

Ronda clínica y epidemiológica

Oscar Osío Uribe¹, John Jairo Zuleta Tobón²

La predicción en medicina: modelos estadísticos para predecir

INTRODUCCIÓN

Nostradamus (1503-1566), o Michel de Nôtre-Dame, fue un médico y consultor astrológico francés que ha sido considerado uno de los más renombrados autores de profecías y predicción de eventos futuros. Su obra profética "Las verdaderas centurias astrológicas y profecías", fue publicada en 1555, anunciando todas las catástrofes del mundo, desde su época hasta el año 3797, fecha en que supuso que acontecería el fin del mundo. No se va a discutir aquí si estos augurios son exactos o son el resultado de traducciones o interpretaciones tendenciosas, pero sí a resaltar que algunas de las técnicas estadísticas que hoy se usan para predecir sucesos en medicina resultan, para algunas personas, de prácticas tan esotéricas como los vaticinios del médico francés.

El análisis de regresión (AR) es un área de la estadística que intenta predecir o estimar el valor de una variable dependiente (o de respuesta o de desenlace) a partir de los valores conocidos de una o más variables independientes (o explicativas). Variable es el término estadístico que se utiliza para referirse a cualquier característica del paciente. Si el análisis se hace a partir de una sola variable explicativa, el procedimiento se llama análisis de regresión simple (ARS) y cuando se hace usando una combinación de variables independientes se denomina análisis de regresión múltiple (ARM). Existen muchos modelos para hacer estas predicciones. Cuando la variable de respuesta es categórica binaria (por ejemplo tener o no tener infarto) el procedimiento más empleado en los estudios clínicos es el análisis de regresión logística (ARL). Cuando la variable de respuesta es continua (peso corporal o nivel de colesterol, por ejemplo) la técnica se denomina análisis de regresión lineal (ARlin). Tanto los ARL como los ARlin pueden ser simples o múltiples. Cuando se quiere predecir el tiempo que transcurrirá para producir un desenlace clínico, el modelo empleado con mayor frecuencia es el análisis de riesgos proporcionales (ARP) o análisis de regresión proporcional de Cox, que es un tipo de análisis de supervivencia usado cuando se quiere evaluar la relación existente entre dos o más variables explicativas continuas o categóricas y una sola variable continua de respuesta (el tiempo para la ocurrencia del evento clínico de interés). Como el desenlace no necesariamente ocurre en todos los sujetos durante el tiempo de observación, la técnica "censura" esas observaciones.

CONSTRUCCIÓN DEL MODELO

Los investigadores luego de recoger los datos de una o varias variables explicativas seleccionan las características que están más fuertemente asociadas con la variable de respuesta y escogen el modelo matemático que mejor explique la situación, expresándolo en una ecuación de regresión. El propósito básico de un ARM construido con fines predictivos es identificar cuál combinación de

¹ Especialista en Medicina Interna, Magíster en Epidemiología Clínica, Director de la Corporación Académica de Patologías Tropicales, Universidad de Antioquia, Medellín, Colombia. cpt_udea@yahoo.com, www.riesgoscervasculares.com

² Profesor de Ginecología y Obstetricia, Magíster en Epidemiología Clínica. Universidad de Antioquia, Medellín, Colombia.

variables predice mejor la variable de respuesta. Otra función de los AR es controlar los efectos de las variables explicativas que puedan confundir los resultados en las variables de respuesta; por ejemplo, pueden separar el efecto que tienen las variables "edad" y "sexo" de los pacientes en la supervivencia luego de una cirugía. Los modelos de regresión están estrechamente relacionados con los análisis de varianza (ANOVA), pero el modelo de regresión lineal es más fácil de desarrollar y de interpretar. En general, las regresiones lineales se emplean en los estudios de observación y los ANOVA más en los ensayos clínicos. Los AR también se usan para crear puntajes de riesgo que son capaces de convertir una ecuación matemática compleja en una sencilla suma de puntos, para explicar cómo las características de un individuo se combinan para producir un desenlace o para calcular la probabilidad de desarrollar una enfermedad o condición.

Generalmente, el médico asigna el pronóstico del paciente a partir de factores demográficos, hallazgos clínicos y tratamientos, y esto lo hace de una manera no cuantitativa e informal, mediante la mezcla de su experiencia y la información publicada. Antes de asumir un método más riguroso, el empleo de los modelos estadísticos predictivos, es necesario entender y evaluar sus características con el fin de emplearlos de una manera efectiva y evitar los errores más importantes que se cometen con ellos. Este objetivo se logra resolviendo ocho preguntas que se explican a continuación, tomadas de las recomendaciones de Braitman y Davidoff.¹

¿CÓMO EVALUAR ESTUDIOS QUE PRESENTEN UN MODELO PREDICTIVO?

¿Hubiera sido el paciente elegible para participar en el estudio que permitió generar el modelo?

Es obvio que no se pueden extrapolar los resultados de un estudio si las características del paciente difieren de las de quienes participaron en la investigación. Si el modelo se hizo para predecir la evolución de pacientes aún sin enfermedad cardiovascular, es imposible calcular con él las probabilidades de un paciente con antecedente de infarto. De manera similar, si de un estudio de estimación del riesgo futuro de enfermedad cardiovascular se excluyen pacientes con diabetes, no se podrá emplear el modelo para tomar decisiones con respecto a un paciente diabético.

¿El resultado evaluado en el modelo estadístico refleja el desenlace clínico que se quiere predecir?

Es importante tener en cuenta no solo que el resultado que se quiere evaluar para el paciente coincida con el seleccionado en el modelo sino que se debe tener en cuenta el tiempo con el que se trabajó en el modelo. Los modelos hacen una predicción de la ocurrencia de un evento a un determinado tiempo, que depende de la duración del estudio. Es lógico que si un modelo se generó para predecir enfermedad cardiovascular no se pueda emplear para predecir enfermedad cerebrovascular, aunque compartan factores. Igualmente, si el modelo se hizo para hacer predicciones a 5 años, no se puede emplear para estimar lo que va a suceder en 10 años.

¿Se cuenta con todas las características necesarias para aplicar el modelo?

Los modelos pueden tomar diferentes características de los pacientes y es necesario que se cuente con toda la información pertinente del paciente antes de aplicarlo. Lo ideal es que los modelos se construyan con características de los pacientes que sean fáciles de medir en la práctica diaria; por eso es mejor uno que incluya información clínica que uno que incluya múltiples mediciones de biomarcadores, en ocasiones difíciles de obtener, por ejemplo, marcadores tumorales, polimorfismos y mediciones bioquímicas complejas.

¿La estimación de la probabilidad de aplicar el modelo ayudará para el pronóstico, selección del tratamiento u otros aspectos importantes para el paciente?

Los modelos predictivos deben proporcionar información útil para tomar una decisión clínica, por ejemplo, iniciar o no un tratamiento, asumir o no los riesgos inherentes a un tratamiento. En este sentido, los modelos deben ser útiles para el médico, el paciente, la familia o el sistema.

¿Es suficientemente pequeña la incertidumbre de la probabilidad calculada para ayudar a hacer un pronóstico específico?

Al igual que con cualquier estimación estadística tomada de una muestra, es necesario conocer la precisión o incertidumbre de ese estimado puntual.

Dependiendo del tamaño de la muestra y de la variabilidad del resultado estudiado, las probabilidades serán más o menos precisas. Esto se evalúa con el intervalo de confianza, que proporciona unos límites, superior e inferior, entre los cuales se espera que esté la probabilidad real. Es mejor un modelo que prediga una probabilidad puntual del 35% con intervalo de confianza del 95% entre 37 y 41% que uno en que tal intervalo esté entre 5 y 63%. Es más fácil tomar una decisión en el primer caso que en el segundo cuando, por ejemplo, se vaya a decidir si se inicia un medicamento con efectos secundarios frecuentes e importantes y el intervalo informa que la probabilidad de la complicación a 5 años puede ser tan baja como 5 o tan alta como 63%.

Estas cinco primeras preguntas tienen que ver directamente con el paciente. Si la respuesta a alguna de ellas no se cumple, la recomendación es no aplicar el modelo. Si el paciente cumple con todas las condiciones para aplicárselo, lo siguiente es evaluar el modelo en sí y su construcción estadística. Esto se hace con las siguientes tres preguntas.

¿Estuvieron las probabilidades estimadas con el modelo lo suficientemente cerca de los resultados reales observados en el grupo de estudio?

Los modelos son construcciones teóricas y abstractas, pero se sustentan en datos objetivos y reales. Se debe evaluar qué tanto se acercan o se alejan las estimaciones calculadas a partir de los datos de lo que realmente sucedió en el grupo de estudio. En el estudio original se conoce el desenlace real que tuvo cada uno de los pacientes incluidos y, con base en el modelo estadístico construido, a cada paciente se le calcula la probabilidad teórica de tener el desenlace en estudio; a partir de la diferencia entre esta probabilidad y lo que sucedió en la realidad con los pacientes, se estima qué tanto se ajusta el modelo, análisis que se conoce como "prueba de bondad de ajuste". Posiblemente la más utilizada para este fin es la de Hosmer-Lemeshow y se debe tener cautela cuando los autores no la reportan.

¿Se evaluó el modelo con datos de un seguimiento prospectivo?

Los modelos se deben probar con datos de estudios prospectivos e idealmente, pero no es indispensable, se debe confirmar su utilidad y validez con un segundo grupo de pacientes, diferente al que sirvió para generarlo.

¿Es el modelo más preciso que el azar y que otros métodos tradicionales de predicción?

Es importante conocer el margen de error del modelo. Esto se logra con el índice c' que representa el porcentaje de pares de pacientes (uno con y otro sin el desenlace) en que el modelo predice correctamente una mayor probabilidad de tener el desenlace en los pacientes que realmente lo tuvieron, por ejemplo, la muerte o la recurrencia de una enfermedad. Un valor de 0,5 (50%) dice que el modelo no aporta nada, es igual a asignar el pronóstico por azar y en la medida en que más se acerque a uno, mayor es su poder predictor. El límite inferior del intervalo de confianza de este índice debe ser superior a 0,5. Por último, si existen otros modelos predictivos previos del mismo desenlace, es importante confirmar que el nuevo es efectivamente mejor que ellos, antes de fundamentar las decisiones y estimaciones en el nuevo.

Una última consideración: lo ideal es validar y confirmar la utilidad del modelo en las poblaciones en que se vaya a aplicar, porque lo que puede ser muy eficiente en una determinada población, podría no serlo en otra. Una limitación grande para el cumplimiento de esta recomendación es la dificultad de hacerlo, por factores económicos, de capacitación y, ante todo, por la exigencia de grandes tamaños de muestra.

A continuación se presenta un ejemplo de utilización de los modelos predictivos en la práctica clínica diaria.

MODELOS DE PREDICCIÓN DEL RIESGO CARDIOVASCULAR

Ancel Keys (1904-2004) entre 1920 y 1930 desarrolló ecuaciones matemáticas de regresión logística para predecir el largo de los peces con base en el peso. Esta fue la génesis de las ecuaciones para predecir el riesgo cardiovascular que se desarrollaron décadas más tarde. Gran parte de lo que sabemos de la enfermedad cardiovascular y de sus factores de riesgo proviene del estudio que viene realizando el Servicio de Salud Pública de los Estados Unidos en Framingham, una pequeña población del estado de Massachussets. En el Estudio Framingham del Corazón (*Framingham Heart Study*) se han enrolado, desde hace 60 años, tres generaciones de habitantes:

La primera en 1948, con una cohorte de 5.209 personas; la segunda, en 1971, con 5.124 descendientes de los primeros; en 2002 el seguimiento de 3.900 personas de la tercera generación. Antes del Estudio Framingham del Corazón no se había entendido bien el papel del colesterol y de otras condiciones en la enfermedad cardiovascular. El estudio Framingham señaló la fuerte asociación existente entre el colesterol y la enfermedad coronaria y el efecto protector del colesterol HDL. También fue este estudio el primero que contradujo la creencia existente hasta los años 70 del siglo pasado de que la hipertensión arterial (HTA) no era perjudicial para las mujeres y que los hombres ancianos podían tolerarla mejor. Además demostró que la presencia de filtro en el cigarrillo no protegía contra el daño que causaba su exposición. Framingham nos ha permitido conocer los factores de riesgo de la enfermedad coronaria, enumerados a continuación, mencionando en cada caso el año en el cual se estableció la asociación: cigarrillo (1970), colesterol, hipertensión arterial, alteraciones electrocardiográficas (1971), diabetes mellitus (1974), menopausia (1976), triglicéridos (1977), factores psicosociales (1978), personalidad tipo A (1988) y la hipertensión sistólica aislada (1988).

TABLAS Y PUNTAJES PARA ESTIMAR EL RIESGO CARDIOVASCULAR DE LAS PERSONAS

En los últimos años se han efectuado reuniones periódicas de expertos para revisar la jerarquía de las investigaciones publicadas sobre el tema del riesgo cardiovascular y luego dar recomendaciones en forma de guías de manejo de los factores de riesgo cardiovascular, principalmente del colesterol sanguíneo (tres consensos) y de la hipertensión arterial (siete paneles de expertos). La estrategia básica actual en el mundo es hacer un enfoque del riesgo cardiovascular global de los individuos, para lo cual existen varios métodos o escalas de riesgo cardiovascular. La ecuación derivada del estudio de Framingham es la usada por el *Tercer Panel de Tratamiento de Adultos (ATP III) del Programa Nacional Educativo de Colesterol de los EEUU*,² cuya más reciente versión fue publicada en 2001; es la más usada por los médicos colombianos y hace la predicción del riesgo mediante el uso de las variables de edad, sexo, colesterol total, colesterol HDL, presión arterial sistólica, consumo de cigarrillo, hipertensión arterial, con diferente asignación de puntos si está controlada o si está en un

tratamiento no efectivo. La ecuación de Framingham determina el riesgo a 10 años de los desenlaces duros de la enfermedad coronaria: infarto del miocardio fatal o no fatal. La primera versión del ATP fue publicada en 1988, la segunda cinco años más tarde, recomendando desde ese entonces que los pacientes con enfermedad coronaria ya manifestada clínicamente deberían llegar a cifras de colesterol LDL menores de 100 mg/dL.

El *Tercer Consenso de Expertos del Programa Nacional Educativo del Colesterol de los Estados Unidos* determinó que existen tres grandes categorías de riesgo cardiovascular, que los médicos debemos determinar para poder establecer cuál es la cifra “meta” de colesterol LDL a la cual debemos llevar los pacientes y poder escoger así el hipolipemiente que garantice la reducción de la fracción lipídica:

Riesgo alto (riesgo absoluto superior al 20% en los siguientes 10 años): es el de los diabéticos, las personas que ya han tenido una primera expresión de enfermedad coronaria, los insuficientes renales crónicos y los pacientes con “equivalentes coronarios” (isquemia cerebral transitoria, aneurismas de la aorta, enfermedad arterial periférica). Estos pacientes deben ser llevados a niveles de colesterol LDL menores de 100 mg/dL, e incluso algunos autores proponen reducirlo a cifras por debajo de 70 mg/dL si el paciente cae en la categoría de “muy alto riesgo”. Esta categoría de “muy alto riesgo” está dada por la coexistencia de varios factores mayores de riesgo, especialmente cuando coexisten con diabetes mellitus, cuando existen factores de riesgo muy pronunciados, sobre todo tabaquismo fuerte o cuando el paciente tiene un síndrome coronario agudo.³

Riesgo bajo (riesgo absoluto menor del 10% en los 10 años siguientes): se da en muy pocas personas; son los afortunados que no tienen ninguna condición de riesgo o tienen un solo factor que aumente su probabilidad de tener problemas cardiovasculares.

Riesgo moderado (riesgo absoluto entre 10 y 20% en los próximos 10 años): se presenta en los individuos en quienes coexisten dos o más factores de riesgo cardiovascular. Es la categoría en la cual caen los pacientes con síndrome metabólico y la mayoría de las mujeres de edad mediana. Se han hecho esfuerzos en los últimos años para mejorar la capacidad de las ecuaciones para predecir el riesgo en los individuos con dos o más situaciones de peligro; por ejemplo la Asociación Americana del Corazón

(American Heart Association) publicó en 2007 la *Guía para la prevención de la enfermedad cardiovascular en la mujer*,⁴ sin cambios en las categorías de riesgo alto y riesgo óptimo (bajo), pero creando una categoría nueva de “mujeres en riesgo”, constituida por aquellas que no caen en ninguna de las dos categorías anteriores y que previamente se clasificaban como de riesgo intermedio:

- Mujeres con aterosclerosis subclínica (por ejemplo: calcificaciones coronarias), síndrome metabólico, mala tolerancia al ejercicio o recuperación inadecuada de la frecuencia cardíaca luego de suspender la actividad física.
- Presencia de por lo menos uno de los siguientes factores de riesgo: antecedentes familiares de enfermedad cardiovascular prematura, tabaquismo, dieta inadecuada, hipertensión arterial, dislipidemia u obesidad, especialmente central.

Otra expresión del interés existente en mejorar los algoritmos de predicción en las mujeres mediante la incorporación de nuevos marcadores de riesgo son las tablas de riesgo de Reynolds. Mediante el seguimiento durante 10,2 años de 24.558 mujeres sanas mayores de 45 años, se estudiaron 35 factores nuevos de riesgo cardiovascular, evaluando la incidencia de diferentes expresiones de la enfermedad cardiovascular; se demostró que con este algoritmo se lograba reclasificar el riesgo en 40 a 50% de las mujeres de riesgo intermedio hacia categorías de menor o mayor riesgo que el predicho por Framingham. Los marcadores incluidos fueron: hemoglobina glicosilada, proteína C reactiva, lipoproteína a, apolipoproteínas A1 y B100, además de los antecedentes familiares.⁵ Recientemente⁶ Ridker demostró la validez de la incorporación de los nuevos marcadores de riesgo también en hombres. En 10.724

hombres no diabéticos la adición de los criterios “niveles de proteína C reactiva de alta sensibilidad” y “carga genética” (antecedentes familiares de infarto del miocardio en los padres menores de 60 años) mejora de manera importante el algoritmo predictor, logrando reclasificar el riesgo en el 20% de los hombres, mejorándolo o empeorándolo, con una mayor seguridad predictora y con una calibración superior.

REFERENCIAS BIBLIOGRÁFICAS

1. Braitman LE, Davidoff F. Predicting clinical states in individual patients. *Ann Intern Med* 1996; 125: 406-412.
2. Executive Summary of the Third Report of The National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III). *JAMA* 2001; 285: 2486-2497.
3. Grundy SM, Cleeman JJ, Merz CN, Brewer HB, Jr., Clark LT, Hunninghake DB, et al. Implications of recent clinical trials for the National Cholesterol Education Program Adult Treatment Panel III guidelines. *Circulation* 2004; 110: 227-239.
4. Mosca L, Banka CL, Benjamin EJ, Berra K, Bushnell C, Dolor RJ, et al. Evidence-based guidelines for cardiovascular disease prevention in women: 2007 update. *Circulation* 2007; 115: 1481-1501.
5. Ridker PM, Buring JE, Rifai N, Cook NR. Development and validation of improved algorithms for the assessment of global cardiovascular risk in women: the Reynolds Risk Score. *JAMA* 2007; 297: 611-619.
6. Ridker PM, Paynter NP, Rifai N, Gaziano JM, Cook NR. C-reactive protein and parental history improve global cardiovascular risk prediction: the Reynolds Risk Score for men. *Circulation* 2008; 118: 2243-2251.

