

# **Modelos de riesgo cardiovascular.**

## **Estudio de Framingham. Proyecto SCORE.**

### **Introducción**

Los datos no tienen significado por sí mismos, sino en relación a un modelo conceptual del fenómeno que los produce. Las manzanas venían cayéndose de los árboles mucho antes de que Newton estableciese las ecuaciones del movimiento de caída libre de los cuerpos; la circulación de la sangre se comportaba de igual forma antes y después de Harvey; las estrellas y planetas se movían en el espacio de igual manera en los tiempos de Ptolomeo, de Galileo, de Copérnico, de Kepler, del mismo Newton, o de Einstein. Los datos relativos al movimiento de los planetas recogidos en las diferentes épocas, salvo por diferencias en la capacidad de la instrumentación existente en el momento, eran similares; lo que fue evolucionando a lo largo del tiempo es el modelo conceptual establecido por cada científico para explicar el fenómeno.

Aunque en la historia de la ciencia, desde un punto de vista formal, durante mucho tiempo solo existieron modelos de tipo determinista, en los que el fenómeno se expresaba mediante leyes matemáticas perfectamente formuladas, de tal manera que conocidas las variables que intervenían en el modelo y su valor, el resultado quedaba completamente determinado, y solo más recientemente se empezó a plantear modelos de tipo probabilístico, en los que conocidas las variables únicamente se calculaba la probabilidad de aparición de un resultado, y en los que se introducía por tanto un margen de incertidumbre, realmente este tipo de modelos probabilísticos ha venido siendo utilizado desde siempre, sin una formulación matemática expresa, por el ser humano, quién en su toma de decisiones se ha basado en la estimación de la probabilidad de que algo ocurra en base a lo que ha observado que ocurrió con anterioridad en situaciones similares, aunque lo que a veces nos dicta lo que conocemos como "sentido común" se ha revelado en muchas ocasiones como contrario a la realidad científica.

La construcción de modelos de riesgo de aparición de un suceso es de gran importancia en medicina, tanto para intentar conocer las variables que influyen en que se presente ese suceso, como para analizar el mecanismo que lo produce y para predecir su aparición. En el primer caso, el conocimiento de las variables que influyen nos permitirá establecer medidas preventivas o terapéuticas, y en el segundo mediante el modelo podemos efectuar cálculos relacionados con la aparición del suceso, por ejemplo para determinar las necesidades de recursos. Precisamente la teoría matemática para el cálculo de modelos de riesgo tiene su origen probablemente en este último aspecto, y más concretamente en el campo de la ingeniería, donde la demanda creciente de equipos que funcionen cada vez mejor y a un menor coste lleva aparejada la necesidad de disminuir las probabilidades de fallo de éstos, lo que posibilít

el estudio y desarrollo de modelos probabilísticos para analizar la naturaleza de esos fallos y minimizar así la probabilidad de que ocurran, lo que en ingeniería se conoce como [teoría de la fiabilidad](#) (*reliability*) y en medicina habitualmente como [análisis de supervivencia](#) (*survival analysis*).

Cualquier construcción matemática, por sencilla que ésta sea, constituye un modelo y como tal una simplificación de la realidad, útil pero simplificación al fin y al cabo. Así en términos de supervivencia es habitual utilizar la mediana como dato resumen. La mediana es el valor que deja el 50% de los datos a cada lado de la distribución. Evidentemente en este caso se trata de una simplificación tremenda y probablemente un mejor modelo de la realidad lo constituya una estimación de la supervivencia a lo largo del tiempo, quizás obtenida mediante el [método de Kaplan-Meier](#), lo cual no obstante sigue siendo una gran simplificación, ya que en ese modelo para el cálculo de la supervivencia solo interviene el tiempo y ninguna característica del paciente, las cuales sin ninguna duda pueden influir decisivamente en el resultado, por lo que el modelo se podrá mejorar incluyendo en el mismo el efecto de variables que se cree pueden afectar a la probabilidad de aparición del evento. No debiéramos olvidar nunca que, a diferencia de los modelos deterministas propios de las leyes físicas, los modelos biológicos son en su gran mayoría modelos probabilísticos, sujetos a incertidumbre, que además se trata de simplificaciones de la realidad y que efectúan cálculos generales para valores promedio, mientras que la práctica clínica se ejerce sobre pacientes concretos con sus características individuales.

En línea con esa idea recomiendo la lectura del artículo "[The median isn't the message](#)", escrito por [Stephen Jay Gould](#), prestigioso biólogo, que fue profesor en la Universidad de Harvard, a quien le diagnosticaron un cáncer abdominal, para el que en su inquietud comprobó en la literatura del momento que la mediana de supervivencia era de 8 meses. En sus reflexiones plasmadas en ese artículo, plantea que a la mayor parte de la gente si se le dice que la mediana de supervivencia en la enfermedad que padece es de 8 meses, lo que realmente oírán es que probablemente van a morir en 8 meses, y sin embargo, tal y como agudamente comenta en su artículo, hay que analizar el asunto con más detenimiento, puesto que de entrada hay al menos ya una versión positiva del dato: la mitad de los enfermos viven más de ocho meses. Precisamente hay que conocer más sobre la distribución de probabilidad, y sobre todo cómo es la cola del lado derecho de la distribución. Y también qué características individuales pueden influir en el resultado. En el caso concreto del Dr. Gould, vivió 20 años más después del diagnóstico y tratamiento de ese cáncer.

Los modelos matemáticos constituyen sin ninguna duda valiosísimas herramientas para el conocimiento, interpretación y en su caso modificación de los fenómenos, pero casi siempre se trata de modelos transitorios, sujetos a verificación y perfeccionamiento, y como todo en el mundo de la ciencia solo pueden ser aceptados con una cierta dosis de escepticismo y con una mentalidad crítica.

## Modelos de riesgo cardiovascular

Puesto que las enfermedades cardiovasculares constituyen una de las principales causas de mortalidad y morbilidad en los países desarrollados, es lógico que sea de gran interés el desarrollo de modelos de predicción del riesgo de padecer este tipo enfermedades, tanto para intentar conocer los posibles mecanismos que afectan al aumento del riesgo, como para poder intervenir precozmente, mediante campañas preventivas, o en su momento con tratamientos terapéuticos. Precisamente uno de los factores de riesgo que se asocian con la probabilidad de desarrollar una enfermedad cardiovascular es la presencia de hipertensión.

Para el cálculo de la probabilidad de aparición de un suceso dicotómico (enfermedad NO, SI) el modelo matemático más habitual se basa en la utilización de la [regresión logística](#), que produce una ecuación en la que conocidos los valores de los diferentes factores de riesgo se puede evaluar la probabilidad de aparición de la enfermedad. Resulta evidente que en muchos procesos dicha probabilidad depende del tiempo de exposición, aumentando a medida que éste transcurre, por lo que o bien el tiempo interviene en la ecuación como factor de riesgo, o bien se utiliza un modelo específico en el que se tenga en cuenta esta característica, calculando ahora la probabilidad de que el suceso ocurra en un momento de tiempo determinado. Esto es precisamente lo que se hace en los modelos probabilísticos de supervivencia, siendo el método más conocido el denominado [modelo de riesgos proporcionales o modelo de Cox](#). Sin embargo no es la única alternativa posible, existiendo otros posibles métodos de modelado denominados [paramétricos](#), debido a que suponen un tipo concreto de ecuación matemática para la función de riesgo, y que aunque en la industria son muy utilizados, sin embargo no es tan normal encontrarlos en la literatura médica, aunque como veremos más adelante, precisamente han sido utilizados recientemente en el modelo de riesgos propuesto por el proyecto [SCORE](#). Entre éstos métodos paramétricos se encuentran los modelos basados en la [función de Weibull](#).

Aunque existen gran número de trabajos relativos al estudio de los riesgos de enfermedad cardiovascular, el conocido como [estudio de Framingham](#) constituye un pilar básico, y en diferentes formas es ampliamente utilizado para la toma de decisiones terapéuticas en base a la estimación de riesgo proporcionada por el modelo al introducir las características de riesgo del paciente concreto. Es tan popular que incluso existen calculadoras de bolsillo que implementan el algoritmo, y también diferentes páginas Web en las que se puede efectuar dicho [cálculo](#).

Sin embargo [se venía observando](#) que el modelo de Framingham sobreestimaba en gran medida el riesgo absoluto de enfermedad cardiovascular cuando se utilizaba en países europeos, caracterizados por una baja incidencia de eventos cardiovasculares respecto al lugar de origen del estudio en Framingham Massachusetts USA, lo que podía influir al utilizar ese modelo en la decisión de tratar un exceso de pacientes en países como España o Italia, en base a una

sobreestimación del riesgo real. Como es lógico esta inquietud crea la necesidad de desarrollar un modelo más adecuado para este entorno, y así recientemente se ha publicado un trabajo correspondiente a la estimación del riesgo de desarrollar en 10 años una enfermedad cardiovascular fatal en países de Europa ([proyecto SCORE](#)).

Existen otros modelos además de los dos citados, entre los que cabe quizás destacar otro trabajo anterior que incluye pacientes de Norte América y Europa, basado en el [proyecto INDANA](#) (Individual Data Analysis of Antihypertensive Intervention Trials), y que también dispone de una [calculadora on-line](#) de riesgo.

### **Fórmulas para el cálculo del riesgo según el modelo de Framingham**

No vamos a entrar aquí a comentar las implicaciones clínicas que se derivan de estos modelos de riesgo, remitiendo al lector a los [enlaces](#) que más abajo se incluyen, pero dado que aquellos que están efectuando un trabajo de investigación les puede interesar disponer de las fórmulas matemáticas del modelo para programarlas en su base de datos u hoja de cálculo y poderlas aplicar a sus datos, vamos a proporcionar una rápida referencia del procedimiento de cálculo, extraído de los artículos correspondientes, que en el caso de Framingham corresponde a [Circulation. 1998;97:1837-1847](#).

---

#### ■ Cálculo del riesgo mediante el modelo de Framingham que utiliza el valor del **colesterol total**

Las variables que intervienen son el *SEXO*, la *EDAD* en años, el *COLESTEROL* sérico en mg/dl, fracción de colesterol ligado a lipoproteínas de alta densidad *HDL*, *PRESION SISTOLICA*, *DIABETES* (No, Sí), *FUMADOR* (No, Sí).

En primer lugar hay que calcular el valor de la siguiente expresión:

$$\begin{aligned} \text{Para los hombres } L_H &= b_{E1} \cdot EDAD + b_C + b_H + b_T + b_D + b_F \\ \text{Para las mujeres } L_M &= b_{E1} \cdot EDAD + b_{E2} \cdot EDAD^2 + b_C + b_H + b_T + b_D + b_F \end{aligned}$$

donde los coeficientes *b* son diferentes para hombres y mujeres y los obtenemos a partir de la siguiente tabla:

<b>Coefficientes para el modelo de Framingham (Colesterol total)</b>		
<b>Coefficiente</b>	<b>Hombres</b>	<b>Mujeres</b>
$b_{E1} \times \text{Edad}$	0.04826	0.33766
$b_{E2} \times (\text{Edad})^2$	0	-0.00268

<b>b<sub>C</sub> Colesterol mg/dl</b>		
< 160	-0.65945	-0.26138
160-199	0	0
200-239	0.17692	0.20771
240-279	0.50539	0.24385
≥ 280	0.65713	0.53513
<b>b<sub>H</sub> HDL-Col mg/dl</b>		
< 35	0.49744	0.84312
35 - 44	0.24310	0.37796
45 - 49	0	0.19785
50 - 59	-0.05107	0
≥ 60	-0.48660	-0.42951
<b>b<sub>T</sub> Tensión arterial mmHg</b>		
PAS < 120 PAD < 80	-0.00226	-0.53363
PAS <130 PAD < 85	0	0
PAS <140 PAD < 90	0.28320	-0.06773
PAS < 160 PAD < 100	0.52168	0.26288
PAS ≥160 PAD ≥100	0.61859	0.46573
<b>b<sub>D</sub> Diabetes</b>		
NO	0	0
SI	0.42839	0.59626
<b>b<sub>F</sub> Fumador</b>		
NO	0	0
SI	0.52337	0.29246

Una vez calculado el valor correspondiente de  $L$ , se le resta la cantidad  $G$  (función evaluada para los valores medios de las variables en el estudio) diferente para hombres o mujeres:

$$G_{\text{Hombres}} = 3.0975$$

$$G_{\text{Mujeres}} = 9.92545$$

Exponenciamos ese valor calculado  $B = \exp(L - G)$  y determinamos el valor de la expresión  $1 - S^B$ , donde  $S$  es (función de supervivencia base a 10 años), que es diferente para hombres y mujeres:

$$S_{\text{Hombres}} = 0.90015$$

$$S_{\text{Mujeres}} = 0.96246$$

También existe la posibilidad de calcular el riesgo mediante otro modelo que utiliza el valor de LDL-col en lugar del colesterol total, siendo la mecánica de cálculo similar aunque lógicamente varían los coeficientes. El procedimiento está también descrito en el artículo de [Circulation. 1998;97:1837-1847](#)

Obviamente estas expresiones no son para calcularlas cada vez sino para programarlas en una hoja de cálculo o en otro dispositivo.

**Ejemplo:** Valoraremos según este modelo el riesgo de una mujer de 57 años, con 238 mg/dl de colesterol, HDL 52 mg/dl, 150/92 mmHg, no fumadora y no diabética.

$$\begin{aligned}L &= 0.33766*57 - 0.00268*57^2 + 0.20771 + 0 + 0.26288 + 0 + 0 \\L &= 11.00989 \\B &= \exp(11.00989-9.92545) = \exp(1.08444) = 2.95778 \\R &= 1-0.96246^{2.95778} = \mathbf{0.107}\end{aligned}$$

Por lo tanto la probabilidad de evento cardiovascular a los 10 años según este modelo es de casi el **11 %** .

A efectos de comparación, en la siguiente tabla se proporciona el riesgo **medio** a 10 años de acuerdo al estudio de Framingham

Edad	Mujeres	Hombres
30 - 34	< 1 %	3 %
35 - 39	< 1 %	5 %
40 - 44	2 %	6 %
45 - 49	5 %	10 %
50 - 54	8 %	14 %
55 - 59	12 %	16 %
60 - 64	13 %	21 %
65 - 69	9 %	30 %
70 - 74	12 %	24 %

También podemos calcular el **riesgo ideal** para una edad y sexo, efectuando los cálculos en el modelo con esa edad y sexo y los mejores valores del resto de factores.

El proyecto [SCORE](#) se inició para desarrollar un sistema de evaluación de riesgo cardiovascular para su uso en Europa. En él se calcula, mediante un modelo basado en la función de Weibull, el riesgo de enfermedad cardiovascular fatal en 10 años, estimándose dos ecuaciones diferentes para enfermedades coronarias y no coronarias, y se dispone de dos métodos de evaluación que se diferencian en que en uno de ellos se utiliza el colesterol y en el otro la relación colesterol/HDL, aunque según los autores no hay ventajas aparentes en la utilización de uno u otro método. Aunque se trata de un modelo similar al de Framingham los conceptos utilizados difieren sensiblemente.

Hay que tener en cuenta que se calculan por separado los riesgos de enfermedad coronaria (EC) y no coronaria (ENC), por lo que el riesgo cardiovascular total corresponderá a la suma de ambos. Además se estiman dos modelos diferentes, denominados de alto y bajo riesgo, correspondientes a países con poblaciones con alto y bajo riesgo de enfermedad cardiovascular, estando en este último grupo de bajo riesgo países como España, Italia o Bélgica.

Aquí el procedimiento de cálculo es el siguiente:

En primer lugar se calcula la probabilidad de supervivencia base para EC y ENC para la edad actual del paciente y a los 10 años, de acuerdo a la siguiente ecuación:

$$S_0(Edad) = \exp\{-\exp(a) \cdot (Edad - 20)^p\}$$

$$S_0(Edad + 10) = \exp\{-\exp(a) \cdot (Edad - 10)^p\}$$

Los coeficientes  $a$  y  $p$  se obtienen de la siguiente tabla:

		Enf. Coronaria		Enf. No Coronaria	
		a	p	a	p
<b>Población de riesgo bajo</b>	<b>Hombre</b>	-22.1	4.71	-26.7	5.64
	<b>Mujer</b>	-29.8	6.36	-31.0	6.62
<b>Población de riesgo alto</b>	<b>Hombre</b>	-21.0	4.62	-25.7	5.47
	<b>Mujer</b>	-28.7	6.23	-30.0	6.42

Se calcula el valor de la siguiente ecuación para enfermedad coronaria y no coronaria:

$$w = b_C(\text{Colesterol} - 6) + b_T(\text{TAS} - 120) + b_F$$

donde los coeficientes se obtienen de la siguiente tabla

	Enf. Coronaria	Enf. No Coronaria
<b>b<sub>c</sub> Colesterol [mmol/l]</b>	0.24	0.02
<b>b<sub>T</sub> TAS [mmHg]</b>	0.018	0.022
<b>b<sub>F</sub> Fumador = SI</b>	0.71	0.63

El colesterol viene dado en **mmol/l**. Para convertir el valor de mg/dl a mmol/l basta con multiplicar por **0.02586**.

El siguiente paso consiste en calcular la probabilidad de supervivencia con esos factores de riesgo a esa edad y a 10 años:

$$S(Edad) = S_0(Edad)^{\exp(w)}$$

$$S(Edad + 10) = S_0(Edad + 10)^{\exp(w)}$$

Ahora para cada tipo de enfermedad calculamos la probabilidad de supervivencia a los 10 años condicionada a la supervivencia a la edad actual:

$$S_{10}(Edad) = S(Edad + 10) / S(Edad)$$

Siendo entonces el riesgo a 10 años:

$$Riesgo_{10} = 1 - S_{10}(Edad)$$

Así obtenemos dos valores de riesgo  $REC_{10}$  para **Enfermedad Coronaria**,  $RENC_{10}$  para **Enfermedad No Coronaria**. El riesgo total corresponderá a la suma de ambos.

**Ejemplo:** Valoraremos según este modelo el riesgo de una mujer de 57 años, con 238 mg/dl de colesterol = 6.15 mmol/l, 150 mmHg de TAS, y no fumadora, perteneciente a un país clasificado como de riesgo alto. Los coeficientes para el cálculo de la función base de supervivencia  $S_0$  son:

	a	p
<b>EC</b>	-28.7	6.23
<b>ENC</b>	-30	6.42

En las siguientes tablas damos los valores obtenidos en cada uno de los pasos

<b><math>S_0(57)</math></b>	<b><math>S_0(67)</math></b>
-----------------------------	-----------------------------



<b>EC</b>	0.99798	0.99107
<b>ENC</b>	0.99891	0.99493

	<b>w</b>
<b>EC</b>	0.576
<b>ENC</b>	0.663

	<b>S(57)</b>	<b>S(67)</b>	<b>S<sub>10</sub>(57)</b>	<b>Riesgo</b>
<b>EC</b>	0.99641	0.98417	0.98772	0.01228
<b>ENC</b>	0.99789	0.99018	0.99227	0.00773

Por lo tanto el riesgo total será  $0.01238+0.00773$ , lo que corresponde a un riesgo total de enfermedad cardiovascular a 10 años del **2 %**.

## Epílogo

Es evidente que aunque ambos modelos tienen por objeto valorar el posible riesgo de enfermedad cardiovascular de un sujeto de acuerdo a una serie de características, sin embargo presentan notables diferencias, no sólo en cuanto a las poblaciones estudiadas, sino también en cuanto a la clasificación del tipo de evento cuya probabilidad se pretende calcular, por lo que cualquiera que esté interesado en su utilización debiera como mínimo leer cuidadosamente los artículos originales que describen el modelo.

Personalmente me resulta curioso que un modelo como el de Framingham se haya popularizado tanto, hasta el punto de que, como ya se ha comentado, incluso existan calculadoras de bolsillo para determinarlo, similares a las que aparecieron con el cambio de moneda al euro. Por supuesto en dichas calculadoras no se explica cómo ha sido determinado el modelo, ni para qué tipo de evento cardiovascular se calcula la probabilidad, ni la sobreestimación que se obtiene para países como el nuestro del área mediterránea.

Para la correcta utilización de un modelo, es conveniente saber cómo ha sido obtenido, a qué población se aplica, sus posibilidades y también sus deficiencias. En esta época de alta tecnología sigue siendo útil mantener una mentalidad analítica y crítica sobre nuestros conocimientos y recordar que casi

siempre los modelos, sobre todos si son estadísticos, sólo son un paso más en el proceso de conocimiento. Quizás si miramos un poco al pasado podamos aprender de nuestros errores, y ello sin retroceder demasiado.

No hace demasiado tiempo que las operaciones de amígdalas eran una práctica habitual, cuando yo era niño. Pues bien, en 1930 la American Child Health Association efectuó un estudio en el que 1000 niños de 11 años elegidos al azar serían examinados por dos pediatras elegidos al azar de un grupo de 20 pediatras. De los 1000 niños 611 (61 %) ya habían sido operados de amígdalas antes de los 11 años. De los restantes 389 niños, los pediatras indicaron que era conveniente operar de amígdalas a 174 niños (45 %) y no era necesario en 215 niños. Estos 215 niños fueron sometidos a una segunda inspección por otro pediatra elegido al azar y resultaron elegidos para ser operados 99 niños (46 %), por lo que sólo 116 niños quedarían sin ser operados, y fueron a su vez sometidos a una tercera opinión, en la que de nuevo a 51 de ellos (44 %) les fue indicada la operación, quedando ya sólo 65 niños sin tratar.

¿Existía quizás una ley natural que determina que el 45 % de los niños vistos por un pediatra deberían ser operados de amígdalas? Lo que más me sorprende es que en 1930 yo no había nacido y mi hermana que es bastante más pequeña que yo fue operada de anginas alrededor de los 10 años. Evidentemente los conocimientos médicos en 1930 no son los actuales, pero el hábito excesivo de operar las amígdalas se mantuvo durante bastante tiempo. Es verdad que actualmente los conocimientos se difunden con rapidez, pero aun así es conveniente mantener una mente crítica con respecto a lo que creemos conocer. En concreto los modelos estadísticos no nos explican el por qué de los sucesos, si acaso podemos edificar una teoría a partir de ellos y por supuesto utilizarlos para predecir sucesos (como por ejemplo en meteorología) y emplear esas predicciones para anticiparnos a las mismas o si es posible corregirlas.

Los modelos estadísticos de regresión se utilizan para calcular la probabilidad de que un suceso ocurra, pero es difícil sin un fundamento teórico determinar qué variables se debe incluir en ese modelo. Es importante conocer hasta qué punto el modelo permite predecir el suceso conocidos los valores de los covariantes. En los modelos de regresión múltiple el coeficiente de correlación múltiple elevado al cuadrado  $R^2$  constituye una medida de la proporción de variabilidad total explicada por el modelo. El concepto de **variación explicada** se puede extender a otros modelos de regresión como la ganancia en precisión predictiva obtenida al incorporar las variables al modelo, frente a no utilizarlas. En el caso de modelos predictivos como los que nos han ocupado en este caso, esta medida tiene incluso más importancia que los parámetros de bondad de ajuste, ya que podemos tener modelos con un buen ajuste a nuestros datos pero de poca capacidad predictiva.

Ignoro si quienes emplean estos modelos de riesgo cardiovascular conocen su capacidad predictiva.