

Revisión

El análisis multivariante en la investigación con antimicrobianos

F. Hervás Maldonado

Servicio de Microbiología, Hospital Central de la Defensa Gómez Ulla, Madrid

RESUMEN

Estudiar simultáneamente numerosas variables requiere un método especial con dos objetivos concretos, que son el análisis de todos los datos de manera conjunta y la asistencia lógica en la toma de decisiones. Se relacionan las técnicas descriptivas y explicativas más comunes. Se indican los siete pasos más comunes que lo desarrollan. Se propone la creación de asesorías metodológicas para la vigilancia y la racionalización de los procesos de investigación. También se destaca la necesidad de más inventores, así como de investigadores más asistidos con mejores presupuestos.

Palabras clave: Análisis multivariante - Variable independiente - Variable dependiente - Métodos estructurales

Multivariate analysis used in the investigation of antimicrobials

SUMMARY

Multivariate analysis techniques are reviewed. The study of variables requires special methods with two precise objectives, the simultaneous analysis of all the inputs and logical support when making decisions. The most common descriptive and explanatory techniques are discussed. The seven common steps usually developed for multivariate tests implementation are shown. The creation of an advisory board for methodology used by research entities is proposed for the rational control and performance of research processes. The need for more inventors, as well as more financial assistance for researchers, is highlighted.

Key words: Multivariate analysis - Independent variable - Dependent variable - Structural models

INTRODUCCIÓN

Con frecuencia vemos artículos de magnífica factura en los que se alude al análisis multivariante como modelo justificativo de un hecho presuntamente probado (1-26). Sin embargo, no siempre es realmente así,

por muchas razones, pese a la gran pureza del método estadístico seguido.

En primer lugar, hemos de ser humildes y reconocer que los modelos matemáticos estandarizados suelen cojear por algún lado en los casos concretos (27-

29). Esto no es válido para pequeñas consideraciones de datos, por muchas variables que les afecten, pero sí lo es cuando el número de datos crece de manera espectacular (30-34). Por ejemplo, trabajar con dos mil individuos no es lo mismo que trabajar con cincuenta, como es lógico, porque siempre existen desvíos significativos de la ley normal, por mucho que nos empeñemos en meter todos los datos en un mismo “saco” estadístico. Pero es que desde el punto de vista de los antimicrobianos no nos interesan tanto los cincuenta individuos como los dos mil, como es lógico, pues nuestra labor es de un espectro amplio y no reducido. Tal vez, si nos concentrásemos en la pura asistencia de una área sanitaria, esto no sería así, pero las nuevas circunstancias de movilidad cambiante de las poblaciones, la variedad de sus orígenes y el interés prospectivo de nuestra asistencia hacen que la variabilidad sea inmensa. Por ejemplo, en nuestro laboratorio de microbiología no podemos considerar una procedencia limitada de cada especie microbiana, sino que más bien hemos de considerarla en evolución. Por tanto, intentar valorar nuestros datos con un método estadístico convencional suele resultar inadecuado. No obstante, si tratamos previamente los datos en la forma conveniente, entonces sí podemos recurrir al método indicado (35-41). Esto, lamentablemente, no suele hacerse por tres razones: premura, falta de preparación y defectos presupuestarios.

En segundo lugar, los modelos utilizados suelen referirse a programas estadísticos concretos, orientados por lo general a tipos de muestras muy determinados, salvo cuando los epidemiólogos utilizan los procedimientos descriptivos recogidos en programas especiales, como el EPI-INFO y otros parecidos (42-62). Aunque, como ya hemos dicho, son descriptivos y algo analíticos, en esto les falta cierta profundidad y, sobre todo, robustez en la prospectiva, que es más bien limitada. Por otra parte, las asociaciones interdependientes suelen ser deficitarias, no tanto las dependientes y sí mucho las estructurales.

Por último, insistir en la necesidad de un soporte consecuente (31, 32, 63-96) de esta clase en cualquier posible instituto de investigación, con personal y medios adecuados, bien financiados y entrenados, con el apoyo de una estructura matemática de suficiente envergadura, como la que proporcionaría el adecuado departamento al respecto. Sin ese apoyo, nuestro proyecto está claramente lastrado y es muy poco prometedor, porque la voluntad –en este caso– no puede suplir con su esfuerzo la falta de medios y de conocimiento. Los medios son claros: ordenadores especia-

les, con muchas gigas de memoria RAM, para poder aunar gráficos y datos múltiples, con posibilidad de adaptar sensores y lectores de todo tipo, y con *software* muy abierto, modificable para cada caso. Porque investigar es caro, no cabe duda.

El análisis multivariante viene a ser un conjunto de métodos estadísticos que analizan muchos datos de muchas variables (más de dos, en sentido purista, varios cientos de forma habitual). Estas variables (36, 59, 97-104) pueden ser dependientes e independientes, o bien sólo estas últimas. Por otra parte, los datos suelen ser miles para cada variable. Llamaremos *inputs* a estos datos y *outputs* a las conclusiones (o conclusión única) que generen. La secuencia de trabajo se definirá mediante un algoritmo, que previamente consensuaremos; pensemos en una ecuación cuando digamos “algoritmo” (29, 105-107), una ecuación que representamos mediante un sistema convencional (y no libre) de grafos. En demasiadas ocasiones vemos algoritmos representados de cualquier manera, pero eso no son algoritmos, son otra cosa (108-111), generalmente árboles de decisión ampliados.

Los dos principales objetivos del análisis multivariante son el análisis conjunto (103) de los datos multivariantes y la asistencia lógica (107) en nuestra toma de decisiones. Dicho de otro modo: el estudio de una muestra grande y su inferencia poblacional. Al fin y al cabo, como en cualquier otro problema estadístico, pero con un método algo diferente.

TIPOS DE ANÁLISIS MULTIVARIANTE

Podríamos dividir los análisis multivariantes, según su finalidad, en dos grandes grupos:

- Las técnicas multivariantes descriptivas, en las cuales no se consideran más que variables independientes. Pueden ser pruebas sincrónicas o diacrónicas. Dentro de ellas, siguiendo la pauta del programa SPSS, por ejemplo, el tratamiento de grupos puede hacerse de manera diversa.
- Técnicas explicativas, según posean una o más variables dependientes.

La clasificación, según el método, podríamos sistematizarla así:

- I. Métodos de dependencia, según la variable o variables dependientes sean:
 - a) Con una sola variable dependiente:
 - Variable dependiente cuantitativa o métrica y:

- A. Variables independientes cuantitativas:
- Una sola variable independiente:
 - Coeficiente de correlación de Pearson (36, 101, 102, 104).
 - Varias variables independientes:
 - Regresión múltiple (36, 102).
 - Análisis de supervivencia (61, 103).
 - MANOVA (36, 61, 102).
 - Correlación canónica (36, 100, 102).

- B. Variables independientes cualitativas:
- Una sola variable independiente:
 - Para dos grupos *t* de Student (36).
 - Para más de dos grupos ANOVA (36, 61, 102).
 - Varias variables independientes:
 - ANOVA factorial (36, 61, 102)
 - Análisis conjunto (61, 103).

- Variable dependiente cualitativa o no métrica y:
 - Variables independientes cuantitativas:
 - Análisis discriminante (61, 103).
 - Variables independientes cualitativas:
 - χ^2 (36, 101).
 - Modelos log-lineales (61, 103).
 - Análisis conjunto (103).
 - Regresión logística (103):
 - * Con variable dependiente de *outputs* dicotómicos (binaria).
 - * Con variable dependiente de *outputs* politémicos (categorizada): modelos LOGIT y PROBIT.
 - * Con variable dependiente de *outputs* censurados (o truncada).

- b) Con varias variables dependientes y:
- Variables independientes cuantitativas:
 - Correlación canónica (36, 100, 102).
 - Modelos causales (61).
 - Variables independientes cualitativas:
 - Una sola variable independiente:
 - Dos grupos: coeficiente r^2 , Hotelling (36, 62, 102).
 - Más de dos grupos: MANOVA (36, 61, 102).
 - Varias variables independientes:
 - MANOVA (36, 61, 102).

II. Métodos de interdependencia, o modelos descriptivos, donde no hay variable dependiente conocida:

- 1) Modelos con descripción sincrónica:
- a) Variables con *inputs* cuantitativos o métricos:
- Determinación de componentes principales (100).

- *Cluster* análisis (29, 111).
- Análisis factorial (36, 100).
- Escalas multidimensionales (36, 100).

- b) Variables de *inputs* cualitativos o no métricos:
- Escalas multidimensionales (36, 100).
 - Análisis de correspondencias (61, 62).
 - Modelos log-lineales (61, 103).

- 2) Modelos con descripción diacrónica:
- a) Series temporales (61).
- b) Análisis de supervivencia (61, 103).

III. Métodos estructurales, que se diseñan de manera personalizada para cada caso (61, 103).

Algunos de los procedimientos indicados tienen las siguientes aplicaciones en la investigación antimicrobiana:

Regresión múltiple

Se utiliza cuando diversas variables cuantitativas dependen de otras varias también cuantitativas e independientes. Por ejemplo, calcular la previsión de gasto en material, las necesidades de personal y el consumo de antimicrobianos por zonas en un proceso epidémico a partir de la incidencia de casos, la prevalencia zonal, las edades de la población, la previsión de temperatura y lluvia, la densidad de población por zonas, las tasas migratorias y la evolución del IPC (índice de precios al consumo) interanual por meses.

Análisis de supervivencia

Es similar al anterior, pero con la diferencia de que la única variable dependiente es el tiempo de supervivencia de un individuo u objeto a partir de diversas variables independientes. Por ejemplo, cuánto tiempo puede permanecer un centro logístico de reparto sin efectuar suministro antibiótico a partir de la tasa de incidencia de las diversas enfermedades infecciosas, ajustando consumo por infección, teniendo en cuenta las caducidades medias de los fármacos, las condiciones de conservación, el personal para su administración y la demanda de asistencia.

ANOVA o MANOVA

El análisis de la varianza (ANOVA) o análisis multivariante de la varianza (MANOVA) se realiza cuando confrontamos grupos de una o más variables inde-

pendientes cualitativas con diversas variables dependientes cuantitativas. Tiene por objeto determinar las diferencias significativas entre los grupos con respecto a las variables cuantitativas. Por ejemplo, averiguar si hay diferencias en la tasa de cepas de *Staphylococcus aureus* resistentes a la meticilina aisladas en catéter por sexos en grupos de edad, o si afecta también el tipo de patología, la planta de hospitalización o el tipo de vestimenta del personal asistencial, así como la influencia de la ubicación geográfica del hospital.

Correlación canónica

Se trata de relacionar entre sí, simultáneamente, diversas variables cuantitativas dependientes e independientes, mediante la determinación de las diferentes correlaciones lineales de cada conjunto de variables, las cuales logren maximizar la correlación existente entre los dos conjuntos de variables, dependientes e independientes. Como ejemplo, podemos ver la relación que hay entre minutos diarios de práctica deportiva, ingesta calórica diaria, capacidad pulmonar y peso, con días de afectación por una determinada enfermedad infecciosa, tiempo de recuperación medio, dosis ajustada (dosis real de antibiótico por kilo de peso), temperatura axilar, tensión arterial mínima y pulsaciones por minuto en el momento del diagnóstico.

Análisis discriminante

Proporciona reglas óptimas de clasificación de *inputs*, desconociendo su grupo de procedencia y basándose en los valores de las variables independientes, que son todas cuantitativas. Un ejemplo puede ser establecer puntos de corte en nuevos antimicrobianos, que sirvan para discriminar toxicidad de no toxicidad en enfermos renales. Nosotros sólo sabemos la tasa de eficacia por kilo de peso, “acoplándola” con la distribución de *inputs* metabólicos encontrada en estos pacientes.

Modelos de regresión logística

La variable dependiente es cualitativa. Es una alternativa al análisis discriminante, cuando la distribución no sigue una ley normal. Por ejemplo, en dos grupos de sanos y enfermos, ver cómo se comportan pruebas positivas y negativas de diagnóstico. O bien, en un grupo de enfermos con varias categorías (muy graves, gra-

ves y leves) vemos cómo se comportan diversas pruebas diagnósticas positivas y negativas. E incluso en un grupo de enfermos de una edad, frente a otro de enfermos de otra edad, ver cómo se comportan dichas pruebas positivas y negativas, condicionándolas según el tratamiento antimicrobiano recibido en los siete días previos.

Análisis conjunto

Esta técnica estudia el efecto de *inputs* de variables independientes cualitativas sobre variables dependientes cuantitativas o cualitativas. La diferencia con el análisis de la varianza (ANOVA) se basa en que las variables dependientes pueden ser cualitativas y en que los *inputs* de las variables independientes no son aleatorios, sino fijados por el investigador. Es lo que se conoce como diseño de experimentos. Por ejemplo, un investigador desea saber la respuesta de distintos antimicrobianos en un grupo de animales de experimentación. Utiliza cuatro grupos de ratas Whistar y administra combinaciones de ellos, valorando su acción metabólica, así como la respuesta en una infección experimental. Para la primera variable dependiente (cuantitativa) somete los *outputs* a un punto de corte, por encima o por debajo de un valor x medido en suero. Para la segunda variable dependiente (*output* cuantitativo) mide la positividad o negatividad de un cultivo de control. Las variables independientes son dos. En la primera, los *inputs* poseen cuatro categorías: dosis estándar de un antimicrobiano, dosis estándar con vehículo metabólico añadido, dosis estándar con vehículo metabólico añadido y un segundo antimicrobiano, y dosis estándar sólo con el segundo antimicrobiano añadido. La segunda tiene dos categorías de *inputs*, parenteral u oral.

Análisis factorial y análisis de componentes principales

Se interrelacionan un número elevado de variables cuantitativas, todas ellas independientes. Sus interrelaciones vienen dadas por un número menor de variables que llamamos factores (si no son observables) o componentes principales (si son observables). Como ejemplo de componentes principales podemos considerar la estandarización de valores biométricos para elaborar una tabla de incapacidades laborales, en la que estudiamos una población amplia y medimos, en

cada paciente sometido a tratamiento con un antimicrobiano determinado, urea, glucosa, colesterol, GOT, GPT, GGT, bilirrubina, presión sistólica, presión diastólica, densidad de la orina, pH urinario, etc. Habrá que buscar la relación entre todas las variables observadas para llegar al número y tipo de ellas a considerar. Otro ejemplo es cuando un intensivista desea definir los factores que determinan el pronóstico de una determinada enfermedad a partir de las respuestas de sus pacientes a un grupo de parámetros diagnósticos, en un ajuste terapéutico asociado, utilizando un análisis factorial para resolverlo.

Escalas multidimensionales

Tienen por objeto representar en un espacio multidimensional las preferencias o semejanzas. Construiremos un mapa, dibujando las posiciones de los objetos, comparando aquellos que percibimos como similares con sus más próximos, y los que percibimos como distintos con sus más lejanos. Por ejemplo, analizar entre un grupo de pacientes los síntomas más percibidos y sus parámetros diagnósticos, según el tratamiento recibido, haciendo un seguimiento de cada paciente durante un mes, para estudiar los efectos a medio plazo de su tratamiento. Nuestro objetivo es ajustar lo mejor posible la duración de los tratamientos.

Cluster análisis o análisis de conglomerados

Tiene como finalidad la clasificación de variables cuantitativas en grupos cada vez menores, basándonos en la similitud y la diferencia de algunos de sus estadísticos, recogida mediante coeficientes de comparación y representada gráficamente mediante una figura denominada “dendrograma” (por parecerse a las dendritas de las neuronas). Difiere del análisis discriminante en que se desconocen el número y la composición de los distintos grupos. Por ejemplo, determinar la probabilidad de que una determinada secuencia genética obtenida defina un gen de resistencia a un determinado grupo antibiótico. Esta prueba (*cluster*) puede utilizarse también con variables cualitativas o mezcla. Los métodos de análisis de conglomerados son muchos (Tablas 1 y 2), como en otras técnicas multivariantes, por lo que conviene elegir el más adecuado y, ante la duda, pedir asesoría a expertos.

Análisis de correspondencias

Se puede utilizar con *inputs* de variables independientes cualitativas, pero no cuantitativas. Se utiliza en las tablas de contingencia multidimensionales. Se diferencia de las escalas multidimensionales en que se representan simultáneamente las filas y columnas de las tablas de contingencia, pero se busca un objetivo similar en ambos procedimientos. Por ejemplo, estudiar la prevalencia de neumonía comunitaria en los ancianos, teniendo en cuenta el tamaño de la ciudad (pequeña, mediana, grande), la zona geográfica de ubicación (norte, centro o sur, montaña o costa, ciudad o rural), si ha habido movimientos demográficos notables en los últimos dos años, sexo analizado y tratamiento prescrito. Es decir, relacionar todo con la presencia o ausencia de neumonía en cada caso y sumar.

Tabla 1. Procedimientos de clasificación de grupos en SPSS.

Clasificación de grupos	Cluster análisis discriminante	Quick cluster
		Cluster
		Discriminante

Tabla 2. Clasificación de grupos de *inputs* mediante conglomerados.

Criterios de selección	
Agrupación por distancias (conglomerados o <i>clusters</i>) con criterios de similitud/diferencia	
1. <i>Métodos jerárquicos</i>	
1.1. Algoritmos jerárquicos aglomerativos	
1.1.1. Algoritmos de contractura espacial	– Método de la ligadura simple
1.1.2. Algoritmos de dilatación espacial	– Método de la ligadura completa
1.1.3. Algoritmos de conservación espacial	– Método de la ligadura media
	– Método de la ligadura media ponderada
	– Método del centroide
	– Método de la mediana
	– Método de la varianza o de Ward
1.2. Algoritmos jerárquicos disociativos	
2. <i>Métodos particionales</i>	
2.1. Algoritmo convergente de K - medias o de Hartigan	
2.2. Otros algoritmos	
2.2.1. Análisis de densidad y moda de Wishart	
2.2.2. Algoritmo líder	

Modelos log-lineales

Se aplican en las tablas de contingencia multidimensionales, buscando establecer un modelo de relación dependiente de tipo multidimensional entre los *inputs* observados de las variables, con objeto de poder explicar las frecuencias calculadas. Es un método logarítmico similar al anterior, pero más orientado a problemas complejos que requieren una mayor precisión. Por ejemplo, analizar la tuberculosis en una población, según las mismas variables del caso anterior y con modelos simulados de progresión de la enfermedad sobre la base de los movimientos migratorios.

Métodos estructurales

Estudian las relaciones existentes entre grupos de variables expresadas mediante sistemas de ecuaciones que se valoran a un tiempo, considerándose que cierto tipo de ellas (llamadas “constructos”) son medidas erróneas en cierto grado, a partir de otras variables observables que se denominan “indicadores”. Estos modelos constan de dos partes: una estructural (que indica las relaciones de dependencia existentes entre las ecuaciones constructos latentes) y otro modelo de medida (que define cómo los indicadores se relacionan con sus constructos correspondientes). Por ejemplo, analizar cómo se relacionan las imágenes obtenidas mediante un estudio de PET con las percepciones de enfermedad del médico y del paciente, así como con otras pruebas diagnósticas sugestivas y su evolución de acuerdo con un tratamiento antibiótico.

SISTEMÁTICA

Existen muchos enfoques posibles a la hora de efectuar un análisis multivariante (29, 30, 61, 103), de manera que cualquier método general de trabajo es simplemente orientativo, aunque si se basa –como es natural– en la experiencia aplicada a la sanidad (o a cualquier otro campo más concreto del conocimiento), el rango de variabilidad metódica se estrecha. Por tanto, nuestra opción está un poco más clara.

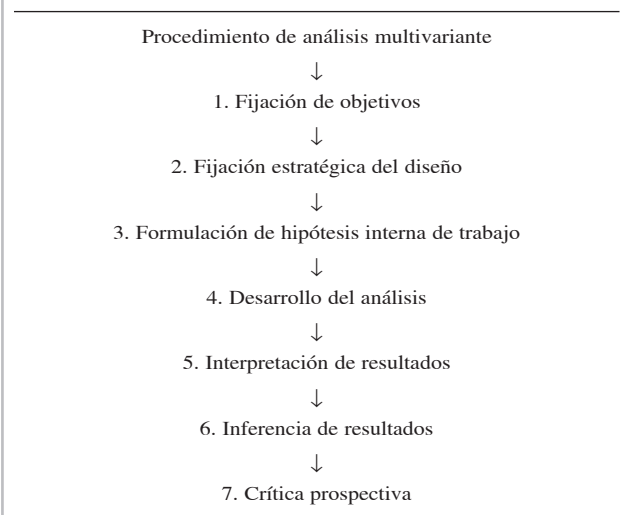
Siete fases (Tabla 3) presiden un estudio de este tipo:

1) Fijación de objetivos del análisis: junto con ellos habremos de fijar las técnicas de análisis multivariante que los posibiliten (29, 30). Por tanto, el investigador tendrá que establecer la relación conceptual de lo que quiere investigar. Definirá los conceptos en liza y cómo se afectan o pueden afec-

tar mutuamente. Es decir, habrá que establecer las relaciones de dependencia o de interdependencia de las variables a considerar, para saber si son dependientes o independientes y si son cualitativas o cuantitativas.

- 2) Establecer el tamaño muestral (36, 59, 60, 100), de acuerdo con el presupuesto y la casuística, en primer lugar, la estimación de ecuaciones previsibles (si es procedente), así como las distancias a calcular (igualmente si procede) y las técnicas empleables para la estimación. A partir de ahí se recogen los *inputs* o datos.
- 3) Elaborar las hipótesis internas de trabajo (36, 59-61, 100, 102). Evaluaremos las posibles hipótesis de normalidad, linealidad, independencia, homocedasticidad (cuando la varianza de los errores estocásticos –de azar– de la regresión no se ve modificada por el incremento de las observaciones)... Igualmente decidiremos lo que hemos de hacer con el ajuste imputable a los datos perdidos (*missing inputs*).
- 4) Efectuar el análisis. Estimamos el modelo y ajustamos a él los datos (36, 59-61, 100, 102, 107). Si nos apareciesen observaciones atípicas (conocidas técnicamente como *outliers*) o que influyan anormalmente sobre la estimación y su ajuste, habremos de considerar la utilidad real de su inclusión. Esto no quiere decir que podamos “maquillar” nuestra prueba, sino todo lo contrario. Lo que deberemos hacer es impedir que nos la “maquillen” des-

Tabla 3. Fases de un análisis multivariante.



de fuera terceras razones no conocidas que hayan modificado previamente los datos observados, pero ha de comprobarse que tales terceras razones existen.

- 5) Interpretar los resultados (60, 61). Esto no es nada sencillo. A veces los resultados nos obligan a volver a la tercera fase y reconsiderar nuestras hipótesis, e incluso a retomar la cuarta, repitiendo los análisis con el mismo u otro método.
- 6) Habremos de validar nuestros análisis mediante su expresión sobre la población a que se refiere la muestra estudiada. Es decir, nuestra inferencia ha de tener un claro fundamento de razón para que la utilidad de nuestro estudio sea patente (39, 60-62).
- 7) Por último, fijaremos la línea o líneas de posible continuidad (32, 107) en nuestro trabajo, mediante la sugerencia de enfoques futuros, siempre a la luz de una autocrítica.

Pondremos un ejemplo: queremos ver cómo se asocian dos determinados genes de resistencia a una familia de antimicrobianos con cuatro fármacos de esa familia en dosis crecientes, en seis especies bacterianas a las que se añaden cantidades ínfimas de antimicrobiano en medio líquido y se comprueba su crecimiento, secuenciando los aislamientos con la ayuda de *microarrays*. Ése es nuestro objetivo. Para ello se propone utilizar un método asociativo de análisis de conglomerados y seleccionamos –por razones que no hacen al caso– el método de Ward, de estimación mediante coeficientes de varianzas relacionadas (29, 30). Por tanto, tenemos una variable cualitativa independiente (gen de resistencia, con dos categorías), otra variable cualitativa independiente (especie bacteriana, con seis categorías), una variable cuantitativa de antimicrobiano independiente (dosis por ml de antimicrobiano, repetidas tras un número variable de ciclos de crecimiento), una más independiente cualitativa (tipo de antimicrobiano añadido), una dependiente cuantitativa (UFC o, si queremos, densidad por ml, tras el último de los ciclos de crecimiento expuesto) y otra dependiente cuantitativa (porcentaje de homología genética con cada gen de resistencia). Es la primera fase.

En segundo lugar, fijamos el tamaño de la muestra (básicamente el presupuesto nos lo va a fijar, en este caso, pero podríamos recurrir a otros procedimientos, aunque al ser inmensa la población bacteriana van a ser poco útiles) y la forma de medir (*microarrays*, *E-test*[®], densitometría...). Recogemos los datos y a continuación los tabulamos.

La tercera fase conlleva revisar nuestro trabajo y comprobar si todo lo hemos hecho de la misma manera, así como si existen factores externos de discordancia que no han sido considerados en alguna de las observaciones. Reajustamos la muestra, desechando lo que observamos que no procede.

En cuarto lugar, efectuamos la prueba de análisis de conglomerados (método de Ward), de manera que vemos cómo se agrupan los elementos y las variables, estableciéndose las distancias entre unos y otros o entre unas y otras, mediante valoración de similitudes y diferencias.

Interpretamos, en quinto lugar, los resultados, viendo cómo se asocia la presencia de un gen de resistencia con la exposición repetida *in vitro* a un antimicrobiano o no.

A continuación (sexta fase) inferimos sobre población en virtud de la frecuencia e intensidad encontradas en esa asociación.

Por último, proponemos un modelo sencillo de seguimiento futuro de resistencias y una posible causa de esta asociación que ha de ser investigada más profundamente en una determinada línea de trabajo.

CONCLUSIÓN

No debemos conformarnos con el estudio somero de los problemas, sino que el estado del conocimiento matemático actual requiere profundizar más en los diseños (29, 39, 61, 102, 103). Es preferible que las investigaciones sean menos y de más calidad, porque así se rentabiliza mucho mejor el presupuesto para ello. Además, la posibilidad de un abordaje cada vez más complejo hace recomendable plantearse asociaciones de trabajo un poco más estables, con unas líneas suficientemente definidas y respaldadas. Es muy conveniente que en estos grupos se establezcan asesorías metodológicas profesionales, lideradas por expertos en el diseño, la producción y la difusión de la investigación, así como en el tratamiento adecuado de los datos obtenidos y difundidos. En el caso concreto de los institutos de investigación, la necesidad de asesoría es mayor, aunque sólo sea para racionalizar esfuerzos y unificar criterios.

Hoy en día, el análisis multivariante no es algo inalcanzable, pues disponemos de magnífico *software* para poder desarrollarlo. Por eso, conviene ser cautos y desconfiar de *softwares* que no presenten una clara dimensión de calidad internacionalmente reconocida, como es el caso de EPI-INFO, SPSS, Statgraphics o

Mathematics, por citar los cuatro más utilizados. Hay otros también muy buenos, dependiendo del problema que nos ocupe. Pero tampoco seamos excesivamente críticos, porque ni éstos ni ningún *software* sirven para todos los casos. En situaciones muy concretas (29) nos lo habrán de desarrollar *ex profeso*.

Por último, pensemos un poco en términos de rentabilidad. ¿Necesitamos más investigadores o inventores? Porque no es la misma cosa ni genera los mismos beneficios. Los inventores generan grandes beneficios a corto plazo y con bajo coste. Los investigadores pueden generar también grandes beneficios –o tal vez no, porque su trabajo se desarrolla en situaciones de incertidumbre, al menos relativa–, pero siempre con un coste muy elevado y a un plazo más largo. Debe haber investigadores, sin duda, pero mejor asistidos con buenos presupuestos. Faltan cantidades ingentes de inventores, que son los que realmente proporcionan los fondos necesarios para costear el gasto en la investigación (112). Su trabajo desarrolla el de los investigadores y lo rentabiliza, merced a las patentes, con inversiones muy ajustadas a demanda. Al menos así es como se costea la investigación en los países de PIB más ventajoso.

Correspondencia: fhermal@oc.mde.es - fhervas@terra.es

BIBLIOGRAFÍA

- Lang, E.S., Boivin, J.F., Leger, R. y cols. *Do electronic linkages between the emergency department and primary care physicians reduce resource utilization in the emergency department? Results of a randomized controlled trial*. Acad Emerg Med 2003; 10: 547-548.
- Saitz, R., Horton, N.J., Sullivan, L.M. y cols. *Addressing alcohol problems in primary care: A cluster randomized, controlled trial of a systems intervention. The screening and intervention in primary care (SIP) study*. Ann Intern Med 2003; 138: 372-382.
- Croudace, T., Evans, J., Harrison, G. y cols. *Impact of the ICD-10 primary health care (PHC) diagnostic and management guidelines for mental disorders on detection and outcome in primary care. Cluster randomised controlled trial*. Br J Psychiatry 2003; 182: 20-30.
- Sheikh, K., Smeeth, L., Ashcroft, R. *Randomised controlled trials in primary care: Scope and application*. Br J Gen Pract 2002; 52: 746-751.
- Bierma-Zeinstra, S.M., Bohnen, A.M., Bernsen, R.M., Ridderikhoff, J., Verhaar, J.A., Prins, A. *Hip problems in older adults: Classification by cluster analysis*. J Clin Epidemiol 2001; 54: 1139-1145.
- Campbell, M.J. *Cluster randomized trials in general (family) practice research*. Stat Methods Med Res 2000; 9: 81-94.
- Kendrick, D., Marsh, P., Fielding, K., Miller, P. *Preventing injuries in children: Cluster randomised controlled trial in primary care*. BMJ 1999; 318: 980-983.
- Cotter, B., Bremer, V., Stroh, G. y cols. *Assessment of neonatal tetanus elimination in African setting by lot quality assurance cluster sampling (LQA-CS)*. Epidemiol Infect 2003; 130: 221-226.
- Heegaard, W.G., Plummer, D., Dries, D. y cols. *Ultrasound in Helicopter EMS*. Acad Emerg Med 2003; 10: 449.
- McLachlan, R.I., Baker, H.W., Clarke, G.N. y cols. *Semen analysis: Its place in modern reproductive medical practice*. Pathology 2003; 35: 25-33.
- Dhatt, G.S., Pum, J.K., Viljoen, A., Jenner, W. *Evaluation of HbA1C results in an external quality assessment scheme in South Africa*. Clin Chem Acta 2003; 331: 147-151.
- Graf, J. *Do you know the frequency of errors in your intensive care unit?* Crit Care Med 2003; 31: 1277-1278.
- Mason, I., Brady, C. *Benchmarking and nutrition*. Nurse Times 2003; 99: 49-50.
- Murray, M., Bodenheimer, T., Rittenhouse, D., Grumbach, K. *Improving timely access to primary care: Case studies of the advanced access model*. JAMA 2003; 289: 1042-1046.
- Whitten, K.L., Rein, M.F., Reppucci, N.D., Turkheimer, E. *The emotional experience of intercourse and sexually transmitted diseases: A decision-tree analysis*. Sex Transm Dis 2003; 30: 348-356.
- Delaney, G., Barton, M., Jacob, S., Jalaludin, B. *A model for decision making for the use of radiotherapy in lung cancer*. Lancet Oncol 2003; 4: 120-128.
- Birkmeyer, J.D., Liu, J.Y. *Decision analysis models: Opening the black box*. Surgery 2003; 133: 1-4.
- Hunsberger, B., Bagwell, C.B., Herbert, D., Bray, C., Langweiler, M. *Effects of resolution reduction on data analysis*. Cytometry 2003; 53A: 103-111.
- Tan, C.K., Law, N.M., Ng, H.S., Machin, D. *Simple clinical prognostic model for hepatocellular carcinoma in developing countries and its validation*. J Clin Oncol 2003; 21: 2294-2298.
- Redleaf, M.I., Bauer, C.A. *Topical antiseptic mouthwash in oncological surgery of the oral cavity and oropharynx*. J Laryngol Otol 1994; 108: 973-979.
- Hawrylyshyn, P.A., Bernstein, P., Papsin, F.R. *Risk factors associated with infection following cesarean section*. Am J Obstet Gynecol 1981; 139: 294-298.
- Raz, R., Okev, N., Kennes, Y. y cols. *Demographic characteristics of patients with community-acquired bacteriuria and susceptibility of urinary pathogens to antimicrobials in northern Israel*. Isr Med Assoc J 2000; 2: 426-429.
- Ruza, F., Alvarado, F., Herruzo, R. y cols. *Prevention of nosocomial infection in a pediatric intensive care unit (PICU) through the use of selective digestive decontamination*. Eur J Epidemiol 1998; 14: 719-727.
- Herruzo-Cabrera, R., García, J.I., García-Magán, P., Del Rey-Calero, J. *Nosocomial infection in a neonatal intensive care unit and its prevention with selective intestinal decolonization. A multivariate evaluation of infection reduction*. Eur J Epidemiol 1994; 10: 573-578.
- El tratamiento del dolor de garganta con antibióticos en atención primaria*. Rev Panam Salud Pública 2001; 10: 258-259.

26. Rubio, L.A., Sedeño, C., Fernández, S.A., Arbesú, M.A., Puig, L. *Grado esperado de uso de los antimicrobianos aminoglucósidos en pacientes oncológicos*. Rev Cubana Farm 2002; 36: 42-49.
27. Varo, J. *Gestión estratégica de la calidad en los servicios sanitarios*. Díaz de Santos, Madrid 1994.
28. Pineault, R., Daveluy, C. *La planificación sanitaria: Conceptos, métodos, estrategias*. 2ª ed. Masson, Barcelona 1992.
29. Hervás, F. *Procedimientos de inteligencia artificial en el estudio de las enfermedades infecciosas*. Díaz de Santos, Madrid 1999.
30. Vogt, W., Nagel, D. *Cluster analysis in diagnosis*. Clin Chem 1992; 38: 182-198.
31. Carrasco, J.L. *El método estadístico en la investigación médica*. 5ª ed. Editorial Ciencia, Madrid 1986.
32. Hulley, S.B., Cummings, S.R. *Diseño de la investigación clínica*. Doyma, Barcelona 1993.
33. Milton, J.S., Toscos, J.O. *Estadística para biología y ciencias de la salud*. Interamericana Mc Graw Hill, Madrid 2001.
34. Eppinger, S.D. *Innovation at the speed of formation*. Har Bus Rev 2001; 79: 149-158, 178.
35. Colton, T. *Estadística en medicina*. Salvat, Barcelona 1979.
36. Martín Andrés, A., Luna del Castillo, J.D. *Bioestadística para las ciencias de la salud*. 4ª ed. Norma, Madrid 1993.
37. Sackett, D.L., Haynes, R.B., Guyatt, G.H., Tugwell, P. *Epidemiología clínica. Ciencia básica para la medicina clínica*. 2ª ed. Médica Panamericana, Madrid 1994.
38. Lasky, F.D., Boser, R.B. *Designing in quality through design control: A manufacturer's perspective*. Clin Chem 1997; 43: 866-872.
39. Kelsey, I.L., Thompson, W.D., Evans, A. *Methods in observational epidemiology*. Oxford University Press, New York 1986.
40. Featherstone, D., Brown, D., Sanders, R. *Development of the global measles laboratory network*. J Infect Dis 2003; 187 (Suppl. 1): S264-269.
41. Larson, E. *Status of practice guidelines in the United States: CDC guidelines as an example*. Prev Med 2003; 36: 519-524.
42. Shih, S.C., Bost, J.E., Pawlson, L.G. *Standardized health plan reporting in four areas of preventive health care*. Am J Prev Med 2003; 24: 293-300.
43. Williams, I.D., Bird, A. *Public perceptions of air quality and quality of life in urban and suburban areas of London*. J Environ Monit 2003; 5: 253-259.
44. Sweeney, J., Brooks, A.M., Leahy, A. *Development of the Irish national perception of quality of care survey*. Int J Qual Health Care 2003; 15: 163-168.
45. Cardew, P.T. *A method for assessing the effect of water quality changes on plumbosolvency using random daytime sampling*. Water Res 2003; 37: 2821-2832.
46. Peterson, A.T., Vieglaiss, D.A., Andreasen, J.K. *Migratory birds modeled as critical transport agents for west nile virus in north america*. Vector Borne Zoonotic Dis 2003; 3: 27-37.
47. Philippe, P. *Sartwell's incubation period model revisited in the light of dynamic modeling*. J Clin Epidemiol 1994; 47: 419-433.
48. Piccardi, C., Lazzaris, S. *Vaccination policies for chaos reduction in childhood epidemics*. IEEE Trans Biomed Eng 1998; 45: 591-595.
49. Cavalieri, L.F., Kocak, H. *Chaos: A potential problem in the biological control of insect pests*. Math Biosci 1995; 127: 1-17.
50. Olsen, L.F., Truty, G.L., Schaffer, W.M. *Oscillations and chaos in epidemics: A nonlinear dynamic study of six childhood diseases in Copenhagen, Denmark*. Theor Popul Biol 1988; 33: 344-370.
51. Koella, J.C., Doebeli, M. *Population dynamics and the evolution of virulence in epidemiological models with discrete host generations*. J Theor Biol 1999; 198: 461-475.
52. Colimon, K.H. *Fundamentos de epidemiología*. Díaz de Santos, Madrid 1990.
53. Greenberg, R.S. *Medical epidemiology*. 3rd ed. Lange, New York 2001.
54. Jenicek Cleroux, R. *Epidemiología. La lógica de la medicina moderna*. Masson, Barcelona 1996.
55. McMahon, B., Trichopoulos, D. *Epidemiología*. 2ª ed. Marban, Madrid 2001.
56. Piédrola Gil, G. y cols. *Medicina preventiva y salud pública*. 10ª ed. Masson-Salvat, Barcelona 2001.
57. Rey Calero, J., Herruzo Cabrera, R., Rodríguez Artalejo, F. *Fundamentos de epidemiología clínica. Síntesis*, Madrid 1996.
58. Rothman, K.J., Greenland, S. *Modern epidemiology*. 2nd ed. Lippincott-Raven, Philadelphia 1998.
59. Lwanga, S.K., Lemeshow, S. *Determinación del tamaño de las muestras en los estudios sanitarios*. OMS, Ginebra 1991.
60. Goldberg, M. *L'Epidémiologie sans peine*. 2ème éd. Editions Frison-Roche, Paris 1990.
61. Kleinbaum, D.G., Kupper, L.L., Morgenstern, H. *Epidemiologic research. Principles and quantitative methods*. Van Nostrand Reinhold, New York 1982.
62. Szklo, M., Nieto, J. *Epidemiología intermedia. Conceptos y aplicaciones*. Díaz de Santos, Madrid 2003.
63. Ovreteit, J., Gustafson, D. *Using research to inform quality programs*. BMJ 2003; 326: 759-761.
64. Mills, A.E., Rorty, M.V. *Total quality management and the silent patient*. Bus Ethics Q 2002; 12: 481-504.
65. Murray, M., Berwick, D.M. *Advanced access: Reducing waiting and delays in primary care*. JAMA 2003; 289: 1035-1040.
66. Korting, H.C. *Total quality management in dermatomycology in Germany*. Mycoses 2003; 46: 2-4.
67. Holl, R.W., Grabert, M. *The quality circle: How to improve the outcome of paediatric diabetes care*. Horm Res 2002; 57 (Suppl. 1): 105-109.
68. Tausch, B.D., Harter, M.C. *Perceived effectiveness of diagnostic and therapeutic guidelines in primary care quality circles*. Int J Qual Health Care 2001; 13: 239-246.
69. Forster, D.H., Krause, G., Gastmeier, P. y cols. *Can quality circles improve hospital-acquired infection control?* J Hosp Infect 2000; 45: 302-310.
70. Henderson, A.R. *Assessing test accuracy and its clinical consequences: A primer for receiver operating characteristic curve analysis*. Ann Clin Biochem 1993; 30: 521-539.
71. The eighteenth annual Arnold O. Beckman conference in clinical chemistry. *Pursuing medical relevance: A crucial role for the clinical laboratory scientist - From a focus on analytical performance to one on medical outcome*. Clin Chem 1995; Special issue, 8B.

72. Wie, S., Shanklin, C.W., Lee, K.E. *A decision tree for selecting the most cost-effective waste disposal strategy in food-service operations.* J Am Diet Assoc 2003; 103: 475-482.
73. Mandelbrot, B.B. *Fractal geometry of nature.* W.H. Freeman, Nueva York 1983.
74. Skinner, J.E., Molnar, M., Vybiral, T., Mitra, M. *Application of chaos theory to biology and medicine.* Integra Physiol Behav Sci 1992; 27: 39-53.
75. Tsonis, P.A., Tsonis, A.A. *Chaos: Principles and implications in biology.* Comput Appl Biosci 1989; 5: 27-32.
76. Philippe, P. *Chaos and public health: Implications for an epidemic.* Can J Public Health 1992; 83: 165-166.
77. Olsen, L.F., Schaffer, W.M. *Chaos versus noisy periodicity: Alternative hypotheses for childhood epidemics.* Science 1990; 249: 499-504.
78. Cross, S. *Fractals in pathology.* J Path 1997; 182: 1-8.
79. Kaitala, V., Ylikarjula, J., Heino, M. *Dynamic complexities in host-parasitoid interaction.* J Theor Biol 1999; 197: 331-341.
80. Davey, H.M., Davey, C.L., Woodward, A.M. y cols. *Oscillatory, stochastic and chaotic growth rate fluctuations in permissively controlled yeast cultures.* Biosystems 1996; 39: 43-61.
81. Lehning, M., Shonnard, D.R., Chang, D.P., Bell, R.L. *An inversion algorithm for determining area-source emissions from downwind concentration measurements.* Air Waste 1994; 44: 1204-1213.
82. Renshaw, E. *Chaos in biometry.* IMA J Math Appl Med Biol 1994; 11: 17-44.
83. Skinner, J.E. *Low-dimensional chaos in biological systems.* Biotechnology (NY) 1994; 12: 596-600.
84. Nieminen, H., Takala, E.P. *Evidence of deterministic chaos in the myoelectrical signal.* Electromyogr Clin Neurophysiol 1996; 36: 49-58.
85. Markham, F.W. *A method for introducing the concepts of chaos theory to medical students.* Theor Med Bioeth 1998; 19: 1-4.
86. Griffiths, F., Byrne, D. *General practice and the new science emerging from the theories of chaos and complexity.* Br J Gen Pract 1998; 48: 1697-1699.
87. Sedivy, R. *The potential role of apoptosis (programmed cell death) in a chaotic determined carcinogenesis.* Med Hypotheses 1996; 46: 455-457.
88. Lindberg, C., Herzog, A., Merry, M., Goldstein, J. *Health care applications of complexity science. Life at the edge of chaos.* Physician Exec 1998; 24: 6-20.
89. Kirkwood, T.B., Bangham, C.R. *Cycles, chaos, and evolution in virus cultures: A model of defective interfering particles.* Proc Natl Acad Sci USA 1994; 91: 8685-8689.
90. Matioli, G.T. *On the unreliability of the axiom of randomness: The case of random mutations in biology.* Med Hypotheses 1997; 49: 181-182.
91. Grenfell, B.T., Kleczkowski, A., Gilligan, C.A., Bolker, B.M. *Spatial heterogeneity, nonlinear dynamics and chaos in infectious diseases.* Stat Methods Med Res 1995; 4: 160-183.
92. Heindl, B., Schmidt, R., Schmid, G. y cols. *A case-based consiliarius for therapy recommendation (ICONS): Computer-based advice for calculated antibiotic therapy in intensive care medicine.* Comp Methods Prog Biomed 1997; 52: 117-127.
93. Dharan, M. *Control de calidad en los laboratorios clínicos.* Reverté, Barcelona 1980.
94. *Manual basic de control de qualitat dels laboratoris clínics.* 1ª ed. Departament de Sanitat i Seguretat Social, Generalitat de Catalunya, Barcelona 1986.
95. Trillas, E., Alsina, C., Terricabras, J.M. *Introducción a la lógica borrosa.* Ariel, Barcelona 1995.
96. Varo, J. *Gestión estratégica de la calidad en los servicios sanitarios. Un modelo de gestión hospitalaria.* Díaz de Santos, Madrid 1994.
97. Leguina, J. *Fundamentos en demografía. Siglo XXI,* Madrid 1989.
98. Martínez-González, M.A., De Irala, J., Faulin, F.J. *Bioestadística amigable.* Díaz de Santos, Madrid 2001.
99. Doménech, J.M. *Bioestadística. Métodos estadísticos para investigadores.* Herder, Barcelona 1980.
100. Solanas, A., Salafranca, L.I., Fauquet, J., Núñez, M.I. *Estadística descriptiva en ciencias del comportamiento.* Thomson editores, Madrid 2004.
101. Pascua, M. *Metodología bioestadística para médicos y oficiales sanitarios.* 2ª ed. Paz Montalvo, Madrid 1974.
102. Sánchez, M., Frutos, G., Cuesta, P.L. *Estadística y matemática aplicadas. Síntesis,* Madrid 1996.
103. Juez, P., Díez, F.J. *Probabilidad y estadística en medicina. Aplicaciones en la práctica clínica y en la gestión sanitaria.* Díaz de Santos, Madrid 1997.
104. Marín-Pliego, J., Riz-Maya, L. *Estadística I. Probabilidad.* Thomson editores, Madrid 2004.
105. Grossman, Z. *Mathematical modeling of thymopoiesis in HIV infection: Real data, virtual data, and data interpretation small star, filled.* Clin Immunol 2003; 107: 137-139.
106. Lorenz, E.N. *Deterministic nonperiodic flow.* J Atmos Sci 1963; 20: 130-141.
107. Aranda, J., Fernández, J.L., Jiménez, J., Morilla, F. *Fundamentos de lógica matemática.* 2ª ed. Sanz y Torres, Madrid 2001.
108. Moldofsky, H. *Sleep and the immune system.* Int J Immunopharmacol 1995; 17: 649-654.
109. Whitmire, D., Cornelius, L., Whitmire, P. *Effects of food on ethanol metabolism.* Drug Metabol Drug Interact 2002; 19: 83-96.
110. Rubinstein, R.Y. *Simulation and the Monte Carlo method.* John Wiley & Sons, New York 1981.
111. Kerry, S.M., Bland, J.M. *The intracluster correlation coefficient in cluster randomisation.* BMJ 1998; 316: 1455.
112. Hervás, F. *Hospital, empresa y sanidad asistencial. Calidad y rentabilidad.* Díaz de Santos, Madrid 2007.