

*Julio 2001  
Sebastián G Martí  
Daniel Preti*

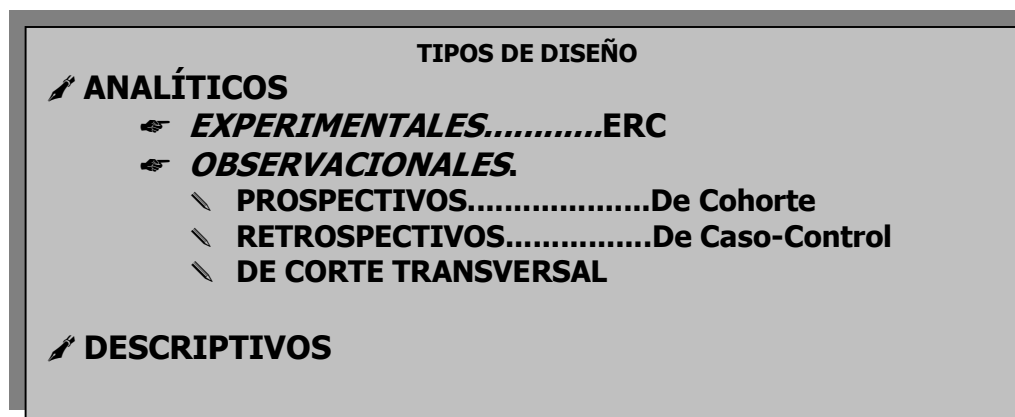
## **RESPONDIENDO PREGUNTAS**

En el momento de tomar una decisión sobre una determinada situación clínica, según la pregunta que nos haremos, existirán estudios que son diseñados para responderla con mayor exactitud y significación que otros. Así: si nuestra pregunta es sobre un **método de diagnóstico**, un estudio de **corte transversal** nos responderá sobre la sensibilidad, especificidad y valores predictivos del mismo; cuando se trata de saber si un **tratamiento** es mejor o peor que otro entonces buscaremos un **estudio randomizado y controlado** (ERC) que compare ambos tratamientos, si el problema que nos preocupa es sobre un **factor de riesgo** o el **pronóstico** de alguna enfermedad la respuesta la hallaremos en una **cohorte o un estudio de caso y control**, pero si nuestro problema es saber sobre la **prevalencia e incidencia** de determinada enfermedad en una región, buscaremos un estudio **descriptivo** que nos lo responda .

Como vemos no existe un único diseño que sea óptimo para todas las situaciones y por lo tanto es importante conocer qué significa cada uno de los diseños más importantes, dado que de ellos dependerá la veracidad de la respuesta que arrojen.

### **DISEÑO:**

**Llamamos diseño a las reglas que establecen las condiciones para la recolección de datos.**



A modo de diccionario terminológico podemos definir:

**Prospectivo:** un trabajo es prospectivo cuando en el momento de plantearse los objetivos, no se tienen los datos aún.

**Retrospectivo:** es aquel que en el momento de formular los objetivos, ya se cuenta con datos de archivo.

**Observacional:** es aquel trabajo en el cual el investigador no interviene manipulando la situación.

**Experimental:** aquí el investigador interviene en la población (grupo experimental) modificando las variables.

**Transversal:** son aquellos en los cuales la medición se realiza una única vez.

**Longitudinal:** la medición en este tipo de estudio se realiza más de una vez.

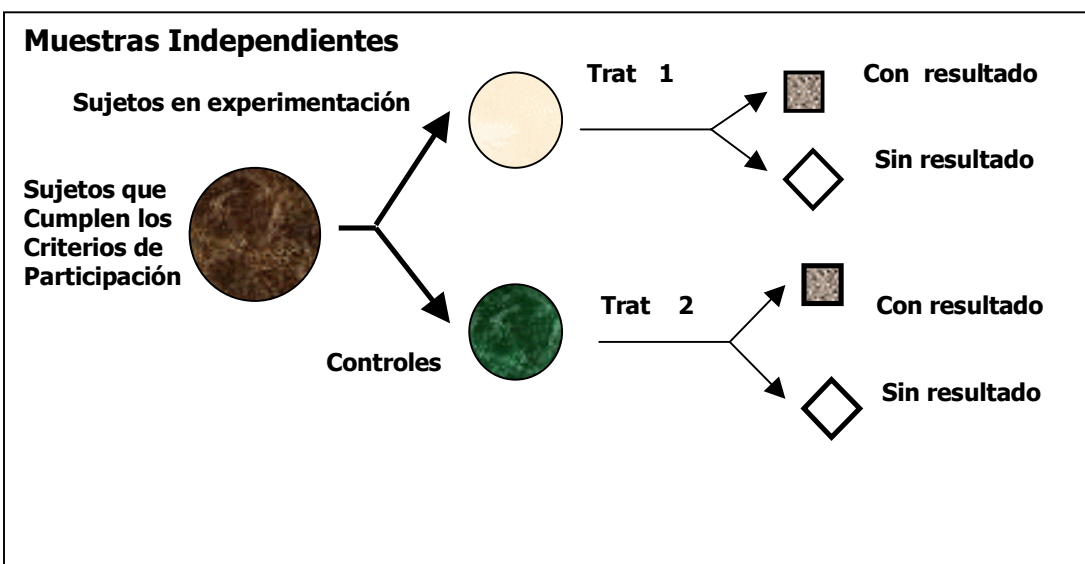
**DISEÑOS ANALÍTICOS:** *Son aquellos diseños que ponen a prueba una hipótesis - contrastan la Hipótesis-*

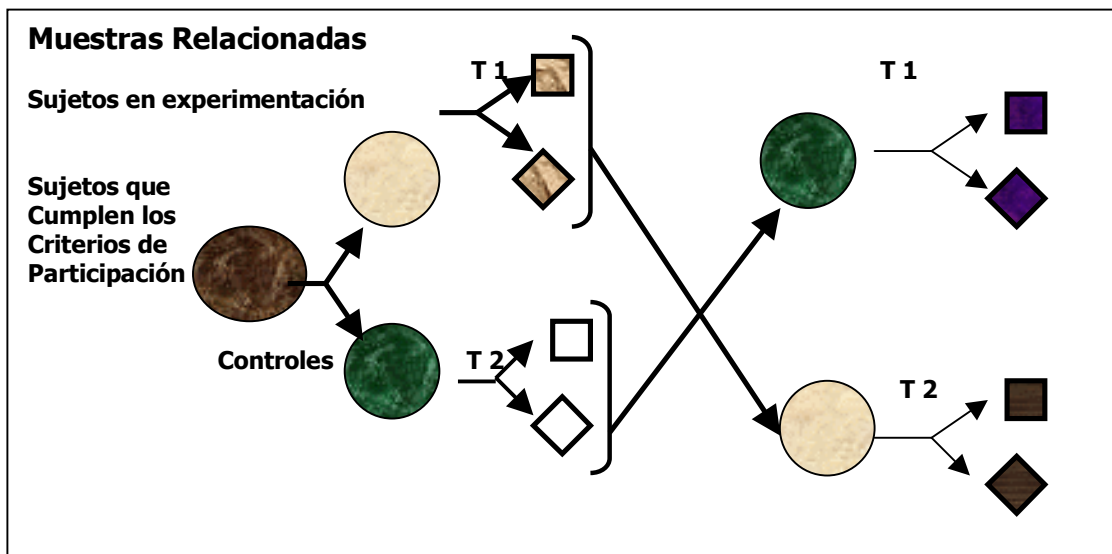
**EXPERIMENTALES:** Los estudios experimentales son fáciles de hallar en las revistas médicas, ya que son muy difundidos en la bibliografía, fácilmente se publica el trabajo completo, ya que los autores tienden a establecer de manera explícita la clase de protocolo usado. En medicina denominamos este tipo de estudio como **Prueba Clínica**, debido a su propósito, el de obtener conclusiones a cerca de un procedimiento o tratamiento en particular.

Estas pruebas clínicas pueden ser: **Controladas** cuando el fármaco o procedimiento es comparado con otra droga o método o con placebo (conocido como **Ensayo Randomizado y Controlado**); o **no controlado**, cuando el investigador describe los beneficios de la droga, pero sin compararlo con placebo u otra terapéutica.

En base a los controles, debe ser la forma de manejar las muestras, y podrán ser según como se relacionen las mismas:

- ☛ **Independientes**
- ☛ **Relacionadas o con Autocontrol**
- ☛ **Únicas**





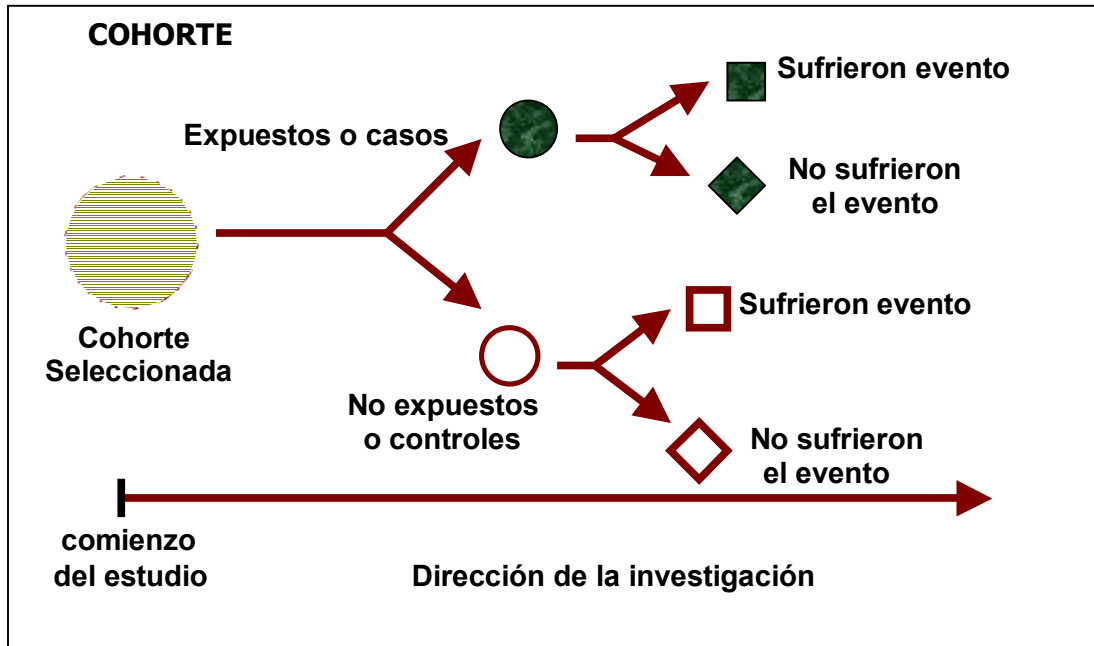
Como podemos ver en las figuras, aquellos ensayos controlados con muestra **INDEPENDIENTE** son los que se realizan con dos grupos de individuos, a los cuales se les asigna un tratamiento a cada uno, y posteriormente se compararán los resultados. En aquellos con muestra **RELACIONADA**, autocontrol o intrasujeto, todos los individuos reciben ambos tratamientos, idealmente en diferentes momentos del estudio, es decir, que en un momento dado se rotan los tratamientos y de esta manera se evitan los sesgos de selección (que tiene que ver con la respuesta individual a determinada terapéutica), ya que todos reciben los 2, de esta manera la aleatorización deja de ser un problema.

Finalmente el caso de muestra **ÚNICA**, son estudios controlados con **muestras históricas** o controles externos, que en general se usan para enfermedades que no tienen cura como el SIDA o enfermedades oncológicas y la comparación se realiza en base a otro trabajo previo, desde el punto de vista metodológico es posible, pero el grado de evidencia que ofrece es de menor calidad.

**OBSERVACIONALES:** Son estudios en los cuales el investigador solo registra las variables a estudiar, pero no interviene para modificar las mismas (solo observa el comportamiento de la variable en estudio).

Estos estudios se pueden realizar de diferente manera:

**PROSPECTIVOS:** Son llamados de "**cohorte**". Una cohorte es un grupo de individuos que tienen algo en común y forman parte de un conjunto durante un periodo largo. En medicina, los sujetos de **estudios de cohorte**, se seleccionan por una o varias características definidas que se sospecha puedan constituir el origen o el factor de riesgo para una enfermedad determinada. Los estudios de cohorte responden a la pregunta "**¿qué pasará?**" y, por lo tanto, su dirección en el tiempo es hacia delante. La figura ilustra el diseño del estudio.



Los investigadores seleccionan sujetos al principio del estudio y luego determinan si poseen el factor de riesgo o se han expuesto. Se observa a estos individuos, expuestos y no expuestos, por cierto tiempo para constatar el efecto de estas características definidas. Debido a que los sucesos que interesan se manifiestan después de iniciado el estudio los llamamos **estudios prospectivos**. Las variables en estudio se miden en reiteradas oportunidades –y cuya frecuencia dependerá de la duración del estudio- (**son longitudinales**) y valora la influencia de un factor de riesgo en el desarrollo de una determinada entidad nosológica generalmente **comparándola** con otro grupo no sometido a dicho factor.- El resultado se medirá como: **Riesgo Absoluto** y **Riesgo relativo**; y es el estudio de mayor poder para estudiar **factores de riesgo**, también son útiles en el estudio de **pronóstico** y en la evaluación de **tratamientos** a largo plazo (por ejemplo drogas antiretrovirales y sobrevida) .

Este tipo de diseño comenzó a partir de los estudios de Keys (en 1948) y posteriormente Dawber (en 1949) en el estudio de los factores de riesgo para enfermedades cardiovasculares, se realizó el reclutamiento de varones y mujeres de 40 a 69 años en una pequeña ciudad industrial semiurbana de 28.000 habitantes, situada a 18 millas de Boston: **Framingham**; aquellos participantes se reclutaron como voluntarios y a los pocos meses se invitó a toda la población de una determinada edad, usando las listas del censo, a los participantes se le realizaron entrevistas y exámenes físicos en el seguimiento cada 2 años. El estudio mencionado ha continuado a lo largo de los años, y a la primera generación formada por aproximadamente 5.000 hombres y mujeres se le sumó otra cohorte algo menos numerosa constituida por la segunda generación (la de los hijos de los primeros participantes), a Thomas Dawber le siguió B.Kannel en la dirección del proyecto y posteriormente W.P.Castelli, constituyéndose en la cohorte más grande y quizá mejor diseñada de la historia.

Cohortes más recientes incluyen el "**Patient Outcomes Research Teams**" (**PORT**) financiado por la US Agency for Health Care Policy and Research, es un estudio

multidisciplinario a gran escala, que analiza resultados y costos de intervenciones alternativas para padecimientos médicos específicos (enf de vías biliares, cardiopatía isquémica, IAM, reemplazo total de rodilla, DBT, tratamiento de cataratas, etc).

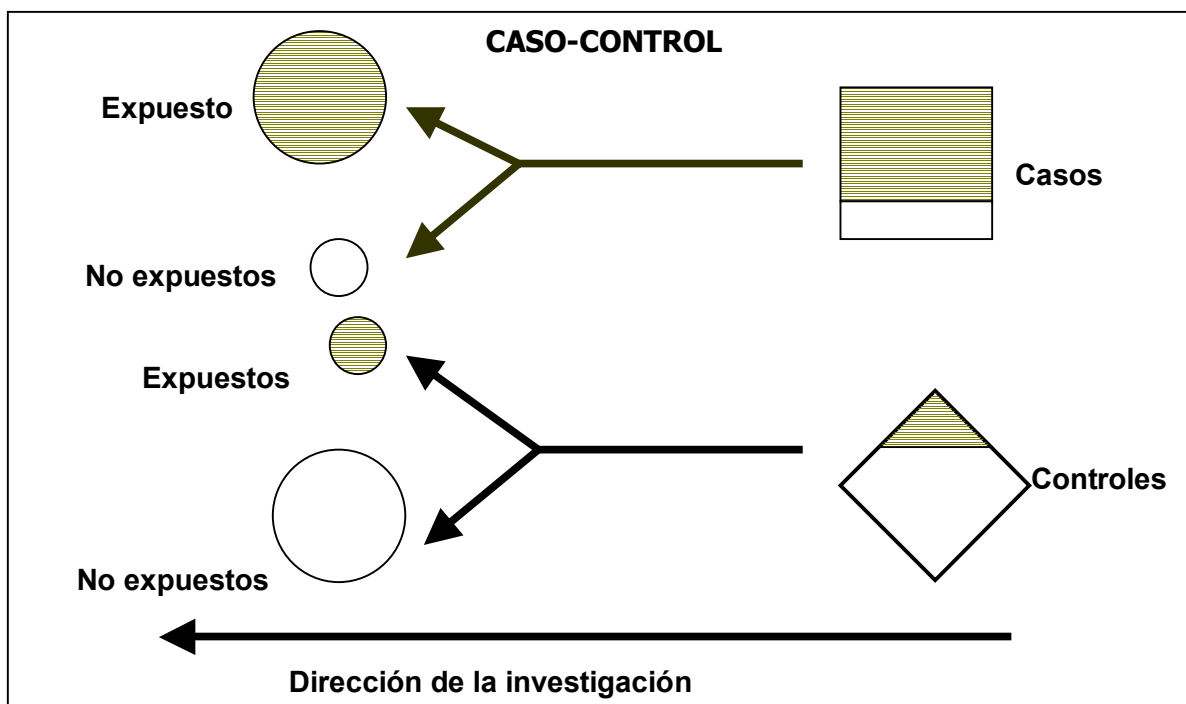
Este al igual que la cohorte de Framingham y muchos ejemplo más, demuestran lo costoso y complejo que es plantear un estudio con éstas características, convirtiéndose el financiamiento en el principal inconveniente.

**RETROSPECTIVOS:** Son conocidos como estudios de **casos y controles**, se inician con la ausencia o presencia de un resultado (que es la **presencia de la enfermedad** que estamos estudiando), y luego se investiga hacia atrás en el tiempo para tratar de determinar las **causas** o los **factores de riesgo** posibles que pueden haber favorecido a la aparición de la misma.

Los **casos** se seleccionan de manera individual, con base en alguna enfermedad o consecuencia (IAM, ACV), los **controles** son individuos sin la enfermedad o consecuencia, que deben ser elegidos respetando las características de semejanza con la población que representa los casos.

Este tipo de estudios responden a la pregunta: " **que pasó ?**", en la figura se ilustra qué sujetos se eligen al principio del estudio, después que se sabe que son casos (es decir que **tienen la enfermedad**: cuadrados), y aquellos que se comportan como controles (**sin la enfermedad**: rombos). La historia de ambos se analiza durante un período previo, en donde se detectará la exposición o no al factor de riesgo en estudio.

Los resultados de los estudios de casos y control, se informan como **Odds Ratio**, (**acepción de múltiples sinonimias, ya que carece de traducción clara: Producto cruzado, razón de probabilidad, razón de productos cruzados, etc.**) que es una medida epidemiológica que se puede parangonar al riesgo relativo, si bien su poder de significación no es de igual jerarquía que aquel.



**TRANSVERSALES:** Son estudios que analizan datos obtenidos de un grupo de sujetos en un momento dado (en lugar de un período). Responden a la pregunta: *“¿qué pasa?”*. La selección de individuos se obtiene en un tiempo breve, y comprenden los estudios que **analizan y comparan métodos de diagnóstico**, habitualmente lo hacen comparando un nuevo método de diagnóstico (mejor o más económico), contra el gold standard, definiendo: sensibilidad, especificidad, VPP y VPN.

		GOLD STANDARD	
		+	-
Nuevo Método	+	<b>Verdaderos Positivos</b>	<b>Falsos Positivos</b>
	-	<b>Falsos Negativos</b>	<b>Verdaderos Negativos</b>

### **VENTAJAS Y DESVENTAJAS DE CADA DISEÑO**

#### **Ventajas y desventajas de los ERC:**

Los **ERC son el standard de oro o referencia en medicina** –es decir el diseño básico contra el que se juzgan los demás diseños- debido que proporciona la justificación máxima y está sujeto al menor número de sesgos. Su utilidad es ampliamente reconocida cuando se evalúan métodos terapéuticos (*comparar fármacos* contra placebo o con otros fármacos).

Los ERC son la mejor clase de estudios para determinar la eficacia de un procedimiento cuando se asignan los grupos al azar (aleatorizadas) a diferentes tratamiento (o brazos, ya que pueden compararse más de dos grupos), y también son los que requieren mayor precisión en cuanto a la metodología aplicada ( para eliminar sesgos, evitar la subjetividad, equiparar las muestras, lograr semejanza en las poblaciones, etc).

Los ensayos, también pueden ser no randomizados o con controles históricos, obviamente poseen una evidencia de menor calidad, pero suelen ser útiles cuando se aplican adecuadamente (en enfermedades sin curación o con una prevalencia muy baja).

#### **Ventajas y desventajas de estudios cohorte:**

Los estudios de cohorte son el planteamiento de **elección para estudiar las causas de un padecimiento**, el curso de una enfermedad o los factores de riesgo, debido a que son longitudinales y llevan el seguimiento de un grupo de sujetos durante cierto periodo. En general, los estudios de cohorte **no pueden probar causas**, ya que son de observación y no incluyen intervenciones; sin embargo, debido a que siguen a una cohorte de pacientes en tiempo futuro, poseen la secuencia correcta para proporcionar **evidencias sólidas de asociación y efectos posibles; entre una variable dada y un evento**, como en la controversia de tabaquismo y cáncer pulmonar. Además, al ser estudios planeados de manera prospectiva, los investigadores pueden controlar muchas fuentes de sesgo relacionadas con la selección de pacientes y el registro de mediciones

(que son de más difícil manejo en casos y controles, ya que dependen de la calidad de los archivos y de la adecuada aleatorización de los controles)

El **tiempo requerido** en un estudio de cohorte depende del problema estudiado, para el caso de enfermedades que se desarrollan durante largo tiempo o situaciones que se presentan como resultado de exposición prolongada a algún agente causal, se necesitan muchos años; los periodos extensos hacen que los estudios sean costosos y también dificultan que los investigadores argumenten causalidad, debido a que otros sucesos que se presenten en el periodo intermedio pueden haber afectado el resultado. Por ejemplo, el largo tiempo transcurrido entre exposición y efecto es una de las razones que dificultan estudiar una relación posible entre agentes ambientales y varios carcinomas. Los **estudios de cohorte que requieren mucho tiempo** para completarse, en especial son vulnerables a problemas relacionados con el seguimiento de pacientes, en particular quienes desisten (individuos que ya no participan en el estudio) y los que emigran (enfermos que se desplazan a otras comunidades).

El tiempo constituye un problema para terminar los estudios de cohorte. **Otra dificultad deriva de si la enfermedad o el resultado son raros**; por ejemplo los tumores carcinoides son neoplasias de crecimiento muy lento y relativamente infrecuentes. Los investigadores que desean analizar mediante un estudio de cohorte las relaciones entre factores de riesgo o características clínicas posibles con respecto a tumores carcinoides, necesitarían examinar un número muy grande de individuos expuestos, efectuar su seguimiento durante un periodo muy largo y probablemente serían muy pocos los enfermos. El tamaño de muestra requerido en un estudio de estas características a menudo vuelve prohibitivo el costo. En estas situaciones, con frecuencia los investigadores deciden utilizar estudios de casos y controles.

### **Ventajas y desventajas de estudios de casos-control:**

Estos estudios son en especial apropiados para observar enfermedades poco frecuentes, examinar trastornos que se desarrollan en un tiempo prolongado e investigar una hipótesis preliminar. En general, son los estudios más rápidos y menos costosos. Por ejemplo, Ross y Roberts (1985) en su estudio sobre factores clínicos relacionados con enfermedad cardíaca carcinoides, analizaron los datos de necropsias y registros médicos de pacientes que habían padecido enfermedad cardíaca carcinoides y de un grupo de sujetos que no la habían padecido. Con este planteamiento, estuvieron en situación de completar el estudio en un lapso y con un número de enfermos relativamente reducidos. Por tanto, **los estudios de casos control son ideales para investigadores que requieren obtener algunos datos preliminares para escribir un protocolo de un estudio más completo, costoso y prolongado**, además, son muy útiles para médicos de familia que necesitan llevar a cabo un proyecto de investigación clínica.

Las ventajas de los estudios de casos control conducen a sus propias desventajas. De todos los métodos de estudio, tienen el número más grande de sesgos o errores posibles, y **dependen por completo de la calidad de los registros existentes**. La disponibilidad de datos requiere en ocasiones compromisos entre aquellos investigadores que desean el estudio y los capacitados para brindar información. Por ejemplo, Berry y colaboradores (1984) deseaban determinar los posibles factores de riesgo para el desarrollo de infección por virus de hepatitis B entre trabajadores del área de la salud. Optaron por usar el número de años desde que un individuo se había graduado de médico como medida del riesgo. Una medida más indicativa podría haber sido el número de veces que un profesional del área de la salud se había expuesto a sangre o sus productos. Sin

embargo, Berry y colaboradores no podían haber obtenido esta información de los registros existentes, de modo que en su lugar tuvieron que emplear una medida sustituta.

**Uno de los mayores problemas en estudios de casos control es la selección de un grupo control apropiado.** Los casos son relativamente fáciles de identificar, pero es más difícil decidir el grupo de sujetos que proporcionan una comparación relevante.

Para ilustrar la importancia -y la dificultad- de elegir un grupo control apropiado, considérese un estudio de casos control planeado para analizar la relación entre el consumo de tabaco, alcohol, té y café con la incidencia de cáncer pancreático (MacMahon, Yen y Trichopoulos, 1981) Los casos eran pacientes hospitalizados que tenían diagnóstico histológico de cáncer pancreático. Para controles, los investigadores seleccionaron pacientes que estaban al cuidado del mismo médico en el mismo hospital y al mismo tiempo. Sin embargo, se excluyeron numerosos enfermos para considerarlos controles, esto es, no se emplearon individuos con enfermedades de páncreas, vías biliares o sistema cardiovascular, ni con diabetes sacarina, cáncer pulmonar, cáncer vesical o úlcera péptica. Los resultados del estudio mostraron que los casos, pacientes con cáncer pancreático, habían consumido cantidades mayores de bebidas con café que los controles.

Una falla mayor de este estudio es la elección del grupo de controles. En este grupo se incluyeron pacientes con cáncer de estómago e intestino y otros con trastornos gastroenterológicos no malignos como gastritis, enteritis y colitis, es bastante probable que dichos enfermos hayan suspendido el consumo de café debido a sus problemas. Incluir estos pacientes en el grupo control hace menos probable que los sujetos en dicho grupo consumieran grandes cantidades de café, lo que incrementa las probabilidades de una relación observada entre café y cáncer pancreático. Un grupo control con mejores características podría haberse integrado con pacientes hospitalizados por problemas no relacionados en absoluto. Debido a los problemas inherentes a la selección de un grupo control en un estudio de casos control, algunos estadígrafos recomiendan el uso de dos grupos control uno semejante en algunas características a los casos (como individuos que están hospitalizados durante el mismo tiempo) y otro de sujetos sanos.

Este tipo de problemas, al igual que la aleatorización de pacientes, es uno de los más complicados fundamentalmente cuando nos referimos a una población hospitalaria (internada), siendo quizá una de las más probablemente sesgada.

### **Ventajas y desventajas de estudios transversales:**

Hemos descrito estudios: ERC, Cohorte y Caso-control, con características decrecientes en cuanto a su grado de evidencia; **los estudios transversales** (sin tiempo, ya que son aquellos en los que se realiza una sola medición) se acercan mucho a los estudios descriptivos, pero a modo didáctico, hemos considerado que los diseños de corte transversal utilizados para **comparar métodos de diagnóstico** como **observacionales** (analíticos) ya que están planteando una comparación entre dos exámenes complementarios, y arriesgando algún grado de análisis (Acá se realiza una sola medición con cada método de diagnóstico, por lo tanto responden a la definición de transversales).

Son por lo tanto diseños de fácil realización, en donde se expone a un grupo de individuos, (sanos y enfermos) a 2 métodos de diagnóstico y se puede al comparar el test en estudio contra un gold standard, calcular la sensibilidad y especificidad de dicho test. En general no requiere de un diseño complicado, ni de tiempos prolongados para su realización.

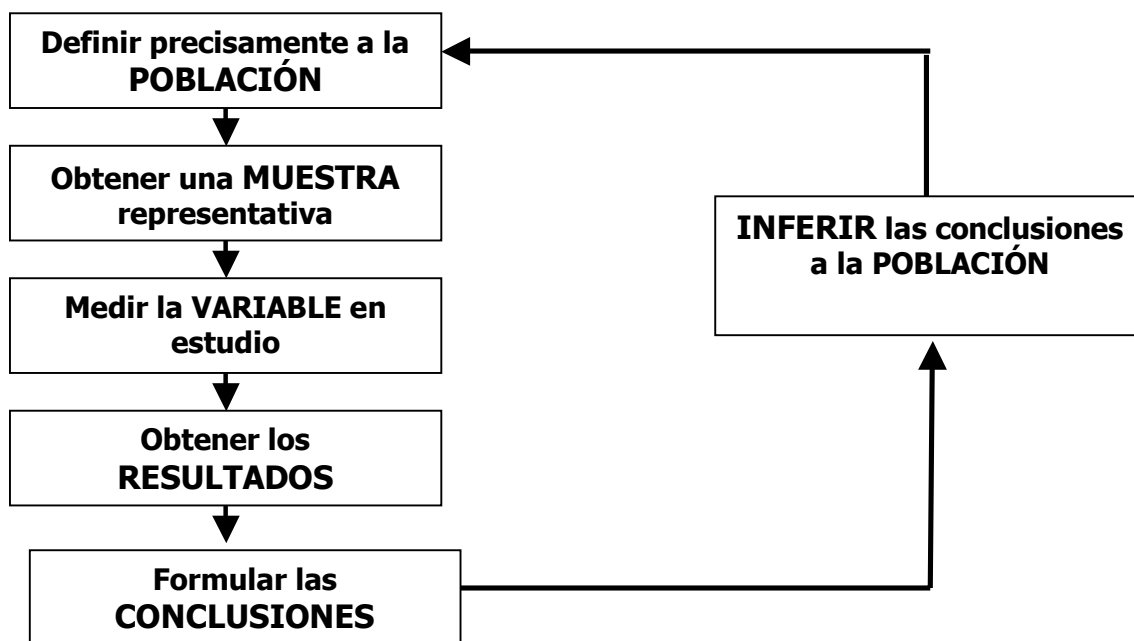
**DISEÑOS DESCRIPTIVOS: *Son aquellos diseños que no ponen a prueba ninguna hipótesis***

Los trabajos descriptivos son los que observan una situación en un momento dado, éste tipo de diseño, no arriesga en general ninguna hipótesis, solo busca identificar la variable en estudio en un grupo poblacional, y representa un INSTANTANEA de esa situación, y responden a las preguntas: **Quiénes son los afectados?, Donde están? Y Cuando aparecen a través del tiempo?**

Estas investigaciones tienen el propósito de describir o explorar, no planteando ninguna hipótesis causal, pero sus resultados son utilizados para elaborar hipótesis de estudios con estos propósitos. Son recomendables cuando se necesita caracterizar un problema de salud para establecer un programa de control, y cuando se desconoce la etiología del fenómeno y necesitamos identificar sus cualidades para enunciar hipótesis causales.

Entran en este tipo de estudios los: trabajos de **prevalencia**, la obtención de **Tasas** de mortalidad, obtención de **valores promedio** en una población (TA, Colesterol, etc.), **encuestas**, etc.

**EL ESQUEMA CONCEPTUAL QUE DEBEMOS REALIZAR PARA UN ESTUDIO DESCRIPTIVO ES EL SIGUIENTE:**



**Si hemos decidido realizar un estudio descriptivo**, (ya que está a nuestro alcance), lo primero es plantearse los objetivos según el problema que queremos responder:

1. Si desconocemos la prevalencia de parasitosis asintomática en los niños de mi Centro, el objetivo será: *estimar la prevalencia de parasitosis en niños de 2 a 10 años, que concurren por cualquier causa al centro de salud durante el mes de febrero 2001.*

2. Si desconocemos cual es el impacto del alcoholismo en nuestro CEMAP, el objetivo será: *“ Estimar la prevalencia de alcoholismo, en los varones de 17 a 60 años, afiliados a la OS, que consulten al CEMAP Capital durante el mes de diciembre ”.*
3. Si desconocemos -realmente- los hipertensos de la OS, el objetivo será: *estimar el porcentaje de hipertensos entre los afiliados que concurren al CEMAP duarnte los meses de enero y febrero del 2001.*

Recordamos siempre que el enunciado de los Objetivos debe contener:

- a) La ecuación matemática que deseo realizar: **porcentaje, tasa, prevalencia**, etc.
- b) La población a la cual está destinado el estudio (**edad, sexo**, etc).
- c) Las coordenadas témporo- espaciales (lugar de donde obtendré la muestra **y período de tiempo que recogeré los datos**)

Para poder avanzar en un estudio descriptivo, debemos conocer algunas reglas de muestreo. Para lo cual enumeraremos los pasos a seguir con el muestreo y los análisis estadísticos posibles:

- Determinar la **Población Blanco**: aquella para la cual se realizará la inferencia de los resultados (destinataria o universo).
- Definir la **Población Accesible** o Muestreada: aquella de la cual se obtendrá la muestra.
- Obtener la **Muestra** con las siguientes características:
  - ☛ **Tamaño** adecuado.
  - ☛ **Representatividad** y **semejanza** (evaluar estratificación)
  - ☛ **Randomización** o aleatorización: igual oportunidad de participar de toda la población

### **MEDIDAS ESTADÍSTICA:**

Una vez obtenida la muestra se realiza la **recolección de datos**, y posteriormente aplicaremos las **estadísticas**.

Llamamos **estadísticas o estadísticos** a la obtención de **un solo valor numérico**, calculado con los **resultados primarios de nuestra muestra**, que pueda **resumir los datos de la manera más representativa y con la mayor precisión posible**.

Por ejemplo: los resultados primarios podrán ser (para el cálculo de la TA media de una población): TAS: 120, 130. 160, 140, 130....., obtenemos la **media** y decimos que la TA media de dicha población es de: **136** que es la estadística obtenida (media o promedio)

Las medidas estadísticas las clasificamos en:

- ✍ **Medidas de Tendencia Central**
- ✍ **Medidas de Dispersión**

### Medidas de *tendencia Central*: (modo; mediana y media)

Son medidas que tratan de agrupar en un solo valor a todo el trabajo, y mediante la cual muchas veces diremos: " el 40% de la población es hipertensa " (**porcentaje**)

o: " La TA **promedio** en los médicos del CEMAP es de 120/85 "

➤ **MODO**.....(valor que más se repite)

1-2-2-3-3-4-4-4-4-5-5-5-5-7-7-9-9-10-12-13-13-15

El **Modo** o **Moda** es aplicable tanto a escalas **Nominales** (donde no existe relación entre los valores: sí- no, blanco- negro, hipertenso- normotenso, etc,) en escalas **Ordinales** (Donde hay un ranking de valores pero no proporcional:: Glasgow, Apgar, etc) como en escalas numéricas **Intervales**.

➤ **MEDIANA**..... (valor central)

1-2-2-3-3-4-4-4-4-5-5-5-5-7-7-9-9-10-12-13-13-15

La mediana solo se puede aplicar en escalas **ordinales** e **intervales**. (ya que requiere de un ordenamiento de los datos que las escala nominales no poseen).

➤ **MEDIA O PROMEDIO**..... (suma de valores dividido el N° de ellos)

$$\text{Promedio} = \frac{\sum x}{n}$$

Esta medida solo es aplicable a las escalas **intervales**.

### **Medidas de Dispersión:**

Son medidas, que tratarán de definirnos, como se distribuye la mayoría de los valores, alrededor de la medida de tendencia central:

En el ejemplo anterior " La TA promedio en los médicos del CEMAP es de 120/85 " deberá ser completado con: "un rango de: TAS:150-100 y TAD: 100-600" y un DS:+/- 15mmHg

➤ **RANGO**.....(extremo de distribución de todos los valores)

➤ **DESVÍO STANDART**.....( Promedio de distribución de los valores alrededor de la media)

$$\text{DS: } \sqrt{\frac{\sum x - \text{promedio}}{n - 1}}$$

### **Maneras de medir Datos en Escalas Nominales:**

Ya mencionamos que las escalas nominales son aquellas en las cuales **no existe relación entre sus valores**, como "sí y no", "blanco y negro" , etc. No existe ninguna

posibilidad de decir que un valor sea superior a otro, incluso existiendo más de dos valores (blanco, negro y verde) . Llamamos escalas binomiales a aquellas en las cuales solo existen 2 valores y son las que más frecuentemente encontramos en los estudios descriptivos. Hay tendencia a **nominalizar** la medición de las variables medidas inicialmente en escala numéricas (**intervalar**) y transformarlas en nominales, por ejemplo, las tensiones arteriales de 1000 individuos, cuando estoy investigando HTA , los datos dejarán de ser "130, 140 150", y se convertirán en hipertenso o normotenso.

Las mediciones usadas en escalas nominales son:

- ✓ **Proporciones o porcentajes:** N° de observaciones con una característica determinada (como la presencia de IAM) dividido el N° total de observaciones. Por ejemplo "la **proporción de IAM en individuos** que ingirieron AAS fue de: **300/11300=0.0265**" (en otras palabras es un **riesgo absoluto** aislado) cantidad de eventos sobre cantidad de expuestos.

$$\text{Proporción: } \frac{a}{a + b}$$

- ✓ **Razones y tasas:** Una **Razón** es: el N° de observaciones con una característica sobre el número de observaciones sin esa característica. Para el caso de los **IAM** en pacientes que ingirieron **AAS** es: **300/11000=0.0272** (IAM vs no IAM) .

$$\text{Razón: } \frac{a}{b}$$

**Tasa:** puede ser una "proporción" o no, pero usa un multiplicador llamado base, que podrá ser 1.000, 10.000. 100.000, etc, según la magnitud del numerador. A lo que se agrega que es obtenida en un período de tiempo y en un lugar determinado, de manera que el denominador suele ser el promedio de población en ese lapso (para mortalidad general por ejemplo). Para el caso previo será: "**26,5 IAM por mil individuos que toman AAS**" o "**27,2 IAM por cada 1000 no IAM que ingieren AAS**".

$$\text{Tasa: } \frac{a}{a + b} \times \text{base}$$

**Así para nuestro trabajo sobre el promedio de TA en nuestra población:** Una vez que hemos obtenido nuestro resultado, podremos decir que tenemos una **conclusión:** "la **prevalencia** de HTA en los afiliados es del **40%**" y que "La **TA promedio** en los afiliados fue de **125/90** con un **rango de TAS: 160-100** y **TAD: 100-50**" y un "**DS: +/-10mmHg**"

Ahora bien: **¿con que certeza afirmaremos esto?:**

Pues bien, si en nuestro trabajo le hubiéramos tomado la TA a toda la población, (es decir no nos hubiese quedado ningún afiliado sin registrar; cosa poco probable) tendremos el 100% de certeza que el 40% de los afiliados son hipertensos y que 125/90 es la TA promedio.

Pero aquí el problema es que tomamos solo una muestra, por lo tanto nunca podremos afirmar con el 100% de certeza que esos números sean los reales, es aquí donde debemos incorporar una nueva figura:

### “ El Intervalo De Confianza ”

Como lo que hicimos es tomar una muestra, ahora nosotros debemos hacer una **INFERENCIA** (ver esquema conceptual) de que los resultados que obtuvimos de la muestra servirán para aplicar a la población; esta inferencia tendrá un margen de error, a esto lo llamamos “ intervalo de confianza ”, que es un rango de valores en torno al estadístico central (promedio, porcentaje, RRR, etc) entre los que se espera pueda encontrarse el verdadero valor del parámetro, con una probabilidad alta conocida, a la que denominamos “Nivel de Confianza” y que habitualmente se acuerda en el **95%**.

En definitiva el **Intervalo de Confianza** se construye con fórmulas, conociendo: **tamaño de la población, tamaño de la muestra, el valor central** (estadístico) y el **grado de dispersión** de los valores (DS).

A menor dispersión de valores y tamaño óptimo de la muestra, el intervalo de confianza tendrá un rango menor, acercándose más al valor central.

Su lectura en síntesis significa, por ejemplo: si digo que la **TAS** promedio es de **135** con un **IC95% de 125-145**, que si repito **100** veces la medición, **95** veces el promedio de TA en la población caerá dentro de esos valores.

### Test de Hipótesis “p”

Todo trabajo de investigación se realiza con el propósito de probar una hipótesis (suposición posible de ser confirmada); por otro lado toda hipótesis posee una hipótesis opuesta a la que se quiere probar que denominaremos Hipótesis nula o **H<sub>0</sub>**.

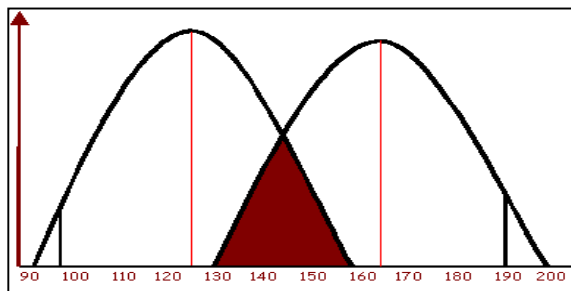
Ahora, supongamos que nosotros queremos diseñar un trabajo que demuestre que esta lloviendo, nuestra hipótesis de trabajo o **H<sub>1</sub>** es que efectivamente esta lloviendo. El trabajo consiste en sacar la mano y fijarse si cae alguna gota, si es así, está lloviendo, pero si no cae ninguna gota obviamente no llueve. Por lo tanto en este trabajo nuestra hipótesis nula **H<sub>0</sub>** es: “no esta lloviendo”, para cada hipótesis de trabajo existe una hipótesis nula, que es opuesta a lo que queremos demostrar.

Ahora, yendo a un ejemplo mas relacionado con la medicina, queremos demostrar que una nueva droga es efectiva para disminuir la tensión arterial, nuestra hipótesis de trabajo (**H<sub>1</sub>**) es que la droga es útil y que efectivamente disminuye la tensión arterial (**H<sub>1</sub>**= “ la droga X es mejor que el placebo para reducir la TA”), nuestra hipótesis nula será que la droga no sirve y por lo tanto no es efectiva para disminuir la tensión arterial (**H<sub>0</sub>**= “la droga X, no es mejor que el placebo para reducir la TA”).

Supongamos que ya hicimos el trabajo, tenemos nuestros datos y obtuvimos un valor promedio de tensión arterial, para el grupo que recibió placebo y otro valor (también promedio) para el grupo que recibió la droga.

Ahora bien, partamos de la base que la droga es igual de efectiva que el placebo con respecto a la disminución de la tensión arterial (o sea que no tiene efecto que es nuestra hipótesis nula) y tenemos que demostrar que sí es efectiva, o sea tenemos que descartar nuestra hipótesis nula, para poder probar nuestra hipótesis de trabajo.

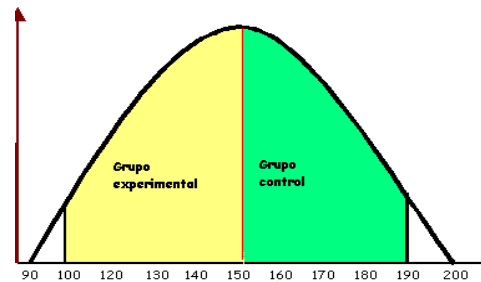
Obtuvimos dos valores promedio de los dos grupos, podemos pensar que estos valores promedio resumen a un conjunto de



valores para cada grupo que asumimos que se distribuyen de forma normal (es decir con forma de campana de Gauss), tenemos una campana para cada grupo.

Nuestros dos grupos se diferencian solamente en que unos recibieron la droga y otros no, por lo que si la droga es igual al placebo, no se van a diferenciar en nada, van a ser iguales. Entonces lo que tenemos que demostrar es que el promedio de cada grupo corresponde a dos grupos distintos (cada uno representado por una campana de la figura) y no se trata de una campana sola en donde cada lado de la misma quedaron por azar el **grupo control y el experimental**.

De esa manera podemos atribuirle la diferencia a la droga, si no lo logramos, la diferencia que obtuvimos es simplemente un valor extremo del mismo grupo, ya que no se diferencian en nada con el otro grupo.



Volviendo a la curva de distribución normal del grupo placebo, sabemos que el área debajo de la curva es igual a 100% o 1 (por definición), por lo que el promedio en una distribución normal me divide a la curva en dos áreas iguales cada una de las cuales ocupa un 50% del área bajo la curva, a medida que nos vamos corriendo hacia los extremos cada vez es menor el área bajo la curva que me queda tanto a la derecha como a la izquierda del valor que voy tomando. Así a medida que me voy corriendo hacia la izquierda, el valor va disminuyendo desde 50% hasta hacerse extremadamente chico a medida que me acerco al extremo izquierdo de la curva.

En algún momento, voy a obtener un valor que deje a la izquierda solo el 5% del área bajo la curva, pero todavía van a ser valores normales, solo que van a estar dentro del 5% inferior.

Ahora lo que tenemos que hacer es demostrar que la diferencia que obtuvimos pertenece a un grupo distinto y no es un valor extremo del grupo placebo, entonces fijamos como punto de corte 5%. Esto significa que si el valor que obtuvimos del grupo intervención es menor que el valor que me divide al área del grupo placebo en un 5% a la izquierda y un 95% a la derecha, digo que ya no es un valor extremo y que la diferencia se la puedo atribuir a la droga.

Por supuesto que nunca voy a poder estar totalmente seguro, bueno, mi grado de error es 5%, tengo un 5% de posibilidades de equivocarme y atribuir la diferencia en las cifras tensionales a la droga cuando en realidad era un valor extremo del grupo placebo.

Ahora, si el valor que obtuvimos en el grupo al cual le dimos la droga, es mayor que el punto de corte que me deja el área del 5% a la izquierda de la curva, o sea que cae dentro del 95% derecho, asumimos que el valor pertenece que no se puede diferenciar del grupo placebo y al asumir que los dos grupos no son distintos, asumimos que la droga y el placebo son iguales, ya que los dos grupos no se diferencian en nada.

O sea, redefiniendo lo que vimos hasta ahora, cuando leemos un trabajo, vemos que asociado a cada valor obtenido vemos un valor de "p". **Este valor de "p", si es menor a 0,05 decimos que la diferencia obtenida es significativa** y si no lo es, que no es significativa.

Y esto a su vez **significa que cuando el valor de "p" es menor a 0,05 la diferencia obtenida tiene menos del 5% de probabilidades de no corresponder a la droga utilizada**, es decir estamos muy seguros de que la causante del efecto fue la droga y por lo tanto decimos que descartamos nuestra hipótesis nula.

Ahora bien si el valor de "p" calculado es mayor a 0,05 tenemos mas de un 5% de probabilidades de que la diferencia obtenida no se deba a la droga, por lo que como ese fue nuestro punto de corte, no podemos descartar la hipótesis nula.

**No decimos que la hipótesis nula es verdadera**, ya que puede ser que otro trabajo la descarte, simplemente decimos que **nosotros no la pudimos descartar**, lo que no implica que no se pueda descartar en el futuro, ya que el valor de "p", no solo depende de la magnitud de la diferencia obtenida, sino también del número de pacientes.

Es decir, si queremos probar que una droga para disminuir la tensión arterial es efectiva pero su efecto no es muy pronunciado, vamos a necesitar muchos pacientes para obtener una "p" significativa, por el contrario si la droga es muy efectiva, con menos pacientes, ya vamos a poder descartar nuestra hipótesis nula, o sea obtener una P menor a 0,05.

Cabe aclarar finalmente, que el test de hipótesis se realiza siempre que en el trabajo se comparen dos grupos, ya sea en un **ERC** que compare droga contra placebo, o dos drogas entre sí, como en una **cohorte** o un **caso control** donde comparemos dos grupos (expuesto y no expuesto) así como en un estudio **transversal** que compara dos métodos diagnósticos. En el caso de los **estudios descriptivos, ya mencionamos que no contrastan ninguna hipótesis**, ya que solo representan una foto instantánea de la situación y no existe hipótesis alguna.

A continuación se describe, el grado de evidencia que arroja cada uno de los estudios mencionados.

**Escala de Evaluación de la evidencia**

<b>Niveles del mayor I al menor-IX</b>	<b>Calidad de la Evidencia</b>	<b>Tipo de diseño del estudio</b>	<b>Condiciones de rigor científico</b>
--	--------------------------------	-----------------------------------	--

<b>I*</b>	<b>Buena</b>	<b>Meta-análisis de ensayos controlados y aleatorizados</b>	<b>Análisis de datos de pacientes individuales Meta-regresión Diferentes técnicas de análisis Ausencia de heterogeneidad Calidad de los estudios</b>
<b>II*</b>		<b>Ensayos controlados y aleatorizados de muestra grande</b>	<b>Evaluación del poder estadístico Multicéntrico Calidad del estudio</b>
<b>III*</b>	<b>Buena regular</b>	<b>Ensayos controlados y aleatorizados muestra pequeña</b>	<b>Evaluación del poder estadístico Calidad del estudio Multicéntrico</b>
<b>IV*</b>		<b>Ensayos prospectivos controlados no aleatorizados</b>	<b>Evaluación del poder estadístico Multicéntrico Calidad del estudio</b>
<b>V*</b>	<b>Regular</b>	<b>Ensayos prospectivos controlados no aleatorizados</b>	<b>Controles históricos</b>
<b>VI*</b> <b>VII*</b>	<b>Regular</b>	<b>Estudios de control Estudios caso-control</b>	<b>Calidad del estudio Multicéntrico Apareamiento</b>
<b>VIII*</b> <b>IX</b>	<b>Baja</b>	<b>Series clínicas no controladas Estudios descriptivos: seguimiento de la enfermedad, vigilancia epidemiológica, registros, bases de datos Comités de expertos, conferencias de consenso Anécdotas o casos</b>	

\* Calidad del estudio evaluado mediante protocolos y condiciones de rigor científico. Adaptado de Jovell AJ. Navarro-Rubio MD. Evaluación de la evidencia científica. Medicina Clínica (Bar) 1995; 105: 740-

## **CONSIDERACIONES SOBRE EL DIAGNOSTICO Y EL TRATAMIENTO**

En todo proceso médico hay dos momentos claves, el diagnóstico y el tratamiento.

Si consideramos el screening como un tipo particular de diagnóstico, ya que en definitiva se trata de diagnosticar una enfermedad en su período presintomático, y el tratamiento como una acción tendiente a que algo no ocurra, ya sea una enfermedad (prevención primaria) o la progresión de una enfermedad ya establecida (prevención secundaria o terciaria), podemos incluir también dentro de estos conceptos los principios de la medicina preventiva.

Dependiendo del diseño con respecto a la dirección en el tiempo (cohortes, casos control o transversal) se les asigna mayor o menor grado de certeza en el momento de analizar la información brindada.

Muchos de los trabajos que analizan el comportamiento de test diagnósticos corresponden a diseños de corte transversal.

Inicialmente vamos a desarrollar conceptos generales de los trabajos de comparación que se ajustan tanto al diagnóstico como al tratamiento, y con posterioridad vamos a dividir el capítulo en una parte tendiente al diagnóstico y otra al tratamiento.

### **Conceptos generales (Tabla de 2x2)**

En muchos trabajos, se realiza una comparación entre dos grupos de pacientes, ya sea porque tienen diferente riesgo de padecer una enfermedad, o porque están sujetos a diferentes intervenciones o en el caso de test diagnósticos se compara un test, contra un gold standard.

Se utiliza un grupo que se denomina grupo intervención (sobre el cual se ejerce algún tipo de intervención) y el otro grupo se denomina grupo control ya que generalmente es un grupo virgen de intervención que debe ser lo más semejante posible al grupo intervención y cuyo objetivo es comparar los resultados del grupo intervención para de esta forma poder evaluar el beneficio de la intervención realizada.

Para analizar los resultados de este tipo de trabajos, se ordenan los datos de una forma estandarizada que se denomina tabla de 2x2.

La tabla de 2x2, tiene dos variantes dependiendo de si se trata de un trabajo de test diagnósticos o de un trabajo que compara dos tratamientos. Los trabajos que analizan factores de riesgo y comparan dos grupos uno expuesto y otro no, también expresan los datos mediante una tabla de 2x2.

En las tablas tendientes al diagnóstico en las filas se agrupan los pacientes de acuerdo al resultado del test en estudio (positivos en la fila superior y negativos en la fila inferior). A su vez se acomodan en las columnas dependiendo del resultado del gold standard (que

es el estudio contra el cual se esta comparando el test a analizar), positivos en la columna de la izquierda y negativos en la columna de la derecha.

De esta manera nos quedan todos los pacientes divididos en cuatro grupos, de acuerdo a dos variables con solo dos resultados posibles, positivo o negativo.

	<b>Enfermedad presente</b>	<b>Enfermedad Ausente</b>
<b>Test positivo</b>	Verdaderos (+) <b>a</b>	Falsos (+) <b>B</b>
<b>Test negativo</b>	Falsos (-) <b>c</b>	Verdaderos (-) <b>D</b>

Ahora cuando estamos analizando datos tendientes al tratamiento, los pacientes tambien se dividen en cuatro grupos de acuerdo a dos variables, pero esta vez en las filas se agrupan de acuerdo a la intervención, así en la fila superior me quedan los pacientes del grupo intervención y en la fila inferior los del grupo control. Al comparar dos tratamientos se mide la ocurrencia o no de un evento que es el que se quiere evitar con el tratamiento, por lo que en las columnas los pacientes se agrupan de acuerdo a la ocurrencia o no del evento en cuestión, en la columna de la izquierda los que presentaron el evento y en la columna de la derecha los que no lo sufrieron, por lo que otra vez quedan los pacientes divididos en cuatro grupos de acuerdo al resultado de dos variables con solo dos resultados posibles.

	<b>Evento presente</b>	<b>Evento ausente</b>
Grupo tratamiento	a	B
Grupo control	c	D

- a) Nro. de pacientes en el grupo de los tratados que tuvieron el evento
- b) Nro. de pacientes en el grupo tratamiento que no tuvieron el evento
- c) Nro. de pacientes en el grupo control que tuvieron el evento
- d) Nro. de pacientes en el grupo control que no tuvieron el evento

Después de estas consideraciones generales vamos a analizar los aspectos relacionados con el diagnóstico y posteriormente con el tratamiento.

### **Diagnóstico**

Comenzamos planteando un caso clínico que nos sirva de guía para desarrollar los conceptos a analizar:

Eva de 54 años, casada, contadora, con 2 hijos, actualmente vive con su marido y nos cuenta que desde hace 8 años realiza terapia de reemplazo hormonal, indicada por una médica que ella veía desde hacía bastante tiempo.

Actualmente se cambió de obra social y nos consulta porque tuvo hace 2 semanas, dos episodios de metrorragias leves separados por 3 días uno del otro, que no se volvieron a repetir, esto nunca le había sucedido con anterioridad.

Nos refiere que nunca estuvo internada salvo para tener a sus dos hijos, fuma aprox. 5 cigarrillos por día, no tiene otros antecedentes de importancia.

A su vez, nos cuenta que a una amiga que le sucedió algo parecido le hicieron un raspado para descartar la presencia de Cáncer de endometrio y que estuvo con bastante molestia durante un tiempo.

Su inquietud es saber si es necesario realizar un raspado y si no hay otra alternativa, para descartar la presencia de Cáncer de endometrio.

¿Qué le vamos a contestar a Eva?

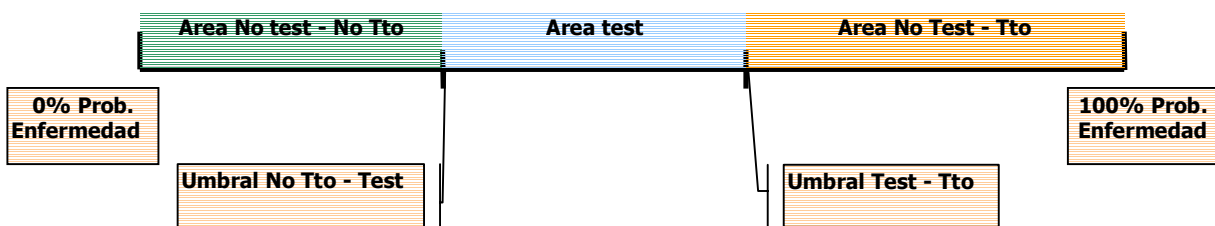
### **Teoría de los umbrales y probabilidad previa**

Como comenzamos diciendo al comienzo, dos son los momentos importantes a la hora de tomar decisiones dentro del proceso médico.

A esto se puede aplicar un modelo que ayuda a facilitar el proceso de toma de decisiones denominado teoría de los umbrales.

La teoría de los umbrales, considera que la ocurrencia o no de enfermedad se puede graficar en un continuo de probabilidad que va de 0 a 100, donde 0 es la ausencia de enfermedad y 100 es la absoluta certeza de enfermedad. Dentro de este continuo existen en el medio infinitas posibilidades de acuerdo a la probabilidad de enfermedad.

La teoría considera dos momentos claves dentro del continuo que son el de solicitar un test diagnóstico y el de implementar un tratamiento, fijando así dos umbrales, y delimitando el continuo en tres áreas.



En el momento en que un paciente entra al consultorio, antes de que nosotros le preguntemos algo, lo único que sabemos es si es varón o mujer y la edad, por lo que tiene la misma probabilidad que cualquier persona de su sexo y edad de tener cualquier enfermedad. Esto significa que la probabilidad previa (antes de hacer algo) es igual a la prevalencia de la enfermedad en cuestión en la población general.

Ahora, con el interrogatorio y el examen físico, nosotros vamos ajustando la probabilidad previa (antes de pedir un test) para determinada enfermedad, a esa persona. Es decir vamos recabando datos, tanto del interrogatorio como del examen físico que nos acercan o nos alejan de la probabilidad de enfermedad. Este proceso, que se repite en cada consulta y para cada dato obtenido, culmina en una probabilidad que puede caer en alguna de las tres áreas antes mencionadas.

Si la probabilidad cae en el área no test, no tratamiento, estamos lo suficientemente seguros de la ausencia de enfermedad, como para no necesitar confirmarlo y por lo tanto tampoco instituímos un tratamiento. Por el contrario si cae dentro del área no test, tratamiento, tenemos la suficiente certeza de presencia de enfermedad, para no necesitar confirmarlo, y comenzamos un tratamiento sin necesidad de solicitar un test diagnóstico.

Ahora bien si nuestra probabilidad previa o pre-test cae dentro del área test diagnóstico, necesitamos pedir un test, para poder correr la probabilidad tanto al área no test, no tratamiento y descartarla, como al área no test, tratamiento y comenzar con el mismo.

Esto significa que un test que no nos modifique la probabilidad lo suficiente como para salir del área del medio, no nos fue de gran ayuda, ya que seguimos en la misma situación que antes de solicitarlo. Mas adelante veremos como las características de los tests, nos permiten predecir el comportamiento del mismo y poder calcular una probabilidad pos-test que va a ser la probabilidad final.

Los valores de los umbrales que dividen las áreas, dependen mas que nada de la gravedad de la enfermedad (queremos estar muy seguros de que alguien no tiene cáncer por ejemplo) y de los efectos adversos del tratamiento (de igual forma no vamos a comenzar quimioterapia en alguien si no estamos lo suficientemente seguros de la presencia de la enfermedad). Existen fórmulas que toman en cuenta estos aspectos y permiten calcular los umbrales con mayor precisión, pero exceden los objetivos propuestos.

En el caso de Eva, la probabilidad previa de Cáncer de Endometrio, aunque es baja (5%) igual al ser el Cáncer una enfermedad potencialmente grave, queremos que la probabilidad de ocurrencia sea aún mas baja, por lo que por las características de su persona, caería dentro del área test diagnóstico.

Es importante destacar que la finalidad de pedir o no un test diagnóstico depende de la capacidad del test para modificar nuestra conducta posterior al resultado del mismo; si después de haberlo pedido nada va a cambiar nuestro proceder no fue de mucha utilidad. Es útil antes de pedir un test, pensar que voy a hacer si da positivo y que voy a hacer si da negativo, para poder objetivar el propósito de la solicitud del mismo.

Resumiendo, tenemos una probabilidad previa que se modifica de acuerdo a los datos obtenidos mediante el interrogatorio y el examen físico y después de acuerdo a esta probabilidad pre-test, evaluaremos la posibilidad de pedir un test para después del mismo corregir nuestra probabilidad a una nueva probabilidad pos-test que va a determinar nuestra conducta.

Es importante entender que la probabilidad pos-test no solo depende de las características operativas del test (sensibilidad y especificidad) sino que también depende de la probabilidad pre-test, como ya veremos más adelante al adentrarnos un poco más en las fórmulas que nos permiten calcularla.

### **Características operativas de los tests: sensibilidad y especificidad**

Mediante la sensibilidad y la especificidad de un test, podemos predecir el comportamiento del mismo, es decir podemos calcular una probabilidad pos-test, si conocemos la probabilidad pre-test, la sensibilidad y la especificidad. Este es el escenario que habitualmente se presenta dentro del consultorio, estamos ante la posibilidad de pedir un test, buscamos en un libro las características operativas del mismo y de acuerdo a la probabilidad previa, calculamos una probabilidad pos-test.

Ahora bien en algún momento alguien tuvo que calcular la sensibilidad y la especificidad de este test, entonces ahora planteemos otro escenario que nos va a ser más útil para entender estas características, que es el del investigador clínico que tuvo que calcular estos valores.

Definimos a la sensibilidad como el porcentaje de pacientes enfermos de una determinada enfermedad en los cuales realizamos el test y este da positivo. Ej: de cada 100 pacientes hipertiroideos, 100 tienen taquicardia, por lo tanto la taquicardia es 100% sensible para el diagnóstico de hipertiroidismo, esto significa que no existe el hipertiroidismo sin taquicardia. La sensibilidad me habla de los pacientes enfermos.

Por otro lado, la especificidad es el porcentaje de pacientes sanos en los cuales el test me da negativo, en realidad para ser más precisos en la definición es el porcentaje de test negativos en personas que no tienen la enfermedad en cuestión, pueden estar sanas o tener otra enfermedad distinta de la que estamos analizando. De esta forma de cada 100 personas que no tienen enfermedad de Wilson (deficiencia de cobre) ninguna tiene el signo de Kaiser-Fleischer (anillo ocre alrededor de los ojos), esto significa que si encontramos este signo, podemos afirmar que alguien tiene Enf. de Wilson, ya que no se encuentra en ninguna otra persona, por lo que el anillo de Kaiser-Fleischer es 100% específico para la deficiencia de cobre.

Como vemos, para poder determinar la sensibilidad o especificidad de un test, es necesario saber con anterioridad si la persona está sana o enferma, entonces retomando al investigador clínico que quiere calcular las características de un nuevo test, vemos que necesita primero saber si esta persona tiene o no la enfermedad, esto se logra con lo que se denomina un gold standard que es un test de referencia, contra el cual vamos a comparar el nuevo test. Si o si, necesitamos un test de referencia, para poder comparar nuestro nuevo test.

Necesitamos un test de referencia (gold standard) y una población a la cual realizarle a todos los dos test, tanto el gold standard como el nuevo test a analizar, para después

poder confeccionar una tabla de 2x2 y calcular las características operativas del test en cuestión. Este tipo de trabajo se encuadra dentro del diseño de corte trasversal ya que en un corto período de tiempo se le realizan los dos test a toda la población.

Retomando el caso de Eva evaluamos la capacidad de la ecografía transvaginal para diagnosticar el cáncer de cuello de endometrio en comparación con la biopsia endometrial que era el procedimiento diagnóstico habitual antes de su aparición y a su vez el que le realizaron a la amiga de Eva.

Queremos saber como se comporta un nuevo test (ecografía transvaginal) y tenemos un gold estándar que es el resultado anatomopatológico de una muestra de endometrio tomada mediante un legrado. Tenemos que tener en cuenta que la ecografía mide el endometrio, por lo tanto los valores son continuos y que los datos en una tabla de 2x2, se ordenan de acuerdo a resultados si-no, por lo que es necesario tener un punto de corte, es decir, mas allá de 5 mm consideramos el resultado positivo y menos de 5 mm lo consideramos negativo, el punto de corte, me permite transformar la forma de presentación de los datos. Mas adelante vamos a hacer algunas consideraciones con respecto a los puntos de corte, ya que se modifican las características operativas de los test, si modificamos los puntos de corte.

Entonces a 1000 mujeres les realizamos tanto una ecografía como un legrado y acomodamos los datos en una tabla de 2x2, de acuerdo a los resultados.

Utilizamos como Gold Standard el legrado.

	<b>Legrado positivo (50)</b>	<b>Legrado negativo (950)</b>
<b>Ecografía positiva (520)</b>	<b>45</b> a	<b>475</b> b
<b>Ecografía negativa (480)</b>	<b>5</b> c	<b>475</b> d

(datos modificados de Transvaginal ultrasonography compared with endometrial biopsy for the detection of endometrial disease. PEPI. N Engl J Med 1997 Dec 18 ;337(25) :1792-8)

Sensibilidad =  $\mathbf{a/a+c}$  =  $45/45+5$  =  $45/50$  = 0,90      Sensibilidad es de 90%  
Especificidad =  $\mathbf{d/d+b}$  =  $475/475+475$  =  $475/950$  = 0,50      Especificidad es de 50%

V. Pred. Pos.=  $\mathbf{a/a+b}$  =  $45/45+475$  =  $45/520$  = 0,08      V. Pred. Pos. es de 8%  
V. Pred. Neg.=  $\mathbf{d/c+d}$  =  $475/475+5$  =  $475/480$  = 0,98      V. Pred. Neg. es de 98%

Habíamos dicho que la sensibilidad era el porcentaje de pacientes enfermos en los cuales el test daba positivo, si tomamos como enfermedad al resultado positivo del legrado, vemos que de las 1000, 50 mujeres están enfermas, lo que nos permite calcular una prevalencia o probabilidad pre-test del 5%. Ahora bien de estas 50, 45 tienen una

ecografía positiva para Ca. Endometrio por lo que podemos calcular una sensibilidad del 90% para la ecografía con respecto al Ca de Endometrio.

Ahora vayamos a la especificidad. Tenemos 950 mujeres que están sanas, y de estas la ecografía fue negativa en 475, por lo que calculamos una especificidad del 50% para la ecografía con respecto al Ca. de Endometrio.

Vemos que tanto para calcular la sensibilidad como la especificidad se utilizan las columnas de la tabla de 2x2. La sensibilidad la calculamos en la columna de la izquierda y la especificidad en la columna de la derecha.

### **Probabilidad pos-test y valores predictivos**

Como habíamos dicho inicialmente, estamos realizando todo este análisis, para poder calcular una probabilidad de enfermedad después de haber realizado un test diagnóstico y así poder evaluar la utilidad de solicitar el mismo.

Una vez solicitado el test y obtenido el resultado, es necesario evaluar la nueva probabilidad de enfermedad después de este.

Definimos al valor predictivo positivo de un test, como la probabilidad de ocurrencia de la enfermedad cuando este test da positivo, esto lo que nos dice es que si un test da positivo, como hay falsos positivos no estamos 100% seguros de la presencia de enfermedad, en otras palabras, el valor predictivo positivo es el porcentaje de verdaderos positivos con respecto a todos los positivos.

De la misma forma analizamos el valor predictivo negativo, que es la probabilidad de enfermedad cuando el test da negativo y que corresponde al porcentaje que ocupan los verdaderos negativos con respecto a todos los negativos.

Volviendo a Eva vemos que la Ecografía TV, cuando se realiza a 1000 personas, con el objetivo de encontrar Ca. Endometrio, resulta positiva en 520, pero de estas solo 45 personas realmente tenían Ca. Endometrio por legrado, por lo que el valor predictivo positivo es  $45/520$ , o sea 0,08, el valor predictivo positivo es del 8%.

Por otro lado con respecto al valor predictivo negativo, haciendo el mismo análisis con respecto a los negativos y los verdaderos negativos, tenemos que  $475/480$ , va a ser el valor predictivo negativo, o sea 98%, lo cual implica que después de realizar una ecografía TV con el objetivo de encontrar Ca Endometrio, utilizando como punto de corte 5 mm de grosor, si esta da negativa, estamos 98% seguros de que la persona no tiene Ca Endometrio.

Así como la sensibilidad y la especificidad la calculábamos en las columnas, los valores predictivos los calculamos en las filas. En la fila superior el valor predictivo positivo y en la inferior el negativo.

Los valores predictivos, se modifican de acuerdo a la probabilidad pre-test o sea a la prevalencia de la enfermedad. Supongamos un test con una sensibilidad del 90% y una

especificidad del 80%, para una prevalencia de 20%, la tabla de 2x2, me queda de la siguiente forma:

	<b>Enfermos (200)</b>	<b>Sanos (800)</b>
<b>Test (+)</b>	<b>180</b>	<b>160</b>
<b>Test (-)</b>	<b>20</b>	<b>640</b>

$$\begin{aligned} \text{Sensibilidad} &= 180 / (180+20) = 90\% \\ \text{Especificidad} &= 640 / (640+160) = 80\% \\ \text{Val. Pred. Posit} &= 180 / (180+160) = 52\% \\ \text{Val. Pred. Neg.} &= 640 / (640+20) = 94\% \end{aligned}$$

Ahora veamos el mismo ejemplo, para una prevalencia del 5% :

	<b>Enfermos (50)</b>	<b>Sanos (950)</b>
<b>Test (+)</b>	<b>45</b>	<b>190</b>
<b>Test (-)</b>	<b>5</b>	<b>760</b>

$$\begin{aligned} \text{Sensibilidad} &= 45 / (45+5) = 90\% \\ \text{Especificidad} &= 760 / (760+190) = 80\% \\ \text{Val. Pred. Posit} &= 45 / (45+190) = 19\% \\ \text{Val. Pred. Neg.} &= 760 / (760+5) = 99\% \end{aligned}$$

Ahora bien, hasta aca, ya vimos como se calculan la sensibilidad, la especificidad y los valores predictivos si conocemos la prevalencia o probabilidad pre-test mediante la realización de una tabla de 2x2. Realizar este proceso cada vez que tenemos que solicitar un test, sería mas que engorroso, entonces se diseñó una forma que nos permite calcular la probabilidad pos-test, teniendo los mismos datos (prevalencia, sensibilidad y especificidad) pero sin tener que hacer una tabla de 2x2.

### **Coefficientes de probabilidad**

Los coeficientes de probabilidad, relacionan las características operativas (sensibilidad y especificidad) de forma tal que al multiplicar este valor posteriormente con la prevalencia o probabilidad previa, nos permite calcular la probabilidad pos-test, facilitando así el proceso de cálculo.

Cuanto mayor es el valor del coeficiente de probabilidad, significa que mayor es la probabilidad de enfermedad después del resultado del test. Es decir, si un test, tiene un coeficiente de probabilidad positivo mas alto que el de otro test, para iguales prevalencias, si los dos test dan positivos, hay mayor probabilidad de enfermedad o probabilidad pos-test con el que tenga el coeficiente mas alto. Con el coeficiente negativo se realiza el mismo análisis, pero la probabilidad de enfermedad disminuye en vez de aumentar.

El coeficiente de probabilidad positivo se calcula, dividiendo la sensibilidad sobre la tasa de falsos positivos ( $CPP = \text{sens} / (1 - \text{esp})$ ), para el coeficiente de probabilidad negativo se divide la tasa de falsos negativos sobre la especificidad ( $CPN = (1 - \text{sens}) / \text{esp}$ ).

Ahora bien, para poder utilizar los coeficientes de probabilidad, se necesita hacer una conversión matemática en la probabilidad pre-test para poder multiplicarla por los coeficientes de probabilidad y después hay que deshacerla para poder obtener la probabilidad pos-test de forma tal de poder interpretarla.

Esta conversión matemática nos permite transformar la probabilidad en una razón de probabilidad. Esto significa que cuando decimos que la probabilidad de determinada enfermedad es del 50%, estamos diciendo que tiene una probabilidad de 1:1, o sea, por cada enfermo hay un sano, tenemos la misma cantidad de sanos que de enfermos. Ahora bien si la prevalencia es del 25%, estamos diciendo que por cada enfermo hay 3 sanos.

Para calcular la razón de probabilidad se divide el porcentaje de enfermos, sobre el porcentaje de sanos.

$$\text{Razón de probabilidad} = \text{Prevalencia} / (100 - \text{prevalencia})$$

$$\begin{aligned} \text{Para una prevalencia } 50\% &= 50 / (100 - 50) \\ &= 1 \end{aligned}$$

$$\begin{aligned} \text{Para una prevalencia } 25\% &= 25 / (100 - 25) \\ &= 25 / 75 \\ &= 0,33 \end{aligned}$$

Cuando la prevalencia es mayor al 50%, significa que hay mas enfermos que sanos, por lo que la razón de probabilidad es mayor a 1. Para una prevalencia del 80%, por cada 4 enfermos tenemos 1 sano (Razón probabilidad=80/20=4).

Una vez convertida la probabilidad, en razón de probabilidad, multiplicamos este valor por el coeficiente de probabilidad del test y obtenemos una nueva razón de probabilidad pero esta vez, pos-test.

$$\text{Razón de probabilidad pos-test} = \text{Razón de probabilidad pre-test} \times \text{Coeficiente de Prob.}$$

Para finalizar este proceso, debemos reconvertir la razón de probabilidad pos-test a una probabilidad expresada en porcentaje para que sea mas fácil la visualización del dato, para esto tenemos que:

Prob. Pos-test expresada en porcentaje=Razón de prob./ (Razón de prob. + 1)

Así, utilizando los mismos ejemplos anteriores vemos que una razón de probabilidad de 1 se convierte a un porcentaje de la siguiente forma:

$$\begin{aligned} \text{Prob. en porcentaje} &= 1 / 1+1 \\ &= 1 / 2 \\ &= 0,5 \text{ ó } 50\% \end{aligned}$$

y para una razón de 0,33

$$\begin{aligned} \text{Prob. en porcentaje} &= 0,33 / 0,33 + 1 \\ &= 0,33 / 1,33 \\ &= 0,25 \text{ ó } 25\% \end{aligned}$$

De esta forma, vemos, como convirtiendo la probabilidad pre-test en una razón de probabilidad y multiplicando dicho valor por el coeficiente de probabilidad tanto positivo como negativo, obtenemos una razón de probabilidad pos-test, que al convertirla nuevamente en porcentaje es el equivalente al valor predictivo para dicho resultado, pero nos obviamos la realización de la tabla de 2x2.

Retomando el caso de Eva, podemos calcular la probabilidad pos-test sin realizar la tabla de 2x2, utilizando la probabilidad pre-test y el coeficiente de probabilidad negativo de la ecografía TV de la siguiente forma:

Probabilidad pre Test Ca. Endometrio = 5%

Si lo pasamos a Razón de Probabilidad =  $5 / (100-5) = 5 / 95 = 0,05$

Ahora lo tenemos que multiplicar por el coeficiente de probabilidad positivo para obtener la probabilidad pos test . El coeficiente de probabilidad positivo es igual a sensibilidad / 1- especificidad, por lo que CPP =  $90 / 50 = 1,8$   
Entonces :

$$\begin{aligned} \text{Razón prob. pos test} &= \text{Razón prob. pre test} \times \text{CPP} \\ &= 0,05 \times 1,8 \\ &= 0,09 \end{aligned}$$

Ahora debemos pasar la razón de prob. a porcentaje para poder interpretarla :

$$\begin{aligned} \text{Prob en porcentaje} &= \text{Razón prob.} / \text{Razón prob} + 1 \\ &= 0,09 / 1,09 \\ &= 0,08 \end{aligned}$$

Lo que implica que el valor predictivo positivo es igual a 8% (muy bajo)

Hacemos lo mismo para el valor predictivo negativo que es el que mas interesa, ya que estamos tratando de descartar la posibilidad de Ca. Endometrio mediante la Ecografía. El coeficiente de probabilidad negativo es igual a  $(1-\text{sens}) / \text{esp}$ .

$$\begin{aligned} \text{CPN} &= (1-\text{sens}) / \text{esp} \\ &= 10 / 50 \\ &= 0,5 \end{aligned}$$

$$\begin{aligned} \text{Razón prob. pos test} &= \text{Razón prob. pre test} \times \text{CPN} \\ &= 0,05 \times 0,5 \\ &= 0,025 \end{aligned}$$

Ahora al igual que antes, tenemos que pasar esta probabilidad pos. test a un porcentaje para interpretarla

$$\begin{aligned} \text{Prevalencia pos test} &= \text{Razón prob.} / 1 + \text{Razón prob.} \\ &= 0,025 / 1,025 \\ &= 0,02 \end{aligned}$$

Esto lo interpretamos como que estamos un 98% seguros de que Eva no tiene Ca. Endometrio de resultar negativa la Ecografía TV.

Hay nomogramas, que nos permiten agilizar este cálculo dentro del consultorio sin necesidad de realizar tantas operaciones matemáticas.

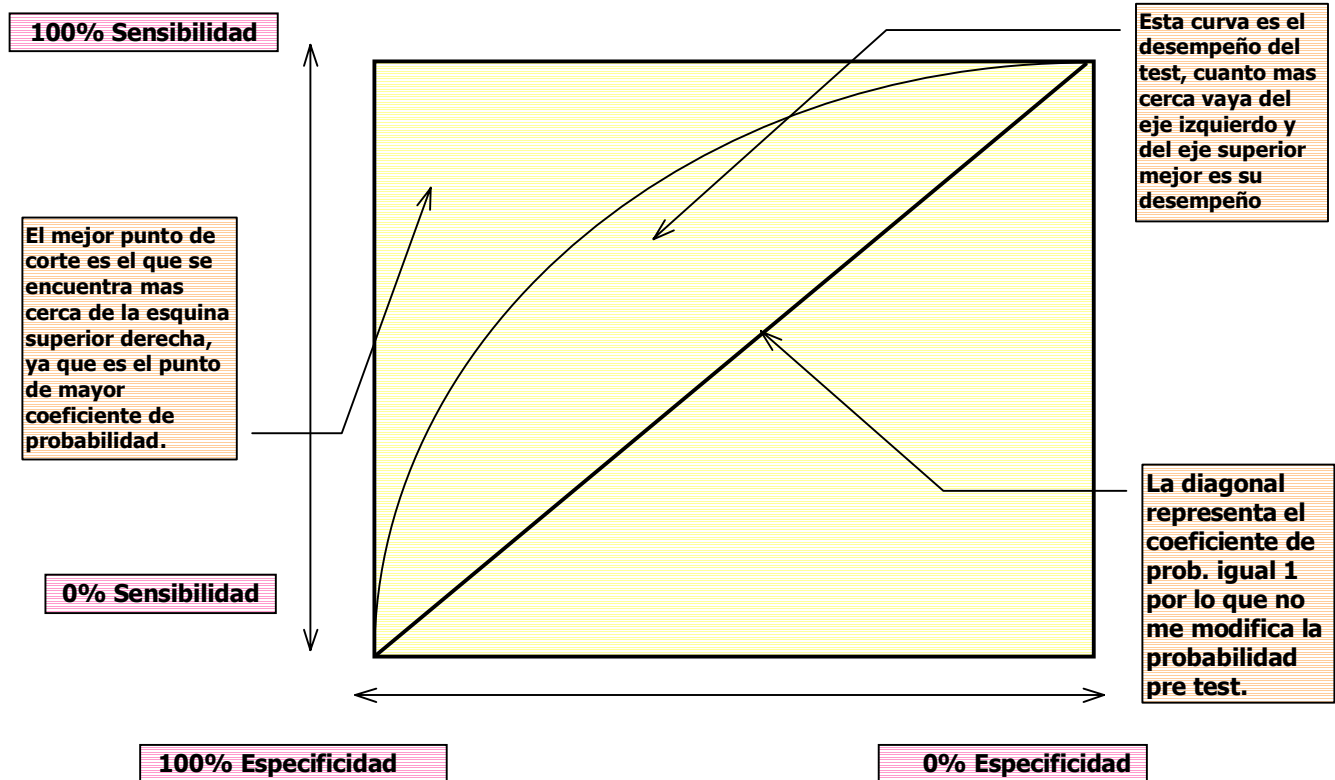
### **Curvas ROC**

Las curvas ROC, obtienen su nombre de la sigla en inglés Receiver Operator Characteristic, y surgieron durante la segunda guerra mundial en Inglaterra en un área bastante alejada de la medicina, ya que tenían relación con la capacidad de un operador de radar de predecir un ataque aéreo.

En las curvas ROC se grafica, la sensibilidad en el eje Y y la tasa de falsos positivos en el eje X, vemos que estamos utilizando los dos componentes del coeficiente de probabilidad positivo uno en cada eje.

De esta forma, podemos determinar el punto de corte, que obtiene la mejor relación sensibilidad/tasa de falsos positivos, o sea el mejor coeficiente de probabilidad. También se utilizan para evaluar el desempeño de un test y para comparar dos test diagnósticos.

No es nuestra intención ahondar demasiado en las curvas ROC, simplemente es gráfico ver como utilizando los dos componentes del coeficiente de probabilidad se puede construir una curva que nos muestra el desempeño de un test y nos permite evaluarlo.



## Tratamiento

El objetivo del tratamiento, es prevenir la ocurrencia de algo, ya sea una enfermedad en prevención primaria o la progresión de la misma tanto en prevención secundaria como terciaria.

Por lo que analizado de esta forma, podemos entender que el objetivo del tratamiento es siempre disminuir el riesgo de algo.

Esta forma de entender el tratamiento nos es muy útil como enfoque de la evaluación del mismo por medio de tablas de 2x2.

Como analizamos al comienzo del capítulo, en los trabajos tendientes a la evaluación de un tratamiento, se compara un grupo intervención (el cual recibió algún tipo de tratamiento) contra un grupo control (que puede haber recibido un placebo u otro tratamiento) para determinar si el tratamiento evaluado es mejor que no hacer nada o que el otro tratamiento. La evaluación se realiza mediante la medición de la ocurrencia de lo que nosotros queremos evitar (evento), ya sea la aparición de una enfermedad o algún evento que nos indique la progresión de la misma. A su vez este dato, lo debemos

expresar de forma tal que nos facilite la comparación y que nos permita tomar decisiones dentro del consultorio.

El diseño que mayor potencia tiene al momento de evaluar la eficacia de un tratamiento es el estudio clínico randomizado controlado contra placebo, ya que disminuye al máximo la posibilidad de factores confundidores.

De la tabla de 2x2, se derivan los conceptos que nos van a permitir evaluar la disminución o el aumento del riesgo de padecer un evento.

	<b>Evento Presente</b>	<b>Evento ausente</b>
<b>Grupo Tratamiento</b>	<b>a</b>	<b>b</b>
<b>Grupo Control</b>	<b>c</b>	<b>d</b>

- Riesgo absoluto Grupo Tratamiento =  $a / a+b$
- Riesgo absoluto Grupo Control =  $c / c+d$
- Riesgo Relativo = Riesgo abs. tto. / Riesgo abs. Control
- Reducción Riesgo Relativo =  $1 - \text{Riesgo Relativo}$
- Reducción Riesgo absoluto = Riesgo abs. Control - Riesgo abs. Tto.
- Numero Necesario para Tratar =  $1 / \text{Reducción Riesgo Absoluto}$
- Odds Ratio =  $(a \times d) / (b \times c)$

Para facilitar la comprensión de los conceptos, tomemos un ejemplo real, con respecto al tratamiento de la hipertensión arterial.

El riesgo de padecer un evento cardiovascular a 10 años de una persona hipertensa de 70 años, tabaquista, diabético e hipercolesterolémico es de 45%.

Esto quiere decir que de cada 100 personas como esta, casi la mitad en 10 años, va a haber padecido o un ACV o un evento coronario.

Analizemos una tabla, en la cual comparamos grupo no tratado vs. Grupo tratado.

	Evento presente	Evento ausente
Grupo Tratamiento	30	70
Grupo Control	45	55

Vemos como en el grupo control cada 100, 45 sufren un evento y en el grupo tratamiento solo 30 lo padecen.

### Riesgo absoluto

El riesgo absoluto, es la probabilidad de ocurrencia del evento a medir haciendo un análisis en forma independiente para cada grupo, de esta forma, tenemos un riesgo absoluto para el grupo intervención y un riesgo absoluto para el grupo control.

Lo obtenemos, dividiendo la cantidad de personas del grupo que sufrieron el evento sobre la cantidad de integrantes de dicho grupo.

A mayor riesgo absoluto, mayor es la probabilidad de ocurrencia del evento en el grupo en estudio.

$$\begin{aligned} \text{En nuestro ejemplo RA para grupo tratamiento} &= 30 / 30+70 = 0,30 \\ \text{y RA para el grupo control} &= 45 / 45+55 = 0,45 \end{aligned}$$

### Riesgo Relativo

El riesgo relativo es una medida de riesgo que relaciona los dos riesgos absolutos de cada grupo, ya que se obtiene dividiendo el riesgo absoluto del grupo intervención sobre el riesgo absoluto del grupo control. De esta forma cuando el riesgo relativo es igual a 1, significa que los riesgos absolutos del grupo intervención y control son iguales, o lo que es lo mismo decir que los dos tratamientos son igual de efectivos.

Por el contrario, cuando el RR es menor a 1, significa que el grupo intervención tiene menor riesgo de padecer el evento que el grupo control, por lo que el tratamiento es mas efectivo que el tratamiento empleado en el grupo control, lo inverso sucede cuando el RR es mayor a 1.

$$\text{En nuestro ejemplo el RR es} = 0,30 / 0,45 = 0,66$$

### Reducción del Riesgo Relativo

La reducción del riesgo relativo nos permite cuantificar cuanto disminuyó (o aumentó cuando el tratamiento es perjudicial) la probabilidad de ocurrencia del evento a medir expresada en porcentaje del riesgo. Se calcula restando a 1 el RR ( $RRR=1-RR$ ).

Esto significa que una reducción del riesgo relativo del 50% indica una disminución a la mitad de la probabilidad de ocurrencia del evento a medir.

En nuestro ejemplo  $RRR = 1 - 0,66 = 0,33$

Una inconveniencia de la reducción del riesgo relativo es que no discrimina grandes variaciones del riesgo de variaciones menores que mantienen la misma proporción.

Por ejemplo si la probabilidad de un evento cae del 80% al 40% la reducción del riesgo relativo será del 50%, de la misma forma que si cae del 4 al 2% también la reducción será del 50%. Esto no nos permite cuantificar en valor absoluto la disminución del riesgo y nos puede confundir llevándonos a pensar que dos tratamientos son similares en cuanto a su eficacia, cuando en realidad la disminución del riesgo en su valor absoluto es muy diferente. Para esto es que se utiliza la reducción del riesgo absoluto.

### **Reducción del Riesgo Absoluto**

La reducción del riesgo absoluto, la obtenemos restando al riesgo absoluto del grupo control, el riesgo absoluto del grupo tratamiento ( $RRA = RA \text{ grupo control} - RA \text{ grupo tratamiento}$ ). De esta forma, tomando nuevamente el ejemplo anterior con respecto al RR, vemos como cuando el riesgo absoluto disminuye de 80 a 40, el RRA es de 40; y cuando el riesgo cae de 4 a 2, el RRA es de 2, mientras que se mantiene la misma reducción del 50% en el riesgo relativo en los dos ejemplos.

En nuestro ejemplo el RRA es  $= 0,45 - 0,30 = 0,15$

El RRA nos permite cuantificar la disminución del riesgo en valores absolutos y de esta forma tener una idea de la cantidad de disminución del riesgo, no del porcentaje como nos muestra el RRR.

### **Número Necesario para Tratar (NNT)**

El NNT, es una medida muy útil en su aplicación clínica, ya que me indica a cuántos pacientes tengo que tratar para que un paciente no padezca el evento que queremos evitar.

El NNT es la inversa del RRA, es decir lo obtenemos dividiendo 1 por el RRA ( $NNT = 1/RRA$ ), esto significa que cuanto mayor es el RRA, menor es el NNT.

Vemos que cuanto mayor es la cantidad de riesgo disminuida por un tratamiento es decir cuanto mayor es el RRA, menor es la cantidad de pacientes que tengo que tratar para que una persona no sufra el evento a evitar.

En nuestro ejemplo el NNT es  $= 1 / 0,15 = 7$

De esta forma podemos comparar dos tratamientos, viendo a cuantos pacientes es necesario tratar para disminuir un evento y cruzar esta información con el costo del tratamiento o con los efectos adversos del mismo, esto nos permite evaluar cuanto cuesta disminuir en 1 la cantidad de eventos, o que probabilidad tengo de ocurrencia de efectos adversos en esa cantidad de tratamientos necesarios para evitar la ocurrencia de un evento.

Para evaluar factores de riesgo se utilizan estudios de cohorte en los cuales se construyen tablas de 2x2, de la misma forma, con la diferencia que al no haber aleatorización, existe mayor posibilidad de factores confundidores, pero la evaluación de la probabilidad de ocurrencia de un evento se realiza de la misma forma.

En los trabajos de cohortes se utiliza el RR a diferencia de los trabajos casos control en los cuales se utiliza el Odds Ratio.

El Odds Ratio, que también se conoce como razón de producto cruzado debido a que se dividen los productos de los extremos de la tabla de 2x2, tiene la misma significación que el RR, pero en trabajos casos control, por lo que tiene menor potencia debido a limitaciones del diseño casos control (retrospectivo) con respecto al diseño de cohortes que es prospectivo. Es decir, significan lo mismo pero cuando existe alguna diferencia entre un RR o un OR que miden lo mismo, el riesgo relativo es mas preciso.

Resumiendo, en la tabla de 2x2, lo que hacemos es comparar dos tipos de tratamiento o factores de riesgo. Mediante los indicadores de riesgo podemos evaluar el desempeño de cada uno de los tratamientos o el peso de los factores de riesgo tanto por separado como en forma comparativa. El RR como la RRR, nos hablan de la variación del riesgo en términos de porcentaje de aumento o disminución, a esto es necesario agregarle el RRA y el NNT cuando nos referimos a tratamientos, ya que nos permiten cuantificar esta variación del riesgo en valores absolutos lo que nos permite tomar decisiones en forma mas precisa.

#### **BIBLIOGRAFÍA RECOMENDADA:**

- .. **Barker. Principles of ambulatory medicine. Ed 2000**
- .. **Sackett D. L. Medicina Basada en la Evidencia. Ed 2000**
- .. **Sackett D. L. Epidemiología Clínica. Ed. 1991**
- .. **Dawson-Saunders. Bioestadística Médica. Ed1994**
- .. **<http://www.fisterra.com/material/investiga/indice.htm>**