

Hablemos de...

Análisis multivariable en estudios pronósticos: problemas y propuestas

JUAN GONZÁLEZ ABRALDES^a Y JUAN TURNES^b

^aLaboratorio de Hemodinámica Hepática. Servicio de Hepatología. Institut d'Investigacions Biomediques August Pi i Sunyer (IDIBAPS). Centro de Investigación Biomédica en Red de Enfermedades Hepáticas y Digestivas (Ciberehd). Hospital Clínic. Barcelona. España.

^bServicio de Aparato Digestivo. Complejo Hospitalario de Pontevedra. Pontevedra. España.



Roger Ballabrera

Puntos clave

El análisis multivariable es un instrumento clave en los estudios pronósticos, ya que permite evaluar la influencia ajustada de diferentes variables en la probabilidad de desarrollar un evento.

La intención del análisis multivariable (explicativo o predictivo) condiciona la estrategia del análisis.

Las técnicas automáticas de selección de variables pueden dar resultados anómalos si no se introducen restricciones en el análisis, especialmente en casos en los que hay colinealidad entre las variables.

No se recomienda la selección de variables mediante análisis univariado.

El método más eficiente para la validación interna de un modelo es el *bootstrapping*.

¿Por qué utilizamos el análisis multivariable en estudios pronósticos?

Cuando un clínico (p. ej., un hepatólogo) establece el pronóstico de un paciente (p. ej., con cirrosis), habitualmente dispone de un gran número de variables (bilirrubina, albúmina, creatinina, protrombina, sodio, colesterol, ascitis, encefalopatía, etc.) que pueden tener relación con la probabilidad de desarrollar un evento (p. ej., que el paciente muera). El hecho de considerar a la vez todos estos factores pronósticos puede parecer intuitivamente fácil, pero se pueden cometer errores importantes. Una fuente de error es que varias de estas variables están, comúnmente, correlacionadas, lo cual puede hacer que una vez considerado el valor pronóstico de una de ellas, una segunda variable carezca de valor pronóstico adicional. Por ejemplo, en un paciente con cirrosis, la presencia de una bilirrubina de 30 mg/dl se asocia a un muy mal pronóstico. Es posible que conocer el valor de la albúmina sérica del paciente (que habitualmente covaría de forma inversa con la bilirrubina en pacientes con cirrosis, es decir, cuanto mayor es la bilirrubina más baja es la albúmina, en términos estadísticos tiene colinealidad) no aporte información pronóstica adicional. El paciente probablemente tendrá albúmina baja. Si el clínico considera que el valor predictivo de la albúmina y el de la bilirrubina son independientes, puede sobrestimar o subestimar el mal pronóstico del paciente.

Para describir de forma adecuada el valor pronóstico *independiente* de estas variables, se necesita considerarlas de forma simultánea en modelos matemáticos. Estos modelos se derivan con análisis de regresión logística para una variable dependiente binaria (p. ej., vivo/muerto, regresión lineal con variables dependientes continuas, regresión de Cox para datos de supervivencia).

Volviendo al ejemplo inicial, si realizamos un análisis univariado de regresión para conocer la relación entre la bilirrubina, la albúmina (variables independientes o "exposición") y la probabilidad de haber muerto a los 3 meses (variable dependiente o "respuesta"), este análisis nos dará unos coeficientes que reflejan la relación "no ajustada" (es decir, considerada aisladamente) de cada una de estas variables sobre la probabilidad de morir. Por ejemplo, este análisis puede dar como resultado que, por cada unidad de aumento de la bilirrubina, la probabilidad de morir se multiplica por 2, y por cada unidad de disminución de la albúmina, la probabilidad de morir se multiplica por 3. Estos coeficientes no tienen en cuenta la colinealidad de las variables. Si realizamos una regresión multivariable, incluidas las 2 variables juntas, podría ser que sólo una de las 2 tuviese valor predictivo independiente. Por ejemplo, el coeficiente de la bilirrubina podría seguir siendo 2, pero el de la albúmina, uno (es decir, sin valor predictivo). Esto querría decir que, si se conoce la bilirrubina, la albúmina no añade valor pronóstico.

El análisis multivariable predictivo y explicativo

En un estudio de factores pronósticos, el análisis multivariable se puede utilizar con 2 intenciones. En el análisis con intención predictiva lo que se busca es predecir la variable dependiente (el evento) a partir de un conjunto de variables independientes (exposiciones). Esto se realiza generando un modelo (ecuación) de regresión. La otra vertiente, la explicativa, busca medir el efecto ajustado de una o más variables independientes (exposiciones) después de eliminar los efectos de un conjunto de variables de control. El análisis predictivo es útil para generar modelos pronósticos que permitan tomar decisiones en la práctica clínica. El análisis explicativo tiene un objetivo más conceptual, y busca, sobre todo, conocer mejor las enfermedades a partir de la descripción de factores pronósticos y posibles dianas terapéuticas.

El ejemplo siguiente trata de ilustrar estas diferencias. En un estudio sobre el pronóstico de la hemorragia aguda por varices esofágicas, uno de los objetivos puede ser elaborar un modelo pronóstico con variables disponibles de forma inmediata en urgencias para estratificar a los pacientes según su riesgo de fracaso terapéutico. Este modelo (que se obtendrá con un análisis multivariable predictivo) podría utilizarse para definir actitudes, como el ingreso en la unidad de cuidados intensivos o aplicar tratamientos diferentes según el riesgo. En este modelo pronóstico, probablemente no se incluirá la medición del gradiente de presión venosa hepático (GPVH), porque no es posible obtenerlo de forma inmediata a la llegada del paciente a urgencias. Pero otro objetivo del estudio puede ser conocer la relación entre el GPVH y el pronóstico. Esto puede ayudar a conocer mejor la fisiopatología de esta complicación, y puede establecer que la presión portal es una diana terapéutica importante en la hemorragia por varices y, por lo tanto, indicar que puede ser útil disminuir la presión portal. En este segundo análisis (regresión explicativa), se trataría de establecer la influencia del GPVH en el pronóstico, ajustada por otras variables pronósticas relevantes. Por ejemplo, un análisis univariado nos puede indicar que tener un GPVH > 20 mmHg se asocia con el riesgo de fracaso 8 veces superior a no tener este GPVH. Sin embargo, cuando realizamos un análisis multivariable, en el que se introducen variables que reflejen la gravedad de la hemorragia, muy probablemente tendremos un riesgo ajustado más bajo, por ejemplo de 4. Esto se explica porque los pacientes con GPVH > 20 tienen tendencia a tener hemorragias más graves. La *odds ratio* ajustada nos indica la asociación del GPVH > 20 con el riesgo de fracaso en pacientes con igual gravedad de la hemorragia. Es decir, uno de los principales objetivos del análisis explicativo es controlar en el análisis los factores de confusión.

La diferencia entre el análisis explicativo y predictivo es un matiz importante que condiciona la estrategia de análisis multivariable. Un buen ejemplo de estos matices se puede encontrar en el estudio de Ripoll et al¹.

En un estudio de factores pronósticos, el análisis de regresión multivariable permite considerar de forma simultánea el valor predictivo independiente de un conjunto de variables.

El análisis predictivo es útil para generar modelos pronósticos que permitan tomar decisiones en la práctica clínica. El análisis explicativo tiene un objetivo más conceptual, y busca sobre todo conocer mejor las enfermedades a partir de la descripción de factores pronósticos y posibles dianas terapéuticas.

Estrategias para construir modelos de regresión multivariables

Un concepto fundamental en el desarrollo de modelos pronósticos con análisis multivariable es que el análisis de una misma base de datos puede dar resultados diferentes, según la estrategia de análisis empleada, y que hay más de una estrategia y solución válidas (con frecuencia muchas). Las premisas que se deben cumplir son que el diseño del estudio y la recogida de datos hayan sido correctos, y que la estrategia de análisis sea estadísticamente correcta, clínicamente razonable, y que sea capaz de proporcionar una respuesta a la pregunta planteada en el estudio.

Regresión por pasos, hacia delante, hacia atrás

Uno de los métodos más populares y comúnmente empleados para construir modelos pronósticos es el método de regresión por pasos (*stepwise regression*). Con este método, se introducen todas las variables candidatas a ser predictores en un análisis automático que el programa estadístico realiza sin ninguna restricción. En el método por pasos hacia delante, el predictor más poderoso entra primero en el modelo. La siguiente variable que aporta más significación al modelo entra después, y así sucesivamente, hasta que no hay ninguna variable que haga una aportación significativa al modelo.

En el método por pasos hacia atrás, todas las variables se introducen inicialmente en el modelo. El programa estadístico prueba todos los modelos retirando cada vez una variable. La variable que explica menor proporción de variancia (la menos significativa) sale del modelo. El proceso se repite hasta que no se puede retirar ninguna variable, sin que ello produzca una reducción significativa en la proporción de variancia explicada por el modelo.

Este método, aunque ha sido ampliamente empleado para generar modelos pronósticos, tiene problemas importantes². El método de selección de variables por pasos es automático, lo realiza el programa estadístico sin supervisión del investigador y sin sentido clínico, es decir, sin restricciones. Ello puede ocasionar resultados anómalos, especialmente en los casos en que las variables introducidas en la regresión por pasos están muy correlacionadas (colinealidad), un hecho extremadamente frecuente en los estudios pronósticos. En un estudio reciente de nuestra unidad, nos encontramos con 2 ejemplos de este problema³. En este estudio buscábamos, por una parte, un modelo pronóstico en pacientes con cirrosis y hemorragia aguda por varices (la parte *predictiva* del estudio). Si introducíamos todas las variables disponibles sin ninguna restricción, el programa seleccionaba a la vez la creatinina y el nitrógeno ureico en sangre (BUN, del inglés *blood urea nitrogen*) (altamente correlacionadas) como variables pronósticas independientes. Llamativamente, los coeficientes indicaban que a mayor BUN, peor

pronóstico, y a mayor creatinina, mejor pronóstico. Este tipo de resultados, sin ningún sentido clínico, pueden ocurrir con frecuencia si se introducen, sin restricciones, variables con alta colinealidad.

Otro hallazgo del mismo estudio ilustra bien otro aspecto de este problema. En la parte *explicativa* del análisis, pretendíamos explorar si el grado de hipertensión portal (medido como GPVH) se asociaba con el pronóstico en un grupo determinado de pacientes. Un problema es que se conoce bien que el GPVH se correlaciona con el grado de insuficiencia hepática y, por lo tanto, con albúmina, bilirrubina, Child-Pugh, etc. Podría ocurrir que el GPVH tuviese valor predictivo independiente de la función hepática, pero que en una muestra particular de datos (siempre trabajamos con muestras “particulares” de datos), el azar hiciese que diversas variables que reflejan la función hepática se relacionasen de forma mínimamente más fuerte con el pronóstico de la hemorragia que el GPVH. Por tanto, en el análisis de regresión por pasos, estas variables podrían dejar fuera del modelo al GPVH, teniendo éste valor predictivo independiente. Este problema se puede solucionar, en parte, poniendo restricciones a la regresión (como se explica en el apartado siguiente). De hecho, metodólogos de gran prestigio sostienen que en estudios pronósticos no se debe realizar análisis multivariable por pasos sin poner restricciones a la regresión^{2,4}.

Regresión con restricciones

En esta estrategia de análisis, la regresión se realiza con restricciones, y se otorga una jerarquía a las variables basadas en la información previa disponible, en la plausibilidad biológica, o simplemente en la voluntad del investigador. Por ejemplo, en el caso anterior, hay datos en animales de experimentación y en estudios en humanos que indican que en la hemorragia por varices (causada por hipertensión portal), el grado de hipertensión portal influye en el pronóstico. Si el objetivo es evaluar el valor predictivo independiente del GPVH en la hemorragia por varices, podemos “forzar” que GPVH entre primero en el modelo, antes que cualquier prueba de función hepática que pueda estar correlacionada con él. Otras razones que pueden llevar a establecer una jerarquía entre variables, a la hora de confeccionar un modelo pronóstico (generalmente en análisis *predictivos*), pueden ser de índole práctica. Por ejemplo, si tenemos 2 variables pronósticas A y B muy correlacionadas, y B es más fácil de medir y más barata que A, podemos forzar la entrada en el modelo de B antes que A, para que, en caso de que sea un predictor igual de potente, la regresión favorezca la presencia en el modelo de B frente a A. Otra solución adoptada en ocasiones es definir a priori diversos modelos con diferentes variables candidatas, y evitar introducir las variables con colinealidades en el mismo modelo⁵. La mayoría de los programas estadísticos per-

Hay más de una estrategia y solución válidas (con frecuencia muchas) para analizar una misma base de datos. El método automático de selección de variables por pasos (*stepwise regression*) puede ofrecer resultados anómalos, especialmente en los casos en que las variables introducidas en la regresión estén correlacionadas. Es recomendable poner restricciones a la regresión, y otorgar algún tipo de jerarquía a las variables.

Al menos debe haber 10 episodios por cada variable independiente introducida en el análisis multivariable. La reducción del número de variables mediante un análisis univariado previo puede excluir variables pronósticas importantes, y no permite evaluar la presencia de variables de confusión. Una opción aceptable para reducir el número de variables candidatas es agruparlas en categorías.

miten realizar regresión multivariable con restricciones mediante la entrada de variables independientes por bloques.

Tamaño de la muestra y número de eventos

La mayoría de los autores coinciden en que al menos debe haber 10 eventos (no pacientes) por cada variable independiente introducida en el análisis multivariable^{6,7}. En el caso de que las variables analizadas presenten colinealidades importantes, el número de eventos por variable requerido sería mayor. Los problemas de disponer de pocos eventos en el análisis son, por una parte, que puede haber variables potencialmente importantes, pero que no alcanzan significación estadística y, por tanto, no se seleccionan como predictores independientes. Por otra parte, el hecho de introducir muchas variables con pocos eventos puede dar lugar a modelos inestables (es decir, mínimos cambios en los datos podrían cambiar radicalmente los resultados del análisis).

Disponer de pocos eventos y muchas variables pronósticas es la norma, más que la excepción, en los estudios pronósticos. Para limitar el impacto de este hecho en los resultados de la regresión, se puede recurrir a varias estrategias. Una de ellas es seleccionar primero un número reducido de variables a través de un análisis univariado inicial. Las que alcanzan un cierto grado de significación estadística ($p < 0,20$, $p < 0,10$ o $p < 0,05$) posteriormente se introducen en el análisis multivariable. Este método se desaconseja⁸, ya que no permite controlar apropiadamente los factores de confusión en el análisis, y las variables que no alcanzan significación en el análisis univariado pueden ser muy significativas y relevantes en el multivariable⁸. En el caso de recurrir a este método es recomendable, una vez generados los modelos finales, forzar todas las variables potencialmente importantes y evaluar si los coeficientes de regresión del modelo o sus intervalos de confianza cambian de forma relevante⁸⁻¹⁰. Por otra parte, muchas veces esto no soluciona el problema, ya que no es extraño que haya múltiples variables significativas en el análisis univariado.

Otra estrategia es agrupar las variables que reflejen información parecida en categorías². Después de ello, se realiza una regresión por separado en cada una de estas categorías, y se confía en que quede un número reducido de variables represente a cada grupo, y posteriormente se realiza el análisis multivariable final con los representantes de cada categoría. Por ejemplo, en un estudio en el que se evaluaba el pronóstico de la hemorragia por varices, las variables pronósticas se agruparon en variables relacionadas con la función hepática, variables asociadas con la gravedad de la hemorragia y variables demográficas³.

Reproducibilidad de los modelos

Uno de los problemas de los modelos pronósticos es que, por lo general, funcionan peor cuando se aplican a muestras independientes de pacientes que en la muestra origi-

nal. Esto ocurre incluso con los modelos que se han generado con una técnica estadística impecable. Esto responde a 2 tipos de causas: *a)* puede ocurrir que las asociaciones observadas entre las variables predictivas y el episodio ocurran por azar en la muestra concreta en la que se ha generado el modelo, y, por lo tanto, no representan asociaciones reales replicables, y *b)* los predictores obtenidos pueden ser idiosincrásicos o particulares de la población o el medio en el que se ha generado y no “funcionan” en un medio diferente.

El primero de los problemas (que representa la validez interna del modelo) se puede disminuir con técnicas estadísticas. La más simple es dividir la muestra en 2 partes: *a)* una para generar el modelo (generalmente dos tercios de la muestra), y *b)* otra (un tercio) para validarlo. Sin embargo, hay otros métodos más eficientes¹¹, que aunque anteriormente se consideraban muy engorrosos, resultan sencillos con las posibilidades de los ordenadores actuales. Éstos se conocen globalmente como técnicas de remuestreo y se tratan en la sección siguiente. Estos métodos validan internamente el modelo, pero si hay alguna idiosincrasia en la muestra, no solucionan el problema de la validez externa. Ésta sólo puede evaluarse probando el modelo en muestras de poblaciones independientes.

Validación interna de modelos con técnicas de remuestreo

Las 2 técnicas de remuestreo más utilizadas son el *jackknife* y el *bootstrapping*. En el *jackknife*, en una muestra de n pacientes se realizan n análisis con $(n-1)$ pacientes, y en cada uno de ellos se excluye un individuo distinto. Los resultados de cada análisis se validan en el paciente excluido.

La técnica más eficiente, sin embargo, es el *bootstrapping*¹¹. A partir de la muestra original de n pacientes, se genera un número determinado de muestras distintas (p. ej., 500) de tamaño n , y en cada ocasión se realiza un muestreo con reemplazamiento. El reemplazamiento implica que cada una de las 500 muestras será diferente, aunque provenga de la misma muestra original, ya que un paciente puede estar representado varias veces en una de las muestras de *bootstrapping*, y ninguna en otras. Posteriormente, se lleva a cabo el análisis multivariable de regresión en cada muestra. Es decir, la regresión se realiza 500 veces sobre muestras distintas, pero que proceden de la muestra original. Se considera que los predictores que se seleccionen en más del 50% de los modelos finales son predictores robustos, es decir, fiables¹². Aunque los resultados del análisis se estabilizan, generalmente, a partir de 100 repeticiones¹², con la potencia del software y de los ordenadores actuales no hay motivos para no realizar 500 o 1.000 análisis. Con esta técnica, también se pueden calcular coeficientes de regresión del modelo con sus intervalos de confianza, que son más realistas que los calculados en una única muestra.

Este método elimina muchas de las arbitrariedades del análisis multivariable, y proporciona un dato muy importante, como es la fiabilidad de una variable predictiva³. Por lo tanto, este método permite presentar los resultados del análisis multivariable no sólo en términos de magnitud del efecto, intervalo de confianza y significación de una variable predictiva, sino también su *fiabilidad* en el *bootstrapping*.

Las técnicas de remuestreo permiten realizar una validación interna de un modelo pronóstico. Entre ellas, el *bootstrapping* es la más eficiente. La validación externa sólo puede conseguirse a partir de la evaluación de su rendimiento en una muestra independiente de pacientes.

Bibliografía



●● Muy importante

■ Epidemiología

1. Ripoll C, Bañares R, Rincon D, Catalina MV, Lo Iacono O, Salcedo M, et al. Influence of hepatic venous pressure gradient on the prediction of survival of patients with cirrhosis in the MELD Era. *Hepatology*. 2005;42:793-801.
2. ●● Guyatt G. Determining prognosis and creating clinical decision rules. En: Haynes B, Sackett DL, Guyatt G, Tulliez M, editors. *Clinical Epidemiology*. Lippincott Williams & Wilkins, 2006. p. 323-55.
3. Abrales JG, Villanueva C, Bañares R, Aracil C, Catalina MV, Garcia-Pagan JC, et al. Value of GPVH predicting 5-day treatment failure in acute variceal bleeding. Comparison with clinical variables. *J Hepatol*. 2008;48:229-36.
4. Guyatt G, Haynes B, Sackett DL. Analyzing data. En: Haynes B, Sackett DL, Guyatt G, Tulliez M, editors. *Clinical Epidemiology*. Lippincott Williams & Wilkins; 2006. p. 446-60.
5. Ripoll C, Groszmann R, Garcia-Tsao G, Grace N, Burroughs A, Planas R, et al. Hepatic venous pressure gradient predicts clinical decompensation in patients with compensated cirrhosis. *Gastroenterology*. 2007;133:481-8.
6. ●● Concato J, Feinstein AR, Holford TR. The risk of determining risk with multivariable models. *Ann Intern Med*. 1993;118:201-10.
7. Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR. A simulation study of the number of events per variable in logistic regression analysis. *J Clin Epidemiol*. 1996;49:1373-9.
8. Sun GW, Shook TL, Kay GL. Inappropriate use of bivariable analysis to screen risk factors for use in multivariable analysis. *J Clin Epidemiol*. 1996;49:907-16.
9. Abrales JG, Tarantino I, Turnes J, Garcia-Pagan JC, Rodes J, Bosch J. Hemodynamic response to pharmacological treatment of portal hypertension and long-term prognosis of cirrhosis. *Hepatology*. 2003;37:902-8.
10. Turnes J, Garcia-Pagan JC, Abrales JG, Hernandez-Guerra M, Dell'Era A, Bosch J. Pharmacological reduction of portal pressure and long-term risk of first variceal bleeding in patients with cirrhosis. *Am J Gastroenterol*. 2006;101:506-12.
11. Steyerberg EW, Harrell FE Jr., Borsboom GJ, Eijkemans MJ, Vergouwe Y, Habbeema JD. Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. *J Clin Epidemiol*. 2001;54:774-81.
12. ●● Blackstone EH. Breaking down barriers: helpful breakthrough statistical methods you need to understand better. *J Thorac Cardiovasc Surg*. 2001;122:430-9.



Actualización Hemorragia digestiva alta no varicosa

Epidemiología, factores pronósticos y causas de
hemorragia digestiva alta no varicosa

Faust Feu Caballé

Medidas generales y tratamiento farmacológico

Virginia Hernández-Gea y Càndid Villanueva

Tratamiento endoscópico

Rafel Campo y Enric Brullet

Otras alternativas terapéuticas. Cirugía y angiografía

Silvia Rofin y Eduard M. Targarona

Implicaciones clínicas de la investigación básica

Interacción entre el hígado y el intestino en la cirrosis

Leticia Muñoz y Agustín Albillos

El lugar en terapéutica de...

Tratamiento de la hiponatremia en la cirrosis. Nuevos fármacos

Andrés Cárdenas y Pere Ginès

Revisión técnica diagnóstica

Tomografía computarizada en el diagnóstico de la obstrucción intestinal

Anna Darnell y Julio Martín

Ensayos clínicos y práctica clínica

Valor del gradiente de presión en el pronóstico de la cirrosis compensada

Salvador Agustín y Joan Genesca

Prevención de...

¿Deben usarse los antiinflamatorios no esteroideos profilácticamente
en los pacientes con poliposis adenomatosa familiar?

Ángel Ferrández Arenas y Sonia Mostacero Tapia

Ventana a otras especialidades

Aplicabilidad de los *scores* en medicina intensiva

José M. Nicolás y Pedro Castro

Hablemos de...

Síndrome de abstinencia en pacientes alcohólicos

María-Vega Catalina Rodríguez y María Senosiain Labiano