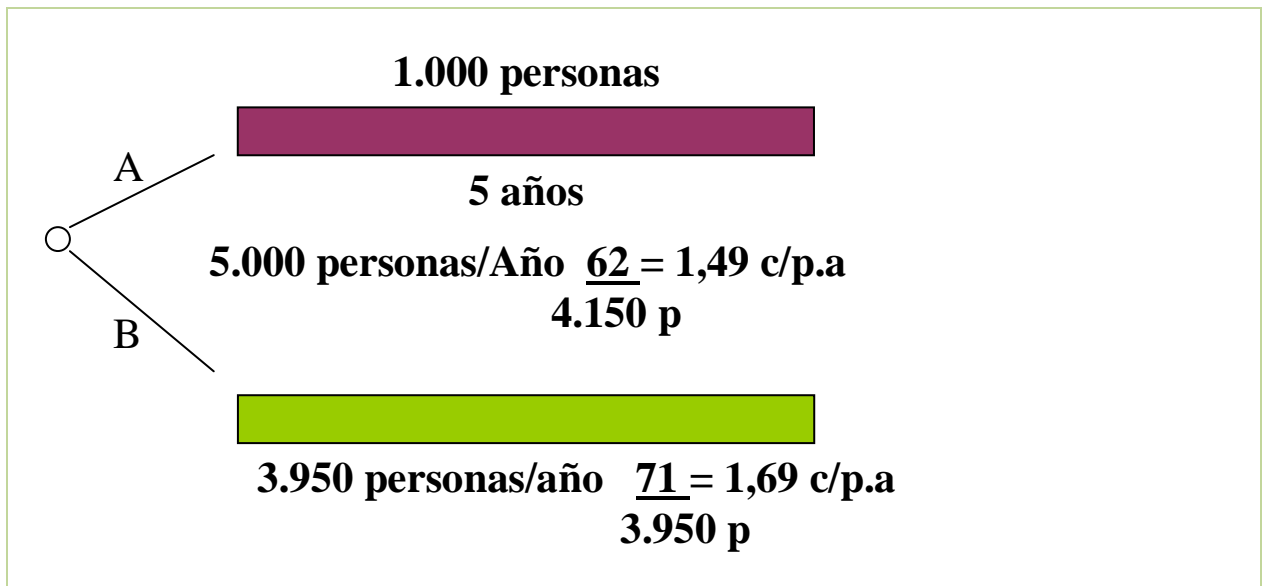
Tasa de incidencia: (¿a qué velocidad enferman?) Es la tasa de incidencia.



$$TI = \frac{\text{No enfermos}}{\sum t_{IE}} \text{ (Tiempos individuales de exposición).}$$

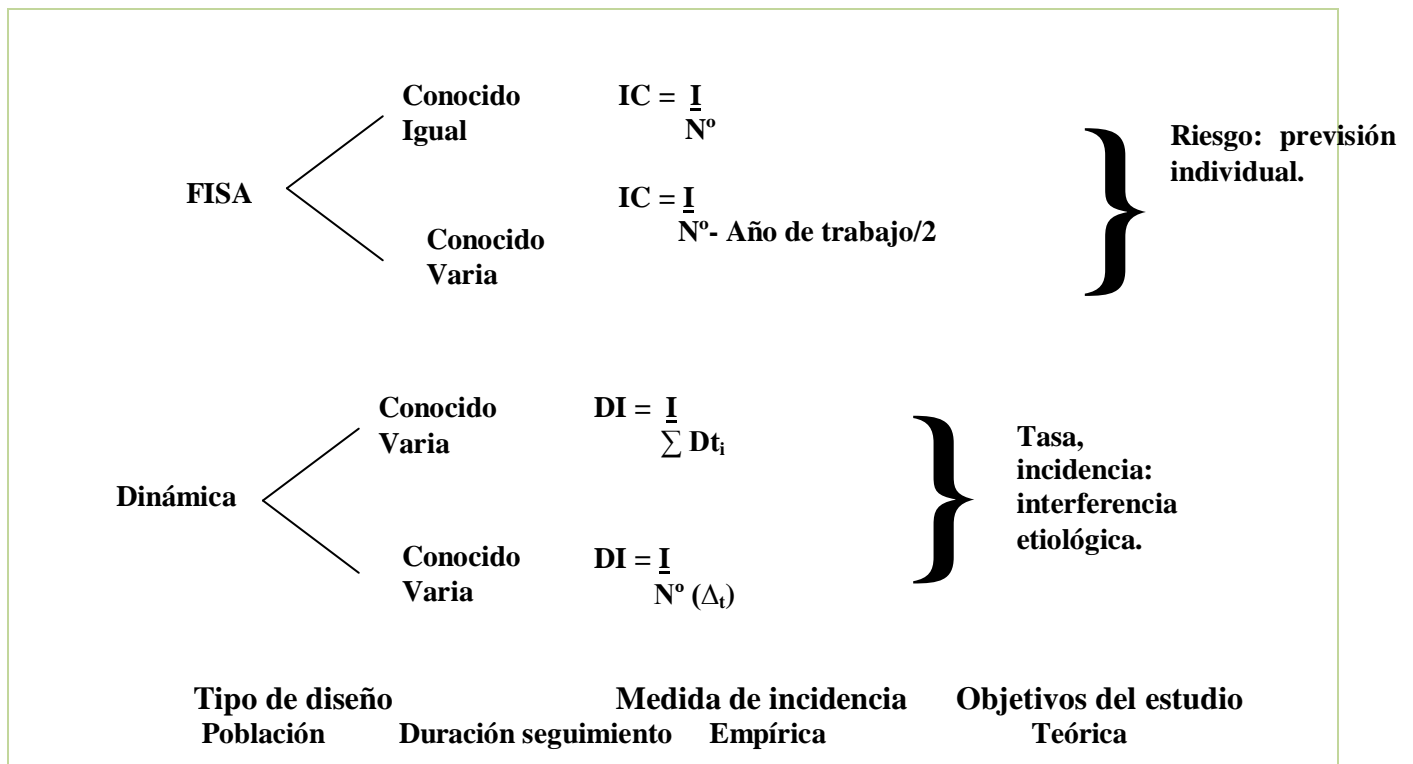
	INCIDENCIA	PREVALENCIA
Definición	Frecuencia de aparición de casos nuevos en un periodo dado.	Número de casos existentes en un momento dado.
Diseño	Estudios longitudinales	Estudios transversales
Expresión	Análisis de los factores de riesgo.	Análisis de la situación de salud.
Utilidad	Evalúa la interferencia del tratamiento, la aparición de sucesos...	Planificación, organización de recursos.





LO IMPORTANTE ES LA CALIDAD DE LA INFORMACIÓN

- Riesgo relativo: compara la probabilidad, es una medida de asociación.
- Es la razón de la incidencia de dicho resultado entre los expuestos frente a la incidencia entre los individuos no expuestos.
- Riesgo relativo: no hay asociación >1(más de 1) factor de riesgo menos de 1(es un factor protector).
- Con 2 incidencias se averigua el riesgo relativo.
- Para calcular la incidencia se calcula el tiempo y el número de personas.



EVALUACIÓN DE LA MEDICIÓN:

1) DEFINICIÓN DE CASOS:

- *Por eliminación clínica:*
 - o Rigor de los criterios diagnósticos.
 - o Puesta a punto de un test diagnóstico.
- *Revisión del historial clínico:*
 - o Información veraz y exhaustiva.
- *Entrevista personal:*
 - o Sesgos de recuerdos.

2) CONTANDO EL NÚMERO DE CASOS:

- Contar personas /casos.
- Cuando empezamos a contar.

3) DEFINICIÓN DE LA POBLACIÓN:

- Denominador: excluir libre de riesgo.
- Población de estudio con población general.
- Muestra aleatoria siempre que se pueda.

NOTA: SIEMPRE SE DEBE DEFINIR BIEN LA PATOLOGÍA.

MEDIR ASOCIACIONES:

El riesgo relativo mide la probabilidad de resultados entre individuos.

RR = Incidencia de enfermedad entre expuestos y no expuestos.

	Casos E+	Casos E-		Lumbalgico	No lumbalgico
F +	<i>a</i>	<i>b</i>	a + b	30	70
F -	<i>c</i>	<i>d</i>	c + d	10	90
				40	160

$$N = a + b + c + d$$

$$RR = \frac{P(E+/F+)}{P(E+/F-)} = \frac{a/(a+b)}{c/(c+d)} = \frac{30/100}{10/100} = 3$$

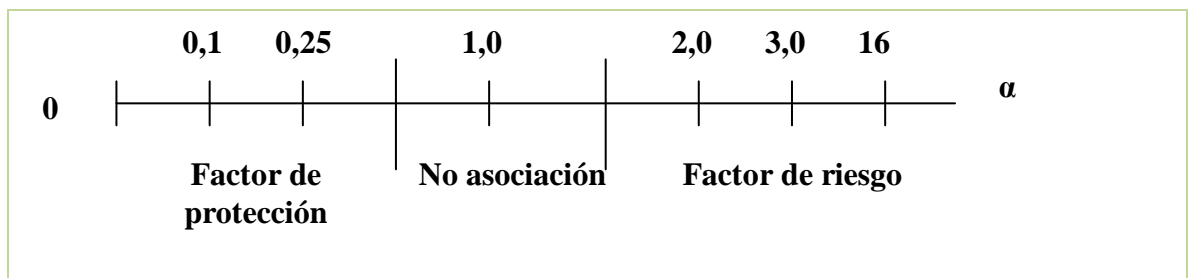
La exposición a un trabajo duro aumenta por 3 el riesgo de lumbalgia.

MEDIR ASOCIACIONES:

1) MEDIDAS PARA:

- Odds ratio en estudio caso-control.
- Razón de mortalidad estandarizada en cohortes (SMR).
- Razón de mortalidad proporcional (PMR).

2) ESCALA DE RR (RIESGO RELATIVO)



Un factor de riesgo de 2-3 ya es muy importante.
Un factor de protección de 0,3 es muy importante.

Los estudios que permiten estudiar un riesgo son:

- Los estudios de cohorte.
- Los estudios de ensayo clínico.

V – ESTUDIOS DE CASOS DE CONTROLES:

No hay periodo de seguimiento.

Se utilizan Casos (Grupo de estudio) y Controles (Grupo testigo).

En los estudios de casos y controles se elige un grupo de individuos que tienen un efecto o una enfermedad determinada (casos), y otro en el que está ausente (controles). Ambos grupos se comparan respecto a la frecuencia de exposición previa a un factor de riesgo (factor de estudio) que se sospecha que está relacionado con dicho efecto o enfermedad.

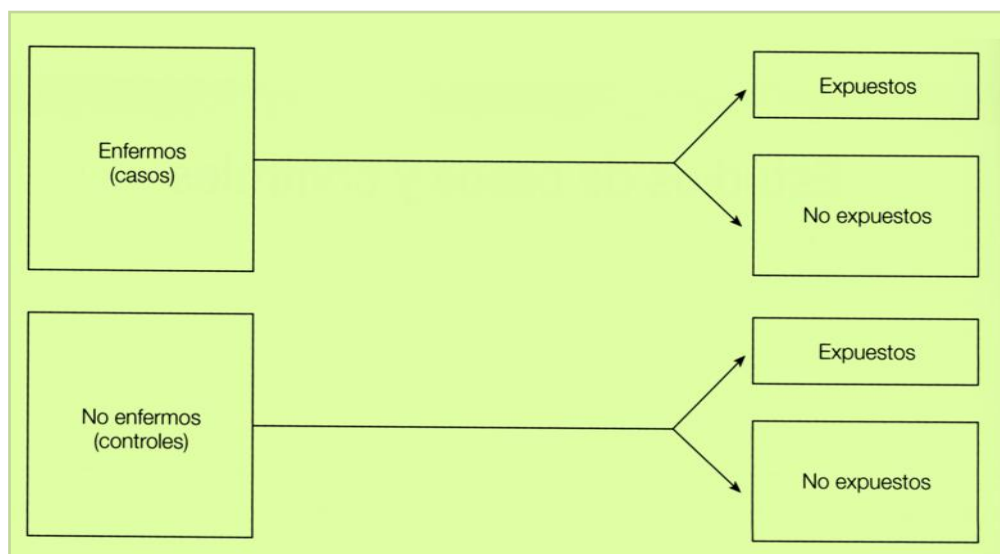


Figura Estructura básica de un estudio de casos y controles.

La necesidad de un grupo control es evidente. El hallazgo, por ejemplo, de que el 47% de las mujeres diagnosticadas de fibroadenoma ha utilizado anticonceptivos, suscita de inmediato la siguiente pregunta: ¿Esta proporción de exposición es superior, igual, o inferior que la esperada? La función del grupo control es estimar la proporción de exposición esperada en un grupo que no padece la enfermedad.

Las estimaciones que se obtienen son la proporción de casos y de controles expuestos a un posible factor de riesgo. También es de interés la intensidad y duración de la exposición en cada uno de los grupos. La medida de asociación o del riesgo de padecer un determinado problema de salud asociado a la presencia de una exposición es la odds ratio (OR).

El hecho de que los casos y los controles provengan de poblaciones diana distintas y de que la información se recoja de forma retrospectiva facilita la introducción de sesgos.

A - SELECCIÓN DE LOS CASOS

Para identificar los casos deben establecerse, de forma clara y explícita, la definición de la enfermedad y los criterios que deben cumplir aquellos que la presenten para ser incluidos en el estudio. Por otro lado, los criterios de selección deben estar dirigidos a que sólo se incluyan sujetos que potencialmente han podido estar expuestos al presunto factor de riesgo. Estos criterios deben aplicarse por igual a casos y a controles.

Lo ideal sería que los casos seleccionados fueran una muestra aleatoria de todos los individuos de la población de estudio que presentan la enfermedad. En los estudios de cohortes la enfermedad se busca de forma sistemática en todos los participantes, pero en los estudios de casos y controles, los casos se obtienen de una serie de pacientes en quienes la enfermedad ya ha sido diagnosticada y están disponibles para el estudio, por lo que pueden no ser representativos de la población de estudio.

B - DEFINICIÓN DE CASO

Los investigadores pueden estudiar formas leves y/o graves de la enfermedad. Si se incluyen casos con todo el espectro de gravedad, existe el riesgo de clasificar mal a individuos sin la enfermedad como casos leves y viceversa, dado que en muchas enfermedades es muy difícil conseguir un diagnóstico de certeza. Si se incluyen sólo formas graves, el riesgo de clasificar mal disminuye, aunque obtener el número de individuos necesario puede ser difícil y, además, se limita la generalización de los resultados.

Ejemplo .El síndrome del túnel carpiano puede presentarse con sintomatología diversa, no siempre fácil de diferenciar por criterios únicamente clínicos, de las acroparestesias nocturnas. Los signos de Tinel y Phalen positivos pueden ser de gran ayuda al reproducir la sintomatología, pero sólo tienen valor de sospecha diagnóstica. Según la gravedad del cuadro existe una mayor o menor afectación sensitiva o motora, no siempre objetivable a la exploración física. Por esto, para conseguir un diagnóstico preciso y fiable de la afectación del nervio mediano, deberán llevarse a cabo pruebas electrofisiológicas.

En el caso de que no sea posible aplicar en todos los casos las mejores técnicas para el diagnóstico, porque son demasiado caras o no están disponibles, es conveniente clasificarlos en categorías tales como «probable» o «definitivo». La elección de unos criterios más o menos estrictos para definir la enfermedad tiene implicaciones en los resultados del estudio. Si se aplican criterios diagnósticos poco sensibles y específicos, se hallarán muchos falsos positivos y negativos, lo que tiende a diluir el efecto que se está estudiando. Según este esquema es de esperar que en el grupo con diagnóstico «definitivo» se encuentren menos falsos positivos que en el de «probable», lo que permitirá analizar con más detenimiento los posibles errores de una mala clasificación.

C - IDENTIFICACIÓN DE LOS CASOS

Una vez se ha elaborado una definición conceptual del caso, el siguiente paso es desarrollar una definición operativa para identificarlos. El objetivo es conseguir una identificación en la que todos los casos tengan la misma probabilidad de ser incluidos en el estudio y válida, es decir, que no se incluya ningún individuo que no pertenezca al caso.

D - SELECCIÓN DE LOS CONTROLES

La elección del grupo control es la clave de este tipo de estudios ya que debe ser comparable al de casos (los controles deben tener la misma probabilidad de haber estado expuestos que los casos). En la mayoría de estudios, los casos y los controles proceden de dos poblaciones

distintas (desde un punto de vista epidemiológico, no desde un punto de vista geográfico), pero lo que debe buscarse es que puedan ser considerados como dos grupos procedentes de una misma población. En realidad, lo importante es que los controles sean representativos de la población de donde provienen los casos.

Conceptualmente esta población la forman los miembros de una cohorte subyacente definida por unos criterios de inclusión y exclusión marcados por el investigador.

Cuando la detección de todos los casos de una población no es posible, no puede definirse temporal y geográficamente la cohorte subyacente de dónde proceden. Esto sucede con frecuencia cuando se trata de enfermedades con síntomas leves que no llevan al individuo a solicitar atención médica como, por ejemplo, en un estudio sobre las causas de la infertilidad en los hombres, ya que ésta sólo habría sido detectada en aquellos que hubieran tratado de tener descendencia y buscado atención médica.

En esta situación, los controles deberían ser representativos de aquellos individuos que, en el supuesto de desarrollar la enfermedad, hubieran sido detectados como casos. Si no se restringe la población de acuerdo a estos supuestos, se podría cometer un sesgo de selección relacionado con alguna variable ligada al hecho de buscar atención médica.

E - CRITERIOS DE SELECCIÓN DE LOS CONTROLES

La selección de los controles no está relacionada con la exposición en estudio. En la práctica, hay que excluir del grupo control a los pacientes que tengan enfermedades relacionadas positiva o negativamente con la exposición en estudio.

Igualmente, con el fin de evitar que los controles puedan concentrarse en alguna enfermedad relacionada con la exposición, conviene escoger controles con distintas enfermedades, con el fin de minimizar un posible sesgo debido al muestreo.

F - NÚMERO DE CONTROLES POR CASO

Cuando existe un número de casos suficiente se suele seleccionar un control por cada uno. Cuando el número de casos es limitado se puede aumentar la potencia estadística del estudio para detectar un efecto determinado y seleccionar más de un control por cada caso.

G - NÚMERO DE GRUPOS CONTROL

Para valorar la posibilidad de que se haya cometido un sesgo en la selección de los controles, algunos autores recomiendan, si es posible, utilizar dos grupos control. Si se obtiene la misma estimación al comparar con cada uno de ellos se podría asumir que no ha existido un sesgo de selección.

Ventajas y desventajas de controles de la población general o la población demandante:

Población demandante:

- Más fáciles de identificar
- Tienden a recordar mejor su historia de exposición
- Más cooperativos, por lo que el número de no respuestas es menor
- Al estar enfermos es más fácil que tengan algún factor de riesgo relacionado con la enfermedad de estudio
- Más fácil de mantener al entrevistador ciego al grupo al que pertenece cada individuo

Población general:

- Más tiempo y dinero para su identificación
- Más probabilidad de que se produzca un sesgo de memoria
- Tienden a ser menos cooperativos
- Probabilidad muy remota
- Difícil de mantener el ciego

H - INFORMACIÓN SOBRE LA EXPOSICIÓN

La secuencia temporal causa-efecto de los estudios de casos y controles obliga a la obtención de información sobre la historia de exposiciones pasadas, es decir, de forma retrospectiva. Por esta razón, las encuestas, ya sean personales, telefónicas o mediante cuestionarios autoadministrados, son el método más utilizado para recoger dicha información. También se emplean medidas biológicas y se asume su estabilidad a lo largo del tiempo.

La elección de la fuente de información depende de la exposición que se quiere medir. Si, por ejemplo, el interés está en determinar el número de cigarrillos y el tiempo que hace que fuma, entonces lo mejor será preguntárselo directamente al individuo, ya que esta información no suele constar en las historias clínicas. Además, si el consumo de cigarrillos está relacionado con la enfermedad en estudio, probablemente se registre mucho mejor esta variable en los casos que en los controles.

En otras ocasiones, es mejor obtener la información de las historias clínicas, en especial cuando se trata de los resultados de pruebas realizadas a los pacientes.

I - SESGO DE MEMORIA

Los estudios de casos y controles han sido criticados muy a menudo a causa de que la información sobre la exposición se recoge retrospectivamente y, de este modo, se facilita la posibilidad de incurrir en un sesgo de memoria. Este sesgo suele ocurrir en enfermedades graves y/o cuando suponen un fuerte trauma psicológico, como las malformaciones congénitas.

En estas circunstancias, es muy posible que los casos recuerden sus antecedentes personales con mucho más detalle que los controles, al estar más sensibilizados por la enfermedad y porque sus médicos les habrán preguntado con insistencia por ellos. La presencia y magnitud de este sesgo varía según la exposición.

J - SESGO DEL ENTREVISTADOR

Puede aparecer siempre que el encuestador (u observador, en general) tenga conocimiento del grupo al que pertenece el sujeto al que está entrevistando y del objetivo del estudio. A la mayoría de investigadores les gusta obtener resultados positivos y, de forma involuntaria, pueden preguntar con más detalle e insistencia a los casos que a los controles. Por ello, cuando la información se recoge a través de un cuestionario y siempre que los recursos lo permitan, es preferible que el encuestador sea alguien ajeno al equipo que ha diseñado el protocolo.

La forma de evitar este problema es que el encuestador no sepa si entrevista a un caso o a un control.

K – CONCEPTO DE ODDS:

$$\text{Odds} = \frac{\text{Proceso}}{1 - \text{Proceso}}$$

Ejemplo: 4/1 (4 por 1).

Ratio de Odds: Evita calcular la incidencia, se porta como el riesgo relativo y se analiza igual (Se utiliza para estudios que se llega al RR).

$$\text{Odds Ratio} = \frac{a \times d}{b \times c}$$

L - CARACTERIZACIÓN OPERATIVA DE LAS PRUEBAS DIAGNÓSTICAS:

1) PRINCIPIO 1:

- Es el principio de la probabilidad: es una representación útil de la incertidumbre diagnóstica. Toma de decisiones en condiciones de incertidumbre.
- 0---1.
- Evita adjetivos y juicios de valor.

Observador A (Prueba A)	2	Observador B (Prueba B)
(I _{Investigador})	I	(P _{Referencia})

1 - Validez de criterio: exactitud y precisión, decidir de la prueba de referencia para ver su exactitud y precisión. Porque concedo, otorgo o elijo. Esta es la prueba estándar para ver su precisión.

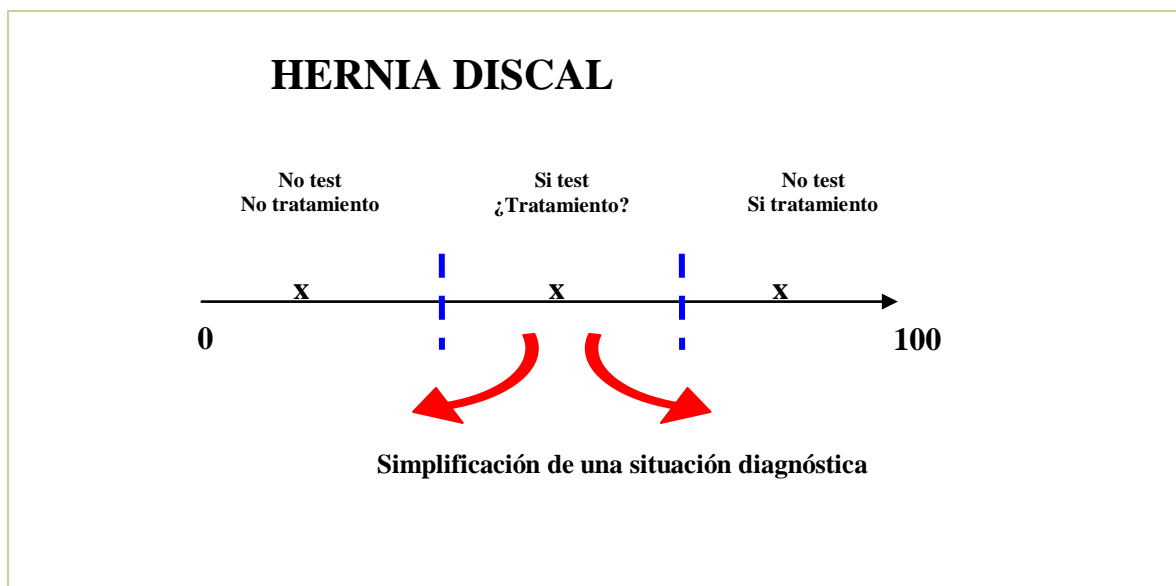
A veces no hay patrón de referencia y no se puede medir con exactitud concordancia e intercambiabilidad.

2 – SI NO HAY PRUEBA DE REFERENCIA, SE ESTUDIA LA RELACIÓN DE ACUERDO CON OTRAS PRUEBAS.

La intercambiabilidad se va a expresar en probabilidades.

2) PRINCIPIO 2:

Las pruebas diagnósticas sólo sirven si permiten la modificación del manejo de un problema.



La utilidad de una prueba diagnóstica depende de la prevalencia del trastorno y de la calidad interna de la prueba (sensibilidad), es decir de su validez.

Se entiende por Prueba diagnóstica (PD) cualquier procedimiento para obtener informaciones clínicas en un paciente.

Se debe expresar en exactitud diagnóstica.

- Validez operativa: Propiedad de clasificar correctamente los sujetos en subgrupos clínicamente pertinentes.
- Exactitud diagnóstica: Grado de calidad de la información otorgada.
- Utilidad diagnóstica: Valor y exactitud de la prueba.

M – CARACTERÍSTICAS DE LA PRUEBA DIAGNÓSTICA:

BASE METODOLÓGICA PARA LA TOMA DE DECISIONES EN LA PRÁCTICA CLÍNICA: Se parte sabiendo que un individuo sano y otro enfermo se pueden distinguir de manera válida (exacta) y reproducible por una prueba determinada diagnóstica.

TIPOS DE ESCALAS EN LAS P.D.:

a. Cualitativa:

METODOLOGÍA E INVESTIGACIÓN

- Presencia/Ausencia de signo para clasificar a los pacientes sanos o enfermos.
- Por ejemplo, existe o no el pulso.

b. Cuantitativa:

- Clasifica a un paciente sano o enfermo debajo de un criterio valor umbral, punto de corte valor de referencia, etc. si el valor de la PD cae por encima o por debajo, el punto de corte según donde se ponga cambiará la prueba, valor de referencia.
- Definir el valor de corte (umbral), el criterio positivo.

CARACTERÍSTICAS DE LA PRUEBA DIAGNÓSTICA:

		Enfermedad			
		Presente	Ausente		
Prueba	positiva	Verdadero positivo <i>a</i>	Falso positivo <i>b</i>	570 Verdadero positivo <i>a</i>	150 Falso positivo <i>b</i>
	negativa	Falso negativo <i>c</i>	Verdadero negativo <i>d</i>	30 Falso negativo <i>c</i>	850 Verdadero negativo <i>d</i>
				600	1.000

a) *Sensibilidad:*

Ocurre en la columna de los enfermos: es la probabilidad que un enfermo tenga un resultado positivo, una tasa de proporción de verdaderos positivos.

$$P = (P_{\text{Positiva}}/E_{\text{Presente}}) = 570/600 = 95$$

Enfermos con Prueba positiva / todos los enfermos.

b) *Especificidad(en la columnas de los sanos)*

Es la tasa de verdaderos negativos.

$$P = (P_{\text{Negativa}}/E_{\text{Ausente}}) = 850/1000 = 85$$

Es la suma de Pruebas negativas / todos los sanos.

Sensibilidad y especificidad son conceptos mutuamente excluyentes.

c) *Proporción de falsos negativos:*

Probabilidad que un enfermo tenga un resultado negativo.

$$PFN = P (P_{Negativa})/ Enfermedad_{Presente} = 30/600 = 0,05$$

Es la suma de los enfermos con prueba negativa/ todos los enfermos.

d) *Proporción de falsos positivos:*

Es la probabilidad que un paciente sano tenga una prueba positiva.

$$PFP = P (P_{Positivo})/ Enfermedad_{Ausente}$$

e) *Prevalencia:*

Es la probabilidad de que un individuo de la población tenga la enfermedad.

$$P (Enfermedad_{Positiva})$$

Es la proporción de individuos enfermos /número total de individuos.

f) *Probabilidad pretest o prevalencia (Probabilidad a priori).*

g) *Probabilidad postest (Probabilidad a posteriori).*

Puede ser positiva o negativa.

h) *Valor predictivo positivo:*

$$VPP = P (Enfermedad_{Presente})/ Prueba_{Positiva}$$

Es la proporción de enfermos con prueba positiva /todos los enfermos con prueba positiva.

VPP es el valor predictivo.

$$\frac{\text{Verdaderos positivos}}{\text{Verdaderos positivos} + \text{Falsos positivos}}$$

$$570/720 = 0,79$$

i) Valor predictivo negativo:

Es la probabilidad que un paciente sano con resultados negativos no tenga enfermedad.

$$VPN = P(\text{Enfermedad}_{\text{Ausente}}) / \text{Prueba}_{\text{Negativa}}$$

$$850/880 = 0.97$$

Es la proporción de pacientes sin enfermedad con prueba negativa / Todos las pruebas negativas.

Lo más importante es que el test tenga poco falsos positivos:

- Relatividad de la sensibilidad y especificidad.
- Independencia de la prevalencia.

j) Conclusión:

Es importante definir el umbral de positividad en el test diagnóstico para no tener demasiado falsos positivos, ni falsos negativos.

Se admite que una prueba es buena si:

- Existe un consenso de los profesionales.
- Existe pruebas históricas.
- Existe criterios anatómicos.
- Existe criterios bioquímicas.

Las imágenes (Rx, IRM, TAC) se basan en un consenso de profesionales y pruebas anatómicas.

COEFICIENTE (COCIENTE) DE PROBABILIDAD:

Un test tiene:

- P (prevalencia) de 5% ,sensibilidad de 90,frecuencia de 80: ¿Cuál es el TVP+?

$$\text{TVP+} = \frac{P(T+/E+)}{P(T+/E-) + 1 - \text{Especificidad}} = \frac{\text{sensibilidad}}{1 - \text{FP (falsos positivos)}} = \frac{\text{VP (verdaderos positivos)}}{1 - \text{FP (falsos positivos)}}$$

- Eje de valor de un test :

METODOLOGÍA E INVESTIGACIÓN

$$CP+ = \frac{P(T+/E+)}{P(T+/E-)} = \frac{S}{1-E} = \frac{VP}{FP} = 1 \text{ (TEST MALÍSIMO).}$$

$$CP+ = \frac{P(T+/E+)}{P(T+/E-)} = \frac{S}{1-E} = \frac{VP}{FP} = 5 \text{ (TEST MUY BUENO).}$$

- Odds : $\frac{P}{1-P} ; \frac{0,05}{0,95} = 0,05$

$$CP+ = \frac{S}{1-P} ; = \frac{0,90}{0,20} = 4,5$$

$$\text{Odds pre-prueba} \times CP+ = \text{Odds post-prueba} \\ = 0,225$$

$$\text{Probabilidad} = \frac{\text{Odds}}{1 + \text{Odds}} = P = \frac{0,225}{1,225} = 0,18$$

El TVP+ es de 18 %, el test no es bueno.

Cuando se hace la maniobra se debe ver el trastorno que hay:

1-Calidad del patrón de referencia (anatomía, probabilidades químicas).

2- Procedimientos independientes.

3-Comparaciones ciegas.

4-Precisión (fiabilidad).

5-Exactitud (validez).

- Error sistemático (sesgo)
- Mala clasificación.

6 -Utilizar una muestra que abarque el espectro más completo de enfermedad.

OTRO EJEMPLO DE TEST:

$$\text{Odds ratio} = 0,05 \\ CP+ = 25.$$

$$\text{Odds pre-prueba} \times CP+ = \text{Odds post-prueba} \\ = 1,25$$

$$\frac{\text{Odds}}{1 + \text{Odds}} = P = \frac{1,25}{225} = 0,55$$

El TVP+ es de 55 %, es test es bueno.

OTRO EJEMPLO DE TEST:

$$\text{Odds} = \frac{0,4}{0,6} = 0,66$$

$$\text{CP+} = 4,5$$

$$\text{Odds pre} \times \text{CP+} = \text{Odds post-test}$$
$$0,66 \times 4,5 = 3$$

$$\frac{\text{Odds}}{1 + \text{Odds}} = P = \frac{3}{4} = 0,75$$

El TVP+ es de 75 %, es test es MUY bueno.

El valor del test depende de la prevalencia y de la calidad del test.
Cuando se hace una maniobra se estudia calidad y frecuencia del trastorno que hay debajo.

COMO VALORAR UN TEST:

1. Calidad del patrón de referencia (anatomía, pruebas químicas...).
2. Procedimientos independientes (el examinador no sabe si el paciente tiene la enfermedad o no).
3. Comparación a ciegas.
4. Precisión (fiabilidad): depende del número de pacientes (70 está bien, mínimo 40). El error aleatorio hace disminuir sensibilidad y especificad.
5. Exactitud (validez).
 - a. Aumenta el riesgo de sesgos.
 - b. Mala clasificación.
6. Muestra que abarque el espectro más completo.
7. Descripción completa y suficiente.
8. Aclarar la definición del “normal”.
9. Claras ventajas.

METODOLOGÍA E INVESTIGACIÓN

Para confirmar una enfermedad, la prueba tiene que ser sensible.

Para eliminar una enfermedad una enfermedad, la prueba tiene que ser válida.

VI – LOS TIPOS DE ESTUDIOS: Experimental y no Experimental

Por diseño de un estudio se entienden procedimientos, métodos y técnicas mediante los cuales el investigador selecciona a los pacientes, recogen una información, la analiza e interpreta los resultados. El diseño es, pues, la conexión entre la hipótesis y los datos.

A -CRITERIOS DE CLASIFICACIÓN

Los criterios para clasificar los diferentes tipos de estudio se sustentan en cuatro ejes:

- Finalidad del estudio: analítica o descriptiva.
- Secuencia temporal: transversal o longitudinal.
- Control de la asignación de los factores de estudio: experimental u observacional.
- Inicio del estudio en relación a la cronología de los hechos: prospectivo o retrospectivo.

B - FINALIDAD: DESCRIPTIVA O ANALÍTICA

Se considera analítico todo estudio que evalúa una presunta relación causa-efecto. El presunto agente causal puede ser tanto un factor que se sospecha que puede conducir etiológicamente a una enfermedad como un tratamiento para prevenir o mejorar una situación clínica.

Se considera descriptivo todo estudio no enfocado en una presunta relación causa-efecto, sino que sus datos son utilizados con finalidades puramente descriptivas. Este tipo de estudios es útil para generar hipótesis etiológicas que deberán contrastarse posteriormente con estudios analíticos.

C - DIRECCIÓN TEMPORAL: TRANSVERSAL O LONGITUDINAL

Se consideran transversales los estudios en los que los datos de cada sujeto representan esencialmente un momento del tiempo. Estos datos pueden corresponder a la presencia, ausencia o diferentes grados de una característica o enfermedad, como ocurre, por ejemplo, en los estudios de prevalencia de una enfermedad en una comunidad determinada, o bien examinar la relación entre diferentes variables en una población definida en un momento de tiempo determinado.

Dado que las variables se han medido de forma simultánea, no puede establecerse la existencia de una secuencia temporal entre ellas y, por tanto, estos diseños no permiten abordar el estudio de una presunta relación causa-efecto. Así pues, los estudios transversales son por definición descriptivos.

Se consideran longitudinales los estudios en los que existe un lapso de tiempo entre las distintas variables, de forma que puede establecerse una secuencia temporal entre ellas. Pueden ser tanto descriptivos como analíticos. En los estudios analíticos debe tenerse en cuenta si la secuencia temporal es de causa hacia desenlace (estudios experimentales y estudios de cohortes), o bien de desenlace hacia causa (estudios de casos y controles).

Algunos autores consideran longitudinales sólo los estudios en los que los sujetos son seguidos en el tiempo desde una línea basal hasta un desenlace y hacen sinónimo este concepto del de cohorte. Sin embargo, según la definición dada en el párrafo anterior, se considera que un estudio es longitudinal si las observaciones se refieren a dos momentos en el tiempo, aun cuando la recogida de información se ha realizado de forma simultánea.

Hay que señalar que, si las distintas observaciones se han recogido en un mismo momento en el tiempo, para que el estudio pueda considerarse longitudinal, se debe asumir una secuencia temporal entre ellas.

D - ASIGNACIÓN DE LOS FACTORES DE ESTUDIO: EXPERIMENTAL U OBSERVACIONAL

Se consideran experimentales los estudios en los que el equipo investigador asigna el factor de estudio y lo controla de forma deliberada para la realización de la investigación, según un plan preestablecido. Estos estudios se centran en una relación causa-efecto (analíticos) y en general evalúan el efecto de una o más intervenciones preventivas o terapéuticas.

Se definen como observacionales los estudios en los que el factor de estudio no es asignado por los investigadores, sino que éstos se limitan a observar, medir y analizar determinadas variables, sin ejercer un control directo sobre el factor de estudio.

La exposición puede haber sido «escogida» por los propios sujetos (p. ej., el consumo de tabaco) o decidida por el profesional sanitario dentro del proceso habitual de atención sanitaria (p. ej., los actos terapéuticos ordinarios). En otras ocasiones, esta exposición viene impuesta (p. ej., el sexo o la raza).

E - INICIO DEL ESTUDIO EN RELACIÓN A LA CRONOLOGÍA DE LOS HECHOS: PROSPECTIVO O RETROSPECTIVO

Los términos prospectivo y retrospectivo son ambiguos y pueden conducir a confusión, ya que pueden (y suelen) aplicarse también a la dirección temporal de las observaciones, de forma que algunos autores consideran el término prospectivo como sinónimo de cohorte o incluso de longitudinal.

Se consideraran retrospectivos estudios cuyo diseño es posterior a los hechos e datos, de forma que los datos se obtienen de archivos o de los datos que los sujetos o los médicos refiere. Cuando existe una combinación de ambas situaciones, los estudios se clasifican como ambispectivos.

CLASIFICACIÓN DE LOS TIPOS DE DISEÑO

ESTUDIOS DESCRIPTIVOS TRANSVERSALES.

- Estudios de prevalencia
- Series de casos transversales
- Evaluación de pruebas diagnósticas
- Estudios de concordancia
- Estudios de asociación cruzada
- Otros estudios transversales descriptivos

ESTUDIOS DESCRIPTIVOS LONGITUDINALES:

- Estudios de incidencia
- Descripción de los efectos de una intervención no deliberada
- Descripción de la historia natural.

ESTUDIOS ANALÍTICOS OBSERVACIONALES:

- Secuencia causa-efecto: estudios de cohortes:
 - Prospectivos
 - Retrospectivos
 - Ambispectivos
- Secuencia efecto-causa: estudios de casos y controles
- Estudios híbridos

ESTUDIOS ANALÍTICOS EXPERIMENTALES:

- Ensayos controlados:
 - Ensayos clínicos en paralelo
 - Ensayos clínicos cruzados
 - Ensayos comunitarios.
- Ensayos no controlados:
 - Ensayos sin grupo control
 - Ensayos con control externo

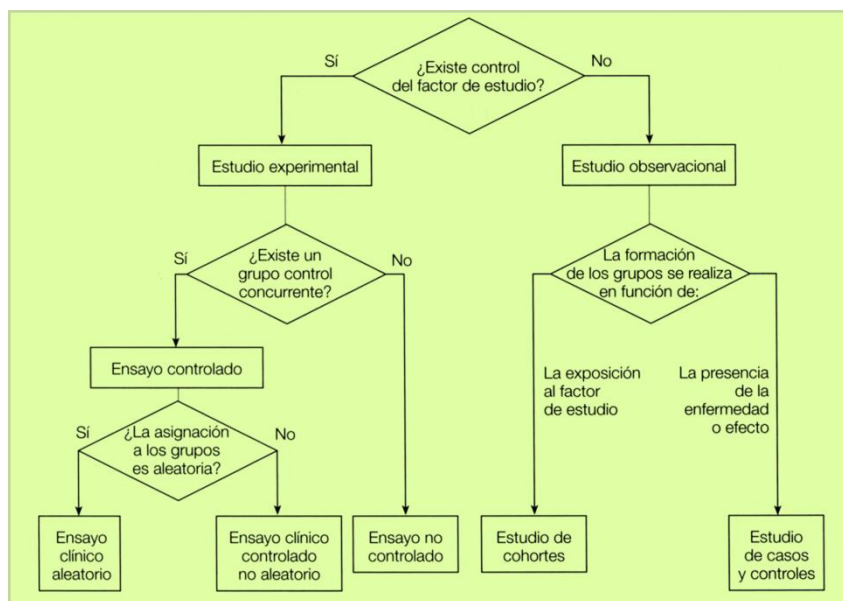


Figura Algoritmo de clasificación de los estudios analíticos.

Un estudio tiene que ser aleatorio.

Intervención (I)
Distribución aleatoria (DA)

Experimental (I+) (A+):

- Ensayo de laboratorio.
- Ensayo clínico.
- Comunitario.

Casi experimental (I+) (A-): Ensayo comunitario o clínico no aleatorio.

Observacional (I-) (A-):

- *Descriptivo:*

- Serie de casos (casos interesantes, a propósito de...dan mucho juego. Se describe maravillosamente bien lo que ves).
- Transversales
- Poblacional.
- No hay estadísticas.
- Es útil para formular hipótesis.

- *Analítico :*

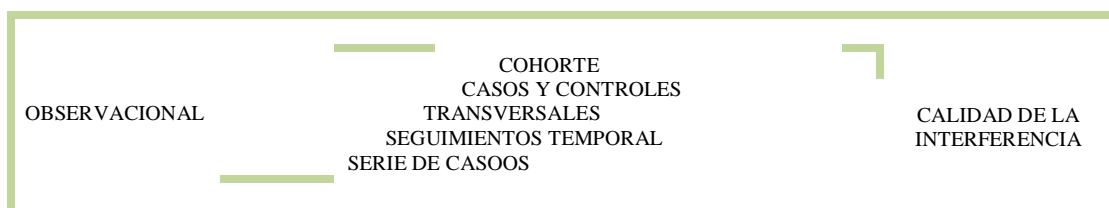
- Cohortes.

METODOLOGÍA E INVESTIGACIÓN

- Casos y controles.

	Intervención	Distribución aleatoria	
Experimental: Técnica osteopática.	+	+	Ensayo clínico. Ensayo comunitario.
Casi-experimental: no hay grupo de comparación	+	—	Ensayo no aleatorio.
Observacional : Test diagnóstico	—	—	Ensayo descriptivo o analítico.

Experimental → ensayo clínico



E - ESTUDIOS DESCRIPTIVOS

Las principales finalidades de estos estudios son describir la frecuencia y las características de un problema de salud en una población, describir la asociación entre dos o más variables sin asumir una relación causal entre ellas y generar hipótesis razonables que deberán ser contrastadas posteriormente mediante estudios analíticos.

Los estudios descriptivos son observacionales, tanto pueden ser transversales como longitudinales, y pueden clasificarse según su objetivo. A continuación se presentan las características principales de los más habituales.

F – SERIE DE CASOS:

Las series de casos transversales consisten en la enumeración descriptiva de unas características seleccionadas observadas en un momento del tiempo en un grupo de pacientes con una enfermedad determinada o en un grupo de sujetos que tienen una determinada condición en común. Por ejemplo, describir las cifras actuales de presión arterial, colesterol y otros factores de riesgo cardiovascular en los diabéticos del centro de salud.

METODOLOGÍA E INVESTIGACIÓN

Muestra de la población	n	Hipercolesterolemia	
		n	(%)
Participantes	600	90	(15)
No participantes	400	120	(30)
Total	1.000	210	(21)

TABLA Resultados de un estudio hipotético sobre la prevalencia de la hipercolesterolemia.

En las series de casos longitudinales o seguimiento de una cohorte se describe la evolución temporal de determinadas características observadas en un grupo de pacientes con una enfermedad o en un grupo de sujetos que tienen una determinada condición en común. Existe una secuencia temporal definida, pero el estudio no evalúa ninguna relación causa-efecto entre las variables.

- Descripción de un caso inusual o interesante.
- Descripción de varios casos donde no se intenta verificar hipótesis ni se comparan los resultados con otro de grupo de casos.
- Describen la experiencia de un paciente o de un grupo de pacientes.
- Documentan la presencia de nuevas enfermedades, efectos adversos, situaciones excepcionales o infrecuentes.

1. *Diseño: direccionalidad*

- Selección de muestra.
- Temporalidad.
- Selección no exhaustiva; oportunista de casos.
- Recogida sistemática de signos y síntomas.
- Observaciones y mediciones limitadas por circunstancias clínicas, geográficas, sociales...

2. Análisis:

- Descripción narrativa-frecuencias, “Dan mucho juego”.

3. Ventajas:

- Útiles para formular hipótesis, pero no sirven para evaluar estadísticas.
- Sencillas, baratas.
- Falacia: es atribuir características de grupo a un individuo.

4. Defectos:

- Interpretación de los datos.

G – ESTUDIOS DESCRIPTIVOS TRANSVERSALES POBLACIONALES:

Son diseños transversales que tienen como finalidad principal la estimación de la prevalencia de una enfermedad o una característica en una población. Se basan en la definición precisa de una población de estudio y la obtención de una muestra representativa de ella, en la que se determina la frecuencia de individuos que presentan la característica o enfermedad de estudio.

La validez de los resultados dependerá de la representatividad de la muestra, de la calidad de los datos obtenidos y de que el número de no respuestas sea pequeño y sus motivos no estén relacionados con la enfermedad.

Los estudios transversales no son útiles en enfermedades raras ni de corta duración. No permite calcular incidencia ni riesgo.

No establecen relación de causalidad (sólo asociación).

Distribución peligrosa de factores de confusión.

Describen la frecuencia de una enfermedad, el factor de riesgo u otra característica en una población y tiempos definidos.

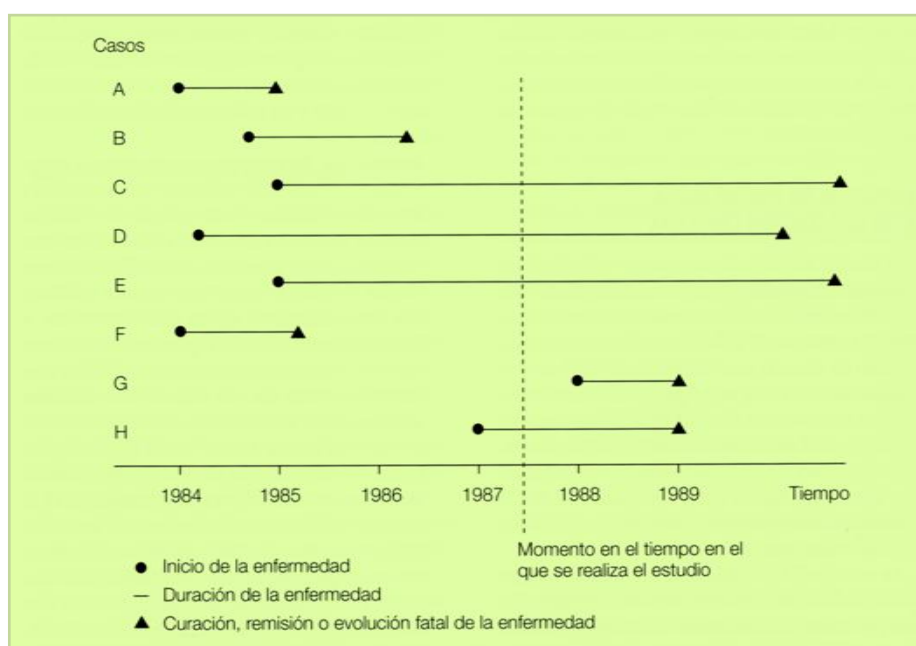


Figura Estudio transversal y su relación con la duración de los casos.

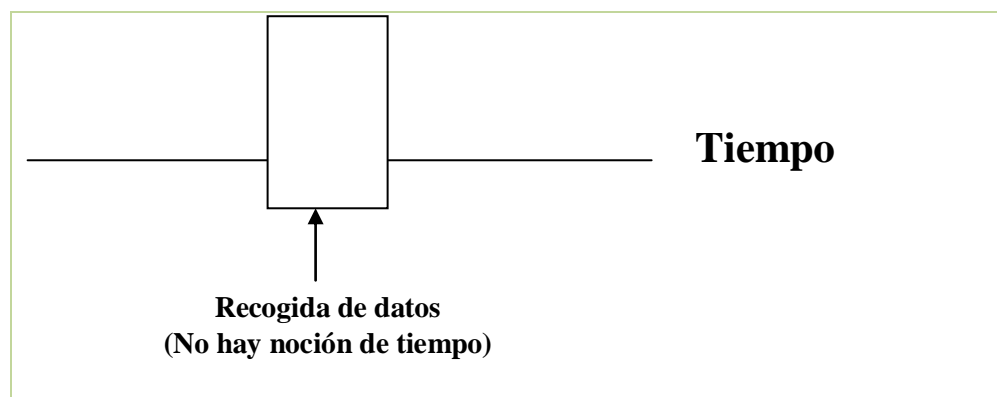
Las variables estudiadas deben ser fácilmente medibles a través de cuestionarios o exámenes médicos simples y seguros, ya que se estudia una muestra de la población que en su mayoría no tendrá la enfermedad, por lo que no pueden usarse métodos que supongan algún riesgo para los participantes. Las pruebas deben ser lo más sensibles y específicas posible, para evitar la clasificación incorrecta de sujetos en situación de remisión o en tratamiento.

– *Proposición:*

- Afirma la prevalencia de características o problemas de salud.
- La pregunta es:

- ¿Prevalencia de la enfermedad X en la población?

Estudio Transversal.



– *Inconvenientes:*

- No permiten calcular el riesgo.
- No establecen una relación de causalidad.
- Muestras no representativas, hay riesgo de confusión.
- Alteración de los grupos.
- Sesgos de selección, de recuerdo, etc.

– *Ventajas:*

- Útiles para enfermedad de larga duración.
- Estudian la prevalencia.
- Útiles para planificación sanitaria.
- Resultados más fácilmente generalizables.

Ejemplo: Observar las disfunciones osteopáticas en una patología.

H – ESTUDIOS ECOLÓGICOS:

Describen una enfermedad en la población en relación con variables como consumo de nutrientes, medio ambiente.

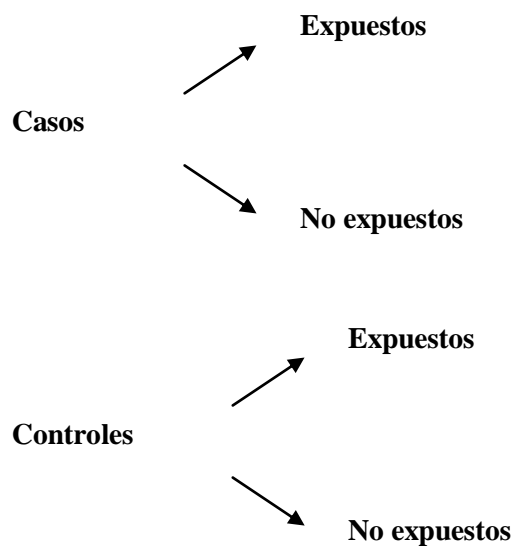
I – ESTUDIO DE CASOS Y CONTROLES:

Las características de las personas con una enfermedad (casos) son comparados con aquellos que presentan otras personas seleccionadas sin la enfermedad (controles, testigos, referentes).

Estudios observacionales analíticos: permiten seguir la relación causa-efecto. Se debe separar la base de datos.

METODOLOGÍA E INVESTIGACIÓN

	Número de pacientes	Tensión arterial	Hematocrito
AGC 23	23	12,6	X
FR 24	24	13,8	X
EM 25	25	18,2	X



¿Cómo elegir el grupo control? Es difícil.

- Misma edad.
- Mismo sexo.
- Misma patología.

$$\text{Odds Ratio (OR)} = \frac{a \times d}{c \times b}$$

	E+	E-
Expo +	a →	b
Expo -	c →	d

1) IDENTIFICACIÓN DEL CASO:

1º. DEFINIR EL CASO: CRITERIOS DIAGNÓSTICOS.

- Especificidad.
- Sensibilidad.

2º. SELECCIÓN DE CASOS: EXPOSICIÓN

2) SELECCIÓN DE LOS CONTROLES (TESTIGO):

- Iguales todos, excepto el factor de estudio.
 - Ningún criterio.
 - Restricción total (diferente de inclusión/exclusión), parcial (estratificación, apareamiento, matching).
- Información de la exposición:
 - Personal.
 - Registro (Historial clínico).

Se utiliza un análisis multivariante:

3) VENTAJAS:

- Útil en enfermedades raras o de latencia prolongada.
- Requieren menos sujetos que otros estudios.
- No supone riesgo para el enfermo.
- Relativamente fácil de diseñar.

Ejemplo: Test de movilidad de la ATM con grupo sano y grupo con síndrome cráneo mandibular.

- Importancia de la calidad de información.

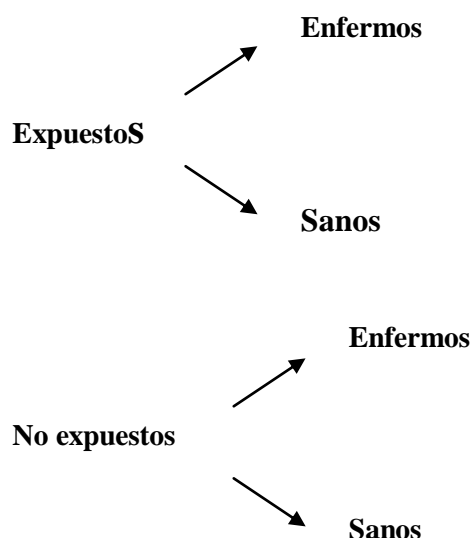
4) LIMITACIONES:

- Errores de clasificación de los sujetos.
- Fallo al ajustar variables de confusión.
- Interpretación errónea del concepto de causalidad.
- Riesgo importante de sesgos +++.
- Validez de la información de exposición difícil.
- No permite el cálculo de los riesgos.
- No informa sobre los mecanismos de producción de la enfermedad.

J - ESTUDIO DE COHORTE:

Es el estudio de las mismas personas durante un periodo de tiempo.

METODOLOGÍA E INVESTIGACIÓN



El término cohorte se utiliza para designar a un grupo de sujetos con una característica o conjunto de características en común (generalmente la exposición al factor de estudio), que son seguidos en el transcurso del tiempo.

Un estudio de cohortes es un diseño observacional analítico longitudinal en el que se comparan dos cohortes que difieren por su exposición al factor de estudio para evaluar una posible relación causa-efecto. Si en un estudio se sigue una cohorte con la única finalidad de estimar la incidencia con que aparece un determinado problema de salud (desenlace o efecto), o describir su evolución, se trata de un diseño descriptivo longitudinal.

En un estudio de cohortes, los individuos, todos ellos inicialmente sin la enfermedad de interés, se clasifican en función de su exposición o no al o los factores de estudio y son seguidos durante un período de tiempo, comparando la frecuencia con que aparece el efecto o respuesta en los expuestos y no expuestos.

La estructura de los estudios de cohortes es parecida a la de los ensayos clínicos aleatorios, con dos importantes diferencias: no existe asignación aleatoria de los sujetos ni control del factor de estudio por parte de los investigadores.

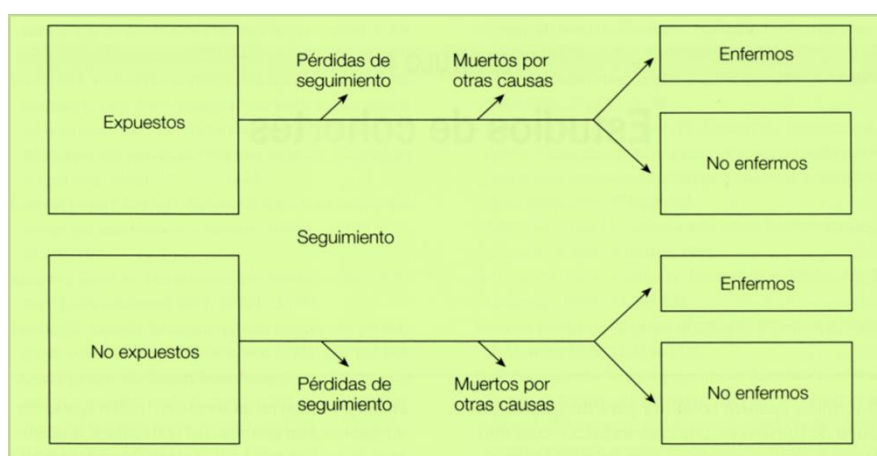


Figura Estructura básica de un estudio de cohortes.

1) TIPOS DE ESTUDIOS DE COHORTES

Según la relación cronológica entre el inicio del estudio y el desarrollo de la enfermedad de interés, los estudios de cohortes se pueden clasificar como prospectivos o retrospectivos.

- En un estudio prospectivo de cohortes, el investigador comienza con un grupo de sujetos presuntamente libres de la enfermedad de interés, los clasifica en expuestos o no a un posible factor de riesgo y los sigue durante un tiempo para determinar. El carácter prospectivo significa que el investigador recoge la información sobre la exposición cuando se inicia el estudio, e identifica los nuevos casos de la enfermedad o las defunciones que se producen a partir de ese momento.
- En los estudios retrospectivos de cohortes tanto la exposición como la enfermedad ya han ocurrido cuando se lleva a cabo el estudio. La identificación de las cohortes expuesta y no expuesta se apoya en su situación en una fecha previa bien definida (p. ej., la fecha de inicio de una exposición laboral), suficientemente lejos en el tiempo para que la enfermedad en estudio haya tenido tiempo de desarrollarse.

En algunas circunstancias, los estudios pueden ser ambispectivos, en los que se recogen datos retrospectiva y prospectivamente en una misma cohorte.

Los estudios retrospectivos se pueden llevar a cabo más rápida y económicamente. En los estudios prospectivos, la información se recoge directamente de los individuos mediante cuestionarios y/o examen físico, lo que facilita que los datos sean de mayor calidad.

2) ESTIMACIONES QUE PUEDEN REALIZARSE EN LOS ESTUDIOS DE COHORTES

Los estudios de cohortes son longitudinales en los que existe seguimiento de sujetos, por lo que permiten realizar las siguientes estimaciones:

- Incidencia de la enfermedad en los sujetos expuestos y en los no expuestos.
- Riesgo relativo como medida de la magnitud de la asociación entre el factor de riesgo y la variable de respuesta, que estima el riesgo de presentar la respuesta de los sujetos expuestos en relación a los no expuestos. Permite también estimar el riesgo relativo, según diferentes características de la exposición y se puede evaluar, por ejemplo, la existencia de un gradiente dosis-respuesta entre la exposición y la incidencia de la enfermedad.
- Fracción o proporción atribuible o proporción de casos de una enfermedad que resulta de la exposición a un factor determinado o a una combinación de ellos.
- Diferencia de incidencias como medida del impacto potencial que tendría la eliminación de la exposición.

3) IDENTIFICACIÓN DE LAS COHORTES

En los estudios de cohortes, dado que el propósito principal es analítico (estudiar la relación entre una presunta causa y una enfermedad), la consideración más importante es la comparabilidad de las cohortes expuesta y no expuesta, así como asegurar que la muestra contenga un número suficiente de sujetos con las características predictoras más importantes y un número suficiente de desenlaces observados durante el estudio, a fin de permitir un análisis válido.

4) IDENTIFICACIÓN DE LA COHORTE EXPUESTA

La cohorte expuesta puede proceder de la población general o de grupos especiales en los que la exposición es frecuente o en los que es fácil efectuar un seguimiento completo.

5) MEDICIÓN DE LA EXPOSICIÓN

Uno de los puntos fuertes de los estudios de cohortes prospectivos es que permiten obtener una información detallada, precisa y objetiva de la exposición en estudio.

6) MEDICIÓN DEL DESENLACE

Los procedimientos para determinar qué sujetos de las cohortes desarrollan el desenlace de interés varían en función de las enfermedades que se estudian y de los recursos disponibles. Esta medición debe ser lo más objetiva posible. Cuando se trata de síntomas o síndromes para los que no existe ningún examen objetivo debe recurrirse a definiciones sobre las que exista un amplio consenso.

Cuando el diagnóstico requiere la utilización de examen físico y exploraciones complementarias, es preferible utilizar criterios estándares reconocidos, lo que permitirá comparar los resultados con los de otros estudios.

No debe olvidarse que la existencia de resultados falsos positivos y falsos negativos puede sesgar los resultados.

7) SEGUIMIENTO

El mayor desafío de un estudio de cohortes es el seguimiento de un elevado número de sujetos durante un prolongado período de tiempo, por lo que la mayoría de esfuerzos se dedican a asegurarlo. Los investigadores deben recoger información sobre los cambios habidos en los factores de riesgo y las variables pronósticas y registrar si alguno de los participantes ha desarrollado el desenlace de interés.

Algunos individuos modificarán su exposición durante el seguimiento. Por ejemplo, si el factor en estudio es el hábito tabáquico, habrá sujetos que empezarán a fumar durante el estudio y otros que, en principio, estaban incluidos en la cohorte expuesta, que dejarán de fumar.

El tiempo de seguimiento debe ser lo suficientemente largo para permitir que aparezca el número suficiente de casos para conseguir la potencia estadística deseada, pero lo suficientemente corto como para no crear graves problemas logísticos y de organización.

El intervalo entre las distintas visitas de seguimiento dependerá de la enfermedad y deberá ser lo suficientemente corto como para detectar su aparición.

8) ESTRATEGIAS PARA REDUCIR LAS PÉRDIDAS

Con la finalidad de reducir el número de pérdidas de seguimiento, suele ser útil, al establecer los criterios de selección, excluir a los individuos que es probable que se pierdan durante aquél por problemas de salud, porque piensan cambiar de domicilio o por cualquier otra razón.

Además, al comenzar el estudio, debe recogerse información que permita localizar a los individuos en caso de que éstos cambien de domicilio o fallezcan.

Los contactos periódicos con los individuos ayudan a mantener el seguimiento, pudiendo también ser útiles para determinar con mayor exactitud el momento en que aparecen los desenlaces de interés. Estos contactos pueden realizarse tanto telefónicamente.

9) VENTAJAS:

- Seguir la evolución de una enfermedad.
- Permite calcular la INCIDENCIA y el riesgo.
- Buen control de los sesgos.
- Facilita la recogida de datos.

10) INCONVENIENTES:

- Inútil en enfermedades raras.
- Difícil de trabajar y de reproducir.
- Necesita mucho tiempo y recursos.
- No vale para estudiar la etiología de una enfermedad.

NOTA: ASIGNACIÓN ALEATORIA

Es el procedimiento más importante porque genera grupos iguales y permite hacer pruebas estadísticas (por esta razón hace falta que grupo de estudio y grupo control tienen que ser iguales).

- Teoría de LAPLACE.
- Principio de las estadísticas: una variable adquiere valor propio.

J - ESTUDIOS EXPERIMENTALES:

El investigador manipula las condiciones de investigación y distribuye aleatoriamente los grupos.

El objetivo de los estudios experimentales es estimar la eficacia de una intervención preventiva, curativa o rehabilitadora. Estos estudios se definen porque los investigadores tienen control sobre el factor de estudio, es decir, deciden qué tratamiento, con qué pauta y durante cuánto tiempo, recibirá cada uno de los grupos de estudio.

La asunción básica de estos diseños es que los grupos que se comparan son similares por todas las características pronósticas que pueden influir sobre la respuesta, excepto por la intervención que se está evaluando. La mejor forma de conseguir grupos comparables es que la asignación de los individuos a los grupos de estudio se realice de forma aleatoria. Si los grupos obtenidos de este modo son comparables y son estudiados con una misma pauta de seguimiento, cualquier diferencia observada entre ellos al finalizar el experimento puede ser atribuida, con un alto grado de convicción, a la diferente intervención a que han sido sometidos los participantes.

Así pues, la gran ventaja de los diseños con asignación aleatoria radica en su alto grado de control de la situación, que proporciona, en el caso de que exista una asociación entre el factor estudiado y la respuesta observada, la mejor evidencia de que dicha relación es causal.

1) INTERVENCIONES QUE SE COMPARAN

El primer aspecto a considerar en el diseño de un ECA es la selección de la alternativa con la que se va a comparar la intervención en estudio. En términos generales, las opciones pueden ser el uso de un placebo o de otro tratamiento o intervención activos.

Cuando existe una opción terapéutica reconocida como eficaz en la situación clínica de interés, ésta debe ser la alternativa con la que comparar la nueva intervención. El uso de un grupo placebo en esta situación presenta limitaciones éticas.

El uso de otro tratamiento o intervención activos como grupo de comparación tiene por objetivo estimar la relación beneficio/riesgo del nuevo tratamiento en una situación clínica concreta. En estos casos, la mejor comparación es el «mejor tratamiento existente» en dicha situación. Ésta no siempre es una elección fácil, ya que en la mayoría de ocasiones

CLASIFICACIÓN DE LOS ESTUDIOS EN FUNCIÓN DEL MOMENTO DEL DESARROLLO CLÍNICO DE UN FÁRMACO.

Características

Fase 1: Primer estadio de la prueba de un nuevo fármaco en el ser humano.

- *Objetivo principal: evaluar la seguridad del fármaco en seres humanos.*
- *Sujetos: voluntarios sanos.*
- *Diseño: en general son estudios no controlados*

Fase 2:

- *Objetivo principal: conocer la farmacocinética del nuevo fármaco, así como su farmacodinamia (búsqueda de dosis, mecanismo de acción farmacológica, relaciones dosis/respuesta).*
- *Sujetos: pacientes potenciales, aunque inicialmente también puede llevarse a cabo en voluntarios sanos.*
- *Diseño: estudios no controlados y ensayos clínicos aleatorios controlados con placebo con criterios de selección muy estrictos.*

Fase 3: Última fase de la evaluación de un medicamento antes de su comercialización

- *Objetivo principal: evaluar la eficacia y relación beneficio/riesgo en comparación con otras alternativas terapéuticas disponibles, o con un placebo si no hay tratamiento disponible. Permiten establecer la eficacia del nuevo fármaco e identificar y cuantificar los efectos indeseables más frecuentes*
- *Sujetos: pacientes*
- *Diseño: ensayos clínicos aleatorios*

Fase 4: Estudios realizados posteriores a la comercialización del fármaco.

- *Objetivo principal: evaluar mejor el perfil de seguridad, las posibles nuevas indicaciones o nuevas vías de administración, la eficacia en las condiciones habituales de uso (efectividad) o en grupos especiales.*
- *Sujetos: pacientes.*
- *Diseño: ensayos clínicos aleatorios y estudios observacionales*

En algunas ocasiones, la finalidad del ECA no es evaluar si un nuevo tratamiento es más eficaz que el de referencia, sino demostrar que es, como mínimo, igual a él, ya que ofrece otras ventajas en cuanto a seguridad, facilidad de administración, etc.

2) SELECCIÓN DE LA POBLACIÓN

Los individuos que participan en el estudio proceden de una población de referencia o diana a la que se quiere extrapolar los resultados. A partir de ella, se especificarán los criterios de selección que darán lugar a una población de estudio (población experimental), que es aquella en la que se desea realizar la experiencia.

Los criterios de inclusión han de definir una muestra de sujetos en los que, a la luz de los conocimientos actuales, las intervenciones que se comparan podrían estar indicadas y, por tanto, potencialmente pueden beneficiarse de ellas.

La utilización de criterios de inclusión y exclusión estrictos, aleja a la población de estudio de la población diana y, por tanto, limita la capacidad de generalización, pero aumenta la validez interna de las observaciones.

3) CONSENTIMIENTO INFORMADO:

Antes de que los individuos de la población experimental sean incluidos en el estudio deben dar su consentimiento informado para ello. El consentimiento informado de los sujetos implica la explicación, tanto oral como escrita, de lo que supone la experiencia y las posibles consecuencias que pueden derivarse de su participación.

La información que se facilite ha de ser el elemento básico para que un paciente decida participar o no en el estudio, y debería cubrir los siguientes puntos:

- a) el objetivo del estudio;
- b) las características básicas del diseño (asignación aleatoria, técnicas de enmascaramiento, uso de tratamiento placebo, etc.);
- c) los posibles efectos secundarios;
- d) los posibles beneficios;
- e) que el paciente tiene el derecho de abandonar el estudio en cualquier momento,
- f) que tiene el derecho de formular cuantas preguntas desee relacionadas con la investigación.

Todos estos puntos deben explicarse en términos comprensibles para el paciente, evitando el lenguaje académico o científico.

Situaciones en las que no es posible o deseable la utilización del doble ciego:

- *Cuando implica riesgos innecesarios para el paciente (p. ej., en el caso de administrar un placebo por vía parenteral de manera repetida y durante un tiempo prolongado).*
- *Cuando no es posible disponer de una formulación galénica adecuada.*
- *Cuando los efectos farmacológicos permiten identificar fácilmente al menos uno de los fármacos estudiados.*
- *Cuando, por cualquier circunstancia, se considera que el diseño de doble ciego puede perjudicar la relación entre médico y paciente.*

4) ESTRATEGIA DE ANÁLISIS

La estrategia de análisis de un ECA es muy similar a la de cualquier estudio analítico que compara dos o más grupos.

5) EVALUACIÓN DE LA EFICACIA DE UNA MEDIDA PREVENTIVA

Existen aspectos que diferencian los ensayos clínicos que evalúan la eficacia de una medida preventiva de los ensayos que estudian la eficacia de un nuevo tratamiento.

La primera diferencia es que, en los primeros, se estudian individuos presuntamente sanos. Ello implica que el número de personas que desarrollarán la enfermedad es muy bajo, y quizá tras un largo período de tiempo, mientras que las complicaciones de una enfermedad se pueden detectar en una alta proporción de individuos enfermos en un tiempo relativamente corto.

METODOLOGÍA E INVESTIGACIÓN

En consecuencia, los ensayos que evalúan una medida preventiva suelen requerir un mayor número de individuos y un seguimiento más largo.

6) ASIGNACIÓN ALEATORIA:

Método para asignar pacientes al grupo de tratamiento. Se utiliza tablas de distribución, tiene que ser ocultado el método de tratamiento.

Evita sesgos atribuibles a la preferencia del investigador o a la naturaleza de los pacientes: produce grupos de comparación similares.

68 68 80 54 00	16 01 92 58 21	65 12 64 64 70	07 28 66 61 59	48 79 74 73 72	08 64 80 91 38
92 36 48 69 45	89 84 05 34 47	09 12 81 93 63	46 13 95 65 96	88 09 31 54 88	97 96 86 01 69
97 50 71 39 79	51 99 98 44 39	99 35 72 61 22	42 36 31 16 59	12 75 10 60 36	80 66 39 94 97
57 93 08 35 69	08 12 60 39 23	89 96 34 22 37	96 18 69 06 30	61 73 84 89 18	26 02 04 37 95
44 71 38 40 37	69 99 47 26 52	89 85 33 22 80	66 10 71 44 05	48 06 30 00 18	03 30 28 55 59
70 72 51 49 73	88 91 28 79 50	81 83 33 98 29	88 77 90 45 59	71 42 14 96 55	98 59 96 01 36
56 42 78 54 06	59 45 27 08 51	68 82 34 08 83	67 98 36 65 56	85 64 23 85 41	64 72 08 59 44
80 61 68 44 19	84 27 17 30 37	62 42 07 12 63	95 39 06 35 63	48 69 49 02 58	98 02 50 58 11
37 79 95 02 66	65 45 53 41 07	61 02 73 36 85	90 54 33 65 84	14 83 46 74 11	76 66 63 60 08
35 36 82 82 59	48 28 01 83 84	61 38 93 73 68	22 30 95 69 72	09 11 21 91 73	97 28 44 74 06
43 40 13 35 45	93 18 31 83 45	80 58 35 06 88	47 19 63 92 75	54 52 62 29 91	53 58 54 66 05
78 36 26 24 06	18 26 32 96 83	74 93 55 39 26	73 87 96 76 23	50 58 45 27 57	14 96 39 64 85
98 84 48 42 92	45 62 63 40 88	60 42 17 18 48	69 63 21 83 41	35 69 34 10 94	32 22 52 04 74
84 65 43 07 30	26 22 59 28 27	59 62 37 95 42	33 56 90 92 57	38 58 22 14 79	24 32 12 38 42
23 79 80 71 37	33 26 74 03 30	49 54 36 85 14	58 61 52 27 03	74 22 19 13 48	30 28 01 92 49
17 79 96 52 35	05 53 15 26 70	37 03 08 98 64	78 35 22 22 88	04 69 22 64 07	04 73 25 74 82
33 48 32 91 54	98 61 70 48 22	53 26 79 20 38	58 70 61 43 97	68 50 64 55 75	42 70 32 09 60
82 17 18 17 14	85 13 41 38 10	95 28 12 73 23	34 78 77 60 25	16 47 61 43 77	83 27 19 70 41
91 95 43 81 14	04 41 66 09 76	84 31 64 64 08	47 42 80 61 03	20 50 73 40 95	24 77 95 73 20
83 59 89 65 06	03 10 16 82 24	21 46 51 44 77	33 11 49 15 16	39 58 20 12 39	82 77 02 18 88
22 32 61 43 75	54 08 18 07 04	12 53 67 51 54	97 30 53 62 38	92 17 63 36 75	33 14 11 11 78
49 48 55 11 39	68 50 33 31 47	16 28 25 82 98	86 62 93 66 71	15 64 88 75 27	04 51 41 61 96
39 07 30 60 77	39 90 89 86 77	64 21 91 15 82	54 80 67 78 66	46 86 86 88 86	50 09 13 24 91
52 49 41 73 46	56 50 45 94 25	12 77 40 70 14	68 47 37 10 84	48 28 48 30 51	60 73 73 03 87
47 10 62 31 28	59 77 64 59 90	44 37 33 53 17	13 26 98 86 29	58 92 62 50 18	93 09 45 89 06
91 67 48 57 10	52 62 24 19 94	25 47 57 91 13	13 50 63 04 23	62 74 29 92 24	64 94 63 15 07
49 92 05 12 07	55 98 78 10 70	47 46 41 90 08	78 66 28 55 80	44 31 52 43 07	19 83 94 62 94
67 95 07 76 30	44 67 32 23 13	12 72 72 27 77	51 57 32 22 27	28 30 62 58 83	13 08 60 46 28
54 50 06 44 75	46 44 33 63 71	56 59 75 36 75	66 86 65 64 60	35 14 82 56 80	59 78 98 76 14
56 23 27 19 03	30 80 05 19 29	65 00 51 93 51	19 83 52 47 53	18 88 26 95 54	48 00 26 43 85
11 28 94 15 52	89 99 93 39 79	34 87 96 23 95	51 78 57 26 17	94 61 47 03 10	88 84 59 69 14
77 55 33 62 02	62 03 55 86 57	70 39 83 66 56	79 68 96 26 60	13 13 38 69 96	43 83 10 13 24
35 96 29 00 45	17 27 27 51 26	48 21 47 74 63	73 52 93 70 50	91 59 11 38 44	05 30 08 46 32
02 84 48 51 97	19 79 95 07 21	09 04 03 35 78	63 99 25 69 02	22 95 30 19 29	20 25 36 70 69
49 90 21 69 74	93 74 21 86 33	71 30 32 06 47	86 28 30 02 35	20 39 84 95 61	99 78 78 83 82
75 27 28 52 13	17 18 16 90 46	56 33 24 87 36	97 96 47 59 97	87 73 19 38 47	83 43 32 26 26
92 42 85 04 31	42 37 13 81 83	97 50 81 79 59	93 41 69 96 07	69 33 17 03 02	42 69 60 17 42
34 59 43 36 96	35 63 02 31 61	40 33 04 46 24	40 24 74 36 42	56 39 59 89 63	78 22 87 10 88
71 63 94 94 33	26 41 77 63 37	18 78 80 36 85	06 06 16 25 98	64 13 09 37 11	31 66 60 65 64
61 63 00 25 92	98 10 39 33 15	39 46 67 21 17	97 81 26 03 89	98 78 80 63 23	20 96 06 79 80
32 15 99 67 43	56 16 88 87 60	90 14 79 61 55	65 99 59 97 84	18 62 85 28 24	39 65 01 73 91
68 49 20 43 29	31 85 33 69 07	60 22 66 72 17	16 91 21 32 41	85 66 48 38 73	48 78 58 08 88
50 93 19 35 56	78 12 03 09 70	35 26 99 18 25	62 03 89 26 32	43 61 00 66 42	33 86 76 71 66
92 20 32 39 67	98 81 99 37 29	37 11 05 75 16	92 27 73 40 38	68 52 16 83 34	48 32 72 26 95
21 68 40 95 79	95 66 39 01 09	00 84 14 36 37	45 51 94 69 04	97 00 12 91 33	83 97 68 95 65
13 81 20 67 58	03 35 63 05 77	12 08 05 75 26	00 81 06 28 48	01 52 48 69 57	02 41 03 89 33
25 76 01 54 03	72 93 78 04 36	60 60 29 99 93	05 06 42 24 07	96 88 22 46 94	60 73 04 84 98
83 79 68 20 66	70 81 10 94 91	60 09 71 87 89	12 68 46 55 89	91 51 63 27 95	88 29 04 79 84
61 58 87 08 05	85 79 76 48 23	67 85 72 37 41	07 79 26 69 61	98 55 83 46 09	41 49 36 83 43
48 84 60 37 65	32 25 34 03 36	62 95 80 97 63	52 16 16 23 56	01 98 00 89 85	10 10 71 19 45

A) ESTIMAR LA CONTRIBUCIÓN DEL AZAR EN LOS DATOS.

Una vez que se han recogido los datos deben utilizarse técnicas estadísticas apropiadas para estimar si las diferencias observadas entre el grupo de tratamiento y de control se deben o no al azar. También pueden utilizarse procedimientos estadísticos específicos para controlar sesgos o fuentes de sesgos conocidos. Sin embargo, ni el más sofisticado de los análisis estadísticos puede rescatar un pobre diseño o un trabajo ejecutado con desgana.

B) TEST DE HIPÓTESIS. P-VALOR

En el contexto de las pruebas de significación, el p-valor representa la probabilidad de que una diferencia (de medias, proporciones, etc) dada se observe en una muestra cuando en realidad esa diferencia no existe en la población relevante. Unos p-valores pequeños proporcionan evidencia para rechazar la hipótesis nula H_0 . Por ejemplo, un p-valor = 0,004 se puede interpretar como una probabilidad de 4‰ de observar una diferencia de determinada (o mayor) magnitud entre dos medias cuando, en la población, los dos grupos comparados tienen la misma media.

Normalmente, se dice que una diferencia es significativa si su p-valor es $\leq 0,05$; sin embargo, es preferible señalar los p-valores exactos a señalar simplemente “ $p < 0,05$ ”. A la hora de analizar la correlación y la regresión, la hipótesis nula que se contrasta es que los coeficientes de correlación o regresión son iguales a 0 (es decir, que no existe relación entre las variables en cuestión). El p-valor puede ser pensado como la probabilidad de que ocurra un error tipo I.

C) INTERVALO DE CONFIANZA

Un intervalo de confianza es un rango de valores entre el que se encuentra el verdadero valor de un parámetro o estimación de un conjunto de observaciones. Los parámetros de interés, también llamadas estimaciones puntuales, son la media, la proporción, la diferencia de medias y de proporciones, los coeficientes de correlación, los riesgos relativos, etc.

El cálculo de los IC se basa en la incertidumbre asociada con el uso de muestras para obtener información de las poblaciones de las que se extraen dichas muestras. Una estimación puntual es muy probablemente inexacta, de modo que el intervalo de confianza al 95% proporciona información adicional sobre el valor poblacional: podemos tener una confianza del 95% en que el valor poblacional se encuentra entre sus límites.

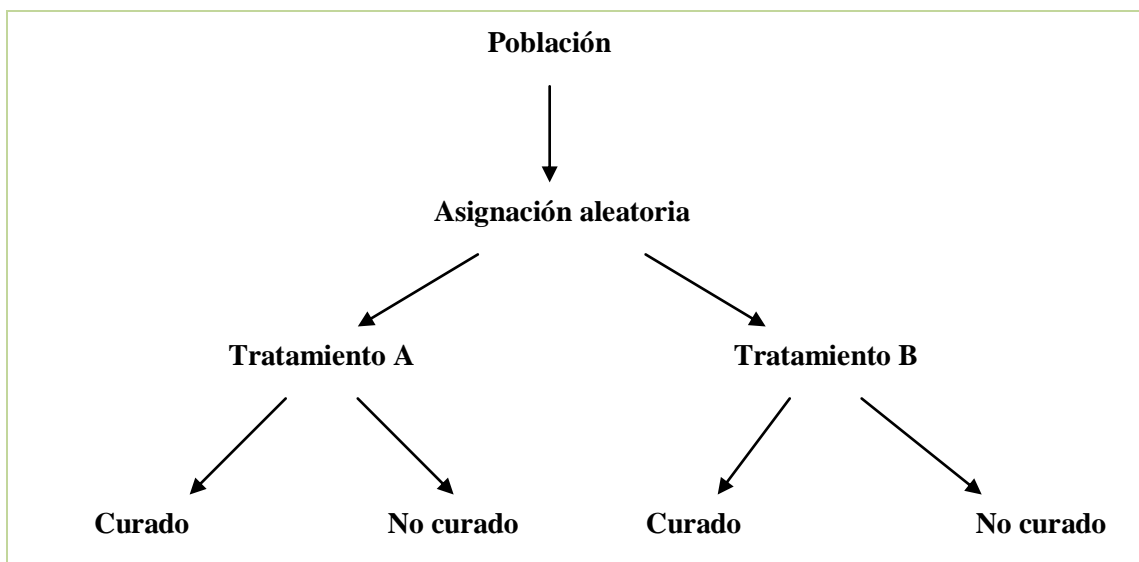
Se pueden establecer niveles de confianza distintos para un intervalo de confianza: los más comunes son 95%, 90%, 99%. Un intervalo de confianza al 99% es más ancho que un intervalo de confianza al 95%. La amplitud del intervalo de confianza depende del tamaño de la muestra: cuanto mayor es la muestra, más estrecho es el intervalo de confianza. Los IC son extremadamente útiles para evaluar la significación clínica de un resultado determinado. Los límites inferior y superior de un intervalo de confianza pueden despertar dudas sobre estimaciones puntuales si estos límites no tienen significación clínica.

Generalmente, un intervalo de confianza se calcula de la siguiente manera: 95% CI = estimador muestral $\pm 1,96 \times EE$. Esta fórmula es válida para muestras grandes. Para muestras más pequeñas (algo menos de 30 sujetos) el EE de la estimación se multiplica por el valor crítico de t, que se puede encontrar en las tablas de la distribución t de Student para los grados de libertad apropiados.

7) EL DISEÑO DEL ESTUDIO DEBE CONTEMPLAR BÁSICAMENTE:

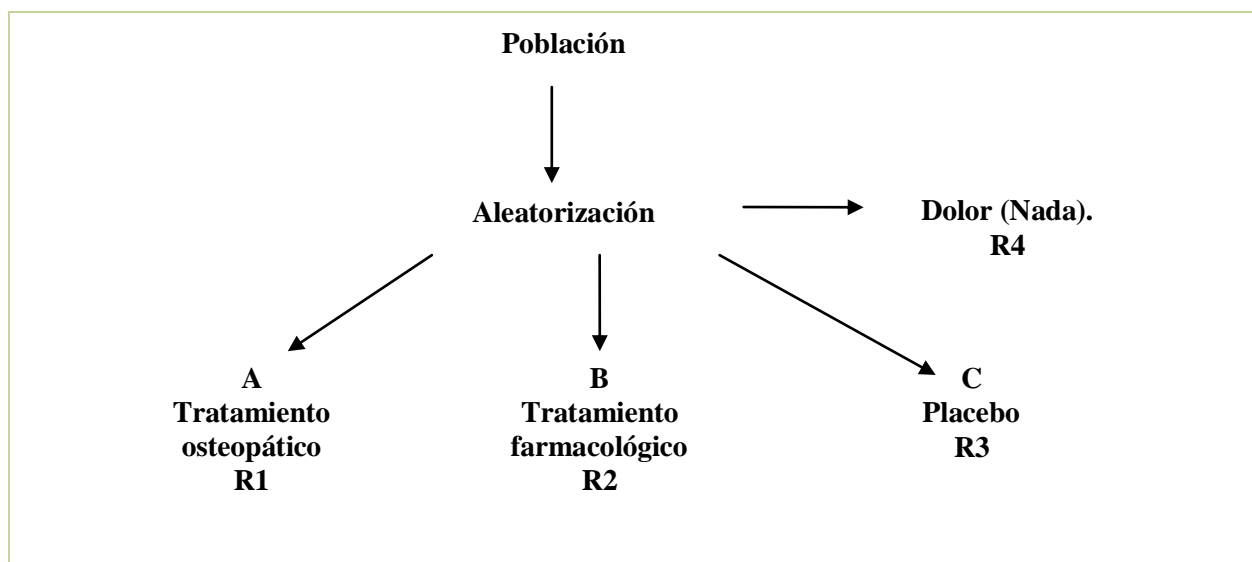
- Ética y justificación del ensayo.
- Población estudiada.
- Selección de los pacientes con consentimiento.
- Proceso de aleatorización.
- Descripción minuciosa de la intervención.
- Definición de la variable final de resultados (describir en detalle todo).

Tipo de diseño simple.



El grupo control puede ser:

- No tratado y se sigue su evolución.
- Tratado por otros medios y se compara su evolución con otra intervención.



NOTA:

- Utilizar un placebo está mal visto, es mejor comparar con el tratamiento estándar habitual.
- En osteopatía es mejor hacer un grupo de “simulación” (se hace otra intervención engañosa sin relación).
 - Grupo de técnica osteopática real.
 - Grupo control de osteopatía simulada.

K - ENSAYO CLÍNICO CRUZADO: +++

A diferencia del ECA en paralelo, donde cada paciente recibe sólo una intervención y los resultados obtenidos en cada uno de los grupos se comparan entre sí, en el ensayo clínico cruzado (cross-over), cada sujeto actúa como su propio control.

En el caso más sencillo, cada individuo recibe aleatoriamente, en un primer período, una de las dos intervenciones y, en un segundo período, la otra. Ambos períodos están separados por una fase de lavado o blanqueo, para permitir que el paciente vuelva a su estadio inicial, por lo que debe ser lo bastante largo para asegurar que el efecto del tratamiento administrado en el primer período ha desaparecido.

PRINCIPALES CARACTERÍSTICAS DIFERENCIALES ENTRE UN ESTUDIO EN PARALELO Y UNO CRUZADO

ESTUDIO EN PARALELO:

- *Cada participante recibe sólo una intervención*
- *Cada participante se asigna a un grupo*
- *El período de blanqueo no es necesario*

ESTUDIO CRUZADO:

- *Cada participante actúa como su propio control*
- *A cada participante se le asigna una secuencia de tratamientos*
- *El período de blanqueo es imprescindible*

El diseño cruzado sólo puede utilizarse en enfermedades crónicas, relativamente estables y en las que los resultados de una intervención desaparezcan de forma rápida.

1) IMPLICACIONES EN EL ANÁLISIS

Se recomienda desglosar el análisis en dos fases. En la primera se evalúa un posible efecto secuencia y si éste no es significativo (se utiliza un criterio poco restrictivo como, por ejemplo, $p = 0,10$), en la segunda se analiza el efecto del tratamiento. Si por el contrario, el efecto secuencia es significativo, el análisis del efecto tratamiento estará sesgado.

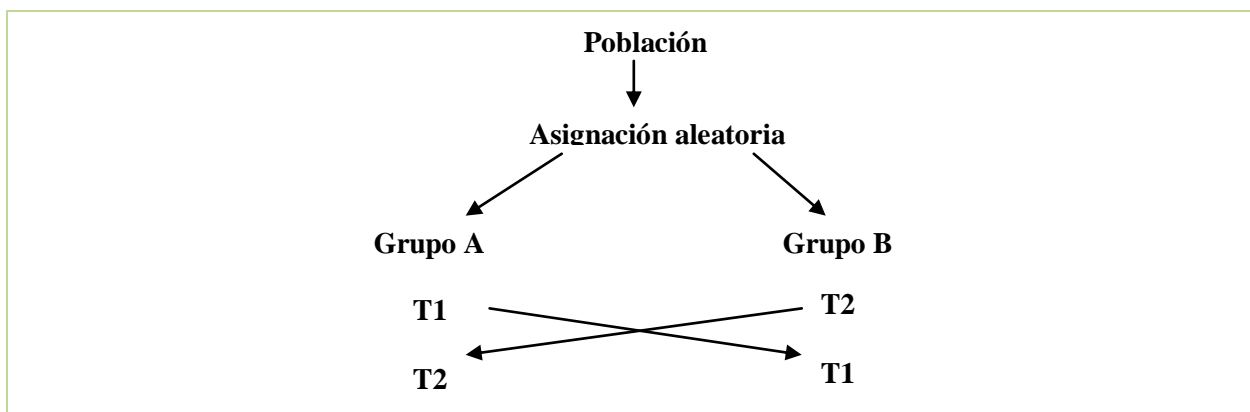
El inconveniente de este tipo de evaluación es que debe llevarse a cabo una vez se han recogido todos los datos, por lo que si se encuentra un efecto secuencia todo el trabajo habrá sido infructuoso. Lo único que se puede comparar son las respuestas observadas en el primer período.

2) VENTAJAS:

Su principal ventaja es su eficiencia ya que cada individuo actúa como su propio control, observándose el mismo número de respuestas que en un estudio en paralelo con la mitad de individuos.

Dado que encontrar el número de sujetos suficientes para completar un estudio es un problema frecuente, y más en Atención Primaria, ésta es una ventaja nada despreciable.

Al evaluar los dos tratamientos en un mismo paciente, la variabilidad intraindividual es menor, lo que permite utilizar pruebas estadísticas para datos apareados que tienen una mayor potencia y, por tanto, se precisa un número todavía menor de sujetos.



Se llama lavado o efecto de arrastre (este ensayo es adecuado para osteopatía). Cuando estemos interesados en reducir la variabilidad Inter. O intra pacientes.

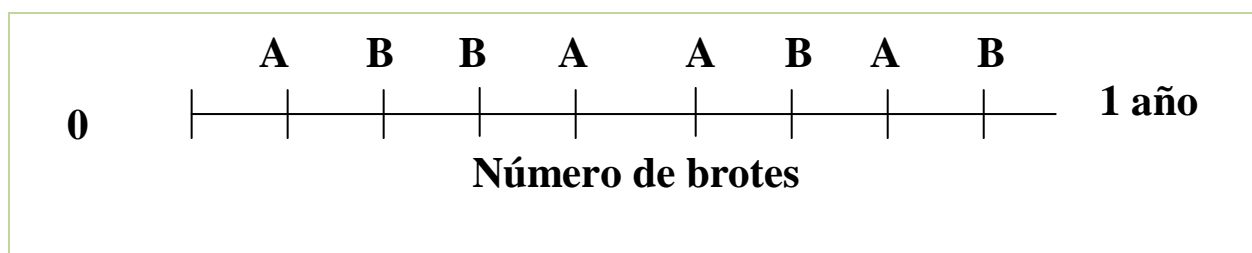
3) INCONVENIENTES:

- Estadísticas complejas.
- Podría ser interesante en osteopatía.
- Se utiliza para disminuir la variabilidad extra e intra individual entre pacientes.

L - ENSAYO = 1 paciente:

Se puede utilizar por ejemplo en la fibromialgia.

Tratamiento osteopático A / Simulación de tratamiento osteopático B.



Organización de los tratamientos:

TOA – TSB- TOA-TOA-TSB...

M – ENSAYO FACTORIAL:

Un ensayo clínico se diseña, en general, para responder a una única pregunta principal. Cuando se desea responder a dos cuestiones en una misma muestra de sujetos, puede utilizarse un diseño factorial, en el que cada paciente es asignado aleatoriamente a una combinación de dos o más tratamientos, o a una combinación de diferentes dosis de un mismo tratamiento.

Este diseño es muy eficiente para evaluar a la vez varios tratamientos con mecanismos de acción y efectos independientes con el mismo número de individuos que hubiera sido necesario para evaluar un solo tratamiento. Por otro lado, es en el diseño de elección cuando se pretende detectar interacciones, aunque en este caso el tamaño muestral se incrementa de forma importante.

La segunda situación en que es útil el diseño factorial para evaluar la existencia de interacción entre dos tratamientos. En este caso, dos grupos reciben uno de los tratamientos cada uno, un

tercer grupo actúa como control, mientras que el cuarto grupo recibe ambos tratamientos conjuntamente, de forma que, en este grupo, puede evaluarse la existencia de interacción.

1) INCONVENIENTE:

- Validez externa: interferencia a otras poblaciones.
- Razón ética.
- Riesgos de sesgos.

2) VENTAJAS: MAYOR CONTROL

VII – MÉTODOS DE VALIDACIÓN DE LOS ENSAYOS CLÍNICOS:

RRR = Reducción Relativa del Riesgo

Tasa de eventos en grupo control: TEC

Tasa de eventos en grupo experimental: TEE

TEC / TEE
15% 5%

$$RRR = \frac{TEC - TEE}{TEC} = \frac{15-5}{15} = 0,66 \text{ (66\%).}$$

RAR = Reducción Absoluta de Recaída

$$TEC - TEI = 10$$

Numero de sujetos necesarios a tratar (NNT)

$$\frac{100}{RAR} = \frac{100}{10} = 10$$

Más pequeña es NNT más eficaz es la intervención.
Es una comparación de medias.

VIII - ESTRATEGIA DE ANÁLISIS:

La finalidad del análisis es extraer la información necesaria a partir de las observaciones realizadas en el estudio para responder a la pregunta planteada en el objetivo. La planificación de la estrategia de análisis forma parte del diseño del estudio y no debe dejarse para después de haberse recogido los datos. Para planificar la estrategia es necesario tener muy presente el objetivo específico del estudio y el marco conceptual que permite establecer las interrelaciones entre las diferentes variables, lo que ayuda a definir cuál es el papel de cada una de ellas en el análisis.

A -REVISIÓN DE LOS DATOS:

Al finalizar el período de recogida de datos, es frecuente que el investigador esté impaciente por responder a la pregunta planteada, y desee realizar los pasos finales del análisis. El problema es que, si no se ha asegurado la calidad de los datos, los resultados del cálculo pueden ser erróneos. Por ello, una fase previa al análisis será la revisión de la matriz de datos.

Hay que prever un análisis de las variables una a una, en busca de valores no habituales o ilógicos, o de errores de transcripción o codificación. Para esta finalidad, son útiles las técnicas de presentación y síntesis de datos que componen la estadística descriptiva.

Una tabla de frecuencias permitirá detectar posibles anomalías, como errores de codificación o transcripción, que obligarán a revisar la hoja de recogida de datos para realizar las correcciones. Al mismo tiempo, podrá comprobarse si la distribución de frecuencias es similar a la esperada, de forma que podrán detectarse errores que de otra forma pasarían inadvertidos. Si se está estudiando una muestra presuntamente aleatoria de población demandante y aparece un 80% de hombres en la misma, podría indicar algún error en la codificación de la variable sexo.

La distribución de frecuencias puede indicar la necesidad de realizar agrupaciones o desagregaciones de determinados valores.

La elección de la medida de incidencia más adecuada viene condicionada por el objetivo del estudio. Si el interés está en predecir el cambio en el estado de salud de un individuo en función de alguna característica, entonces se deberá estimar el riesgo de que este cambio ocurra (incidencia acumulada). Si el tiempo de seguimiento es muy desigual entre los sujetos de la muestra, deberá optarse por la densidad de incidencia.

B - CALCULAR EL INTERVALO DE CONFIANZA DEL PARÁMETRO:

Sea cual fuere la medida de frecuencia utilizada, las estimaciones puntuales deben acompañarse de sus correspondientes intervalos de confianza. De hecho, el valor observado en la muestra estudiada no es más que uno de los posibles valores que hubieran podido obtenerse al estudiar las múltiples muestras que se pueden extraer de una población. Diferentes muestras producirían diferentes resultados. Por ello, será necesaria una medida de la precisión de esta estimación que permita conocer entre qué límites se encuentra el verdadero valor de la

población con una confianza determinada, lo que se hará mediante el cálculo del llamado intervalo de confianza (IC).

El IC proporciona mucha más información que la simple estimación puntual, ya que permite evaluar la precisión con que el parámetro poblacional ha sido estimado, es decir, cuán próxima está la estimación puntual del verdadero pero desconocido valor. Hay que decidir el grado de confianza con que se realizará la estimación, que suele ser del 95%.

C - DESCRIBIR LAS NO RESPUESTAS Y/O PÉRDIDAS DE SEGUIMIENTO Y EVALUAR SU IMPACTO POTENCIAL:

La descripción de las no respuestas y pérdidas deberá incluir, no sólo su número, sino también sus motivos, si son conocidos. El problema principal es que pueden sesgar los resultados.

Deben preverse estrategias para evaluar el impacto potencial del posible sesgo: Comparar los sujetos que responden y los que no lo hacen, por las variables que estén disponibles. Generalmente estas variables son las que se utilizan para seleccionar la muestra de estudio. Si al compararlos por estas variables se observan diferencias entre ellos, este dato sugerirá que las no respuestas han introducido un sesgo de selección.

D - ANÁLISIS DE SUBGRUPOS:

Cuando los investigadores piensan que la frecuencia del problema de salud puede variar en función de diferentes características, puede ser de interés planificar la estimación del parámetro en los diferentes subgrupos de interés.

Si se desea realizar análisis de subgrupos, deberá tenerse en cuenta en el cálculo del tamaño de la muestra y en el método de selección de los sujetos, ya que, en caso de no hacerlo, se perderá precisión en la estimación del parámetro en cada subgrupo en relación a la obtenida cuando se analiza el total de la muestra, ya que el número de sujetos será claramente inferior.

E - ESTRATEGIA DE ANÁLISIS DE UN ESTUDIO ANALÍTICO:

El objetivo de los estudios analíticos es estimar la magnitud del efecto o la asociación entre un factor de estudio y una variable de respuesta. La finalidad de la estrategia de análisis es obtener la estimación más válida, eliminando la influencia de todas las variables que puedan interferir, y lo más precisa posible.

PASOS DE LA ESTRATEGIA DE ANÁLISIS DE UN ESTUDIO ANALÍTICO

0. REVISAR LOS DATOS

1. DESCRIBIR LOS SUJETOS ESTUDIADOS

2. EVALUAR LA COMPARABILIDAD INICIAL DE LOS GRUPOS

3. ESTIMAR LA EXISTENCIA Y MAGNITUD DEL EFECTO O ASOCIACIÓN

4. EVALUAR LA MODIFICACIÓN DEL EFECTO
5. AJUSTAR POR POTENCIALES FACTORES DE CONFUSIÓN
6. ANÁLISIS DE SUBGRUPOS
7. RESPONDER LAS PREGUNTAS SECUNDARIAS

F - DESCRIBIR LOS SUJETOS ESTUDIADOS:

Su utilidad y las características por las que se realiza esta descripción son similares a las comentadas para los estudios descriptivos.

G - EVALUAR LA COMPARABILIDAD INICIAL DE LOS GRUPOS:

El análisis se sustenta en que los grupos de estudio son comparables y no existen sesgos que alteren los resultados. Por ello, antes de determinar si el efecto o la asociación es diferente entre los grupos debe planificarse un análisis de su comparabilidad respecto a las variables que puedan influir sobre la respuesta. Su utilidad es que, si se encuentran diferencias, deberán ser tenidas en cuenta en fases posteriores del análisis como potenciales factores de confusión.

El análisis de la comparabilidad no se sustentará sólo en criterios de significación estadística. No es infrecuente que algunas de estas comparaciones sean estadísticamente significativas sólo por azar. Pueden existir también diferencias no detectadas como estadísticamente significativas por las pruebas estadísticas, sobre todo si el número de sujetos es reducido, pero con una influencia relevante sobre el efecto que se estudia.

H - ELECCIÓN DE LA MEDIDA DEL EFECTO:

Existen diferentes formas para expresar los resultados. La elección de la más adecuada depende del tipo de estudio y su finalidad, y de la escala de medida de las variables.

Existen dos tipos de medidas del efecto: relativas y absolutas. Las medidas relativas más utilizadas son el riesgo relativo (RR) y la odds ratio (OR):

- El RR corresponde al cociente entre las incidencias observadas en el grupo expuesto y no expuesto y, por tanto, sólo puede calcularse directamente en aquellos estudios que permitan determinar incidencias, es decir, en los estudios de cohortes y en los ensayos clínicos. El RR indica el número de veces que es más frecuente la aparición de la respuesta en un grupo respecto al de referencia. Por este motivo, al preparar la estrategia de análisis, es importante decidir cuál será la categoría que se utilizará como referencia y definirla con claridad.
- En los estudios de casos y controles, la medida relativa que se utiliza es la OR. Si el estudio está bien diseñado y se cumplen determinadas asunciones, la OR es una buena estimación del RR, por lo que se interpreta de forma similar a aquél. Cada vez es más

frecuente el uso de la OR como medida relativa del efecto en investigación clínica y epidemiológica, debido a que puede calcularse independientemente del diseño del estudio. Además, las técnicas estadísticas multivariantes de uso habitual, como la regresión logística o el modelo de Cox, facilitan su cálculo.

Otra forma de expresar los resultados de un estudio es a través de medidas absolutas. La más utilizada es la diferencia de riesgos o diferencia entre las incidencias acumuladas observadas en ambos grupos. En los estudios de casos y controles no puede determinarse directamente, ya que no permiten estimar la incidencia acumulada. Respecto a las medidas relativas, tiene la ventaja de que proporciona una idea del impacto que tendría sobre la incidencia la eliminación de un factor de riesgo o la aplicación de una intervención.

Además de medir el efecto en la escala absoluta como diferencia en las respuestas observadas, se estima el efecto en una escala relativa (OR = 1,4; IC 95%: 0,8 a 2,4).

I - VARIABLES A CONTROLAR:

Al planificar la estrategia de análisis, es importante identificar todas las variables implicadas, lo que debe hacerse a partir del marco conceptual elaborado al definir el objetivo del estudio.

Según su papel en la relación causa-efecto que se esté estudiando, pueden distinguirse los siguientes tipos de variables:

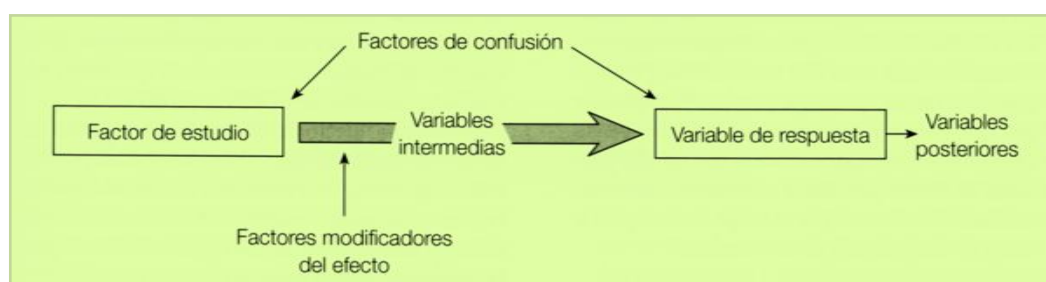


Figura: Papel de las variables en la estrategia de análisis.

- Variable de respuesta. En términos estadísticos, suele denominarse variable dependiente.
- Factor de estudio. Suele denominarse variable independiente principal, ya que la finalidad del análisis es estimar su efecto sobre la variable dependiente, ajustando o controlando la influencia del resto de variables.
- Variables modificadoras del efecto. Cuando el efecto o la asociación entre el factor de estudio y la variable de respuesta varían según la existencia o diferentes grados de una tercera variable, se dice que ésta es modificadora de dicho efecto. Estas variables son importantes porque forman parte del propio efecto del factor de estudio, por lo que, más que controlar por ellas, interesa tenerlas en cuenta en el análisis para poder describir cómo modifican dicho efecto.

- Potenciales factores de confusión. Son variables que están relacionadas tanto con el factor de estudio como con la variable de respuesta de forma que, si no se tienen en cuenta en el análisis, se obtiene una estimación sesgada del efecto. Por tanto, deberá ajustarse la estimación del efecto por estas variables.
- Variables de la cadena causal. Se trata de variables relacionadas con el factor de estudio y con la variable de respuesta, pero que se encuentran situadas en la cadena causal, ya sea antes (variables intermedias) o después de la variable de respuesta (variables posteriores). En principio no debe ajustarse por este tipo de variables, ya que, al formar parte de la cadena causal, la estimación ajustada del efecto del factor de estudio no tendría en cuenta el componente que pasa a través de las variables intermedias o relacionado con las variables posteriores.

J - ELECCIÓN DE LA PRUEBA ESTADÍSTICA:

En la planificación de la estrategia de análisis es necesario prever las pruebas estadísticas que se utilizarán, y que dependen fundamentalmente del tipo de variables implicadas, el número de grupos que se comparan, el tamaño de la muestra y condiciones de aplicación específicas de cada una de las pruebas.

Por otro lado, si es necesario obtener una estimación del efecto ajustando por múltiples variables, debe recurrirse a modelos estadísticos multivariantes. La elección del modelo apropiado depende de la escala de medida de la variable de respuesta (variable dependiente) y de otras condiciones de aplicación específicas. Los modelos más utilizados en investigación clínica y epidemiológica son la regresión lineal múltiple (cuando la variable de respuesta es cuantitativa), la regresión logística (cuando es dicotómica) y el modelo de Cox (cuando la variable de respuesta es el tiempo de aparición de un suceso).

K - DETERMINAR LA PRECISIÓN DE LA ESTIMACIÓN DEL EFECTO:

Una vez calculada la estimación que se considera más válida deberá acompañarse de su correspondiente intervalo de confianza como medida de la precisión.

L - ANÁLISIS DE SUBGRUPOS:

El análisis de subgrupos se realiza cuando interesa evaluar los resultados en algún grupo especial de sujetos, habitualmente en función de su edad, sexo o alguna variable pronóstica.

El análisis por subgrupos se justifica si se ha planeado previamente al inicio del estudio, en función de una hipótesis fundamentada. En algunas ocasiones, aunque no haya sido establecido antes, el análisis de los resultados observados en algún subgrupo de sujetos puede ser razonable. Sin embargo, no hay que olvidar que, si se define un gran número de subgrupos en función de diferentes combinaciones de variables, un análisis indiscriminado puede conducir a la obtención de conclusiones erróneas, al aparecer el problema de las comparaciones múltiples.

PREGUNTAS SECUNDARIAS

A menudo, interesa responder preguntas secundarias. Por ejemplo, si se va a realizar un estudio sobre la eficacia de un programa de educación sanitaria sobre el control metabólico de los pacientes diabéticos, también se desearía responder a preguntas como:

- ¿Qué grupo de población ha asistido a las sesiones?,
- ¿Qué características socioculturales definen a los no cumplidores?,
- ¿Qué individuos han incrementado su grado de conocimientos sobre la enfermedad?,
etc.

Estas respuestas podrían ayudar a enriquecer la controversia e interpretación de los resultados. Sin embargo, están sometidas a los mismos inconvenientes que se han citado para el análisis por subgrupos, por lo que deben ser escasas en número, establecidas a priori y estar debidamente fundamentadas. En su interpretación, debe recordarse las limitaciones de las comparaciones múltiples.

VIII – PAPEL DE LA ESTADÍSTICA SEGÚN ARGIMON PALLÁS Y JIMÉNEZ VILLA*

La estadística debe considerarse como una parte integrante del método científico. Su adecuada utilización permitirá evaluar e incluso cuantificar la variabilidad debida al azar y, si es posible, reducirla, de forma que el estudio tenga las máximas garantías para alcanzar el objetivo deseado. La premisa previa para su uso es que el estudio haya sido diseñado y ejecutado de forma correcta.

Clásicamente, la estadística se diferencia en descriptiva e inferencial.

- La estadística descriptiva comprende la organización, la presentación y la síntesis de la información y es fundamental en la revisión de los datos recogidos en un estudio para asegurar su calidad y la validez del análisis posterior, así como para describir la muestra de sujetos estudiados.
- La estadística inferencial comprende las bases lógicas mediante las cuales se establecen conclusiones relacionadas con poblaciones a partir de los resultados obtenidos en muestras. Su aplicación en la fase de análisis tiene dos finalidades principales: evaluar la variabilidad aleatoria y controlar los factores de confusión.

Fases de una investigación en las que interviene la estadística.

- Selección de la variable de respuesta
- Definición de los criterios de selección de la población de estudio
- Elección de la técnica de selección de los sujetos
- Cálculo del número de sujetos necesarios
- Selección de las variables que deben ser medidas
- Medición de las variables (precisión y exactitud)
- Descripción de la muestra de sujetos estudiados
- Estimación de la magnitud del efecto o respuesta observada
- Comparación del efecto observado en diferentes grupos
- Control de los factores de confusión
- Interpretación de los resultados.

A - LA MUESTRA:

De forma esquemática, en el caso de una variable cuantitativa, cada muestra presentaría una media diferente. Si se representara gráficamente la distribución de estas medias muestrales, podría comprobarse que sigue la ley normal.

Esta distribución de medias muestrales tiene dos características que la hacen especialmente interesante.

- La primera es que su media es la media de la población de la que proceden las muestras, es decir, que las medias muestrales se distribuyen normalmente alrededor de

la media poblacional desconocida que se desea estimar.

- La segunda es que la desviación estándar de dicha distribución tiene el valor σ/\sqrt{n} siendo σ la desviación estándar de la población de origen y n el tamaño de la muestra.

Este valor es conocido como error estándar de la media (EEM), y mide la dispersión de las medias muestrales respecto a la media poblacional. El EEM no debe ser confundido con la desviación estándar σ o s , que son medidas de la dispersión de los valores de la variable en la población y en la muestra, respectivamente. Puede deducirse fácilmente que el EEM disminuye cuando aumenta el tamaño de la muestra n , lo que explica el hecho de que las muestras grandes estimen el valor poblacional con mayor precisión.

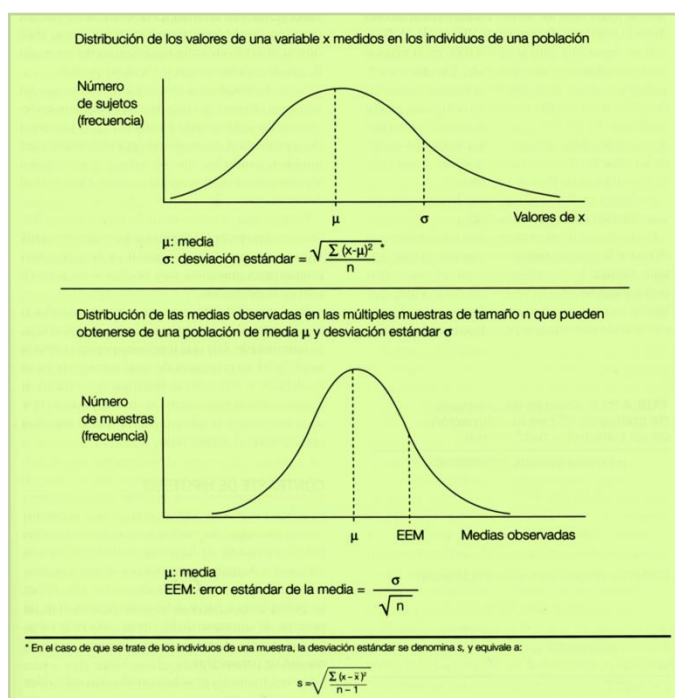


Figura Diferencia entre desviación estándar y error estándar de la media. Distribución de los valores de una variable x medidos en los individuos de una población.

En el caso de las variables cualitativas, la distribución de la proporción sigue la ley binomial. Sin embargo, cuando los productos $n \cdot p$ y $n \cdot (1-p)$ son superiores a 5, se acepta que la distribución se asemeja a la normal. En este caso, las proporciones observadas en las muestras se distribuyen alrededor de la verdadera proporción poblacional. La dispersión de esta distribución se mide mediante el error estándar de la proporción (EEP), cuyo valor es $\sqrt{p \cdot (1-p)/n}$, de características similares al EEM.

B - ESTIMACIÓN DE UN PARÁMETRO POBLACIONAL: INTERVALOS DE CONFIANZA

Un objetivo frecuente en investigación médica es el de estimar un parámetro poblacional a partir de los valores que la variable de interés adopta en los individuos de una muestra. Si la

variable es cuantitativa, la media m y la desviación estándar s observadas en la muestra son la mejor estimación que se dispone de los verdaderos valores de los parámetros poblacionales.

Diferentes muestras conducirían a diferentes resultados. Por ello, se necesita una medida de la precisión de esta estimación, lo que se hace mediante el cálculo del llamado intervalo de confianza (IC).

Si se desea una confianza del 95% en la estimación, se trabaja con un valor α del 5%, que corresponde a un valor Z (distribución normal tipificada) de 1,96. En el ejemplo, aplicando la fórmula, se obtendría un IC del 95% que sería aproximadamente de 150 ± 5 mmHg, lo que significa que la PAS media de la población de referencia está situada entre 145 y 155 mmHg con un 95% de confianza.

El cálculo del IC proporciona mucha más información que la simple estimación puntual, ya que permite evaluar la precisión con que el parámetro poblacional ha sido estimado, es decir, entre qué límites se tiene una determinada confianza de que esté situado su verdadero pero desconocido valor. Si se repitiera el estudio en 100 ocasiones, el IC incluiría el verdadero valor en 95 de ellas. Sin embargo, como no puede descartarse totalmente que se trate de una de las 5 ocasiones restantes, un determinado IC puede contener o no este verdadero valor.

<p>IC de una media (variable cuantitativa)*:</p> $\bar{x} \pm (Z\alpha \cdot EEM)$ <p style="text-align: center;">siendo $EEM = \frac{s}{\sqrt{n}}$</p>
<p>IC de una proporción (variable cualitativa)**:</p> $p \pm (Z\alpha \cdot EEP)$ <p style="text-align: center;">siendo $EEP = \sqrt{\frac{p \cdot (1-p)}{n}}$</p>
<p>\bar{x}: media observada en la muestra. s: desviación estándar observada en la muestra. n: número de individuos de la muestra. EEM: error estándar de la media. p: proporción observada en la muestra. EEP: error estándar de la proporción. $Z\alpha$: valor de la variable normal tipificada correspondiente al valor α para un valor de confianza $(1-\alpha)$.</p> <p>* Este cálculo se basa en la distribución normal. El valor de $Z\alpha$ para un IC del 95% es 1,96. Para muestras de tamaño inferior a 30 individuos, este valor debe sustituirse por el de la distribución de la t de Student-Fisher para $(n-1)$ grados de libertad. ** Las variables cualitativas no presentan una distribución normal. Las fórmulas de la tabla se basan en una aproximación a la normalidad, aplicable cuando los productos $n \cdot p$ y $n \cdot (1-p)$ son mayores de 5. En caso contrario, debe aplicarse una corrección al valor de $Z\alpha$.</p>

TABLA Cálculo del intervalo de confianza (IC) en la estimación de un parámetro poblacional.

De las fórmulas de la tabla se deduce que un aumento del número de sujetos produce un estrechamiento del intervalo y aumenta así la precisión de la estimación. Su amplitud depende también del grado de confianza que se utilice, por ejemplo, aumentando si se incrementa su valor convencional del 95 al 99%.

Siempre que se esté estimando parámetros poblacionales a partir de estadísticos muestrales, los resultados deben expresarse como IC, y no sólo como estimaciones puntuales, para poder valorar la precisión de la estimación.

En el cálculo del IC se asume que se ha estudiado una muestra aleatoria de la población de referencia. Al interpretarlo, hay que tener siempre en cuenta la posibilidad de existencia de otras fuentes de error no debidas al azar (errores sistemáticos o sesgos). Si éstos existen, o si la muestra no es aleatoria, el error de la estimación puede ser mayor que el sugerido por la amplitud del intervalo.

C - CONTRASTE DE HIPÓTESIS:

La aplicación más frecuente de la inferencia estadística en investigación médica son las llamadas pruebas de contraste de hipótesis o de significación estadística. Aunque la situación es similar a la anterior, dado que se pretende descubrir algo sobre las poblaciones a partir del estudio de muestras, las pruebas de contraste de hipótesis valoran la variabilidad debida al azar de forma diferente a la estimación de parámetros.

Existen múltiples pruebas estadísticas aplicables en diferentes situaciones en función del número de grupos que se comparan, la escala de medida de las variables, el número de sujetos analizados, etc.

D - HIPÓTESIS NULA E HIPÓTESIS ALTERNATIVA:

La hipótesis que en realidad se va a contrastar estadísticamente es la de que no existen diferencias entre los porcentajes de hipertensos controlados observados en ambos grupos. La prueba de significación estadística intentará rechazar esta hipótesis, conocida como hipótesis nula (H_0). Si se consigue, se aceptará la hipótesis alternativa (H_a) de que existen diferencias entre ambos grupos.

El primer paso de una prueba de significación es formular la H_0 . A continuación se calcula, mediante la prueba estadística más adecuada, la probabilidad de que los resultados observados puedan ser debidos al azar, en el supuesto de que la H_0 sea cierta. En otras palabras, la probabilidad de que, a partir de una población de referencia, puedan obtenerse dos muestras que presenten unos porcentajes tan diferentes como los observados.

Esta probabilidad es el grado de significación estadística, y suele representarse con la letra p . En tercer lugar, basándose en esta probabilidad, se decide si se rechaza o no la H_0 .

Cuanto menor sea la p , es decir, cuanto menor sea la probabilidad de que el azar pueda haber producido los resultados observados, mayor será la evidencia en contra de H_0 y, por tanto, mayor será la tendencia a concluir que la diferencia existe en la realidad.

En el ejemplo, una vez aplicada la prueba estadística adecuada, se obtiene un valor de p aproximadamente de 0,10. Esto significa que, si H_0 fuera cierta, la probabilidad de que el azar pueda producir unos resultados como los observados es del 10%, o bien, que existe un 10% de

probabilidad de que dos muestras de 30 sujetos obtenidas de una misma población presenten unos porcentajes del 70 y el 50% sólo por variabilidad aleatoria.

Conocida esta probabilidad, se decide si se rechaza o no la H_0 . Para ello, debe haberse fijado previamente un valor de p por debajo del cual se considerará que se dispone de la suficiente evidencia en contra de la H_0 para rechazarla. Este valor es conocido como el valor de significación α . De forma arbitraria, y por convenio, se adopta el valor del 5% o del 0,05.

Si se adopta el valor 0,05 para el de significación, dado que el valor de p obtenido en el ejemplo es de 0,10, se considerará que la probabilidad de haber obtenido estos resultados por azar es demasiado elevada y que, por tanto, no se dispone de la suficiente evidencia para rechazar la H_0 .

Se concluye que no se han encontrado diferencias estadísticamente significativas en el porcentaje de pacientes controlados en ambos grupos. No se concluye que ambos grupos son iguales, sino que no se ha encontrado la suficiente evidencia para decir que son diferentes.

Supongamos ahora que en el grupo que recibió el tratamiento E sólo se hubieran controlado 12 pacientes a los 3 meses, lo que supone un porcentaje observado en este grupo del 40%. Si se repiten los cálculos, se obtiene un valor de p menor a 0,02.

Como este valor es inferior al valor de significación del 0,05, se considerará que la diferencia observada es estadísticamente significativa, ya que es poco probable ($p < 5\%$) que el azar pueda haber producido estos resultados. La respuesta a la pregunta de si esta diferencia es debida al nuevo tratamiento D dependerá del diseño correcto del estudio. El valor de p sólo informa de la existencia de una diferencia entre ambos grupos, y de que muy probablemente no es debida al azar.

El verdadero interés de la p es el de permitir descartar que la diferencia observada es fruto de la variabilidad aleatoria. No es una medida de la fuerza de la asociación. Un estudio en el que se obtenga una $p < 0,001$ no quiere decir que la asociación encontrada sea más fuerte (o la diferencia más importante) que otro estudio en que la p sea del 0,05. Sólo quiere decir que es más improbable que su resultado sea debido al azar.

No hay que ser excesivamente rígido en el límite del valor de significación. Una p de 0,048 es estadísticamente significativa al valor del 5%, y una p de 0,052, en cambio, no lo es, pero en ambos casos la probabilidad de observar el resultado por azar es prácticamente la misma, y muy próxima al 5%. Por ello, es conveniente indicar el valor obtenido de p al informar unos resultados, sobre todo si es próximo al valor de significación fijado, en lugar de limitarse a decir si existe o no significación estadística. De esta forma, el lector podrá valorar adecuadamente los resultados.

E - PRUEBAS UNILATERALES Y PRUEBAS BILATERALES:

En ocasiones, lo que interesa no es determinar si existen o no diferencias entre dos tratamientos, sino evaluar si un nuevo fármaco es mejor que otro. En este caso, la hipótesis alternativa no es que D y E difieren, sino que D es mejor que E. Por tanto, la H_0 que se va a

contrastar es que D no difiere o es peor que E. Dado que sólo interesa un sentido de la comparación, se habla de pruebas unilaterales o de una cola.

¿Cómo afecta este hecho a la prueba de significación? No es la prueba en sí la que está afectada. El cálculo es idéntico al anterior. Lo que se modifica es el grado de significación alcanzado. Como la distribución de Z sigue la ley normal y, por tanto, es simétrica, en las pruebas unilaterales el verdadero valor de p es la mitad del valor α , dado que sólo se está interesado en uno de los extremos. Existen tablas que proporcionan directamente los valores α para pruebas unilaterales.

F - ERROR α a Y ERROR β :

En estadística no puede hablarse de certeza absoluta, sino de mayor o menor probabilidad. Sea cual fuere la decisión que se tome respecto a la hipótesis nula, se corre un cierto riesgo de equivocarse.

La realidad no es conocida, ya que, si lo fuera, no sería necesario realizar el estudio. Si no se rechaza la H_0 , y ésta es cierta, no se comete ningún error.

Si se rechaza y es falsa, tampoco se comete un error. Pero, ¿qué pasa en las otras dos situaciones? En un estudio, puede concluirse que existen diferencias cuando de hecho no las hay. Es decir, puede rechazarse la H_0 cuando es cierta. Si esto ocurre, la decisión es incorrecta y se comete un error, conocido como error tipo 1 o error α .

La probabilidad de cometer este tipo de error es la probabilidad de que, si se concluye que existen diferencias significativas, éstas sean en realidad debidas al azar. Si se hace un símil entre una prueba estadística y una diagnóstica, equivale a la probabilidad de obtener un resultado falso positivo. Esto es precisamente lo que mide la p o grado de significación estadística de la prueba.

Si, por el contrario, se concluye que no existen diferencias estadísticamente significativas, es decir, si no puede rechazarse la H_0 , puede ocurrir que en realidad ésta sea falsa y sí existan diferencias entre ambos grupos, en cuyo caso se comete otro tipo de error, llamado error β o tipo II. Utilizando el símil con la prueba diagnóstica, equivale a la probabilidad de obtener un resultado falso negativo.

Su valor complementario ($1-\beta$), denominado potencia o poder estadístico, indica la capacidad que tiene la prueba para detectar una diferencia cuando ésta existe en la realidad. Lógicamente, cuanto mayor es la diferencia existente entre dos poblaciones y mayor el número de individuos estudiados, mayor capacidad existe para detectarla, es decir, el poder estadístico es mayor y, por tanto, la probabilidad de cometer un error tipo II es menor.

METODOLOGÍA E INVESTIGACIÓN

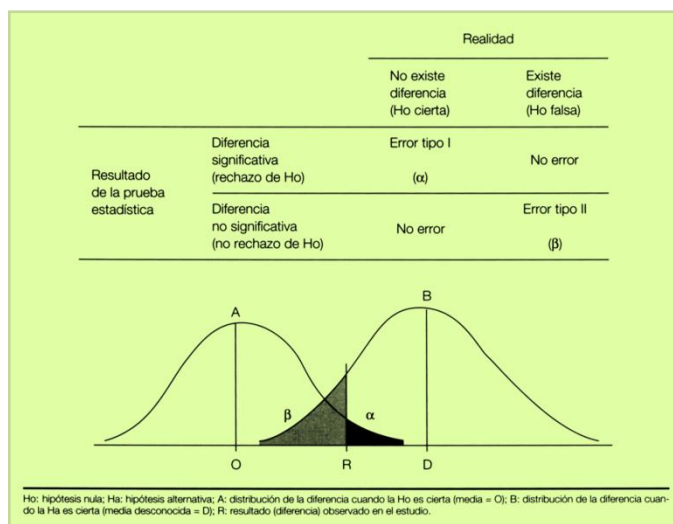


Figura Resultados de una prueba de significación estadística.

Existe una interdependencia entre el grado de significación (p), el poder estadístico ($1-\beta$), el número de individuos estudiados y la magnitud de la diferencia observada. Conociendo tres de estos parámetros, puede calcularse el cuarto. Antes de iniciar un estudio se calcula el número de sujetos necesarios para detectar una determinada diferencia, fijando a priori el valor de significación y el riesgo de cometer un error tipo II que se está dispuesto a asumir. Si una vez acabado el estudio, se concluye que no se han encontrado diferencias estadísticamente significativas, dado que n , p y la diferencia observada son conocidas, puede calcularse el poder estadístico, lo que es fundamental para evaluar de forma adecuada el resultado. No es lo mismo concluir que no se ha encontrado una diferencia estadísticamente significativa entre dos tratamientos cuando se tiene una probabilidad del 90% de haberla detectado si hubiera existido ($\beta = 0,10$), que cuando esta probabilidad es sólo del 50% ($\beta = 0,50$).

¿DIFERENCIA ESTADÍSTICAMENTE SIGNIFICATIVA O CLÍNICAMENTE RELEVANTE?

Un resultado estadísticamente significativo no implica que sea clínicamente relevante. El valor de la p no mide la fuerza de la asociación. Pueden obtenerse valores pequeños de p (por tanto, resultados estadísticamente significativos), simplemente estudiando un número grande de sujetos. Al aumentar el tamaño de la muestra, se incrementa el poder estadístico para detectar incluso pequeñas diferencias.

Número de sujetos estudiados por grupo	Valor de Z*	Valor de p
30	0,41	0,68
100	0,76	0,45
200	1,06	0,29
500	1,69	0,09
700	2,00	0,05
1.000	2,39	< 0,02
2.000	3,38	< 0,001

* Valor de la Z de comparación de dos proporciones.

TABLA Influencia del número de sujetos estudiados sobre el grado de significación estadística de la comparación de los porcentajes de pacientes controlados con dos tratamientos: D (70%) y E (65%).

La diferencia que se considera clínicamente relevante depende de su magnitud y de otros factores, como la frecuencia y gravedad de los efectos secundarios de ambos fármacos, la facilidad de administración o su coste económico.

IC de la diferencia de dos proporciones*

a) Muestras independientes

$$(P_A - P_B) \pm Z_{\alpha} \cdot EED$$

siendo $EED = \sqrt{\frac{P_A \cdot (1 - P_A)}{n_A} + \frac{P_B \cdot (1 - P_B)}{n_B}}$

b) Muestras apareadas

$$(P_A - P_B) \pm Z_{\alpha} \cdot EED$$

siendo $EED = \frac{1}{n} \sqrt{(b+c) \frac{(b-c)^2}{n}}$

P_A, P_B : proporciones observadas en las muestras A y B.
 n_A, n_B : número de sujetos de las muestras A y B.
 b, c: número de casos que presentan valores diferentes en ambas mediciones (series apareadas).
 n: número total de casos.
 EED: error estándar de la diferencia.
 Z_{α} : valor de la variable normal tipificada correspondiente al valor α , para un nivel de confianza $(1-\alpha)$.

* Las variables cualitativas no presentan una distribución normal. Las fórmulas de la tabla corresponden a una aproximación a la normalidad, aplicable cuando todos los productos $n \cdot P_A$, $n \cdot (1 - P_A)$, $n \cdot P_B$ y $n \cdot (1 - P_B)$ son mayores de 5. En caso contrario, deben aplicarse correcciones en el valor de Z_{α} .

Cálculo del intervalo de confianza (IC) de la diferencia entre dos proporciones

G - ESTIMACIÓN FRENTE A SIGNIFICACIÓN ESTADÍSTICA:

Los investigadores están interesados no sólo en saber si una diferencia o asociación es estadísticamente significativa, sino también en determinar su magnitud. El valor observado en el estudio es la mejor estimación puntual de dicha magnitud. Si se repitiera el estudio con otras muestras, podrían observarse resultados de diferente magnitud.

Por tanto, hay que calcular un IC que contenga, con una determinada confianza, la verdadera magnitud de interés, tanto si se trata de una diferencia como de una medida de asociación. Existen fórmulas para el cálculo de los IC en cualquier situación.

Cuando se utiliza como medida del efecto una diferencia, si el IC del 95% incluye el valor 0, que es el valor correspondiente a la H_0 de que no existe diferencia entre ambos grupos, se concluirá que el resultado no es estadísticamente significativo. Si, por el contrario, el IC del 95% excluye este valor 0, se concluirá que la diferencia observada es estadísticamente significativa.

Además de saber si la diferencia es o no estadísticamente significativa, el IC permite conocer entre qué límites es probable que se encuentre la verdadera diferencia, lo que es muy útil en la interpretación de los resultados.

IC de la diferencia de dos medias*

a) Muestras independientes

$$(\bar{x}_A - \bar{x}_B) \pm Z\alpha \cdot \text{EED}$$

siendo $\text{EED} = s \cdot \sqrt{\frac{1}{n_A} + \frac{1}{n_B}}$

y $s = \sqrt{\frac{s_A^2(n_A - 1) + s_B^2(n_B - 1)}{n_A + n_B - 2}}$

b) Muestras apareadas

$$\bar{x}_d \pm Z\alpha \cdot \text{EE}\bar{x}_d$$

\bar{x}_A, \bar{x}_B : medias observadas en las muestras A y B.
 s_A, s_B : desviaciones estándar observadas en las muestras A y B.
 n_A, n_B : número de sujetos de las muestras A y B.
 EED: error estándar de la diferencia.
 \bar{x}_d : media de las diferencias de las dos mediciones en cada individuo (series apareadas).
 $\text{EE}\bar{x}_d$: error estándar de la media de las diferencias individuales.
 $Z\alpha$: valor de la variable normal tipificada correspondiente al valor α , para un nivel de confianza $(1-\alpha)$.

* El cálculo se basa en la distribución normal. El valor de $Z\alpha$ para un IC del 95% es 1,96. Para muestras de tamaño inferior a 30 individuos, este valor debe sustituirse por el de la t de Student para $(n-1)$ grados de libertad. Asimismo, el cálculo requiere que no existan diferencias significativas entre las desviaciones estándar de ambas muestras.

TABLA Cálculo del intervalo de confianza (IC) de la diferencia entre dos medias.

Ejemplo 23.2. Supongamos un estudio que compara la eficacia de dos tratamientos A y B en dos grupos de 30 pacientes. Se observa una diferencia en el porcentaje de éxitos del 20% (70 – 50%) a favor del tratamiento B, que no es estadísticamente significativa ($p = 0,12$). El IC del 95% de la diferencia entre los dos tratamientos es:

$$\text{IC 95\%: } 0,2 \pm 0,24, \text{ es decir, de -4 a 44\%.}$$

La verdadera magnitud de la diferencia está en un intervalo que va desde un 4% a favor del tratamiento A hasta un 44% a favor de B. Dado que una diferencia del 0% también es posible, no puede descartarse que éste sea su verdadero valor, por lo que la prueba estadística da un valor no significativo.

En cambio, el IC informa, además, que también son posibles grandes diferencias a favor de B, y que son improbables grandes diferencias a favor de A. Aunque los resultados siguen sin ser concluyentes, se dispone de más información para interpretarlos de forma adecuada. El IC cuantifica el resultado encontrado y provee un rango donde es muy probable que se encuentre el valor real que se está buscando.

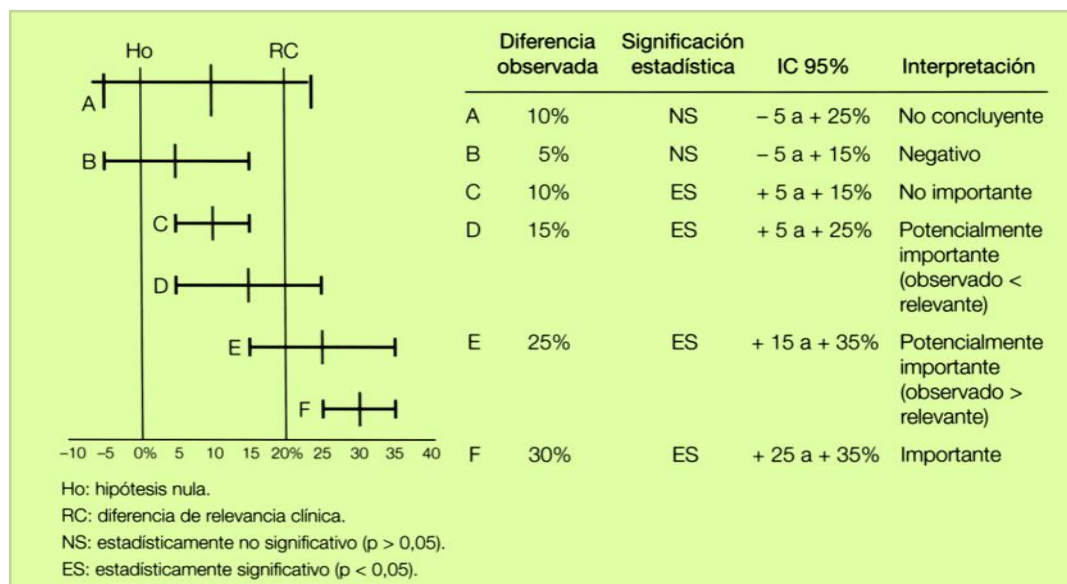
Ejemplo. Supongamos un estudio que compara la eficacia de dos tratamientos A y B en dos grupos de 80 pacientes. Se observa una diferencia del 5% (65-60%) a favor del tratamiento B, que es estadísticamente significativa ($p = 0,04$).

METODOLOGÍA E INVESTIGACIÓN

El IC del 95% de esta diferencia es:

IC 95%: $0,05 \pm 0,047$, es decir, de 0,3 a 9,7%.

Al excluir el valor 0%, la diferencia es estadísticamente significativa. Además, informa que la diferencia es de pequeña magnitud.



La figura muestra los posibles resultados de un estudio que compara dos grupos y utiliza como medida del efecto la diferencia entre los porcentajes observados. Se considera que la mínima diferencia de relevancia clínica es del 20%.

Los IC tienen otra ventaja adicional, y es la de expresar los resultados en las unidades en que se han realizado las mediciones, lo que permite al lector considerar críticamente su relevancia clínica.

Al diseñar un estudio, los autores establecen la magnitud mínima de la diferencia que consideran de relevancia clínica y, según ésta, calculan el tamaño de la muestra necesario. Al acabar el estudio, la interpretación del resultado observado y de su IC debe tener en cuenta también esta magnitud.

Además de determinar si el IC del 95% excluye el valor 0 para saber si el resultado es estadísticamente significativo, también debe determinarse si incluye o excluye el valor de la mínima diferencia de relevancia clínica, para poder evaluar si el estudio es concluyente acerca de la existencia de una diferencia clínicamente importante.

Ejemplo. En la figura se presentan 6 posibles resultados de un estudio que compara dos grupos y utiliza como medida del resultado la diferencia entre los porcentajes observados. Los autores establecieron en el diseño que la mínima diferencia de relevancia clínica era del 20%.

Situación A. Se observó una diferencia del 10% (IC 95: -5 a +25%), que no es estadísticamente significativa, ya que el IC incluye el valor 0. Por otro lado, el valor

20% también es un valor posible, ya que está situado en el interior del IC. Se trata, por tanto, de un resultado que no permite descartar ninguna conclusión.

Situación B. En esta ocasión, se observó una diferencia entre los grupos del 5% (IC 95: -5 a +15%). No es estadísticamente significativa, ya que el IC incluye el valor 0. Pero, por otro lado, excluye el valor 20%. Por tanto, el resultado es negativo en el sentido de que no puede descartarse que los grupos sean iguales, pero, aunque fueran diferentes, es muy improbable que la diferencia fuera de relevancia clínica.

Situación C. La diferencia observada es del 10% (IC 95: +5 a +15%), estadísticamente significativa, ya que el IC excluye el valor 0. Dado que el límite superior del IC no alcanza el valor 20%, se concluye que existe una diferencia, pero que ésta no es de relevancia clínica.

Situación D. La diferencia observada es del 15% (IC 95: +5 a +25%), estadísticamente significativa y potencialmente importante, ya que el valor 20% es un valor posible, incluido en el IC, aunque no es del todo concluyente.

Situación E. Similar a la anterior. Aunque la diferencia observada es mayor del 20%, el resultado no es concluyente a favor de la existencia de una diferencia clínicamente relevante, ya que el límite inferior del IC está por debajo del 20% e indica que la verdadera diferencia podría ser inferior a este valor.

Situación F. La diferencia observada es del 30%. Dado que el límite inferior del IC es mayor del 20%, este resultado es estadísticamente significativo y concluyente a favor de la existencia de una diferencia de relevancia clínica.

Cuando se utiliza para medir el efecto una medida relativa (riesgo relativo, odds ratio, etc.), la H_0 de igualdad entre los grupos corresponde al valor 1. Por tanto, cuando el IC del 95% de una medida relativa incluye el valor 1 no es estadísticamente significativa.

Aunque las pruebas de significación continúan siendo los procedimientos estadísticos utilizados con mayor frecuencia, las ventajas de la utilización de los IC en el análisis e interpretación de los resultados tanto si el objetivo es la estimación de parámetros como el contraste de una hipótesis, hacen que cada vez más revistas recomienden a los autores su utilización.

H - COMPARACIONES MÚLTIPLES:

La controversia anterior se refiere a la aplicación de una prueba estadística considerada de forma aislada. Sin embargo, en cualquier estudio, es frecuente que se lleven a cabo múltiples comparaciones para determinar si, por ejemplo, dos grupos difieren en más de una variable. Ello obliga a la utilización de múltiples pruebas estadísticas, lo que tiene dos grandes inconvenientes:

- La realización de pruebas para cada variable por separado ignora el hecho de que muchas de ellas pueden estar relacionadas entre sí, de forma que el resultado de una

prueba estadística determinada puede estar influido por diferencias en la distribución de otras variables relacionadas.

- Si cada una de las pruebas estadísticas se realiza con el valor de significación prefijado del 5%, en promedio, 5 de cada 100 comparaciones pueden resultar significativas sólo por azar. Las cosas improbables suceden, y lo hacen tanto más cuanto más ocasiones se les den. Por tanto, al realizar múltiples comparaciones, aumenta la probabilidad de obtener algún resultado significativo.

Existen diversas situaciones en las que puede aparecer este problema. Por ejemplo, al inicio del análisis, el investigador está interesado en evaluar la comparabilidad inicial de los grupos, para lo que los compara respecto a las características de la línea basal, realizando múltiples pruebas estadísticas. Al final del estudio, puede aparecer cuando se comparan múltiples variables de respuesta (múltiples hipótesis), o bien en función de diferentes subgrupos.

Para estimar la probabilidad de obtener un resultado por azar tras realizar un número n de pruebas estadísticas, cada una de ellas con el valor de significación α ce, puede usarse la llamada desigualdad de Bonferroni, de la que se deriva la siguiente fórmula:

$$Pr = 1 - (1 - \alpha)^n$$

Donde Pr representa la probabilidad de encontrar un resultado significativo.

TABLA 23.5. Cálculo del intervalo de confianza (IC) de la diferencia entre dos medias

IC de la diferencia de dos medias*

a) Muestras independientes

$$(\bar{x}_A - \bar{x}_B) \pm Z\alpha \cdot EED$$

siendo $EED = s \cdot \sqrt{\frac{1}{n_A} + \frac{1}{n_B}}$

$$y \quad s = \sqrt{\frac{s_A^2 (n_A - 1) + s_B^2 (n_B - 1)}{n_A + n_B - 2}}$$

b) Muestras apareadas

$$\bar{x}_d \pm Z\alpha \cdot EE\bar{x}_d$$

\bar{x}_A, \bar{x}_B : medias observadas en las muestras A y B.
 s_A, s_B : desviaciones estándar observadas en las muestras A y B.
 n_A, n_B : número de sujetos de las muestras A y B.
 EED: error estándar de la diferencia.
 \bar{x}_d : media de las diferencias de las dos mediciones en cada individuo (series apareadas).
 EE \bar{x}_d : error estándar de la media de las diferencias individuales.
 $Z\alpha$: valor de la variable normal tipificada correspondiente al valor α , para un nivel de confianza $(1 - \alpha)$.

* El cálculo se basa en la distribución normal. El valor de $Z\alpha$ para un IC del 95% es 1,96. Para muestras de tamaño inferior a 30 individuos, este valor debe sustituirse por el de la t de Student para $(n-1)$ grados de libertad. Asimismo, el cálculo requiere que no existan diferencias significativas entre las desviaciones estándar de ambas muestras.

Ejemplo: Supongamos que se realizan 8 comparaciones independientes, cada una de ellas con el valor de significación de 0,05. La probabilidad de que alguna de estas comparaciones conduzca a un resultado significativo simplemente por azar es:

$$\text{Pr} = 1 - (1 - 0,05)^8 = 0,336$$

Es decir, que existe un 33% de probabilidades de cometer algún error tipo I en dichas comparaciones, y no el 5% con el que se deseaba trabajar. Puede utilizarse como aproximación el producto del valor de significación por el número de pruebas realizadas ($n \cdot \alpha$). En el ejemplo, este valor sería 0,40, que, como puede observarse, sobreestima ligeramente dicha probabilidad.

La solución más adecuada a este problema es la de reducir en lo posible el número de comparaciones a realizar, y asignar prioridades antes del análisis, decidiendo cuál es la comparación principal y cuáles las secundarias, de forma que estas últimas sean utilizadas para matizar la respuesta a la primera.

Otra solución es la de dividir el valor de significación deseado por el número de pruebas a realizar y obtener así el valor de significación corregido con el que debe realizarse cada una de las pruebas. Este procedimiento se conoce como corrección de Bonferroni. En el ejemplo, si se desea mantener el valor de significación global del 0,05, cada una de las pruebas deberá realizarse con el valor corregido de $0,05/8 = 0,00625$. Esta técnica es muy estricta, ya que exige un nivel de significación muy bajo. Además, la corrección de Bonferroni asume que las comparaciones son independientes, por lo que no resulta adecuado cuando los análisis están mutuamente asociados.

Otra alternativa es la utilización de técnicas multivariantes que tomen en consideración las relaciones que existen entre las variables que se comparan.

Un problema similar puede presentarse en estudios prolongados, cuando el investigador decide analizar sus datos de forma repetida a medida que el estudio progresa, y nuevos pacientes son incluidos. El uso de los valores de significación estadística calculados según el método habitual no resulta apropiado. El cálculo convencional del valor de significación asume que el tamaño de la muestra está fijado, y que los resultados del estudio se analizarán una única vez con los datos de todos los sujetos. Por tanto, si un investigador busca el valor de significación en diversas ocasiones durante la recogida de datos, la probabilidad de alcanzar un valor significativo es mayor que la deseada. Cuando se realiza este tipo de estudios, la opción más deseable es la de utilizar un diseño secuencial, en el que no es necesaria la asunción de que el análisis se realizará una sola vez para calcular los valores de significación, y en el que las normas para finalizar el estudio permiten la evaluación continuada de los datos.

I - ANÁLISIS MULTIVARIANTE:

En muchas ocasiones, interesa considerar la influencia de más de dos variables de forma simultánea. Ello requiere técnicas sofisticadas, basadas en modelos matemáticos complejos, agrupadas bajo el nombre genérico de análisis multivariante.

Existen múltiples técnicas estadísticas multivariantes. En investigación clínica y epidemiológica las más utilizadas son las que analizan la relación entre una variable

dependiente (variable de respuesta) y un grupo de variables independientes (factor de estudio y variables a controlar). Estas técnicas implican la construcción de un modelo matemático. La elección de un modelo u otro dependerá del diseño empleado en el estudio, la naturaleza de las variables y las interrelaciones entre el factor de estudio, la variable de respuesta y las restantes variables incluidas en el modelo (variables a controlar).

Estas técnicas pueden aplicarse con las siguientes finalidades:

Proporcionar una estimación del efecto de una variable independiente principal (factor de estudio) sobre una dependiente (variable de respuesta), ajustada por un conjunto de factores o variables independientes (potenciales factores de confusión y modificadores del efecto).

Las técnicas que relacionan un conjunto de variables independientes con una dependiente derivan del modelo de regresión lineal. Pueden clasificarse en función de la escala de medida de la variable dependiente.

Cuando la finalidad del análisis es predictiva o descriptiva de las relaciones entre las variables, lo que interesa es obtener una ecuación o un modelo lo más sencillo posible, de forma que la selección de las variables que formarán el modelo se basa en criterios de significación estadística. En cambio, cuando la finalidad es estimar el efecto del factor de estudio controlando determinados factores de confusión, no puede confiarse en que la selección automática incluya las variables que interesa, ya que estas variables no siempre están asociadas de forma estadísticamente significativa, por lo que se utilizan otros procedimientos que obligan a la selección por parte del investigador de las variables que desea controlar.

La correcta utilización de la estadística en el análisis de los datos es fundamental para poder interpretarlos de forma adecuada y obtener conclusiones válidas. Por ello, una vez realizado el análisis multivariante, hay que evaluar la bondad del modelo obtenido, ya que, por ejemplo, una ecuación puede ser estadísticamente significativa, pero predecir con poca fiabilidad el riesgo de enfermar de un sujeto. Debe prestarse atención a las medidas globales de bondad del modelo, así como comprobar si se cumplen las condiciones de aplicación de cada una de las técnicas.

El resultado obtenido del análisis es una estimación puntual. Para evaluar su relevancia clínica deberá tenerse en cuenta su intervalo de confianza. Por tanto, todos los comentarios realizados a propósito de la significación estadística y la relevancia clínica en las pruebas estadísticas bivariantes son aplicables a las multivariantes. La principal diferencia es que en el análisis multivariante se tienen en cuenta simultáneamente las relaciones entre múltiples variables, de forma que una ecuación predictiva mejorará su capacidad de predicción al incorporar más de una variable o la estimación del efecto del factor de estudio estará ajustada, es decir, se habrá controlado el efecto simultáneo de diferentes factores de confusión.

En otras ocasiones, no puede diferenciarse entre variables dependientes e independientes, sino que, de acuerdo con el objetivo del estudio, el análisis persigue finalidades diferentes, como la clasificación de variables o individuos aparentemente heterogéneos en grupos homogéneos según un conjunto de datos (análisis de conglomerados o cluster analysis), o la identificación de los factores subyacentes en un conjunto de variables, es

decir, la reducción de la dimensionalidad de los datos, combinando un conjunto amplio de variables observadas en unas pocas variables ficticias (combinaciones lineales de las anteriores) que representen casi la misma información que los datos originales (análisis factorial y análisis de componentes principales). Estas técnicas se utilizan con frecuencia en la investigación sobre servicios sanitarios, pero poco en investigación clínica y epidemiológica.

ANEXO:

1) MEDIDAS DE ASOCIACIÓN:

En los estudios analíticos, no sólo interesa conocer si existe una asociación entre el factor de estudio y la variable de respuesta, sino también su magnitud. Esto se consigue comparando la frecuencia del suceso de interés en un grupo expuesto al factor de estudio con la de un grupo no expuesto.

Para el cálculo de las medidas de asociación, los datos suelen presentarse en forma de una tabla de contingencia 2 x 2 (tabla A2.1). En los estudios en los que se utiliza como medida de frecuencia la densidad de incidencia los datos se presentan en forma algo distinta (tabla A2, 2). Estas tablas se generalizan cuando existen varios grados de exposición.

	Enfermedad	No enfermedad	Total
Exposición	a	b	a + b
No exposición	c	d	c + d
Total	a + c	b + d	a + b + c + d

Las casillas a, b, c y d representan diferentes combinaciones entre la enfermedad y la exposición:
a: número de personas expuestas que tienen la enfermedad.
b: número de personas expuestas que no tienen la enfermedad.
c: número de personas no expuestas que tienen la enfermedad.
d: número de personas no expuestas que no tienen la enfermedad.
Los marginales de las tablas resultan de sumar las correspondientes casillas:
a + b: número total de expuestos.
c + d: número total de no expuestos.
a + c: número total de personas con la enfermedad.
b + d: número total de personas sin la enfermedad.
La suma de las cuatro casillas, a + b + c + d, representa el número total de participantes en el estudio.

TABLA A2.1. Presentación de una tabla 2 x 2 para el cálculo de las medidas de asociación.

	Enfermedad	Persona-tiempo
Exposición	a	P _{Te}
No exposición	c	P _{To}
Total	a + c	P _{Te} + P _{To}

a: número de personas enfermas en el grupo expuesto.
c: número de personas enfermas en el grupo no expuesto.
P_{Te}: unidades de persona-tiempo de observación en el grupo expuesto.
P_{To}: unidades de persona-tiempo de observación en el grupo no expuesto.
En consecuencia, los marginales serán:
a + c: total de personas enfermas.
P_{Te} + P_{To}: total de unidades de persona-tiempo de observación.

TABLA A2.2. Presentación de una tabla para el cálculo de las medidas de asociación cuando el denominador son unidades de persona-tiempo de observación.

2) RIESGO RELATIVO:

El riesgo relativo (RR) estima la magnitud de una asociación e indica el número de veces que es más probable que una enfermedad se desarrolle en el grupo expuesto en relación al grupo no expuesto. En otras palabras, es la razón entre la incidencia en el grupo expuesto (I_e) y la incidencia en el grupo no expuesto (I_o), por lo que también se denomina razón de incidencias o de riesgo:

$$RR = \frac{I_e}{I_o} = \frac{a/(a+b)}{c/(c+d)}$$

Ejemplo A2.1. Supongamos una cohorte de 368 individuos que reciben oxigenoterapia domiciliaria, divididos en dos grupos según si han dejado o no de fumar, que se sigue durante un año para evaluar su mortalidad (tabla A2.3). La medida de frecuencia que se decide utilizar es la incidencia acumulada. En primer lugar, se calcula el riesgo o probabilidad de que un individuo que no tenga la enfermedad (en este caso la muerte) la desarrolle durante el año de duración del estudio.

El resultado es del 12,3% (19/154) en la cohorte expuesta, mientras que para los exfumadores es del 7% (15/214). El riesgo relativo (RR) de los fumadores respecto a los exfumadores es de $12,3/7,0 = 1,8$, lo que significa que un fumador tiene 1,8 veces más probabilidad de morir que un exfumador en un año.

	Número de defunciones	Número de pacientes que sobreviven	Total
Fumadores	19	135	154
Exfumadores	15	199	214
Total	34	334	368

TABLA A2.3. Resultados de un estudio hipotético de una cohorte de 368 individuos que reciben oxigenoterapia domiciliaria (ejemplo A2.1)

Incidencia acumulada en el grupo expuesto: $I_e = 19/154 = 12,3\%$.

Incidencia acumulada en el grupo no expuesto: $I_o = 15/214 = 7,0\%$.

Riesgo relativo: $RR = 12,3/7,0 = 1,8$.

Diferencia de incidencias: $DI = 12,3 - 7,0 = 5,3\%$.

METODOLOGÍA E INVESTIGACIÓN

Cuando la medida de frecuencia es la densidad de incidencia, el RR se calcula de forma similar:

$$RR = \frac{I_e}{I_o} = \frac{a/P_{Te}}{c/P_{To}}$$

Ejemplo A2.2. Supongamos que la tasa de incidencia de cardiopatía isquémica es de 13,0 casos por 1.000 personas-año en hombres fumadores de 50 a 65 años, mientras que en los no fumadores de este mismo grupo de edad sólo es de 6,2 por 1.000 personas-año de observación. El riesgo relativo es el cociente entre ambas tasas de incidencia: $13,0/6,2 = 2,1$.

Este resultado se interpreta como que la tasa de aparición de nuevos casos de cardiopatía isquémica en fumadores es 2,1 veces la que se observa en no fumadores.

Un RR de 1,0 indica que no existe relación entre el factor de estudio y la enfermedad. Si es mayor que donde $\ln RR$ es el logaritmo natural del RR. El IC de $\ln RR$ se calcula del siguiente modo:

$$\ln RR \pm z_{\alpha/2} \times EE(\ln RR)$$

Dado que la fórmula anterior permite obtener los límites del IC del logaritmo del RR, deberán determinarse sus antilogaritmos para conocer el IC del RR.

Ejemplo A2.3. Supongamos un estudio en que se ha seguido una cohorte de 500 sujetos expuestos y otra de 500 sujetos no expuestos a un factor de riesgo, observándose los resultados de la tabla A2.4. La estimación puntual del RR es 4 y los límites de su IC del 95% son 2,1 y 7,4. Dado que el IC excluye el valor 1, existe asociación estadísticamente significativa entre la exposición y la enfermedad.

	Enfermos	No enfermos	Total
Expuestos	48	452	500
No expuestos	12	488	500
Total	60	940	1.000

Estimación puntual del RR: $RR = \frac{48/500}{12/500} = 4$

Error estándar del logaritmo del RR:

$$EE(\ln RR) = \sqrt{\frac{452}{500 \times 48} + \frac{488}{500 \times 12}} = 0,31$$

IC del 95% del $\ln RR$:

$\ln 4 \pm 1,96 \times 0,31$
 $1,39 \pm 0,62$
 0,77 a 2,01

Límites del IC del 95% del RR:

Límite inferior: $\exp(0,77) = 2,15$
 Límite superior: $\exp(2,01) = 7,44$

TABLA A2.4. Ejemplo de cálculo del intervalo de confianza (IC) del riesgo relativo (RR) (ejemplo A2.3)

El RR obtenido en un estudio es una estimación puntual y, por tanto, debe calcularse su intervalo de confianza (IC). Si el IC del 95% no incluye el valor $RR = 1$, existe una asociación estadísticamente significativa entre el factor de estudio y el desenlace. Para calcularlo, debe determinarse en primer lugar el error estándar del logaritmo natural (neperiano) del RR:

$$EE(\ln RR) = \sqrt{\frac{b}{(a + b) \times a} + \frac{d}{(c + d) \times c}}$$

3) ODDS RATIO:

En los estudios de cohortes el RR se estima directamente, ya que se conoce la incidencia de la enfermedad en los individuos expuestos y no expuestos a un factor. Por el contrario, en los estudios de casos y controles no se puede calcular la incidencia, porque la población de estudio se selecciona a partir de individuos que ya han desarrollado la enfermedad.

La razón $a/(a + b)$, por ejemplo, no estima el riesgo de contraer la enfermedad en aquellos expuestos al factor de riesgo, sino que refleja, en parte, decisiones como elegir dos controles por caso.

La medida de asociación que se utiliza en los estudios de casos y controles es la razón de odds u odds ratio (OR). Por odds se entiende la razón entre la probabilidad de que un suceso ocurra y la probabilidad de que no ocurra.

Si la probabilidad de que una persona con la enfermedad esté expuesta es del 0,75, la odds de exposición se calculará dividiendo este valor por la probabilidad de no estar expuesto ($0,75/11 = 0,068$). La OR no es más que la razón entre la odds de exposición observada en el grupo de casos (a/c) y la odds de exposición en el grupo control (b/d).

$$OR = \frac{a/c}{b/d} = \frac{a \cdot d}{b \cdot c}$$

Para explicar intuitivamente de donde proviene la fórmula de la OR se partirá del esquema de un estudio de cohortes.

Supongamos que toda una población se clasifica según la presencia o no de un factor de riesgo, y es seguida durante un tiempo para observar en cuántos individuos aparece la enfermedad y clasificarlos de acuerdo con la combinación exposición-enfermedad que presenten.

El riesgo relativo sería:

$$RR = \frac{a/(a + b)}{c/(c + d)} = \frac{a(c + d)}{c(a + b)}$$

En la mayoría de enfermedades, la proporción de individuos clasificados como enfermos será pequeña, es decir, a será pequeño en comparación con b, por lo que el total de personas expuestas a + b será prácticamente igual a b, y lo mismo sucede con c en comparación con d, por lo que el total de personas no expuestas c + d será igual a d. Por tanto, la anterior ecuación sería aproximadamente:

$$RR = \frac{a/(a + b)}{c/(c + d)} = \frac{a/b}{c/d} = \frac{a \cdot d}{c \cdot b}$$

De este modo, los estudios de casos y controles pueden ser conceptualizados como un estudio en el que el grupo de casos lo formarían todos los individuos que padecen la enfermedad (a + c), y los con-troles se escogerían entre aquellos que no la padezcan (b + d). Lógicamente no se estudiarán todos los individuos, sino una muestra. Si los casos y los controles se han elegido de forma independiente de la historia de exposición previa, la OR puede considerarse una buena estimación del RR. El ejemplo de la tabla A2.5 ilustra este concepto.

	Enfermos	No enfermos	Total
Alcohol ≥ 80 g/día	175	2.825	3.000
Alcohol < 80 g/día	207	6.793	7.000
		RR = 1,97	

TABLA A2.5.

- Estudio de cohortes hipotético diseñado para evaluar el riesgo de úlcera gástrica asociada a la ingesta de alcohol
- Estudio de casos y controles hipotético realizado en la misma población. Los casos son todos los individuos que tienen la enfermedad y, como controles, una muestra (10%) de los no enfermos.

La OR aventaja al RR en que su resultado es independiente de si se expresa en términos de riesgo de contraer la enfermedad, o de no contraerla.

	Casos	Controles	Total
Alcohol \geq 80 g/día	175	282	457
Alcohol $<$ 80 g/día	207	679	886
		OR = 2,03	

Ejemplo A2.4. En el estudio hipotético de la tabla A2.6, el RR de enfermar en la comunidad A respecto a B es $(2/100)/(1/100) = 2$.

La OR es $(2/98)/(1/99) = 2$. Así pues, ambas medidas dan el mismo resultado. Considérese la posibilidad de analizar los datos desde la perspectiva de no enfermar. El RR sería $(98/100)/(99/100) = 1$, es decir, no habría asociación entre vivir en una comunidad y el hecho de no enfermar. La OR sería $(98/2)/(99/1) = 1/2$. Con el RR se obtiene un resultado distinto según se exprese en relación al hecho de enfermar o al de no enfermar.

Sin embargo, con la OR el riesgo de enfermar en la comunidad A es el doble que en la comunidad B, mientras que la probabilidad de no enfermar en esta última comunidad es la mitad. Es decir, proporciona el mismo resultado con independencia del fenómeno que se exprese.

La estimación puntual de la OR debe acompañarse de su correspondiente intervalo de confianza (IC). Un método sencillo y aproximado para su cálculo consiste en calcular el error estándar del logaritmo natural (o neperiano) de la OR:

$$EE(\ln OR) = \sqrt{(1/a + 1/b + 1/c + 1/d)}$$

Donde $\ln OR$ es el logaritmo natural de la OR. El IC de $\ln OR$ se calcula del siguiente modo:

$$\ln OR \pm Z_{\alpha/2} EE(\ln OR)$$

La fórmula permite obtener los límites del IC del logaritmo de la OR. Para conocer los límites del IC de la OR deberemos determinar sus respectivos antilogaritmos.

Ejemplo A2.5. Supongamos un estudio con 200 casos y 200 controles en el que se obtienen los resultados de la tabla A2.7. La estimación puntual de la OR es 4,4 y los límites del IC del 95% son 2,6 y 7,4. La asociación entre la exposición y la enfermedad es estadísticamente significativa ya que el IC del 95% de la OR excluye el valor 1.

La OR se utiliza cada vez con más frecuencia como medida de la magnitud de un efecto o asociación tanto en los estudios de casos y controles, de cohortes y ensayos clínicos, como en revisiones sistemáticas y meta análisis. Habitualmente se interpreta como equivalente del RR, aunque en realidad la OR sólo es una buena aproximación del RR en determinadas

condiciones. Cuando el riesgo basal (frecuencia del resultado de interés en la población de estudio) es bajo (menor del 10%), ambas medidas son muy similares. Cuanto mayor es el riesgo basal, la OR más exagera la magnitud del efecto en relación al RR, es decir, más lo sobreestima cuando es mayor que 1 y más lo infraestima cuando es menor de 1.

De hecho, la diferencia entre la OR y el RR depende tanto del riesgo basal como de la propia magnitud del efecto, de forma que grandes discrepancias sólo se observan en efectos de gran tamaño y riesgos basales elevados. Por ello, en la mayoría de ocasiones, la interpretación de la OR como RR no conduce a cambios en la interpretación cualitativa de los resultados, aunque debe tenerse en cuenta que puede existir cierta sobreestimación de la magnitud del efecto.

En los estudios de cohortes, suele utilizarse la regresión logística para controlar múltiples factores de confusión, obteniéndose una OR ajustada. Zhang

Comunidad	Enfermos	No enfermos	Total
A	2	98	100
B	1	99	100
Total	3	197	200

TABLA A2.6. Comparación del riesgo de enfermarse en dos comunidades (ejemplo A2.4)

	Casos	Controles	Total
Expuestos	70	22	92
No expuestos	130	178	308
Total	200	200	400

Estimación puntual de la OR: $OR = (70 \times 178) / (22 \times 130) = 4,36$
 Error estándar del logaritmo de la OR:
 $EE (\ln OR) = \sqrt{1/70 + 1/22 + 1/30 + 1/178} = \sqrt{0,073} = 0,27$

IC del 95% del ln OR:
 $\ln 4,36 \pm 1,96 \times 0,27$
 $1,47 \pm 0,53$
 0,94 a 2,00

IC del 95% de la OR:
 Límite inferior: $\exp(0,94) = 2,56$
 Límite superior: $\exp(2,00) = 7,39$

TABLA A2.7. Ejemplo de cálculo del intervalo de confianza (IC) de la odds ratio (OR) (ejemplo A2.5)

Y Yu (1998) proponen una fórmula para corregir este valor y obtener una mejor estimación del RR:

$$RR = OR / [(1 - P_0) + (P_0 - OR)]$$

Donde P_0 es la frecuencia del resultado de interés en el grupo no expuesto (riesgo basal). Esta misma corrección puede aplicarse a los límites del IC.

METODOLOGÍA E INVESTIGACIÓN

Las medidas de asociación no miden el impacto potencial sobre la salud de la población de la exposición al factor de estudio o de su eliminación.

IX – ELECCIÓN DE LA PRUEBA ESTADÍSTICA

En este anexo no se pretende explicar las asunciones ni los métodos de cálculo de las diferentes pruebas estadísticas, que están disponibles en múltiples textos, sino proporcionar unas normas sencillas que faciliten la elección de la técnica de análisis más adecuada.

A - PRUEBAS ESTADÍSTICAS BIVARIANTES:

La elección de la prueba estadística depende de (tabla A6.1):

TABLA A6.1. Factores de los que dependen la elección de la prueba estadística más adecuada

- Escala de medida de la variable de respuesta
- Escala de medida del factor de estudio
- Carácter apareado o independiente de los datos
- Condiciones de aplicación específicas de cada prueba:
 - Número de observaciones
 - Asunciones sobre la distribución poblacional de las variables

- La escala de medida de la variable de respuesta. Las pruebas estadísticas tienen una mayor potencia si la variable de respuesta es cuantitativa, ya que contiene más información que si fuera cualitativa.
- La escala de medida del factor de estudio. Puede ser cualitativa dicotómica (tratamiento activo/placebo, exposición/no exposición), cualitativa con más de dos categorías (tres pautas terapéuticas, o diferentes grados de exposición a un factor de riesgo) o cuantitativa (valores de la colesterolemia o la presión arterial).
- El carácter apareado o independiente de los datos. Desde el punto de vista estadístico, se habla de medidas repetidas o apareadas cuando han sido realizadas sobre los mismos sujetos (p. ej., comparación de las cifras de presión arterial obtenidas en los individuos de una muestra al inicio y al final de un determinado período de tiempo). Dado que los sujetos son los mismos, existe una menor variabilidad en las mediciones, lo que permite utilizar pruebas más potentes que tengan en cuenta este fenómeno. En caso de que los grupos que se comparan estén formados por individuos diferentes, se habla de datos independientes.
- Las condiciones de aplicación específicas de cada prueba. Las pruebas estadísticas que utilizan datos cuantitativos suelen realizar determinadas asunciones sobre la distribución de las variables en las poblaciones que están siendo comparadas. Estas pruebas son conocidas como pruebas paramétricas. La mayoría son robustas, es decir,

que toleran relativamente violaciones de estas asunciones, sobre todo si el número de sujetos estudiado es elevado. En muchas situaciones, especialmente cuando las muestras son de pequeño tamaño, no se puede determinar si se cumplen dichas asunciones. En estos casos, se recurre a otras pruebas estadísticas menos potentes, que no requieren asunciones para su aplicabilidad, conocidas como pruebas no paramétricas. Este mismo tipo de pruebas es aplicable cuando se trata de analizar datos ordinales.

De forma esquemática, cuando tanto el factor de estudio como la variable de respuesta son variables cualitativas, la prueba estadística más apropiada para determinar si existe asociación entre ellas es la *ji* al cuadrado, siempre que exista un número suficiente de sujetos en cada una de las casillas de la tabla de contingencia.

Ejemplo A6.1. Se desea comparar el porcentaje de diabéticos controlados en un grupo que ha recibido educación sanitaria grupal con el de otro grupo que ha recibido los cuidados habituales. Se trata de datos independientes, ya que son grupos de sujetos diferentes. El factor de estudio (educación sanitaria) es una variable dicotómica (grupal/habitual), al igual que la variable de respuesta (controlado/no controlado). Por tanto, los datos se presentarían en una tabla 2×2 y se utilizaría una *ji* al cuadrado (o una *Z* de comparación de dos proporciones), siempre que el número de sujetos lo permita. Si no puede utilizarse, se recurriría a la prueba exacta de Fisher.

Cuando se comparan dos grupos (factor de estudio dicotómico) respecto a una variable cuantitativa, la prueba estadística más adecuada es la *t* de Student-Fisher, si se cumplen las condiciones necesarias para su aplicación. En caso contrario, debe recurrirse a una prueba no paramétrica equivalente, como la *U* de Mann-Whitney.

Ejemplo A6.2. Supongamos que en el ejemplo A6.1 se desea comparar la glucemia de ambos grupos de sujetos al finalizar el estudio. En este caso, la variable de respuesta es cuantitativa (glucemia). Se trata, por tanto, de la comparación de dos medias, y la prueba de elección es la *t* de Student-Fisher, si se cumplen las condiciones de aplicación.

Si se comparan más de dos grupos (factor de estudio con más de dos categorías) respecto a una variable cuantitativa, debe utilizarse el análisis de la variancia (ANOVA). Si no se cumplen los criterios de aplicación del análisis de la variancia, debe recurrirse a la prueba de Kruskal-Wallis.

Ejemplo A6.3. Supongamos que en el ejemplo A6.1 se han incluido pacientes procedentes de tres centros de salud, y que desea determinarse si la glucemia de los sujetos difiere según el centro de procedencia. En este caso, el factor de estudio (centro de salud de procedencia) tiene tres categorías, por lo que no puede aplicarse la *t* de Student-Fisher. Si se cumplen las condiciones para su utilización, se aplicaría un análisis de la variancia. Si se observa una diferencia estadísticamente significativa, se realizaría una comparación posterior para determinar qué grupo o grupos difieren, utilizando alguna de las técnicas de comparación múltiple específicamente diseñadas para ello (p. ej., las de Scheffé o de Neuman-Keuls).

Factor de estudio	Variable de respuesta			
	Cualitativa nominal (dos categorías)	Cualitativa nominal (> 2 categorías)	Cualitativa ordinal	Cuantitativa*
Cualitativo (dos grupos)				
Independientes	Z de comparación de proporciones Ji al cuadrado Prueba exacta de Fisher	Ji al cuadrado	U de Mann-Whitney	t de Student-Fisher Prueba de Welch
Apareados	Prueba de McNemar Prueba de Fisher	Q de Cochran	Prueba de los signos Prueba de los rangos signados de Wilcoxon	t de Student-Fisher para datos apareados
Cualitativo (más de dos grupos)				
Independientes	Ji al cuadrado	Ji al cuadrado	Prueba de Kruskal-Wallis	Análisis de la variancia
Apareados	Q de Cochran	Q de Cochran	Prueba de Friedman	Análisis de la variancia de dos vías
Cuantitativo				
	t de Student-Fisher	Análisis de la variancia	Correlación de Spearman Tau de Kendall	Correlación de Pearson Regresión lineal

* Cuando las pruebas estadísticas aplicables a las variables cuantitativas no cumplen las asunciones necesarias para su aplicación, deben utilizarse las pruebas correspondientes como si la variable de respuesta fuera una variable ordinal (pruebas no paramétricas).

TABLA A6.2. Pruebas bivariantes de significación estadística utilizadas con mayor frecuencia.

Si se trata de determinar la posible asociación entre un factor de estudio y una variable de respuesta cuantitativa, la prueba adecuada es la correlación de Pearson o, si no se cumplen las condiciones de aplicación, la correlación de Spearman. En el caso de que pueda asumirse una relación de dependencia lineal de una de las variables respecto a la otra, se habla de regresión lineal simple.

Ejemplo A64. Supongamos que se desea evaluar si existe asociación entre la edad de un grupo de sujetos y sus cifras de presión arterial sistólica (PAS). Dado que ambas variables están medidas en una escala cuantitativa, se trata de un análisis de correlación. Si se deseara determinar en cuánto se incrementa la PAS a medida que aumenta la edad de los sujetos, se realizaría un análisis de regresión lineal, en el que la PAS sería la variable dependiente y la edad la independiente. En cambio, no tendría sentido evaluar en cuánto aumenta la edad de los sujetos al incrementarse su PAS.

B - TÉCNICAS ESTADÍSTICAS MULTIVARIANTES:

En investigación clínica y epidemiológica, las técnicas multivariantes se utilizan habitualmente cuando existe una variable dependiente (variable de respuesta) y múltiples variables independientes (factor de estudio y otras variables a controlar).

Estas técnicas pueden utilizarse tanto con finalidades de predicción (obtener una ecuación que permita, conociendo los valores de un conjunto de variables independientes, predecir el valor de la variable dependiente), de descripción de la relación entre variables (identificar, de entre un conjunto de variables independientes, cuáles están asociadas con la variable dependiente), como de estimación del efecto del factor de estudio (obtener una estimación del efecto del factor de estudio sobre la variable de respuesta y controlar la influencia de variables de confusión).

En función de las escalas de medida de las variables implicadas en el análisis, puede seleccionarse la técnica más adecuada (tabla A6.3).

Una de las técnicas multivariantes más utilizadas es la regresión lineal múltiple, que se aplica cuando tanto la variable dependiente como las independientes son cuantitativas, aunque en la práctica permite que algunas de las variables independientes sean cualitativas. El valor de un coeficiente de regresión lineal múltiple es una estimación del efecto de la variable independiente correspondiente sobre la variable dependiente, ajustado por el resto de variables independientes de la ecuación.

Es decir, representa el cambio esperado de la variable dependiente cuando se incrementa en una unidad el valor de la variable independiente, asumiendo que el resto de variables del modelo se mantienen constantes.

Ejemplo A6.5. En un estudio, se evalúa la asociación entre la edad, el índice de masa corporal (IMC) y la clase social con la presión arterial sistólica (PAS). Dado que la variable dependiente (PAS) es cuantitativa, y que dos de las variables independientes también lo son (edad e IMC), se utiliza una regresión lineal múltiple.

Al final del análisis, se obtiene un coeficiente de regresión de la edad de 0,5, indicando que la PAS es, en promedio, 0,5 mmHg más elevada por cada año más de edad de los sujetos, supuestos constantes la clase social y el IMC, es decir, que este efecto de la edad está ajustado por estas otras dos variables.

Otra técnica es la regresión logística, que es aplicable cuando la variable dependiente es cualitativa. Se utiliza mucho en investigación clínica y epidemiológica, ya que es útil para estudiar variables de respuesta dicotómicas (enfermo/no enfermo, curado/no curado, etc.).

Mientras que en la regresión lineal los coeficientes representan directamente el cambio en la variable dependiente, en la regresión logística estiman medidas relativas, por ejemplo, la odds ratio asociada al factor de estudio y ajustada por el resto de variables independientes. Dado que el modelo logístico no es lineal, sino exponencial, se utilizan transformaciones logarítmicas que hacen que los coeficientes no puedan interpretarse directamente.

Ejemplo A6.6 Un estudio evalúa la relación entre diferentes factores de riesgo y el desarrollo de cardiopatía coronaria en una muestra de hombres adultos. Dado que la variable de respuesta es dicotómica (desarrollo o no de la enfermedad), se realiza un análisis de regresión logística con esta variable como dependiente y el conjunto de factores de riesgo como independientes.

En la ecuación resultante, la variable edad tiene un coeficiente de 0,12. Para poder interpretar este resultado, se calcula su antilogaritmo natural $e^{0.12} = 1,13$, valor que corresponde a la odds ratio asociada al aumento de un año de edad de los sujetos ajustada por el resto de variables de la ecuación.

METODOLOGÍA E INVESTIGACIÓN

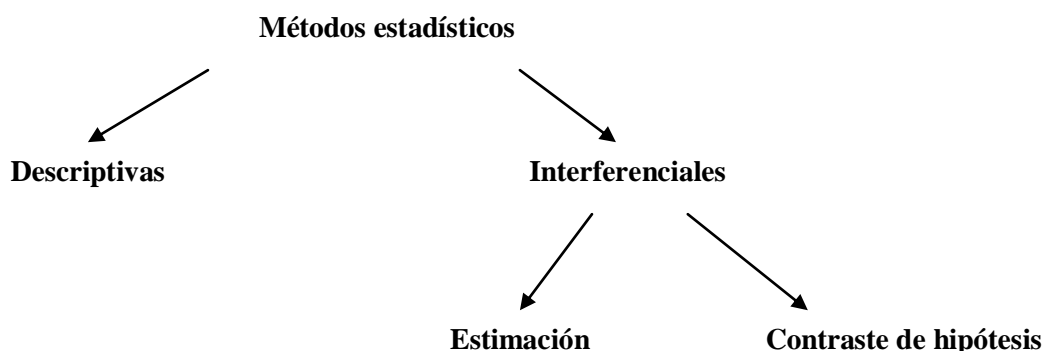
Técnicas inferenciales* (pueden diferenciarse variables dependientes e independientes)		
Variable dependiente	Variables independientes	Técnica multivariante
Cuantitativa	Cualitativas	Análisis de la variancia
Cuantitativa	Cuantitativas (1)	Regresión lineal múltiple
Cualitativa	Cualitativas	Modelo log-lineal
Cualitativa	Cuantitativas (1)	Regresión logística (2)
		Análisis discriminante (3)
		Regresión de Cox (4)
Técnicas descriptivas (todas las variables presentan el mismo estatus)		
Finalidad	Técnica multivariante	
Reducir la dimensionalidad de la matriz de datos (identificar los factores subyacentes en un conjunto de variables)	Análisis factorial Análisis de componentes principales Análisis de correspondencias	
Clasificar individuos o variables en grupos homogéneos	Análisis de <i>clusters</i>	
* Sólo se presentan las principales pruebas aplicables cuando existe una sola variable dependiente.		
(1) Algunas variables independientes pueden ser cualitativas.		
(2) Suele utilizarse cuando la variable dependiente es dicotómica.		
(3) Suele utilizarse cuando la variable dependiente tiene más de dos categorías.		
(4) Se utiliza cuando la variable dependiente incluye la dimensión tiempo (análisis de supervivencia).		

TABLA A6.3. Elección de la técnica multivariante adecuada.

También se utiliza con gran frecuencia el modelo de regresión de Cox, que es útil cuando la variable dependiente es el tiempo de supervivencia o el tiempo transcurrido hasta la aparición de un desenlace determinado.

Las técnicas multivariantes descriptivas se aplican cuando no existe una variable que pueda ser considerada dependiente, y son poco utilizadas en investigación clínica y epidemiológica.

X - MÉTODOS ESTADÍSTICOS:



Las estadísticas descriptivas permiten describir la media de cada grupo.

La inferencia estadística es el proceso de elaboración de conclusión sobre características de una población utilizando la información proporcionada por la muestra.

Parámetros potenciales estimados	Estimación muestral
Media μ	$\bar{X} = \frac{\sum^n x_i}{n}$
Proporción	$P_s = \frac{x}{n}$

- Variable aleatoria (VA):
 - Es la variable que toma ciertos valores con determinadas probabilidades.
- Distribución de probabilidades:
 - Es el modelo matemático que relaciona cada valor posible de una VA con su probabilidad de ocurrencia en la población.
- Tipos de variables aleatorias:
 - Discretas (presencia o no de cierto defecto en un producto).
 - Continuas (toman cualquier valor real en un intervalo de los reales (volumen, peso)).
 - Las variables se distribuyen normalmente o no.
- Función de probabilidad:
 - Variable de una VA.
 - Distribución de POISSON ($x = 0, 1, \dots$).

Para una tesis:

1. originalidad de las ideas
2. calidad de la metodología y desarrollo del proceso: variables.

La estadística es un apoyo a nuestro estudio.

A - POTENCIA ESTADÍSTICA (POWER)

Probabilidad de encontrar una diferencia estadísticamente significativa cuando de hecho esta diferencia existe. Por ejemplo, una potencia del 80% en un ensayo clínico de tamaño N, representa una probabilidad del 80% de detectar una verdadera diferencia en proporciones – igual a algún valor especificado previamente- con un p-valor asociado pequeño.

En muchos estudios, las diferencias observadas pueden ser simplemente debidas al azar, pero las diferencias reales pueden no alcanzar significación estadística en ensayos con tamaños muestrales pequeños. La potencia estadística de un estudio particular crece al crecer su tamaño muestral.

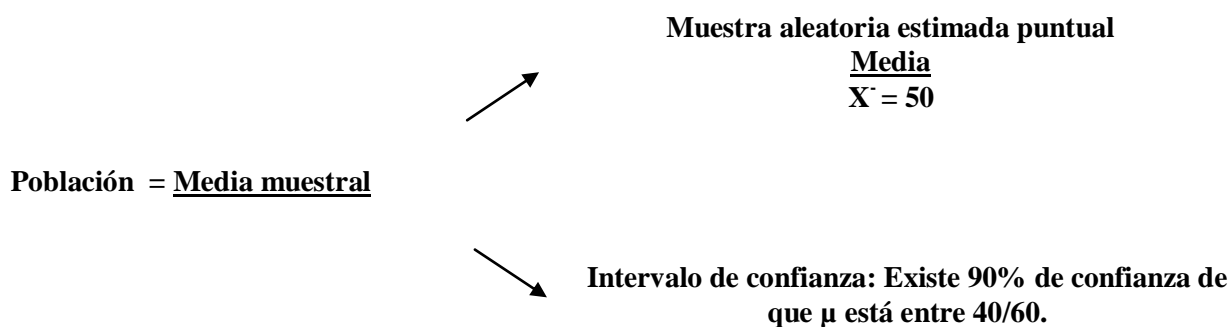
Con efectos cuantitativos (quantitative outcomes), cuanto mayor sea la variabilidad de mediciones individuales, menor será la potencia estadística del estudio. El nivel de significación establecido para los resultados (p-valor o error tipo I) también determina la potencia estadística de un estudio. El complementario de la potencia estadística es el error tipo II (potencia = $100 - \beta$).

B - MÉTODOS DE EVALUACIÓN:

1) EVALUACIÓN PUNTUAL:

media, mediana... (Parámetro de esa muestra).

2) POR INTERVALOS (ESTIMACIÓN DE CONFIANZA):



3) DESVIACIÓN ESTÁNDAR:

- Permite conocer la precisión del estudio.
- Tiene que ser igual al intervalo de confianza.
- No se puede indicar una variable sin su intervalo de confianza, es lo que da precisión (95% es muy bueno, siempre se deja 5 % de error: pero el intervalo tiene que ser estrecho.
- Un estudio piloto de pequeña muestra puede tener un intervalo de confianza de 90%.

4) EVALUACIÓN DEL RESULTADO:

- Medida del efecto clínicamente relevante.
- Intervalo de confianza preciso.
- P significativa (no mide magnitud del efecto).

NOTA: En la escala visual analógica del dolor (EVA), una diferencia es relevante a partir de 5,5 mm.

C – TESTS DE HIPÓTESIS:

- Un test o contraste de hipótesis indica si los resultados se deben al azar o no.
- p.- 5% (Test de Fischer) permite valorar el efecto del azar:
 - o p . inferior a 0,05 = hipótesis nula, se rechaza porque no hay diferencia entre los grupos de estudio.
 - o P. = 0,05; hipótesis positiva, se acepta porque hay diferencia entre los grupos de estudio.
- La P expresa que la probabilidad de los datos obtenidos sean explicados por el azar. No tiene que ver con la magnitud del efecto. Si es cierto la P es que hay efecto, para que sea significativa (40 a =). Y el efecto despreciable necesito muchos pacientes.

		No existe diferencia (No cierto)	Existe diferencia (No falso)
Resultado de la prueba	Diferencia significativa (Rechazada No)	Error tipo I $\alpha = 5 \%$	No error
Resultado de la prueba		No error	Error tipo II $\beta = 20\%$

- Tests de correlación :
 - o Spearman.
 - o Pearson.
 - o Kendall.

METODOLOGÍA E INVESTIGACIÓN

- Wilcoxon.

Permiten calcular la p.

Ejemplo: Escala del dolor

- Variable independiente: Intervención si o no.
- Variable dependiente: mediana, dolor.

	Binario	Nominal	Ordinal	Intervalo
Binario	X^2	X^2	Modelos de log. lineales	Reg. logística
Nominal	X^2	X^2	Correlación de Spearman o Kendall	Correlación de Spearman o Kendall
Ordinal	U-man Whitney	Kruskall Wallys		
Intervalo	T student	ANOVA	Correlación de Spearman o Kendall	Regresión, correlación de Pearson

		Interdependiente	
		A	B
Dependiente	A	λ_A 6,5 mm/Hg	λ_B 6,6 mmHg
	B		

D - DATOS PAREADOS (medidas repetidas):

		VARIABLE INDEPENDIENTE	
VARIABLE DEPENDIENTE	Tipo	Binario	Nominal
	Ordinal	Wilcoxon.	ANOVA Friedman
	Intervalo	"t" de student pareada	ANOVA Medidas repetidas

NOTA: Utilizar mucho intervalo de confianza y NNT para los estudios.

E -LOS TESTS:

1) PRUEBA DE WILCOXON:

Supongamos que se dispone de una muestra de una población y que, sobre cada individuo de la muestra, se miden dos variables en escala al menos ordinal X e Y, cuyos posibles valores son comparables.

La prueba de Wilcoxon se utiliza para contrastar la hipótesis nula de que la muestra procede de una población en la que la magnitud de las diferencias positivas y negativas entre los valores de las variables X e Y es la misma. Es decir, si $m+$ y $m-$ son las magnitudes de las diferencias positivas y negativas, respectivamente, la hipótesis nula que se desea contrastar es:

$$H_0: m + = m -$$

Si el *p-valor* asociado al estadístico de contraste es menor que α , se rechazará la hipótesis nula al nivel de significación α .

La prueba de Wilcoxon, al igual que la de los signos, se utiliza para comprobar que los resultados obtenidos en las dos variables son los mismos, pero con la primera, más que el sentido de las diferencias entre los valores, interesa comparar la magnitud de dichas diferencias, por lo que podría suceder que los resultados de las dos pruebas

2) LOS COEFICIENTES DE CORRELACIÓN DE PEARSON Y SPEARMAN:

El coeficiente de correlación de Pearson mide el grado de asociación lineal entre dos variables medidas en escala de intervalo o de razón, tomando valores entre -1 y 1.

Valores próximos a 1 indicarán fuerte asociación lineal positiva: a medida que aumentan los valores de una de las dos variables aumentan los de la otra; valores próximos a -1 indicarán fuerte asociación lineal negativa: a medida que aumentan los valores de una de las dos variables disminuyen los de la otra, y valores próximos a 0 indicarán no asociación lineal, lo que no significa que no pueda existir otro tipo de asociación.

Su cuadrado puede interpretarse como la proporción de variabilidad de los valores de una de las dos variables explicada por los de la otra.

El coeficiente de correlación de Spearman es una variante del coeficiente de correlación de Pearson en la que, en lugar de medir el grado de asociación lineal a partir de los propios valores de las variables, se mide a partir de la asignación de rangos a los valores ordenados (naturales consecutivos de 1 al número total de observaciones). En consecuencia, se aplica cuando cada valor en sí no es tan importante como su situación respecto a los restantes.

En este sentido es una medida que también puede ser adecuada en el caso de variables en escala ordinal. Por lo demás, sus valores se interpretan exactamente igual que los del coeficiente de correlación de Pearson.

3) PRUEBA JI-CUADRADO EN TABLAS DE CONTINGENCIA:

La prueba Ji-cuadrado se aplica al caso de que se disponga de una tabla de contingencia con r filas y c columnas correspondiente a la observación de muestras de dos variables X e Y, con r y c categorías, respectivamente. Se utiliza para contrastar la hipótesis nula:

$$H_0: \text{Las variables } X \text{ e } Y \text{ son independientes}$$

Si el *p*-valor asociado al estadístico de contraste es menor que α , se rechazará la hipótesis nula al nivel de significación α .

<i>Pruebas Ji-cuadrado de Pearson en tablas de contingencia</i>	
Prueba	Descripción
<i>Ji-cuadrado de Pearson</i>	Determinar si las diferencias entre las frecuencias observadas en la tabla de contingencia correspondiente al cruce de los valores de las dos variables y las frecuencias esperadas, supuesto que las variables son independientes, son estadísticamente significativas.

4) PRUEBA DE KOLMOGOROV-SMIRNOV:

Supongamos que se dispone de una muestra de una población y que, sobre cada individuo de la muestra, se mide una variable continua X. La prueba de Kolmogorov-Smirnov es una prueba de bondad de ajuste que se utiliza para contrastar la hipótesis nula de que la muestra procede de una población en la que la distribución de X es una determinada distribución teórica F_e . Es decir, la hipótesis nula que se desea contrastar es:

$$H_0: F = F_e$$

Si el *p*-valor asociado al estadístico de contraste es menor que α , se rechazará la hipótesis nula al nivel de significación α .

Si denominamos IC a la variable «Índice cardíaco», la hipótesis nula que se desea contrastar es que la distribución de la variable IC es Normal:

$$H_0: F = \text{Normal}$$

La prueba de Kolmogorov-Smirnov sobre la variable IC, con distribución teórica o esperada igual a una Normal.

La prueba consistirá en comparar, para cada valor de la variable, la proporción de casos observados con valor inferior o igual a dicho valor con la proporción de casos esperados bajo la hipótesis nula de distribución Normal de media y desviación típica las muestrales («2,9356 y 1,2962, respectivamente»). El estadístico de contraste se construirá a partir de la máxima diferencia, en valor absoluto, encontrada.

5) PRUEBA Q DE COCHRAN:

Supongamos que se sospecha que el efecto de dos fármacos distintos, A y B, para hacer desaparecer los síntomas de una úlcera es el mismo y que, además, dicho efecto es nulo, en el sentido de que en los casos en los que los síntomas desaparecen con A o con B también desaparecen sin necesidad de utilizar ningún fármaco.

Para comprobarlo, se elige una muestra de 50 pacientes a la que se suministra un placebo, y por cada paciente de la muestra se buscan otros dos de características similares (dos «gemelos»), a los que se suministran los fármacos A y B (uno a cada uno).

Después de un período de observación se comprueba, en cada caso, si los síntomas han desaparecido o no. Para determinar si la probabilidad de que los síntomas desaparezcan es la misma tanto con cada uno de los dos fármacos como con el placebo, la prueba que se aplicará es la Q de Cochran.

Supongamos que se dispone de una muestra de una población y que, sobre cada individuo de la muestra, se miden k variables dicotómicas X_1, \dots, X_k , cuyos posibles valores son los mismos.

La prueba Q de Cochran se utiliza para contrastar la hipótesis nula de que la muestra procede de una población en la que la probabilidad de obtener uno cualquiera de los dos resultados posibles es la misma para las k variables. Es decir, si p_{ij} es la probabilidad del i -ésimo resultado, $i = 1, 2$, en la variable X_j , $j = 1, k$, la hipótesis nula que se desea contrastar es:

$$H_0: p_{i1} = \dots = p_{ik} \quad \text{o, equivalentemente,} \quad H_0: p_{1j} = \dots = p_{2j}$$

Si el p -valor asociado al estadístico de contraste es menor que α se rechazará la hipótesis nula al nivel de significación α .

6) HERRAMIENTAS BÁSICAS EN ESTADÍSTICA DESCRIPTIVA:

<i>Herramientas básicas en la estadística descriptiva</i>		
Variable	Descripción	Estadísticos y gráficos
Cualitativa en escala nominal	Valores no numéricos con ausencia de orden entre ellos	Distribución de frecuencias Moda Diagrama de barras
Cualitativa en escala ordinal	Valores no numéricos con presencia de orden entre ellos	Mínimo Máximo Mediana Cuartiles Percentiles Rango intercuartílico Gráfico de caja y bigotes
Cuantitativa en escala de intervalo o de razón	<ul style="list-style-type: none"> • Cuantitativa discreta: únicamente puede tomar una cantidad finita o numerable de valores numéricos. • Cuantitativa continua: puede tomar cualquier valor numérico en un intervalo. • En escala de intervalo: únicamente tiene sentido la diferencia entre sus valores. • En escala de razón: si, además de la diferencia, tiene sentido la razón entre sus valores 	Media Rango Varianza Desviación típica Coeficiente de variación Coeficiente de asimetría Coeficiente de curtosis Histograma Gráfico de tallo y hojas

7) PRUEBAS “T” DE STUDENT:

<i>Pruebas t de Student</i>	
Prueba t	Objetivo
Una muestra	Determinar si la diferencia entre la media de una variable y un determinado valor teórico es estadísticamente significativa.
Dos muestras relacionadas	Determinar si la diferencia entre las medias de dos variables es estadísticamente significativa.
Dos muestras independientes	Determinar si la diferencia entre las medias de una variable en dos grupos distintos de individuos es estadísticamente significativa.

8) LAS PRUEBAS NO-PARAMÉTRICAS:

<i>Pruebas no paramétricas para muestras relacionadas</i>			
Prueba	Núm. de variables	Variables	Objetivo
McNemar	2	Cualitativas: 2 valores	Determinar si la diferencia entre las distribuciones de frecuencias de los valores de las dos variables es estadísticamente significativa.
Signos	2	En escala al menos ordinal	Determinar si la diferencia entre el número de veces en que el valor de una variable es mayor que el de la otra y el número de veces en que es menor es estadísticamente significativa
Wilcoxon	2	En escala al menos ordinal	Determinar si la diferencia entre la magnitud de las diferencias positivas entre los valores de las dos variables y la magnitud de las diferencias negativas es estadísticamente significativa.
<i>Q</i> de Cochran	$p > 2$	Cualitativas: 2 valores	Determinar si las diferencias entre las distribuciones de frecuencias de los valores de las p variables son estadísticamente significativas.
<i>F</i> de Friedman	$p > 2$	En escala al menos ordinal	Determinar si las diferencias entre las distribuciones de las p variables son estadísticamente significativas.
Concordancia de Kendall	p	En escala al menos ordinal	Medir el grado de concordancia entre los elementos de la muestra respecto a las puntuaciones de las p variables.

9) PRUEBAS DE BONDAD:

<i>Pruebas de bondad de ajuste para una muestra</i>		
Prueba	Tipo variable	Objetivo
Binomial	Cualitativa: 2 valores	Determinar si las diferencias entre las proporciones de cada uno de los dos valores de la variable y unas determinadas proporciones teóricas son estadísticamente significativas.
<i>Ji</i> -cuadrado	Cualitativa: $K > 2$ valores	Determinar si las diferencias entre las frecuencias de cada uno de los valores de la variable y unas determinadas frecuencias teóricas son estadísticamente significativas.
Kolmogorov-Smirnov	Cuantitativa	Determinar si la diferencia entre la función de distribución empírica de la variable y una determinada función de distribución teórica es estadísticamente significativa.
Kolmogorov-Smirnov-Lilliefors	Cuantitativa	Determinar si la diferencia entre la función de distribución empírica de la variable y la función de distribución esperada bajo hipótesis de normalidad es estadísticamente significativa.

10) MEDIDAS DE ASOCIACIÓN PARA ESCALA DE INTERVALO:

<i>Medidas de asociación para escala de intervalo o de razón</i>		
Medida de asociación	Escala de medida	Observaciones
Eta	V.D.: intervalo V.I.: nominal	<ul style="list-style-type: none"> • Los valores de la variable independiente establecen grupos en la población. • Toma valores entre 0 y 1. • Cuanto más próximo a 1 sea su valor más diferenciados estarán los grupos en términos de las puntuaciones de la variable dependiente (mayor será la dependencia de las puntuaciones respecto de los grupos).
Correlación de Pearson	intervalo	<ul style="list-style-type: none"> • Son medidas del grado de asociación lineal entre las dos variables.
Correlación de Spearman	intervalo (u ordinal)	<ul style="list-style-type: none"> • Los coeficientes de correlación de Pearson y de Spearman toman valores comprendidos entre -1 y 1, que indican máximo grado de asociación lineal negativa y positiva, respectivamente.
Asociación lineal Mantel-Haenszel	intervalo	<ul style="list-style-type: none"> • La correlación de Spearman es la correlación de Pearson entre los rangos asignados a los valores ordenados. • La medida de asociación lineal de Mantel-Haenszel se define como el cuadrado del coeficiente de correlación de Pearson multiplicado por $(N-1)$, siendo N el tamaño muestral.

11) PRUEBAS DE INDEPENDIA PARA UNA MUESTRA:

<i>Pruebas de independencia para una muestra</i>		
Prueba	Tipo variable	Objetivo
Rachas	Cualitativa: 2 valores	Determinar si la diferencia entre la secuencia de apariciones de uno y otro valor de la variable y la secuencia correspondiente a apariciones aleatorias es estadísticamente significativa.
Autocorrelación	Cuantitativa	Determinar si la secuencia de apariciones de los distintos valores de la variable corresponde a una secuencia de observaciones independientes entre sí.

12) Pruebas no-paramétricas para muestra independiente:

<i>Pruebas no paramétricas para muestras independientes</i>			
Prueba	Núm. de grupos	Variable dependiente	Objetivo
Mann-Whitney	2	En escala al menos ordinal	Determinar si la diferencia entre el número de veces en que el valor de la variable en un grupo es mayor que en el otro y el número de veces en que es menor es estadísticamente significativa.
Kolmogorov-Smirnov	2	En escala al menos ordinal	Determinar si la diferencia entre las funciones de distribución empíricas de la variable en cada uno de los dos grupos es estadísticamente significativa.
Wald-Wolfowitz	2	En escala al menos ordinal	Determinar si la diferencia entre la secuencia de observaciones pertenecientes a uno y otro grupo y la secuencia correspondiente a observaciones completamente mezcladas es estadísticamente significativa.
Mediana	k	En escala al menos ordinal	Determinar si las diferencias entre las frecuencias de observaciones en cada uno de los k grupos con valor igual o por debajo de la mediana de la variable sin distinguir grupos y las frecuencias esperadas supuesto que la mediana fuera la misma en los k grupos son estadísticamente significativas.
Kruskal-Wallis	k	En escala al menos ordinal	Determinar si las diferencias entre las medias de los rangos (asignados a las observaciones ordenadas) en los k grupos son estadísticamente significativas.

XI – LOS SEGOS (errores)

- Sesgos de clasificación.
- Sesgos de información (mala clasificación).
- Sesgos de confusión (por ejemplo una edad no establecida correctamente).

XII - CONCLUSIONES:

A - ESTUDIO DE VALIDACIÓN O DE CARACTERIZACIÓN DE UN TEST DIAGNÓSTICO.

- Grupos de estudio/Testigo
- Inter observador ciego o doble ciego
- Tipo de Estudio: Estudio Descriptivo transversal
- Tipo de estadísticas: Sensibilidad y especificidad del test con sus intervalos de confianza.

IC = 95% = Sensibilidad y especificidad.

B - ESTUDIO DE LOS EFECTOS DE UNA MANIPULACIÓN.

- Grupos aleatorios Experimental/Testigo de simulación.
- Interobservador ciego o doble ciego.

NOTA:

En el simple ciego el paciente no sabe si se trata o si se simula.

En el doble ciego la persona que evalúa es diferente de la que trata.

- Tipo de Estudio: Ensayo clínico

C - ESTUDIO DEL EFECTO DE UN PROTOCOLO DE TRATAMIENTO OSTEOPÁTICO EN UNA ENFERMEDAD X.

- Grupos aleatorios de estudio/Testigo.
- Muchas variables con muchos subgrupos por cada tipo de protocolo
- Ínter observador ciego o doble ciego
- Tipo de Estudio: Análisis estratificada compleja multivariante.
- Necesita una población muy grande.

D - ESTUDIO DE LAS LESIONES OSTEOPÁTICAS ASOCIADAS A UNA ENFERMEDAD.

- Grupos aleatorios de estudio/Testigo
- Interobservador ciego o doble ciego
- Tipo de Estudio: Descriptivo transversal, Odds Ratio, casos y controles (mínimo 50 casos), estudio de cohorte (100 a 200 pacientes mínimo en cada grupo).

$$RR = \frac{I_A^+}{I_A^-}$$

Existen normas para hablar de efecto causal (Bradford, Hill).

La relación causal de Hill: da estudio de cohortes, verifica más cosas, te ayuda a ver la incidencia. Es más difícil que el caso de control, además te permite establecer asociación pero no la incidencia.

1. Asociación fuerte (Odds Ratio).

AR
Odds fuerte

2. Temporalidad: la causa está antes del efecto.
3. Gradiente de dosis (tratamiento) /efecto (resultado).
4. Consistencia (Sea cual sea la edad, el sexo, el país).
5. Plausibilidad /Se puede explicar).
6. Reproducción experimental.

El estudio de cohorte verifica los puntos 1 a 4, el estudio de casos y controles no permite de calcular la prevalencia.

Nos aproxima a la entelequia de la causa.