

# Fibroadenoma y carcinoma de mama: Análisis discriminante logístico

M. I. Nieto López

## SUMMARY

*The differential diagnosis between solid breast tumors was studied. The semiological value of the observed signs in the anamnesis and the physical examination was calculated independently and jointly. Statistical methods used are discriminant and contingency tables analysis. Discriminant analysis using as variables the different localization of tumors was not resolute. Although fibroadenomas are correct reclassified (87.78%), only 30.23% of breast cancers were identified. The total yield was only 59.66% of proper reclassification. Among other variables only age and the tumor size had a standardized coefficient statistically significant. The family history, number of pregnancies, affected breast and menopause were not relevant. Reclassification matrix showed a correct classification of fibroadenomas (85.56%) and breast cancer (81.40%). The total yield was 83.51%. According the contingency table analysis there was no preferent localization of tumor in both pathologies. However family history of breast cancer was a factor of differentiation between the two pathologies. Menopause was highly significant.*

## Palabras clave

*Fibroadenoma, Cáncer de mama, Análisis logístico discriminante.*

## Key words

*Fibroadenoma, Breast neoplasms, Logistic discriminant analysis.*

## INTRODUCCION

El diagnóstico diferencial entre tumoraciones sólidas mamarias (fibroadenoma y carcinoma como entidades patológicas más frecuentes) es muy importante, ya que varía enormemente su tratamiento y evolución. Es bien sabido que la edad de presentación de ambas patologías es diferente. El fibroadenoma empieza su aparición con las primeras menstruaciones y alcanza un máximo de incidencia a los 20 años, disminuyendo hacia los 30. El carcinoma tiene una distribución en progresión ascendente desde los 20

años hasta edades avanzadas de la vida. De esta manera existe un grupo de edad en que se pueden presentar ambas enfermedades, aunque las incidencias sean diferentes.

En la práctica diaria, posiblemente debido a la falta de sensibilización de las mujeres ante la importancia de la patología mamaria, es fácil encontrar una prevalencia no irrelevante de fibroadenomas en edades avanzadas de la vida a pesar de que su incidencia sea mínima. La posible causa está en que no han sido diagnosticados con anterioridad, ya que la inmensa mayoría de las mujeres en nuestro medio no se

autoexploran ni acuden a controles médicos periódicos.

La incidencia del fibroadenoma es del 7 al 12% de las pacientes que consultan por patología mamaria. Siguiendo en importancia al tumor más frecuente: el carcinoma. La incidencia del fibroadenoma múltiple es del 12-20% de casos. El tratamiento del fibroadenoma es quirúrgico. Sin embargo, en la actualidad se plantea la cuestión de si debe tratarse o no. El principal argumento para la abstención terapéutica es la bajísima tasa de malignización (inferior a la del resto de la glándula) y la posibilidad de un diagnóstico exacto con los métodos exploratorios actuales. A partir de los 35 años se describen tasas de error diagnóstico que varían entre 1,3 y un 6% de tumores catalogados como fibroadenomas que resultaron ser neoplasias. Un argumento a favor de la extirpación quirúrgica es la preocupación que causa la tumoración palpable en la mujer. En trabajos publicados se admite que pocas pacientes aceptan la no extirpación (entre el 7 y el 20%) y de estas últimas una proporción importante solicita la exéresis pasado un tiempo.

El diagnóstico de las tumoraciones mamarias se basa en la combinación de la información proporcionada por la clínica, la exploración radiológica y la citología. La fiabilidad diagnóstica combinada que se alcanza es del 88,5 al 90%.

Para hacer el diagnóstico es imprescindible una anamnesis exhaustiva y una exploración clínica detallada. Radiológicamente el carcinoma se manifiesta mediante signos directos y signos indirectos. Los signos directos representan la traducción radiológica de una alteración de un área más o menos extensa del parénquima mamario. Los signos indirectos representan las alteraciones que la enfermedad produce sobre estructuras más o menos alejadas de la lesión.

El fibroadenoma se manifiesta radiológicamente como un patrón nodular único y generalmente unilateral. La imagen nodular es única y presenta de contornos más o menos nítidos que permiten diferenciarla del parénquima circundante. Ante todo nódulo mamario deberá estudiarse su forma, su contorno, su densidad y su tamaño.

Por lo que hace referencia al estudio citológico, el análisis de las principales series estudiadas permite establecer como una sensibilidad diagnóstica por encima del 70% y una especificidad superior al 98%. Desde la perspectiva citológica, los fibroadenomas se caracterizan por la presencia de grandes placas de

células ductales de características benignas acompañadas por numerosas células mioepiteliales. Los carcinomas, por su parte, presentan células con atipias nucleares, disposición irregular y pérdida de cohesión.

El diagnóstico de las tumoraciones ha de hacerse combinando los métodos exploratorios para obtener una mayor fiabilidad, ya que la ausencia de un método totalmente efectivo hace que el diagnóstico final deba hacerse por anatomía patológica.

En este estudio se ha calculado el valor semiológico de los signos observados en la anamnesis y en la exploración clínica de forma separada y conjunta. Este estudio es importante para caracterizar individualmente cada maniobra o variable exploratoria; en segundo lugar para poder compararlos entre sí y por último como herramienta de trabajo de cálculo del coste beneficio.

El valor semiológico de un signo o síntoma clínico se evalúa mediante 2 medidas: especificidad y sensibilidad. La característica clínica medida puede constituir una variable continua, como la dimensión del tumor, o una variable discreta, como, por ejemplo, la edad medida en años. Un caso extremo de variable discreta son las variables binarias, en que ésta sólo adquiere 2 valores, positivo o negativo, por ejemplo, la existencia o no de antecedentes patológicos. La especificidad es la capacidad de distinguir entre una población afectada por una enfermedad, por ejemplo, el carcinoma de mama, y otra población no afectada con la que se establece el diagnóstico diferencial. En este caso la población contrapuesta es la afectada de fibroadenoma de mama. La especificidad de una variable *continua* se hace en función de declarar como *positivo* los valores que, por ejemplo, superen un determinado valor *d* denominado discriminante y *negativo* cuando sean inferiores a *d*:

$$\text{especificidad} = Es = \text{probabilidad } (y = 0 \mid \bar{E}, d)$$

siendo *y* el resultado observado en la exploración de un paciente, «0» la conclusión negativa de tal exploración, porque por ejemplo no supere el valor discriminante *d*, y  $\bar{E}$  representa la ausencia de enfermedad (de carcinoma). Según esta definición, la especificidad de los antecedentes patológicos es alta: 0,98.

Otra característica importante es la sensibilidad. Se define como la proporción de resultados positivos entre la población caracterizada como enferma. Formalmente:

$$\text{especificidad} = Se = \text{probabilidad } (y = 1 \mid E, d)$$

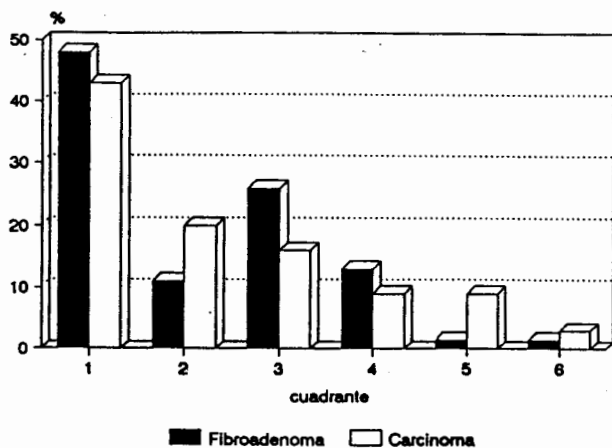


Fig. 1. Distribución de las tumoraciones por cuadrantes.

siendo «1» la conclusión positiva de la exploración, porque por ejemplo supere el valor discriminante  $d$ , y  $E$  es la existencia de la enfermedad. Según esta definición, la sensibilidad de la presencia de antecedentes patológicos es muy baja: 0,19.

Desafortunadamente no existen pruebas diagnósticas con sensibilidad y especificidad perfecta, 1,00, debiéndose proceder a una selección de pruebas atendiendo a ambas características. Cuando se trata de un problema diagnóstico del tipo de un *screening* es importante una sensibilidad lo más alta posible que garantice el menor número de falsos negativos, aun a costa de introducir falsos positivos. Estos últimos serán resueltos en fases posteriores de la investigación clínica. En el caso de que se trate de pruebas de diagnóstico con consecuencias de alto coste o en poblaciones de alta prevalencia, es preferible una alta especificidad, ya que no interesa introducir falsos positivos. Como consecuencia, los antecedentes patológicos no constituyen un dato de utilidad. Su baja sensibilidad lo hacen inadecuado para la utilización como *screening*. De hecho, así lo indica una medida global de sensibilidad y especificidad, denominada índice de Yates:

$$I = (Se + Es)/2$$

Teniendo en cuenta que esta cantidad oscila entre 0,5 y 1, para existencia de antecedentes patológicos es de 0,59.

Estas características, sensibilidad y especificidad

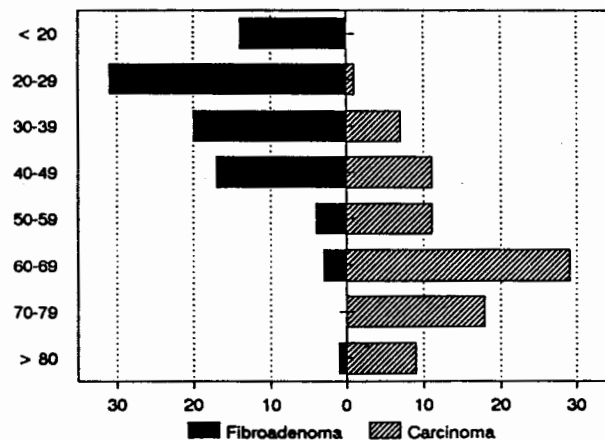


Fig. 2. Distribución de las tumoraciones por grupos de edad.

se calculan fácilmente en una variable binaria o continua dicotomizada en función del valor discriminante  $d$ , claro que en este caso implica una pérdida de información.

En la práctica clínica, sin embargo, la preocupación del médico no es, en principio, la proporción de resultados positivos ( $y = 1$ ) ciertos, o sensibilidad de una maniobra exploratoria, ni la proporción de resultados negativos ( $y = 0$ ) ciertos, o especificidad. El clínico desea saber esencialmente si un paciente sometido a una exploración padece la enfermedad ( $E$ ) que sospecha, cuando el resultado de tal maniobra exploratoria es *positivo*, es decir, la probabilidad  $P$  de que padezca la enfermedad  $E$  cuando el resultado es positivo ( $y = 1$ ):

$$P(E | y = 1)$$

Alternativamente, cuando el resultado de la exploración es *negativo*, hasta qué punto puede descartarse la enfermedad  $E$ . Es decir cuál es la probabilidad de que un sujeto cuyo resultado de la exploración es negativo ( $y = 0$ ) no padezca la enfermedad,  $\bar{E}$ . Formalmente:

$$P(\bar{E} | y = 0)$$

Estas probabilidades se denominan *valores predictivos* de un resultado positivo y de un resultado negativo, respectivamente. Para calcularlas, el teorema de Bayes establece:

$$P(\mathbf{E} | y = 1) = \frac{P(\mathbf{E})P(y = 1 | \mathbf{E})}{[P(\mathbf{E})P(y = 1 | \mathbf{E}) + (P(\bar{\mathbf{E}})P(y = 1 | \bar{\mathbf{E}}))]}$$

$$P(\mathbf{E} | y = 1) = P(\mathbf{E})Se/[P(\mathbf{E})Se + (P(\bar{\mathbf{E}})\{1-Es\})]$$

Donde  $P(\mathbf{E})$  indica la probabilidad de padecer la enfermedad  $\mathbf{E}$ . Esta probabilidad *antes de realizar* la prueba o exploración se denomina también *probabilidad a priori*, mientras que la probabilidad resultante después de practicar la exploración o valor predictivo,  $P(\mathbf{E} | y)$ , se denomina *probabilidad a posteriori*. Un ejemplo de probabilidad *a priori* puede ser simplemente la prevalencia de la enfermedad  $\mathbf{E}$  (número de enfermos en una población determinada), pero en realidad puede ser cualquier valor razonable. Mientras que la probabilidad *a priori* no depende de ninguna elección de un valor discriminante, la probabilidad *a posteriori* dependerá de los valores discriminantes, siempre que deba hacerse una dicotomización.

Ahora bien, denominando  $L(1)$  el cociente de las probabilidades de observar un resultado positivo en cada tipo de diagnóstico:

$$L(1) = P(y = 1 | \mathbf{E})/P(y = 1 | \bar{\mathbf{E}})$$

$$L(1) = Se/(1-Es)$$

Este cociente se denomina *conciencia de verosimilitud*. Se interpreta como *un resultado positivo es  $L(1)$  veces más frecuente en el paciente de  $\mathbf{E}$  que en aquellos sujetos que no padecen dicha enfermedad,  $\bar{\mathbf{E}}$ .*

Mediante una simple manipulación algebraica, se puede relacionar la probabilidad *a posteriori*  $P(\mathbf{E} | y = 1)$  con  $L(1)$ .

$$P(\mathbf{E} | y = 1) = P(\mathbf{E})L(1)/[P(\mathbf{E})L(1) + \{1-P(\mathbf{E})\}]$$

TABLA I

**COCIENTE, ERROR ESTANDAR Y SIGNIFICACION DE CADA VARIABLE**

Variable	$\beta$	es	z
Ter. ind. ....	-0,8998	0,5863	-1,5349
1 .....	1,0669	0,5969	1,7874
2 .....	0,2177	0,6116	0,3560
3 .....	1,4434	0,6487	2,2271
4 .....	1,2769	0,6983	1,8280
5 .....	-2,6213	1,4552	-1,8013
6 .....	-0,6263	1,3012	-0,4813

Análogamente puede definirse el cociente de verosimilitud de un resultado negativo,  $L(0)$ , que se relacionará con la probabilidad *a posteriori* de un resultado negativo ( $P | y = 0$ ):

$$P(\mathbf{E} | y = 0) = [1-P(\mathbf{E})]/\{[1-P(\mathbf{E})] + P(\mathbf{E})L(0)\}$$

El cociente de verosimilitud puede definirse asimismo para variables continuas y es fácilmente generalizable para el caso de varias variables, multivariado.

El modelo general, que abarca tanto variables dicotómicas como continuas (y/o la combinación de ambas), es la denominada relación exponencial:

$$L(x) = \exp(a_0 + a_1x_1 + \dots + a_kx_k)$$

donde  $a_0$  es un término independiente y  $\{a_1, \dots, a_k\}$  un conjunto de coeficientes a estimar a partir de la muestra de datos extraída de paciente y no-pacientes de la enfermedad  $\mathbf{E}$ ,  $x_i$  representan las diferentes variables incluidas en el modelo. Así, por ejemplo, en el caso de una prueba de resultado binario cuyos coeficientes de verosimilitud son  $L(1)$  y  $L(0)$ :

$$L(x) = \exp(a_0 + a_1x)$$

donde:

$$a_0 = \log[(1-Se)/Se]$$

$$a_1 = \log[SeEs/(1-Se)(1-Es)]$$

### Análisis discriminante

Cuando se trata de establecer un diagnóstico diferencial entre varios estados patológicos la técnica estadística más apropiada es el análisis discriminante. Esta técnica consiste en la deducción de las ponderaciones de una combinación de variables tal que optimicen la separación (en su significado más geométrico) de las poblaciones, de las cuales se han extraído las muestras de pacientes analizadas.

Cuando se trata de variables continuas, gaussianas, la técnica estadística se denomina análisis discriminante lineal porque de él resulta una combinación o ecuación lineal (es decir, que las potencias de las diferentes variables son 1). Cuando la distribución de una variable no es gaussiana puede transformarse por manipulación algebraica. Esta facilidad no

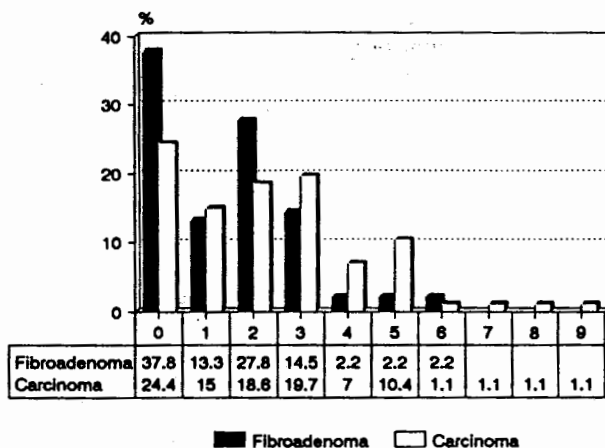


Fig. 3. Número de embarazos.

es tan evidente cuando se trata de una distribución multivariada. Y mucho menos si entre las diferentes variables las hay de tipo dicotómico. El modelo general, aplicable a estas situaciones, no es, por tanto, el modelo lineal, sino el modelo exponencial.

El modelo lineal se utiliza cuando la variable dependiente es una variable real, el modelo logístico se utiliza cuando la variable dependiente es una variable binaria o dicotómica. Existen varias posibles distribuciones cuando se tienen resultados dicotómicos. Sin embargo, el modelo logístico es particularmente importante por cuanto, desde el punto de vista matemático, es extraordinariamente flexible y desde el punto de vista biológico fácilmente interpretable.

La aproximación logística a la discriminación se basa en formular hipótesis sobre los cocientes de verosimilitud ( $L(x)$ ), antes que sobre los valores  $X$ , o sobre la distribución de probabilidad ( $P(x | E_i)$ ). Se parte del supuesto de que los cocientes de verosimilitud tienen una forma o estructura exponencial simple:

$$L_{ij}(x) = \exp(a_{oi} + a_i'x)$$

donde  $i = 1, \dots, g$ ;  $a_{oi}$  es la intercepción y  $a_i' = (a_{i1}, \dots, a_{ik})$  es el vector de ponderaciones asociado a las variables. Se ha demostrado que este modelo es aplicable a una gran variedad de situaciones: variables continuas, discretas y a la mezcla de ambas. La probabilidad posterior será ahora:

$$P(E | x) = \frac{\exp(a_{oi} + a_i'x)}{\sum \exp(a_{oj} + a_j'x)}$$

TABLA II  
ALOCACION DE CADA INDIVIDUO DESPUES  
DE APLICAR LA ECUACION DISCRIMINANTE

	Grupos reales	
	Fibroadenoma	Carcinoma
Grupos distribuidos:		
Fibroadenoma .....	79	60
Carcinoma .....	11	26
Total .....	90	86

Este modelo se denomina *modelo logístico generalizado* porque se trata de una generalización del modelo logístico clásico:

$$y(t) = \frac{\exp(t)}{1 + \exp(t)}$$

para más de 2 grupos.

La estimación de los coeficientes se hace mediante el método de la máxima verosimilitud, un proceso iterativo que requiere algoritmos informáticos relativamente sofisticados (concretamente utilizando el procedimiento de Newton-Raphson).

## MATERIAL Y METODOS

### Pacientes estudiados

Se analizaron todos los casos de tumoración diagnosticados en una unidad de patología mamaria durante un período de 4,5 años en el que fueron atendidas un total de 543 pacientes.

Las pacientes con tumoraciones pertenecen a 2 grupos:

*Grupo I.* Fibroadenomas. Se han incluido 81 mujeres, en 9 de las cuales se detectó la presencia de 2 fibroadenomas (5 casos en la misma mama y en 4 casos bilateral).

*Grupo II.* Carcinomas. Se incluyen 85 mujeres, una de las cuales con un carcinoma bilateral.

Los respectivos diagnósticos se confirmaron en todos los casos por mamografía y citología mediante punción-aspiración con aguja fina. Los fibroadenomas no intervenidos, por decisión de las pacientes, fueron 16 (19,7%). Los carcinomas no intervenidos

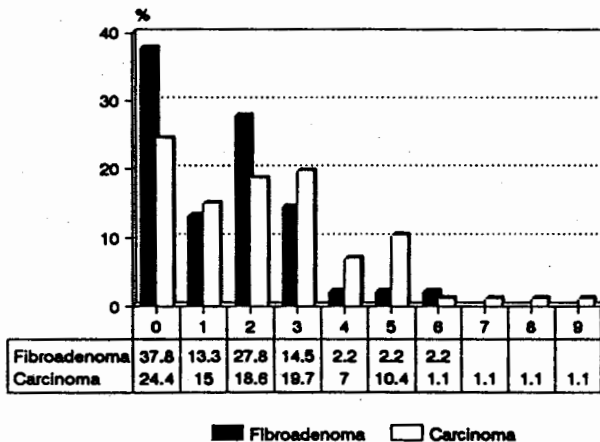


Fig. 4. Tamaño de las tumoraciones.

fueron 10 (11,2%) por corresponder a pacientes en estadio IV con metástasis a distancia.

#### Variables

Se han estudiado las siguientes variables:

- Edad (en años).
- Tamaño del tumor (en milímetros) medido una vez extirpado. En caso de no haberse extirpado a través del estudio radiológico.
- Antecedentes de cáncer de mama en familiares de primer grado.
- Número de embarazos.
- Mama en que se localiza la tumoración.
- Status menstrual: pre o postmenopausia.

Asimismo se ha estudiado la localización en la mama:

- Cuadrante superoexterno.
- Cuadrante superointerno.
- Cuadrante inferoexterno.
- Cuadrante inferointerno.
- Retroareolar.
- Prolongación axilar.

#### Procedimiento estadístico

El análisis de las tablas de contingencia 2x2 se realizó mediante la prueba de chi-cuadrado para 2 colas.

Para el cálculo del discriminante logístico se ha uti-

lizado el programa Logdis, para análisis discriminante logístico multidimensional, escrito en lenguaje Fortran compilado. En un ordenador personal IBM con coprocesador 8087 y 640 kb de RAM.

## RESULTADOS

### Localización de la lesión

En la figura 1 se indican el número de tumores en las diferentes localizaciones. Considerando el valor «1» como *presente* y «0» como *ausente*, se realizó el análisis discriminante logístico con las 6 variables correspondientes a las 6 posiciones descritas. La tabla I presenta los resultados encontrados: el vector de coeficientes  $\beta$ , el error estándar de cada coeficiente y las correspondientes variables estandarizadas:

$$z = \beta/es(\beta)$$

En la tabla II se presentan los resultados de aplicar la ecuación discriminante a los 176 casos. La prueba chi-cuadrado para comprobar la significación de esta tabla de contingencia 2x2 dio un valor igual a 16,35 con 6 grados de libertad, que corresponde a una probabilidad de 0,012.

### Otras variables

Se ha estudiado la influencia relativa de las siguientes variables: edad, tamaño del tumor, antecedentes patológicos, paridad, mama afectada y menopausia.

TABLA III  
COEFICIENTE ( $\beta$ ), ERROR ESTANDAR (es) Y SIGNIFICACION DE CADA VARIABLE (z ES EL COEFICIENTE ESTANDARIZADO)

Variable	$\beta$	es	z
Ter. ind. ....	6,5473	1,3283	4,9290
Edad .....	-0,1046	0,0314	-3,3336
Tamaño .....	-0,0334	0,0175	-1,9133
Anteced. ....	1,4147	1,0184	-1,3892
Paridad .....	-0,1868	0,1414	-1,3212
Mama .....	-0,1954	0,4803	-0,4069
Menopausia .....	-0,6835	0,9127	-0,7489

La edad promedio de las pacientes estudiadas fue de 46,75 años (con una desviación estándar de 20,14 años). El grupo de pacientes con fibroadenoma la edad media fue de 32,2 años y la de los carcinomas de 61,9. En el tamaño medio de los tumores fue de 26,52 mm, el promedio del tamaño de los fibroadenomas fue de 22,28 mm y de los carcinomas de 30,95 mm. En el caso del tamaño de los fibroadenomas, la desviación estándar es de 12,27 mm, mientras que la dispersión del tamaño de los carcinomas fue de 20,78 mm. El resto de las variables son de carácter binario.

La tabla III presenta los resultados encontrados: el vector de coeficientes  $\beta$ , el error estándar de cada coeficiente y las correspondientes variables estandarizadas.

En la tabla IV se presentan los resultados de aplicar la ecuación discriminante a los 176 casos. La prueba de chi-cuadrado para comprobar la significación de esta tabla de contingencia 2x2 dio un valor igual a 129,74 con 6 grados de libertad, que corresponde a una probabilidad inferior a  $10^{-5}$ .

#### Análisis univariado de tablas de contingencia

El análisis de las tablas de contingencia 2x2 de la variable *antecedentes familiares de carcinoma* resultó altamente significativa estadísticamente (chi-cuadrado = 8,882,  $p = 0,0028$ ).

El análisis de las tablas de contingencia 2x2 de la variable *menopausia* resultó altamente significativa estadísticamente (chi-cuadrado = 85,9058,  $p < 000001$ ).

El análisis de las tablas de contingencia 2x2 de la variable *situación del tumor* apenas resultó significativa estadísticamente (chi-cuadrado = 3,905,  $p = 0,0481$ ).

TABLA IV  
ALOCACION DE CADA INDIVIDUO DESPUES  
DE APLICAR EL ANALISIS DISCRIMINANTE

	Grupos reales	
	Fibroadenoma	Carcinoma
Grupos adjudicados:		
Fibroadenoma .....	77	16
Carcinoma .....	13	70
Total .....	90	86

## DISCUSION

Los modelos logísticos han demostrado la utilidad de esta técnica estadística en epidemiología, así como en otras ramas de la ciencia. Así, por ejemplo, en estudios de medicina preventiva se ha presentado como alternativa con ventajas sobre el estudio empírico de los factores de riesgo. En el presente trabajo se ha aplicado el análisis discriminante logístico a la distribución del tumor y en segundo lugar a un conjunto de variables que corresponden a observaciones clínicas ampliamente reconocidas y fácilmente identificables como factores de riesgo para el cáncer de mama.

#### Localización de las lesiones tumorales

La localización de la lesión tumoral se ha distribuido por cuadrantes. Sin embargo, en algunos casos las lesiones ocupaban más de un cuadrante. Entre los fibroadenomas hay uno que ocupa toda la mama y otro que ocupa 2 cuadrantes. Entre los carcinomas hay 4 casos que ocupan 2 cuadrantes y otros 2 que ocupan 3 cuadrantes. Estos casos se codifican como *presencia* en las respectivas variables. La distribución de la localización de los tumores, el predominio en los cuadrantes superoexternos de los carcinomas coincide con estudios anteriores (tabla V).

El análisis discriminante que utiliza como variables las diferentes localizaciones de los tumores presenta unos resultados poco resolutivos. Una forma simple de comprobar la efectividad de la ecuación discriminante es estudiar la matriz o tabla de reclasificación

TABLA V  
DISTRIBUCION DE LOS CARCINOMAS  
POR CUADRANTES

	n	1 (%)	2 (%)	3 (%)	4 (%)	5 (%)
Truscott (1947) .....	836	50,0	12,0	20,0	5,0	13,0
Hamett (1948) .....	2.131	43,0	9,6	13,3	4,4	11,0
Smithers (1952) .....	662	47,7	8,8	14,8	6,0	22,8
Donegan (1967) .....	764	48,0	11,0	16,0	6,0	17,0
Haagensen (1972) ..	711	40,0	10,1	15,8	7,3	26,7
Presente estudio (1990) .....	86	46,0	16,0	20,0	9,0	9,0

TABLA VI  
ANTECEDENTES FAMILIARES DE CARCINOMA

	Fibroadenoma	Carcinoma
Sí .....	2	14
No .....	88	72
Total .....	90	86

obtenida al aplicar a cada caso la ecuación discriminante. Si bien los fibroadenomas son reclasificados aceptablemente (un 87,78% de fibroadenomas son reclasificados como tales), hecho que hace que la significación del análisis chi-cuadrado de la tabla de contingencia 2x2 (tabla II) sea significativo estadísticamente. No obstante, sólo un 30,23% de los carcinomas son reconocidos como tales, mientras que 69,77% de carcinomas son confundidos con fibroadenomas. El rendimiento total es de tan sólo el 59,66% de reclasificaciones correctas.

La localización del tumor en la mama tiene poca importancia en el diagnóstico y el pronóstico de la enfermedad. La importancia en ambos grupos radica en la relación con el abordaje quirúrgico. En los carcinomas la situación del tumor no es un factor relevante en la decisión entre el tratamiento conservador y la mastectomía, aunque sí es interesante en las indicaciones de tratamientos complementarios, como por ejemplo la irradiación de la cadena ganglionar de la mamaria interna en los tumores de cuadrantes internos.

### Otras variables

#### Edad

La distribución por edades dentro de cada grupo de la muestra estudiada no constituye la distribución habitual. Un gran número de fibroadenomas han sido diagnosticados por encima de la edad de aparición habitual. Así, un elevado porcentaje por encima de los 40 años, grupo de edades en que la incidencia se considera baja. Por su parte, entre los carcinomas aparece diagnosticado un considerable porcentaje de mujeres de edades superiores a los 45 y 59 años, grupo de edad en que la incidencia de carcinoma es máxima. Es decir, las curvas de edades de ambos

TABLA VII  
MENOPAUSA

	Fibroadenoma	Carcinoma
Sí .....	7	67
No .....	83	17
Total .....	90	86

grupos se encuentran desplazadas a la derecha. En el retraso de diagnóstico influyen factores referentes a la propia paciente, así como factores de tipo sanitario.

#### Tamaño del tumor

El tamaño del tumor, considerado como un factor aislado, no es orientativo para el diagnóstico diferencial entre ambos grupos, así como tampoco lo es la evolución del tamaño, ya que tanto en procesos benignos como malignos existen tumores de evolución rápida y otros de evolución lenta. La valoración del tamaño es fundamental para la clasificación TNM de los procesos malignos, con las consecuentes implicaciones pronósticas y terapéuticas.

De hecho, en los grupos formados según el tamaño, en ambos diagnósticos se encuentran proporciones muy semejantes (fig. 4). Entre el 80 y el 90% de las tumoraciones tienen un tamaño entre 10 y 50 mm (T1-T2). Entre los fibroadenomas un pequeño grupo de menos de 10 mm y otro de mayores de 50 mm que corresponden a tumores de crecimiento rápido (fibroadenoma filodes, etc.). Entre los carcinomas de más de 50 mm hay algunos de crecimiento rápido (en mujeres jóvenes); sin embargo, la mayoría suelen ser por retraso en el diagnóstico.

#### Menopausia y paridad

Los factores hormonales y reproductivos han sido claramente identificados como factores de riesgo de la aparición de carcinoma de mama. Existen dudas acerca de la importancia de otros factores, como paridad y edad del primer parto, y controversia acerca de factores, como las hormonas exógenas, nutrición y dieta, etc.



TABLA VIII  
SITUACION DEL TUMOR

	Fibroadenoma	Carcinoma
Mama D .....	50	34
Mama I .....	40	52
Total .....	90	86

La falta de significación estadística que aparece en el estudio de estos signos debe ser interpretada en el contexto relativo en que se han incluido. Es decir, que no deben ser objeto de comparación con estudios epidemiológicos en donde se establece diagnóstico diferencial de tumor con ausencia de tumor.

Según se aprecia en la tabla III, tan sólo las 2 primeras variables tienen un coeficiente estandarizado  $z_i$  estadísticamente significativo:  $-3,33$  y  $-1,91$ , respectivamente, para la edad y tamaño del tumor, siendo irrelevante la participación de los antecedentes patológicos, paridad, mama afectada y menopausia.

La matriz de reclasificación (tabla IV) indica cómo los fibroadenomas son clasificados correctamente en un 85,56%, mientras que los carcinomas se reclasifican correctamente en un 81,40%. El éxito global de la aplicación de la ecuación discriminante sería del 83,52%.

### Análisis de las tablas de contingencia

El análisis de las tablas de contingencia se hace para ver si un determinado criterio es capaz de afectar a la proporción de observaciones que se observan en dos categorías distintas. Alternativamente, el análisis de tablas de contingencia pretende ver si 2 muestras extraídas de distintas poblaciones tienen la misma o diferente proporción de observaciones de una cierta categoría. Según esta última explicación, en las 3 variables estudiadas, las categorías eran, respectivamente, la existencia o no de antecedente familiar de carcinoma, el ser o no menopáusica y la situación del tumor en la mama derecha o en la mama izquierda. Las poblaciones son los fibroadenomas y los carcinomas. Se ha utilizado la interpretación de 2 colas porque se estudia las diferencias entre proporciones, sin establecer prejuicios acerca del sentido de tales diferencias.

Según este estudio, obviamente no puede hablarse de preferencia por una determinada situación del tumor, ya que la significación estadística de la tabla de situación del tumor es aproximadamente el límite de significación habitual ( $p = 0,05$ ). A pesar de que una diferencia del orden del 11,1% de los fibroadenomas se sitúa en la mama derecha y la diferencia entre los carcinomas en la mama izquierda era de casi un 20% sobre los carcinomas localizados en la mama derecha.

La existencia de antecedentes familiares de carcinoma como criterio es altamente significativa, demostrando que tal criterio introduce un factor de diferenciación entre ambos grupos patológicos. Este resultado coincide con otros datos de la literatura que otorgan un mayor riesgo a aquellas mujeres que tienen historia familiar de carcinoma.

La situación hormonal de la mujer ha sido considerada por muchos autores como uno de los factores de riesgo más importante de cáncer de mama. Las hormonas sexuales esteroideas, especialmente los estrógenos, han sido implicadas en la etiología del cáncer de mama. La consideración de la menopausia es altamente significativa. Según se aprecia en la correspondiente tabla de contingencia, sólo un 8% de los fibroadenomas fueron descubiertos en pacientes preclimáticas, mientras que un 78% de las mujeres con carcinoma habían pasado el climaterio. Esto se refleja también en la distribución de edades (fig. 2) y explica la importancia relativa de la edad en la segunda ecuación discriminante.

### RESUMEN

Se presenta un estudio sobre diagnóstico diferencial entre tumoraciones sólidas mamarias, calculando el valor semiológico, de forma separada y conjunta, de los signos observados en la anamnesis y en la exploración clínica. Las técnicas estadísticas utilizadas son el análisis discriminante y el análisis de tablas de contingencia.

El análisis discriminante que utiliza como variables las diferentes localizaciones de los tumores presenta unos resultados poco resolutivos. Si bien los fibroadenomas son reclasificados aceptablemente (87,78%), sólo un 30,23% de los carcinomas son reconocidos como tales. El rendimiento total es de tan sólo 59,66% de reclasificaciones correctas.

En el estudio de las otras variables sólo la edad y

el tamaño del tumor tienen un coeficiente estandarizado estadísticamente significativo, siendo irrelevante la participación de los antecedentes familiares, la paridad, la mama afectada y la menopausia. La matriz de reclasificación clasifica correctamente a los fibroadenomas en un 85,56% y a los carcinomas en un 81,40%. El rendimiento total es del 83,51%.

Aplicando el análisis de las tablas de contingencia no puede hablarse de preferencia por una determinada localización del tumor según las patologías. En cambio, la existencia de antecedentes familiares de carcinoma introduce un factor de diferenciación de ambos grupos patológicos. La menopausia también es un factor altamente significativo.

## REFERENCIAS

- Albert A, Harris EK. Multivariate interpretation of clinical laboratory data. New York: Marcel Dekker, 1987.
- Albert A, Anderson JA. On the existence of maximum likelihood estimates in logistic regression models. *Biometrika* 1984; 71: 1-10.
- Anderson JA. Separate sample logistic discrimination. *Biometrika* 1972; 59: 19-35.
- Armitage P, Gehan EP. Statistical methods for the identification of prognostic factors. *Int J Cancer* 1974; 13: 16-33.
- Cuzik J, Wang DY, Bulbrook RD. The prevention of breast cancer. *Lancet* 1986; 2: 83-86.
- Day NE, Kerridge DF. A general maximum likelihood discriminant. *Biometrics* 1967; 23: 313-323.
- Dunn JE. Epidemiology and possible identification of high risk groups that could develop cancer of the breast. *Cancer* 1969; 23: 775-780.
- Dupont WD, Page DL. Risk factors for breast cancer in women with proliferative breast disease. *N Engl J Med* 1985; 312: 146-151.
- Fernández Cid M. Patología mamaria. Barcelona: Salvat, 1982.
- Haagensen CD. Diseases of the breast. 3.ª edición. Philadelphia: Saunders, 1986.
- Helmrich SP, Shapiro S, Rosemberg L. Risk factors for breast cancer. *Am J Epidemiol* 1983; 117: 35-45.
- Herbert JR, Agustine A, Baron EJ, Kabat GC, Kinne DW, Wynder EL. Weight, height and body mass index in the prognosis of breast cancer: early results of a prospective study. *Int J Cancer* 1988; 42: 315-318.
- Hosmer DW, Lemeshow SL. Applied logistic regression. New York: John Wiley, 1989.
- Jawed Iqbal M, Taylor W. Hormonal and reproductive factors. New evidence. En: Stott BA, ed.: Women at high risk to breast cancer. Dordrecht: Kluwer Academic Pub., 1989; 41-46.
- Kelsey JL. A review of the epidemiology of human breast cancer. *Epidemiol Rev* 1979; 74-109.
- Lynch HT, Marcus JN, Watson P, Lynch JF. Familial and genetic factors. New evidence. En: Stott BA, ed.: Women at high risk to breast cancer. Dordrecht: Kluwer Academic Pub., 1989; 27-39.
- Negri E, La Vecchia C, Brui P, Dardanoni G, Decarli A, Palli D, Parazzini F, Del-Turco MR. Risk factors for breast cancer: pooled results from three Italian case-control studies. *Am J Epidemiol* 1988; 128: 1207-1215.
- Negri E, Decarli A, Laveccia C, Marubini E. Identification of high risk groups for breast cancer by means of logistic models. *J Clin Epidemiol* 1990; 43: 413-418.
- Ottman R, Pike MC, King MC, Casagrande JT, Henderson BE. Familial breast cancer in a population-based series. *Amer J Epidemiol* 1986; 123: 15-21.
- Prats M. Tumores benignos. Fibroadenoma. Diploma Universitario de Patología Mamaria. Barcelona: Universidad de Barcelona, 1990.
- Solé M. Citologías técnicas. Diploma Universitario de Patología Mamaria. Barcelona: Universidad de Barcelona, 1990.
- Velasco M. Estudio radiológico de la mama. Semiología radiológica. Diploma Universitario de Patología Mamaria. Barcelona: Universidad de Barcelona, 1990.