



# 17

## Fases del análisis estadístico de los datos de un estudio

---

Antonia Guillén Serra  
Ágata Carreño Serra  
Neus Canal Díaz

### 17.1. Introducción

La finalidad de este capítulo es repasar de forma práctica y didáctica los contenidos que hemos ido viendo a lo largo de este manual, tratando de describir de forma esquematizada cómo realizar un análisis estadístico y dar unas pautas de utilización de los métodos y distintas pruebas utilizadas para medir las variables de un estudio, y así poder contestar a los objetivos que se plantean en el protocolo del mismo.

Es un hecho que la disponibilidad de paquetes informáticos ha contribuido a la extensión del uso de la estadística por profesionales poco expertos en esta materia y a la utilización en ocasiones abusiva de técnicas sofisticadas y novedosas. Sin embargo, no son estas pruebas las que los autores de este manual aconsejamos para los usuarios, sino las más utilizadas y reconocidas, y ante todo las de más fácil comprensión teniendo la certeza de que su utilización no afectará ni la seriedad ni el prestigio del estudio que vayamos a realizar. Debe quedar claro que el esquema de análisis dado en este capítulo es flexible y que deberá adap-



tarse a las características de los diferentes estudios en concreto a realizar, pero nos servirá como referencia ante cualquier análisis estadístico.

## 17.2. Cómo asegurar la calidad de los datos

Es evidente que para que los resultados de un estudio se adecuen a la realidad, los datos deben ser correctos y deben cumplir con unos estándares de calidad. Es por ello que no tener en cuenta una mínima calidad de los datos puede conducirnos a concluir nuestro estudio de forma errónea.

### 17.2.1. Gestión de datos

La gestión de datos comprende diferentes pasos descritos ampliamente en el capítulo 3 de este manual y que recordaremos brevemente, remarcando los consejos más prácticos de esta fase. En primer lugar, se realizará una base de datos con entrada de datos validada que contenga sus filtros y rangos que impedirán que los valores inconsistentes sean grabados. Esta base de datos almacenará los datos que contiene el cuaderno de recogida de datos (CRD) reflejando toda la información aportada por los investigadores para su posterior análisis estadístico. Es importante que el diseño de la base de datos siga el mismo orden que el CRD. Nos deberemos asegurar de que los datos transcritos en el CRD se pasan de forma correcta a la base de datos. Los datos que no cumplan los requisitos serán revisados y corregidos, pudiendo aparecer la necesidad de emitir una “*query*”, que puede ser identificada tanto durante el proceso de entrada de datos como durante la codificación y revisión de los datos introducidos. La respuesta será actualizada en la base de datos cuando se disponga de dicha información.

Los estudios realizados en nuestro campo sanitario (IRC), en donde sus resultados pueden ser aplicados a otros grupos de pacientes afectados de esta patología, nos hace insistir en la importancia de que se registre y documente cualquier cambio realizado en la base de datos para seguir los principios de trazabilidad, imprescindible en esta fase donde intervienen diversas personas implicadas en el estudio como son el investigador, el monitor, el responsable del manejo de los datos y el estadístico, siendo un proceso siempre auditable por la responsabilidad que conlleva.

En cuanto a la codificación de los apartados clínicos del CRD, nuestro consejo es que tanto las enfermedades, como la medicación y acontecimientos adversos sean codificadas con diccionarios específicos para tal fin, asegurando así la homogeneidad y veracidad de la codificación.



### 17.2.2. Revisión de los datos

Existen una serie de procedimientos a realizar en los datos ya introducidos para valorar si contienen valores incorrectos, que nos pueden conducir a conclusiones equivocadas. Para llevar a cabo este proceso utilizaremos las técnicas que componen la estadística descriptiva revisando las variables una por una en búsqueda de valores ilógicos en las mismas o con errores de transcripción. Por ejemplo, una tabla de frecuencias nos permite ver anomalías en la introducción de los datos de los criterios de inclusión de un estudio, en que (1) corresponderá a 'Sí' que los cumple y (2) a 'No' los cumple. Cualquier otro valor implica un error de introducción. Nuestra actuación deberá ser, si es posible, consultar el CRD o la fuente real de información para obtener el verdadero valor, y corregirlo. En las tablas de distribución de cada variable nos fijaremos que sus totales sean iguales al número de pacientes incluidos en el estudio. Un número elevado de datos faltantes puede invalidar nuestros resultados.

Otra posibilidad que nos ofrece esta tabla de frecuencias es valorar si la distribución de determinadas variables como el sexo o el tratamiento, es la que se esperaba. Como por ejemplo, el número de pacientes a los que se les administra un determinado tratamiento (HD/DP/TX) o la proporción de hombres y mujeres que se esperaba en la muestra del estudio, etc. Si la distribución de estas variables no es la esperada puede indicarnos un error sistemático, por lo que deberemos revisar el proceso y corregir los posibles errores.

Otro problema frecuente que detectamos con la tabla de frecuencias y que nos solemos encontrar cuando estamos revisando los datos y preparándolos para el análisis es la necesidad de hacer ciertas agrupaciones o dividir en diferentes categorías algunas variables. Por ejemplo, en el caso de la variable 'estado civil' que se ha agrupado en diferentes categorías soltero/casado/viudo/separado/divorciado nos puede suceder que en alguna de estas categorías haya un número reducido de casos, que nos dificultará la distribución de la variable al no tener el número de casos superior a 30, y por tanto la interpretación y utilización de las técnicas estadísticas que apliquemos. En este caso la solución pasaría por recodificar estos valores en menos grupos, como por ejemplo solteros/casados/otros.

También nos podemos encontrar que una determinada categoría de una variable tenga un gran número de casos y necesitemos dividirla en categorías más pequeñas, como por ejemplo, la situación laboral donde tenemos las categorías activo/parado/otros y esta última contiene muchas opciones como son jubilados, pensionistas, amas de casa y estudiantes. Sin embargo, hemos de tener presente que este proceso siempre se podrá hacer si están especificadas en el cuestionario las categorías a las que corresponden la opción de respuesta 'otros'.



En las variables cuantitativas, si la desviación estándar es muy elevada en relación a la media, nos indica que existe dispersión de los valores o una falta de normalidad en la distribución de los datos, por lo que revisaremos también la distribución de las frecuencias para verificar este hecho, y en el caso de ser verdad, poder utilizar los estadísticos más adecuados.

Otro dato a tener en cuenta es la distribución en percentiles que nos indicarán si existen valores extremadamente altos o bajos que pueden influir sobre la media y afectar la distribución de la variable, por lo que deberemos tenerlo en cuenta en el cálculo de los estadísticos a aplicar. Valores en principio imposibles, como un peso de 12 Kg. para un recién nacido, o un Hto. de 70% ó de 11%, pueden requerir una confirmación de la veracidad del dato.

Existen una serie de chequeos que se realizan en todos los estudios previos al análisis y donde las frecuencias son imprescindibles y nos ayudarán a detectar valores erróneos o imposibles o, en el caso de que comprobemos su veracidad nos obligarán a eliminar a estos pacientes del estudio, como es el caso del cumplimiento y verificación de los criterios de inclusión en los sujetos del estudio. Por ejemplo, el tratamiento administrado a los mismos no puede ser otro que HD, o la edad de los pacientes del estudio debe oscilar entre los 18 y 65 años.

Otra tarea frecuente en esta fase es la recodificación de alguna de las variables cuantitativas en cualitativas mediante la creación de intervalos y donde la distribución de las frecuencias puede ser imprescindible. Por ejemplo establecer diferentes grupos de dosis de eritropoyetina o rangos de TAS/TAD. Generalmente la codificación de las variables cuantitativas a categóricas obedece a dos criterios, el de homogeneidad de los grupos en cuanto a número de pacientes y el de obtener un significado clínico.

También revisaremos la consistencia de diversas distribuciones comparando dos variables (Tabla de contingencia) buscando esos errores improbables para corregirlos, como es el caso de pacientes repetidos o bien una enfermedad de próstata en una mujer.

Estos son algunos de los procedimientos que nos ayudarán a corregir y eliminar gran parte de los errores de la base de datos. Una vez se considere que este proceso ha finalizado y que la base de datos está validada se realizará el cierre de la misma, asegurándonos de que no sufrirá ninguna manipulación posterior y se iniciará el análisis estadístico.

### **17.3. Definición del análisis estadístico**

El análisis estadístico es un apartado dentro del protocolo cuya finalidad es exponer la estrategia de análisis a seguir. Contemplará la definición de cada una



de las variables imprescindibles y secundarias del estudio y nos informará de la técnica utilizada para la medición y análisis para asegurar la respuesta a la pregunta planteada en el estudio.

### **17.3.1. Plan de análisis estadístico**

En el plan de análisis se expone y detalla, de acuerdo a los objetivos propuestos, las medidas de resumen de las variables y cómo serán presentadas (cuantitativas y/o cualitativas), indicando los modelos y técnicas analíticas que se utilizarán para comprobar las hipótesis del estudio. Se recomienda que éste sea previo al análisis para no decantarse por opciones que puedan verse alteradas por los primeros resultados obtenidos.

En el plan de análisis también se describe brevemente el software que será utilizado para el análisis de datos y se establece el nivel de significación estadística para todas las pruebas estadísticas realizadas con las variables de resultados que por lo general suele ser de 0,05. Se nombra en él, la utilización de técnicas estadísticas preliminares para asegurar el cumplimiento de los supuestos estadísticos. En el caso de que no se cumplan los supuestos establecidos se utilizarán pruebas equivalentes que no presenten dichas limitaciones, como por ejemplo pruebas no paramétricas.

Todo el análisis va a ir dirigido a responder preguntas concretas y sencillas, generalmente establecidas en el protocolo previo, dejando claramente definidas las variables principales que se van a utilizar como respuesta, tanto por su significado clínico como por la forma de cálculo que se va emplear, si lo hubiera (por ejemplo, recodificaciones o cálculos específicos).

### **17.4. Consideraciones prácticas de la selección de la prueba estadística adecuada**

Decidir qué prueba estadística es la adecuada para nuestros datos no es un proceso mecánico. Depende de cada situación en particular y de los conocimientos del investigador para justificar el porqué de su decisión. Es muy importante que el investigador se sienta cómodo con la prueba que utilizará, ya que su capacidad para interpretar de forma adecuada el resultado obtenido dependerá de la correcta comprensión de la prueba. En los distintos capítulos de que consta este manual se han ido referenciando las distintas pruebas estadísticas a utilizar según las características de la variable, además de dirigirnos al anexo II donde se expone de forma esquemática una clasificación de las pruebas estadísticas más apropiadas para datos apareados e independientes. Hay que tener en cuen-



ta que este esquema nos orientará pero no es el único ni definitivo. A continuación nombraremos algunas normas que nos pueden facilitar la elección de la prueba estadística más adecuada para nuestros datos.

Para elegir la prueba adecuada debemos tener claras las características de las variables a analizar. Existen las pruebas paramétricas y las pruebas no paramétricas, en función de la distribución de las variables y pruebas para datos independientes o apareados.

Las pruebas que deberemos utilizar serán distintas según las variables a analizar sean cualitativas o cuantitativas. Una variable cuantitativa presenta una mayor información que una variable cualitativa, aunque esta última puede llegar a ser mucho más fácil de interpretar.

Un aspecto que debemos tener claro es si nuestros datos se tratan de datos apareados o independientes. Los datos medidos sobre el mismo individuo serán siempre apareados. Este es el caso, por ejemplo, de la comparación de las cifras de hematocrito de los individuos de una muestra en dos momentos del tiempo, antes y después de una intervención. Para esta situación existen pruebas más potentes, dado que la variabilidad en las pruebas paramétricas de estas mediciones es menor. Si no existe apareamiento, se habla de datos independientes.

Existen unas condiciones de aplicación para ciertas pruebas de análisis estadístico en datos cuantitativos entre las que es importante la distribución de la variable. La mayoría de estas pruebas son robustas y toleran relativas violaciones de estas asunciones sobre su distribución, sobretodo si el número de sujetos estudiado es superior a 30. En caso de que el número sea inferior, deben emplearse otras pruebas, habitualmente de cálculo más complejo y menos potentes, llamadas pruebas no paramétricas.

Cuando deseamos evaluar la posible relación entre variables cualitativas, se utiliza la prueba de la Ji-Cuadrado (capítulo 11) y dependiendo del tipo de estudio se suele utilizar medidas de asociación tales como el Odds-ratio o OR y el riesgo relativo o RR (capítulo 15). Un ejemplo sería analizar en HD, la relación entre grupos de gravedad y las diferentes técnicas de educación sanitaria de dieta.

Cuando comparamos dos grupos respecto a una variable cuantitativa, como por ejemplo el valor del Hto. en dos grupos de pacientes en HD, la técnica a aplicar es la t de Student. En el caso de no cumplirse las condiciones necesarias para su aplicación debe recurrirse a la prueba no paramétrica equivalente, la U de Mann-Whitney (capítulo 12).

Si lo que se desea es comparar una variable cuantitativa en más de dos grupos, por ejemplo, la comparación de la tensión arterial diastólica en tres grupos de pacientes a los que se aplican técnicas sustitutivas distintas HD/DP/TX, debe



realizarse un análisis de varianza (capítulo 12). En caso de que se demuestre la existencia de diferencias significativas en la distribución de los tres grupos, se justifica su comparación posterior para detectar cuál de ellos es el que difiere, utilizando técnicas de comparación múltiple específicamente diseñadas, como son Bonferroni, Tukey o Scheffé. La prueba equivalente no paramétrica del análisis de la varianza es la prueba de Kruskal-Wallis. Esta prueba nos dice si existen diferencias en varios grupos, pero no cual de ellas es la que se difiere, por lo que si nuestro objetivo así lo requiere, hay que realizar comparaciones “dos a dos” con el test no paramétrico, según se trate de datos independientes (Mann-Whitney) o de datos apareados (Wilcoxon). Aunque con la multiplicidad de pruebas deberá ajustarse el nivel de significación.

Si lo que nos interesa es determinar la asociación entre dos variables cuantitativas, por ejemplo, la edad y el colesterol, la prueba adecuada es la correlación de Pearson o, si no se cumplen las condiciones de aplicación, la correlación de Spearman (capítulo 13). 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 (capítulo 14).

#### **Aclaraciones sobre cual es el verdadero valor de la p**

Creemos importante aclarar de nuevo cual es el interés de la p, tantas veces referenciada en los resultados estadísticos de los estudios, y que nos puede llevar a confusión. La expresión resultado estadísticamente significativo podría ser sinónimo de “es improbable que el resultado observado sea debido al azar o rechazo de la hipótesis nula”. O sea, el valor p se refiere a que la diferencia observada en los datos no es fruto de la casualidad.

Tampoco se puede interpretar como una medida de fuerza de la asociación, así pues si nos encontramos con una  $p < 0,001$  no quiere decir que la asociación encontrada sea más fuerte o la diferencia más importante que en otro valor en que la  $p > 0,05$ , sólo nos está diciendo que el resultado obtenido tiene menos probabilidad de que sea debido al azar o causalidad. Por coherencia tampoco deberemos ser muy rígidos en el límite del nivel de significación por ejemplo una p de 0.049 y una p de 0,051 no son lo mismo ya que la primera es estadísticamente significativa al nivel del 5% y la segunda no. Sin embargo la probabilidad de observar el resultado por azar es prácticamente la misma, por ello se recomienda informar del valor obtenido para la p, sobre todo si es muy próximo al nivel de significación fijado, y así poder valorar de forma adecuada los resultados que nos están dando.



## 17.5. Población para el análisis

Todo estudio que se precie debe contener una serie de consideraciones que deben estar justificadas en muchos ámbitos: En primer lugar, deberemos indicar la fuente de la que hemos extraído la muestra (ámbito hospitalario, atención primaria, servicio de nefrología), así como si ha existido algún mecanismo de selección de sujetos (aleatorización, por bloques, muestreo sistemático). Posteriormente, deberemos describir la muestra de pacientes que estamos analizando, teniendo en cuenta las variables recogidas durante el proceso.

Debemos tener en cuenta variables de tipo sociodemográfico tales como la edad, el género, los estudios finalizados, la situación laboral, así como también datos antropomórficos como peso, talla e IMC. De cara a homogeneizar la muestra de pacientes, generalmente se seleccionan aquellos que cumplen ciertos criterios (criterios de inclusión y exclusión). De esta forma, se conocerá la población diana, es decir, la población de la que se quiere extraer resultados y se podrá interpretar de forma más correcta los resultados. Por otra parte, estas características podrán ser tenidas en cuenta para evaluar su influencia en los resultados, evaluando la posibilidad de extrapolar los resultados y permitir la replicabilidad del estudio por otros investigadores.

Seleccionar los pacientes de una muestra también obedece a criterios propios de cada estudio, por lo que es posible que haya más de una población. Generalmente, existe la población formada por todos los pacientes incluidos en el estudio y que hayan cumplido los criterios de inclusión; pero también se analiza el grupo de pacientes que, además de haber cumplido los criterios de selección, han permanecido en el estudio durante por lo menos 6 meses, por ejemplo.

Finalmente, en cualquier estudio debe quedar clara la enfermedad que estamos evaluando, el tiempo de evolución, la gravedad y los criterios diagnósticos utilizados.

## 17.6. Valoración del análisis de grupos de estudio comparables

En cuanto a las características de la muestra, también es importante analizar la situación clínica al inicio del estudio para evaluar la homogeneidad basal entre los grupos de estudio y para dar una idea de la situación de partida de estos pacientes. Para que los resultados de nuestro estudio sean válidos y no tengan sesgos que alteren nuestros resultados, antes de efectuarse cualquier comparación en la variable respuesta, deberemos asegurarnos de que los dos (o más) grupos de estudio son comparables inicialmente en relación a variables que puedan influir en la respuesta. Por ejemplo, si en un estudio que pretende comparar la eficacia de dos





fármacos en los que las mujeres siempre presentan una mayor eficacia, si obtuviéramos diferencias en la distribución de mujeres en ambos grupos, podríamos estar delante de un problema a la hora de realizar el análisis puesto que el género podría interferir en la interpretación de los resultados (posible factor de confusión). Es normal que si disponemos de muchas variables a analizar entre grupos de estudio, alguna resulte significativa; es por ello que siempre deberá tenerse en cuenta cuando clínicamente esté justificada su presencia.

### 17.7. Análisis de resultados

Cuando estimamos un parámetro, por ejemplo, disminución de TAS o porcentaje de pacientes diabéticos que presentan un IMC superior a 35 Kg./m<sup>2</sup>, sólo tenemos un grupo de pacientes y podemos estimarlo a partir de la media y porcentaje, con sus respectivos intervalos de confianza. Así damos el valor observado muestral y el rango de valores entre los que se incluye su valor.

En el caso de que comparemos dos grupos de tratamiento, por ejemplo, debemos dejar bien claro la prueba que hemos aplicado, así como también la prueba estadística que hemos considerado. Para el análisis principal, siempre debe indicarse la respuesta para cada uno de los grupos y el nivel de significación (p-valor). Estos datos crudos sólo hacen referencia a la probabilidad de que el azar haya podido explicar los resultados, pero no dan una magnitud del efecto de un tratamiento. Por lo tanto, si nuestra respuesta es categórica debemos estimar el intervalo de confianza asociado de la diferencia de proporciones, el OR o el RR y si es continua, estimaremos el intervalo de confianza de la diferencia de medias y, si se cree necesario, el intervalo de confianza de cada una de las medias de los grupos.

En nuestro ejemplo hipotético aparecen dos tratamientos de diálisis: diálisis peritoneal (Terapia A) y hemodiálisis (Terapia B). En ambas se mide el porcentaje de pacientes con acontecimientos adversos (AA) o síntomas agudos. Nuestra respuesta será la siguiente:

	Terapia A	Terapia B	P	IC 95%
Presencia de AAs	16/140 (11,4%)	29/142 (20,4%)	0,039	0,5% - 17,5%
Presencia de AAs graves	11/140 (7,9%)	24/142 (9,9%)	0,554	-4,6% - 8,6%

**Tabla 23.** Comparación de la incidencia de acontecimientos adversos de dos terapias



En esta tabla se indica que existen diferencias estadísticamente significativas en el porcentaje de pacientes con AAs entre las dos terapias. En la terapia A presentan AAs el 11,4% de los pacientes, siendo del 20,4% en los pacientes tratados con la terapia B ( $p=0,039$ , es decir,  $p<0,05$ ). Sin embargo, si se consideran los AAs graves, esta diferencia no es estadísticamente significativa ( $p=0,554$ ). En el primer caso, el intervalo de confianza de la diferencia de proporciones no incluye el 0 absoluto, aunque por poco, y en el segundo caso se incluye en un intervalo muy amplio.

## 17.8. Análisis estratificado

Muchas veces, parece indicado realizar un análisis estratificado de nuestros datos, principalmente de cara a responder preguntas secundarias que se nos plantean importantes; pero en muchas ocasiones también puede ser realizado para evaluar la existencia de un factor de confusión que puede influenciar en los resultados así como para evaluar el cambio del efecto producido en nuestra respuesta dependiendo de valores de otras variables que actúan como modificadoras del efecto.

Primeramente, debemos asegurarnos que existe un número de sujetos aceptable en cada estrato, aunque siempre que sea coherente puede realizarse alguna agrupación con sentido clínico. Las variables para las que se puede estratificar deben ser escogidas cuidadosamente, realizando una selección apropiada y no realizándolo con todas las variables disponibles. Cuando se tengan definidos los estratos, podremos realizar pruebas estadísticas sencillas, de acuerdo a nuestros objetivos y tipo de variables a estudiar.

Generalmente deberemos tener los resultados por estratos (por ejemplo fumadores y no fumadores) y por el global de la población, por lo que a partir de los análisis sencillos, podremos controlar factores de confusión y variables modificadoras del efecto.

### 17.8.1. Factores de confusión

Los factores de confusión son variables que interfieren en nuestra variable principal, pudiendo afectar a los resultados que, con tanto esfuerzo se han obtenido durante el análisis. Siempre que podamos debe prevenirse en la fase de diseño, aunque no siempre es posible. Podremos identificar un factor de confusión cuando este factor aparece relacionado con nuestra variable de respuesta y, a la vez, está asociado con otras variables del estudio que se relacionan con nuestro factor, sin ser una variable intermedia entre estas relaciones. Cuando



tenemos en cuenta un factor de confusión, la relación entre el factor y el efecto que estamos evaluando se modifica de la misma forma en todos los grupos.

Por ejemplo, si estamos estudiando la relación entre el consumo de café y el infarto agudo de miocardio (IAM), observaremos que el consumo de café aumenta el riesgo de padecer un IAM. Antes de llegar a esta conclusión, intentaremos comprobar que esta relación se ha debido sólo al efecto del café y no de otras variables, como por ejemplo, el tabaco. Si el hecho de ser fumador se correlaciona con un mayor consumo de café y el hecho de fumar es un factor de riesgo del IAM, estamos hablando de un factor de confusión. La tabla 24 ilustra el ejemplo:

	Consumo de Café	IAM	No IAM	OR
Total	Sí	96	54	3,16
	No	54	96	
	Total	150	150	
No fumadores	Sí	9	19	0,93
	No	41	81	
	Total	50	100	
Fumadores	Sí	81	40	1,07
	No	19	10	
	Total	100	50	

**Tabla 24.** Control de factores de confusión. Incidencia de IAM según el consumo de café estratificado por hábito tabáquico

Los resultados totales muestran una asociación positiva entre el consumo de café y el IAM ( $OR=3.16$ ;  $OR>1$ ), de manera que aparentemente el riesgo de sufrir IAM se ve incrementado por el consumo de café. Sin embargo, esta relación puede deberse a que el consumo de café está relacionado con un mayor número de fumadores. Sospechando de esta relación y, estratificando los resultados por hábito tabáquico, nos encontramos que la relación para cada grupo o estrato desaparece con OR cercanos a 1. Por lo que concluimos que el tabaquismo estaba actuando como factor de confusión, produciendo una aparente relación entre el consumo de café y el IAM.

Puesto que no siempre el análisis “a posteriori” va a ser lo más adecuado, como hemos dicho anteriormente, debemos evaluar durante la fase de diseño si un determinado factor puede enmascarar los verdaderos resultados de nuestro estudio. Su interferencia puede evitarse mediante la recogida de datos, restrin-



giendo la selección de pacientes o bien, realizando un análisis estratificado desde el inicio del estudio.

### 17.8.2. Modificación del efecto

Hemos visto que a pesar de encontrar relaciones entre las variables de nuestro estudio, éstas pueden estar enmascaradas por otras relaciones mucho más fuertes y que actúan como factores de confusión. Sin embargo también puede darse el caso de que una tercera variable influya en el efecto que estamos estimando. Estas variables se identifican si al estratificar por ellas los resultados de nuestras tablas, se obtienen resultados diferentes para cada estrato y en general.

Por ejemplo, imaginemos que queremos comparar la satisfacción que produce en los pacientes la toma de dos fármacos, uno administrado semanalmente (Fármaco A) y otro a diario (Fármaco B). Se observa que, si tenemos en cuenta el total de la tabla, no observamos diferencias en la satisfacción entre los dos fármacos (45,5% vs. 45,4%). Sin embargo, cuando estratificamos por género, sí se aprecian diferencias, puesto que los hombres se decantan por el fármaco A de aplicación semanal (72,7% vs. 26,6%), mientras que las mujeres se decantan claramente hacia la administración diaria del fármaco (10% vs. 72,5%).

	Satisfecho	Fármaco A	Fármaco B	OR
Total	Sí No Total	100 120 220	99 119 218	1,00
Hombres	Sí No Total	80 30 110	29 80 109	7,36
Mujeres	Sí No Total	11 99 110	79 30 109	0,04

**Tabla 25.** Control de variables modificadoras del efecto. Satisfacción con el tratamiento según género

Si nos hubiéramos quedado con la tabla global del análisis bruto, se hubiese concluido que no existen diferencias entre la satisfacción del tratamiento según la periodicidad de administración. Sin embargo, nuestras conclusiones no se hubieran ajustado a la realidad, puesto que los hombres mostraron mayor satisfacción por uno de los tratamientos mientras que las mujeres lo hicieron por el





tratamiento contrario. De esta forma, podemos decir que el género actúa como variable modificadora del efecto, observando resultados distintos en función del grupo que estemos analizando.

### **17.9. Evaluación de resultados de diferentes grupos de sujetos (análisis de subgrupos)**

Hasta ahora, se ha considerado el análisis estratificado como el análisis a realizar cuando se sospecha de que existen variables categóricas confusoras o modificadoras del efecto y, de esta forma, cómo deben ser analizadas. Sin embargo, cuando deben controlarse más de una variable o bien se introducen variables continuas que deben ser controladas para la correcta interpretación de los resultados, es necesario recurrir a análisis multivariantes como la regresión lineal y la regresión logística (capítulo 14).

Debe tenerse en cuenta que el primer paso de cualquier análisis empieza por la relación de la variable más importante o respuesta en función de otras variables. En este primer paso, se pueden establecer análisis posteriores, para llegar a los resultados más específicos. El análisis por subgrupos se justifica si se ha planeado previamente al inicio del estudio, aunque puede ser razonable de cara a los resultados establecer el análisis de otros grupos, por ejemplo, en pacientes con edades entre 20 y 30 años, o con patologías concomitantes específicas.

Por otra parte, no hay que olvidar que si se realizan indiscriminadamente análisis de subgrupos, se puede llegar a conclusiones erróneas al aparecer el problema de las comparaciones múltiples, en la que seguro que aparecerán relaciones significativas única y exclusivamente por azar. Es por eso muy importante que aunque queramos responder a preguntas secundarias que de igual modo nos parecen trascendentales, éstas deben estar debidamente justificadas y con un número limitado de hipótesis.

#### **17.10. Consideraciones importantes**

Hay que tener en cuenta la importancia del diseño del cuaderno de recogida de datos por lo que recomendamos documentarse sobre su elaboración en diferentes publicaciones que nos hablen sobre 'Metodología de la Investigación' donde se aborde este tema. A grandes rasgos nos interesa que el diseño del CRD nos facilite la codificación de las diferentes variables y que se incorporen técnicas que faciliten y den más rapidez en su introducción, como por ejemplo, la utilización de numeración bajo las opciones de respuesta o la incorporación en las opciones de respuesta de 'dato no disponible'.



El análisis estadístico no es una tarea que se improvisa una vez se han procesado los datos. Por el contrario, el análisis surge más del marco teórico trazado que de los datos concretos obtenidos, por lo que se debería tener una idea precisa de cuales serán los análisis principales que se deberán efectuar antes de comenzar a recolectar datos. Se podrá definir así, con suficiente antelación, qué datos serán capaces de rechazar o afirmar una hipótesis, qué resultados indicarán una u otra conclusión. Esta actividad es fundamental para evitar sorpresas lamentables, como por ejemplo la de encontrar que no tenemos suficientes datos al final del procesamiento, o de que los que poseemos no nos sirven en realidad para mucho.

Es cierto que la estadística está siendo cada vez más utilizada en la investigación clínica desde que se introdujo el método estadístico en los estudios de eficacia y seguridad de los fármacos. Sin embargo hay un alto porcentaje de publicaciones en que se aprecian errores, por lo que cada vez más, los editores y evaluadores de las diferentes publicaciones son más estrictos en la selección de los estudios a aceptar y la evaluación de la parte estadística de los mismos que optan a su publicación. Por tanto, consideramos importante y cada vez más imprescindible la formación continuada en bioestadística para profesionales de la salud, y el asesoramiento y colaboración de estadísticos en el desarrollo de nuestros estudios.