Atención a la Salud de la Comunidad

Tema 5. Validez y precisión de las estimaciones: significación estadística (valor de p). Intervalo de confianza. Errores sistemáticos (sesgos): de selección y de información. Causalidad.

Celia Nespral Gaztelumendi

Miguel Santibáñez Margüello

Montserrat Bustamante Fonfría

Dpto. de Enfermería

Este tema se publica bajo Licencia:

Creative Commons BY-NC-SA 3.0

Objetivos:

1) Conocer y diferenciar los conceptos de validez y precisión de las estimaciones.
2) Conocer e interpretar correctamente la significación estadística (valor de p). Intervalo de confianza.
3) Conocer los principales errores sistemáticos (sesgos que afectan a la validez de un estudio).
4) Conocer los principales criterios de causalidad

Material de apoyo:

•P. Fernández et al. **Tipos de estudios Epidemiológicos**. Disponible en :http://www.fisterra.com/mbe/investiga/6tipos_estudios/6tipos_estudios.htm

Para mas información:

M. Delgado et al. **Sesgos**. En Piédrola Gil, "Medicina preventiva y salud pública", Barcelona [etc.] Masson 2001. Pag 133-143.

F. Rodriguez Artalejo et al. **Causalidad en epidemiología**. En Piédrola Gil, "Medicina preventiva y salud pública", Barcelona [etc.] Masson 2001. Pag 133-143.

Rebagliato M, Ruiz I, Arranz M. metodología de investigación en epidemiología. Madrid, Diaz de Santos, 1996.

QUE SEA PRECISO.

- PRECISIÓN: Carencia de ERROR ALEATORIO.
- Para reducir este error lo mas importante es tener el tamaño de la muestra adecuado (a mayor tamaño menor error aleatorio).

QUE SEA VÁLIDO.

- VALIDEZ: Carencia de ERROR SISTEMÁTICO.
- Para no tener estos errores hay que PENSAR, para que en nuestra investigación NO HAYA SESGOS.

VALIDEZ

VALIDEZ INTERNA.

- ERROR SISTEMATICO=SESGO.
- Si NO tengo SESGOS, los resultados que he obtenido en mi estudio SON VERDAD y aplicables a mi población de estudio.

VALIDEZ EXTERNA.

- Grado en que mis resultados son generalizables (se pueden aplicar a todo el mundo).
- Si dudo de que mis resultados sean verdad y no me los creo ni para mi población de estudio, ¡¡como me los voy a creer generalizandolos a todo el mundo!!.

LA VALIDEZ INTERNA ES POR TANTO UN PREREQUISITO PARA QUE PUEDA DARSE LA EXTERNA

PRECISIÓN: Evaluación del papel del azar en nuestros resultados

Al hacer observaciones sobre muestras y no sobre poblaciones completas, corremos el riesgo de que nuestros resultados se hallan producido simplemente por el azar.

El papel del azar en nuestros resultados se puede cuantificar o estimar mediante **pruebas estadísticas** (calculan la probabilidad de que el azar explique los resultados, dándonos un valor p ($p < 0,05$ o un intervalo de confianza).

Todas estas pruebas estadísticas tienen un N en el denominador= Cuánto mayor sea el tamaño de muestra, menor será el error aleatorio, y nuestras estimaciones serán mas precisas.

Test de Hipótesis: Expresa el grado en que nuestros datos son compatibles con los datos estadísticos, o bien se deben al azar.

Hipótesis nula (Ho): hipótesis que se plantea y que sometemos a análisis, mediante test estadísticos, para poder aceptarla o rechazarla. Se refiere a la ausencia de efecto (Ho: las cifras de colesterol son iguales con la dieta A y la dieta B).

Hipótesis alternativa (HA): La afirmación contraria (Ha: las cifras de colesterol son diferentes con la dieta A y la dieta B).

Para CONTRASTAR esta Ho, necesitamos establecer un rango de valores, para poder decidir si aceptamos o rechazamos la Ho: **intervalo de confianza o significación estadística (valor p).**

* No existe seguridad absoluta, sino que se deben fijar unos porcentajes que expresan la probabilidad de equivocarnos al aceptar o rechazar la Ho:

- **Error tipo I ó α:** probabilidad de rechazar la Ho, siendo cierta. Generalmente se fija en un 5%.
- **Error tipo II ó β:** probabilidad de no rechazar (aceptar) la Ho, siendo falsa.
- **Potencia (1 - β):** probabilidad de rechazar la Ho, siendo falsa (aceptar la HA). Suele establecerse en un 80% ó 90%. Indica la capacidad para detectar el efecto.

Valor "p" (nivel de significación):

Probabilidad de obtener valores tan extremos o más como el observado, si la Ho es cierta. Valor que nos permite aceptar o rechazar la Ho.

- Cuando "**p**" < **0.05**: rechazamos la Ho; es decir, las diferencias observadas son estadísticamente significativas y no son explicadas únicamente por el azar.
- Cuando "**p**" > **0.05**: no rechazamos, aceptamos la Ho; es decir, las diferencias encontradas no son estadísticamente significativas y pueden ser explicadas por la variabilidad aleatoria, por el azar.

Valor de p (p<0,05)

Probabilidad, condicionada a la hipótesis nula de no asociación entre exposición y enfermedad (RR=1), de observar una asociación tan extrema o más como la observada en un determinado estudio. Medida de consistencia entre la hipótesis nula y los datos que se tienen.

Tres estudios

RR = 4.45 (p = 0,004)	⇒	Muy poco consistente con H_0
RR = 1,12 (p = 0,89)	⇒	Muy consistente con H_0
RR = 0,49 (p = 0,05)	⇒	Poco consistente con H_0

Intervalo de Confianza

Evaluación de la validez de una asociación

INTERVALO DE CONFIANZA

*Rango de valores dentro del cual la verdadera magnitud del efecto (RR, OR, etc) cae con un determinado grado de seguridad 90%, **95%**, 99%)*

Resultados de los tres estudios:

RR = 4.45 (IC95%: 1,48 - 7,39)

RR = 1,12 (IC95%: 0,58 - 1,98)

RR = 0,49 (IC95%: 0,19 - 0,99)

Evaluación de la validez de una asociación

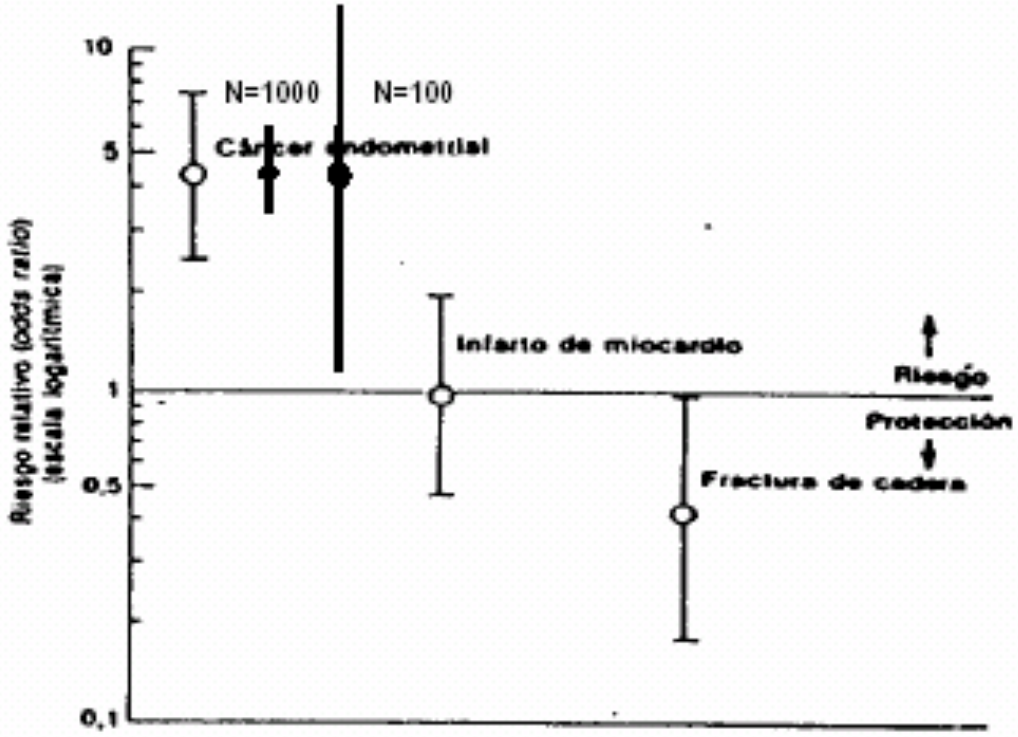

Figura 8.6. Estimadores de punto (0) e intervalos de confianza (I): riesgos y beneficios de los estrógenos exógenos para mujeres posmenopáusicas. (Datos de Antunes CM, Stolley PD, Rosenshein NB, Davies JL y cols.: Endometrial cancer and estrogen use. Report of a large case control study. N Engl J Med 300:9-13, 1979; Rosenberg L, Armstrong B, Jick H: Myocardial infarction and estrogen therapy in postmenopausal women. N Engl J Med 294:1256-1258, 1976; y Paganini-Hill A, Ross RK, Gerkins VR, Henderson BE, Arthur M, Mack TM: Menopausal estrogen therapy and hip fractures. Ann Intern Med 95:28-31, 1981.)

SESGOS

No depende del tamaño de la muestra estudiada. No disminuye al aumentar el tamaño de la muestra.

Son debidos al propio **diseño del estudio** y su eliminación o control se debe llevar a cabo durante la fase de diseño y ejecución del estudio (PARA ELLO HAY QUE PENSAR), o mediante técnicas particulares de análisis en el caso de la confusión.

Podemos distinguir:

1. **Sesgos de selección**
2. **Sesgo de información**
3. **Confusión**

SESGO SELECCIÓN

CASOS Y CONTROLES. Sesgo de supervivencia. No respuesta. **Principio de la población base: tanto los casos como los controles deben ser representativos de la misma población base.**

COHORTES: Pérdidas en el seguimiento. **En los estudios de seguimiento se debe asegurar un seguimiento completo en ambos grupos.**

Ejemplo clásico de Sesgo de Selección: Café y Cáncer de Pancreas; MacMahon et al; (N Eng J Med 1981; 304:630-3).

CASOS: Pacientes con cáncer de páncreas exocrino diagnosticados en 11 hospitales del área metropolitana de Boston y Rhode Island.

CONTROLES: Hospitalarios. Otros pacientes tratados en los mismos hospitales que los casos.

Al final los pacientes con enfermedades gastrointestinales en el grupo control estuvieron sobrerepresentados con respecto a la población general del hospital, y mucho mas con respecto al área Metropolitana de Boston y Rhode Island.

Controles con problemas gastrointestinales, tienen menos probabilidad de beber café. **El grupo control del estudio tiene una probabilidad de exposición menor que la población base.**
Los resultados estarán sesgados en el sentido de aumentar la estimación de la verdadera asociación del café con el cáncer de páncreas.
Es el llamado "***Sesgo de Berkson***", por indebida elección de pacientes hospitalarios como controles, que produce falta de representatividad respecto a la población.

OTROS SESGOS SELECCIÓN

CASOS Y CONTROLES:

Sesgo de supervivencia: Si cogemos casos prevalentes y la supervivencia esta relacionada con la exposición (sobreviven los menos expuestos): tenemos un sesgo?. Los casos deben ser incidentes para que esto no ocurra.

No respuesta: Tasa de participación (participantes/elegidos). Tasa de respuesta (participantes/contactados). Si los que no han participado estan mas o menos expuestos que los que han participado: tenemos un sesgo?. Cuando más bajas sean estas tasas, más probabilidades de que ocurra un sesgo. Hay que obtener tasas de respuestas >=90%.

COHORTES:

Pérdidas en el seguimiento. Si en la cohorte de no expuestos se nos va la gente del estudio y justo esos que se han ido desarrollan la enfermedad: serán válidos nuestros resultados?

Ejemplo Sesgo de Selección: Estudios de Cohortes.

Cohorte experimental. Modelo experimental de Prácticas en Farmacología.

Cohorte CONTROL: Prácticas tradicionales en Farmacología.

Si entre las perdidas en el grupo experimental están sobrerepresentados los que peor valoraban la utilidad del modelo experimental de prácticas: Los resultados estarán sesgados en el sentido de aumentar la estimación de la verdadera utilidad del modelo experimental respecto a las prácticas tradicionales de farma.
El problema fundamental en estos estudios son las pérdidas en el seguimiento.

SESGOS DE INFORMACIÓN

Se producen en el proceso de recogida de datos. El mas importante es el sesgo por mala clasificación de la exposición o la enfermedad, esta mala clasificación puede ser

a) DIFERENCIAL: El error al clasificar u observar se comete de forma distinta en expuestos y no expuestos o entre enfermos y no enfermos.

Sesgos de memoria: los enfermos se "esfuerzan más" en recordar sus antecedentes que los sanos, habrá mas enfermos expuestos, consecuencia: aumenta la asociación. ***Sesgo del entrevistador***: por exceso de motivación, insiste y saca mas información de los enfermos, consecuencia: aumenta la asociación.

Estos sesgos se pueden controlar mediante los estudios "**CIEGOS**", que pueden ser:

SIMPLE CIEGO: cuando el sujeto elegido no sabe a qué grupo pertenece.

DOBLE CIEGO: cuando ni el sujeto ni el investigador saben a qué grupo pertenecen.

TRIPLE CIEGO: ninguno de los anteriores conoce a que grupo pertenecen los sujetos, pero tampoco la persona que analiza los datos.

SESGOS DE INFORMACIÓN: Sesgo por mala clasificación.

B) **NO DIFERENCIAL**: El error al clasificar u observar se comete por igual tanto en casos como en controles, la consecuencia siempre es que el sesgo disminuye la asociación.

Ejemplo: Estudio sobre relación entre Cáncer de Esófago y consumo de alcohol

	E	Sano
Exp	52	105
No exp	59	203

El test con el que medimos alcohol en sangre esta roto, y al 40% de los que consumen alcohol los clasifica como no bebedores

	E	Sano
Exp	31	63
No exp	80	245

Conclusión: La asociación positiva se ha reducido y ha dejado de ser significativa=sesgo hacia la nulidad.

SESGOS DE INFORMACIÓN

Otros sesgos de información:

Regresión a la media: Tendencia a que el valor extremo de la primera medición de una variable continua se aproxime a la media de la población en mediciones posteriores. Ejemplo: Valoramos eficacia de un tratamiento en sujetos seleccionados sobre la base de valores extremos (colesterol elevado o Tensión arterial alta). Las mediciones siguientes en estos sujetos tenderán a ser mas bajas (tenderán a valores cercanos a la media poblacional por el efecto de regresión a la media), independientemente de la eficacia del tratamiento instaurado. Solución: Un grupo control en el que también se de la regresión a la media pero sin tratamiento.

Falacia ecológica: Consiste en obtener conclusiones inadecuadas a nivel individual basándonos en datos poblacionales (Estudios Ecológicos).

CRITERIOS DE CAUSALIDAD

Para identificar si una asociación encontrada en un estudio, es la causa de la enfermedad, se dispone de unas reglas metodológicas denominadas criterios de causalidad.

Ninguno de estos criterios por separado es suficiente para demostrar la causalidad, pero cuantos mas y mejor se cumplan, mas verosimil será que la relación exposición-enfermedad sea causal.

Por ser los mas conocidos y utilizados en epidemiología, comentaremos los criterios que Sir Arthur Bradford Hill publicó en 1965:

1. Temporalidad. *La exposición la factor presuntamente causal debe preceder en el tiempo a la aparición del efecto (enfermedad).*

*Parece obvio, pero se puede garantizar esto de manera sencilla en un estudio de prevalencia, o casos y controles sobre dieta y enfermedad, o sobre enfermedades mentales y consumo de drogas?**

** La dieta a veces cambia tras (y a consecuencia de) las etapas iniciales de la enfermedad que presuntamente causan.*

2. Fuerza. *Es la magnitud o intensidad de la asociación encontrada entre el Factor de Riesgo y la Enfermedad. Normalmente en epidemiológía se mide por el valor de la OR o el RR). Cuanto mas elevados sugieren relaciones causales, pero los RR bajos tambien pueden ser causales.*

3. Consistencia de resultados. *Los resultados han de repetirse en otros estudios epidemiológicos en condiciones similares.*

Ejemplo: Consumo de tabaco-enfermedad coronaria, las asociaciones son relativamente débiles, pero el juicio causal se vió fortalecido cuando tras realizarse un gran número de estudios casos-control y de cohortes en los últimos 40 años en una amplia gama de poblaciones de culturas distintas, observando millones de personas, todos ellos mostraron consistentemente un riesgo aumentado.

4. Gradiente biológico. *Es la relación dosis-respuesta. Supone que si una relación es causal, cuanto mayor es la exposición al Factor de Riesgo mayor es el riesgo de enfermedad.*

5. Plausibilidad biológica. *Se conoce un mecanismo biológico plausible a la hora de explicar la etiopatigenia de la enfermedad.*

Los 5 criterios anteriores son los principales, pero también postuló los siguientes:

6. Especificidad y analogía. *Despreciados por algunos autorers. Especificidad=una causa conduce a un efecto único, no a efectos múltiples (tabaco, solo produce una enfermedad?, si es ciertono obstante que el principal efecto es el Cáncer de Pulmón). Analogía=asociaciones causales similares podrían producir enfermedades similares. Ejemplo: Si un medicamento pude causar malformaciones congénitas, quizá puede ahacerlo otro.*

7. Experimento. *Si un Factor de Riesgo tiene un efecto, este efecto debería cesar cuando se elimina el Factor de Riesgo.*

8. Coherencia. *La asociación no debe entrar en conflicto con lo que se sabe de la enfermedad.*

Aunque se admita que los criterios de Hill son útiles en la práctica, no reunen todos los elementos útiles para la inferencia causal, muchos expertos los complementan con los dos criterios siguientes:

Criterio de independencia: *La asociación ha de ser indendiente, es decir, se ha comprobado que hay ausencia sesgos, con un buen control de la confusión y se ha descartando el azar.*

Calidad de la evidencia: *Las fuentes o tipos de estudios que muestran la asociación han de ser de calidad (Ensayo clínico>cohortes>casosy controles>transversales)*