

Adolescentes y consumo nocivo de alcohol. Chile 2009: mirando a las políticas públicas

MARCELO VILLALÓN C.¹, CAROLL CUELLAR²

Proposing a data mining scaffolding for preventing harmful alcohol consumption in Chilean adolescents

¹Programa Doctorado en Salud Pública. Escuela de Salud Pública. Facultad de Medicina, Universidad de Chile, Santiago, Chile.

²Programa de Magíster en Bioestadística. Escuela de Salud Pública. Facultad de Medicina, Universidad de Chile, Santiago, Chile.

Recibido el 19 de abril de 2012, aceptado el 16 de abril de 2013.

Correspondencia a:
Dr. Marcelo Villalón C.
Av. Independencia 929,
Independencia, Santiago,
Chile.
Teléfono: 56 9 7107 3783
(celular)
Fax: 56 2 7322436
E-mail: mavillal@uchile.cl

Background: Adolescent alcohol and drug consumption are important public health problems in the Chilean young population. **Aim:** The purpose of this study was to examine the potential of a data mining approach in scaffolding policy making, using the particular case of differential risks of harmful alcohol consumption in adolescent students. **Material and Methods:** Index and control groups were composed by 7918 and 7138 participants respectively (drawn from a CONACE survey 2009), aged 16 ± 2 years, 52% male. Heavy drinking at last month was the independent variable. As dependent variables parenting style, peer group influence, age and sex were used. For data analysis, a data mining approach was applied (CART, SPSS version 15). **Results:** The peer group influence was the main discriminant variable in males and the total sample, proving to be the only relevant variable in the case of women. The results suggest how a data mining approach may be useful in order to develop a hard data scaffolding for making and implementing policies in general and policies addressing adolescent alcohol consumption in particular.

(Rev Med Chile 2013; 141: 644-651).

Key words: Adolescent; Alcohol drinking; Data mining.

En Chile el alcohol representa un riesgo y daño para la salud de primer orden. El porcentaje de adolescentes entre 7° básico y 1° medio que alcanza la embriaguez una o más veces al año alcanza hasta 26,2% en la Región Metropolitana (EMSE 2004-2005) con un consumo promedio semanal en población entre 15 y 25 años de 124,92 g de alcohol puro, sin diferencias por sexo (ENS 2009-2010, P.191 Tabla V.2.6.2.). Se constituye así el patrón de consumo de alcohol más nocivo entre los adolescentes, el *binge-drinking*¹⁻⁷ definido como el emborrachamiento debido al elevado consumo de alcohol en un breve período de tiempo⁶.

El problema está siendo reconocido como problema país, como queda refrendado en la inclusión del consumo de alcohol en adolescentes como

patología con Garantías Explícitas en Salud (GES); la formulación de una Estrategia Nacional Para la Prevención del Consumo Nocivo de Alcohol desde una perspectiva de Salud Pública⁴; el nuevo Servicio Nacional de Prevención y Rehabilitación de Drogas y Alcohol (SENDA, ex-CONACE) que refuerza la institucionalidad que se ocupa del consumo de sustancias, incluyendo el alcohol; y la menor tolerancia de los límites de alcohol en sangre para conductores con sanciones más severas.

Chile monitorea el problema del consumo de sustancias desde hace casi dos décadas de manera específica. Dispone de encuestas con representación nacional desarrolladas por SENDA que permiten acercarse de buena forma a conocer la evolución del consumo nocivo de alcohol entre los adolescentes⁸. El desarrollo mediante *Classification*

And Regression Trees (CART), técnica de creciente uso en salud, provee de una herramienta que permite la elaboración de perfiles y su representación gráfica para una comunicación más efectiva de resultados de investigación en esta área⁹ que favorezca una mejor relación entre la generación de evidencia en salud y la toma de decisiones. Como objetivo se establece la identificación de subpoblaciones con riesgo diferencial de consumo nocivo de alcohol en población adolescente como un modo de contribuir a mejorar el diagnóstico del problema y con ello su tratamiento.

Material y Método

Se desarrolló un estudio transversal utilizando la base de datos del VIII Estudio Nacional de Drogas en Población Escolar de Chile, realizado por el SENDA el año 2009. La muestra, representativa de un universo de 969.339 escolares, alcanzó a 1.512 colegios, 2.877 cursos y 48.980 alumnos, con un promedio de 17 alumnos por curso. Otros detalles del diseño metodológico general de esta encuesta y del muestreo han sido descritos y relacionados con alcohol en adolescentes chilenos por Iglesias et al¹⁰.

Luego de explorar los módulos sociodemográfico y de alcohol, se definió como variable respuesta la pregunta 61c del cuestionario aplicado ¿Cuántas veces te has emborrachado en los últimos 30 días?, recodificando la categoría “Ninguna” como “NO” y las restantes alternativas (“1 ó 2 veces”, “3 a 9 veces”, “10 a 19 veces”, “20 a 39 veces”, “Más de 40 veces”) como “SÍ”.

Como variables descriptoras se utilizaron las siguientes preguntas del cuestionario con las recodificaciones que se indican:

Pregunta 1. Sexo (variable dicotómica).

Pregunta 2. Edad (variable cuantitativa medida en años).

Pregunta 7. Después que sales del colegio o durante los fines de semana, ¿Cuántas veces ocurre que tu madre o tu padre no saben dónde estás? Variable categórica: Nunca o casi nunca, A veces no saben, Siempre o casi siempre. Variable no recodificada.

Pregunta 8. ¿Cuán atentos están tus padres (o alguno de ellos) respecto de lo que haces en el colegio?, variable categórica con niveles: Mucho, Bastante, Poco, Nada. Esta variable fue recodificada en dos categorías: SI equivalente a Mucho y Bastante y NO equivalente a Poco y Nada.

Pregunta 9. En general, ¿Tú dirías que tus padres (o alguno de ellos) conocen a tus amigos cercanos bastante, más o menos o poco? Variable categórica con niveles: Bastante, Más o menos, Poco. Esta variable fue recodificada en dos categorías: SI equivalente a Bastante y NO equivalente a Más o menos y Poco.

Pregunta 46. ¿Cuántos de tus amigos toman regularmente alcohol? Variable categórica con niveles: Ninguno, Menos de la mitad, La mitad, Más de la mitad, Todos o casi todos, No sabe. La variable fue recodificada en tres categorías: Menos de la mitad, La mitad o más de la mitad y No sabe.

Fueron excluidos del estudio los datos de quienes no contestaron la pregunta usada como variable dependiente (1,2%) y quienes declararon tener una edad de 12 años o menos y más de 19 años (0,4%).

Respecto de las variables seleccionadas para la elaboración del árbol, se privilegió una selección de variables ampliamente reconocidas como asociadas a la determinación de la conducta consumo de alcohol tipo “*binge-drinking*”, entre las que se reconocen variables asociadas a los padres y su estilo de crianza, grupos de pares-amigos, edad y sexo¹¹.

A continuación, y dado que el porcentaje de estudiantes que declararon haberse emborrachado en el último mes fue de 16% de la muestra global, se constituyó una submuestra equiparando consumidores y no consumidores. Es decir, se seleccionaron los estudiantes de la muestra total que declararon haberse emborrachado en el último mes ($n = 7.918$) y posteriormente, se tomó una muestra aleatoria de tamaño semejante ($n = 7.638$) de entre los estudiantes que declararon nunca haberse emborrachado durante el último mes. Así, la muestra útil quedó conformada por 15.357 estudiantes (30,8% de la muestra total) 52,4% hombres, con una edad promedio de $15,7 \pm 1,5$ DE.

Se utilizó la metodología CART^{12,13} que permite construir árboles de clasificación mediante la representación gráfica de una serie de reglas sobre las decisiones tomadas para asignar un valor de salida a una determinada entrada, mediante divisiones binarias sucesivas. Esto es, se divide cada nodo madre en dos nodos hijos homogéneos mediante la aplicación de respuestas sí/no en cada nodo de decisión. Las etapas para la generación del modelo son: seleccionar la variable que más dis-

crimina de acuerdo con una función de partición de la impureza de los datos, la repetición de esta partición hasta que los nodos hijos se consideren lo suficientemente puros para ser terminales, una poda del árbol resultante para evitar el sobreajuste del modelo y la validación de éste por medio de una muestra de comprobación. Las aplicaciones en salud de esta herramienta han aumentado de modo exponencial y se han empleado para el estudio de alcohol en adolescentes⁹⁻¹⁴.

El procedimiento de análisis contempló el uso del módulo de árboles de decisión del *software* SPSS 15.0, método de crecimiento del árbol por medio de CART, con análisis de sensibilidad basado en índice de diversidad de Gini¹⁵ y sistema de validación por división muestral (70% aprendizaje, 30% comprobación). Los otros parámetros utilizados fueron: control del tamaño del árbol (hoja con registros superiores a 100 y profundidad del árbol igual a 5), control del podado por reducción de error estándar, con una diferencia máxima aceptada de una unidad de error estándar entre un árbol y el podado¹⁵⁻¹⁶.

Resultados

La lectura del árbol se realiza verticalmente de arriba hacia abajo, identificando nodos de partición para las variables explicativas y su importancia relativa. Esto es, la variable que se constituye como primer nodo da cuenta de un mayor peso o relevancia en la explicación de la variable de interés (emborrachamiento) respecto a las más distantes. El indicador cuantitativo de esta distancia es lo que en la Figura se llama "mejora". La Figura 1 se refiere a la muestra total, mientras que las Figuras 2 y 3 corresponden a los árboles construidos para hombres y mujeres por separado.

Este árbol de clasificación predice si un estudiante se ha emborrachado en los últimos 30 días a partir de la cantidad de amigos que toman regularmente alcohol, si los padres no saben donde está y la edad. Se distingue un grupo de adolescentes que se emborracha cuando: (a) "A lo menos la mitad de sus amigos toma regularmente", sin que influyan significativamente otras variables predictoras, lo que establece un nodo terminal, mientras que un segundo grupo lo hace cuando (b) "Menos de la mitad de sus amigos bebe regularmente", "habitualmente sus padres no saben donde está" y tienen 17 ó 18 años. Este modelo clasificó a

los consumidores de alcohol con sensibilidad de 75,7% y especificidad de 72,8% con un nivel de confianza de 95%¹⁶.

A continuación, considerando los diferenciales biopsicosociales en el desarrollo de adolescentes chilenos¹⁷, se construyeron árboles de clasificación para hombres y mujeres (Figura 2 y 3).

La Figura 2 revela dos grupos de adolescentes varones que se han emborrachado en los últimos 30 días: (a) un primer grupo que se emborracha si "el número de amigos que toma regularmente alcohol es la mitad o más de la mitad", mientras que un segundo grupo lo hace (b) si "el número de amigos que toma regularmente alcohol es menos de la mitad", "los padres saben donde están sólo a veces/casi nunca/nunca" y tienen entre 16 y 19 años. Respecto a la precisión, el árbol clasifica correctamente al 58,6% de los adolescentes que NO se han emborrachado durante los últimos 30 días y al 80,4% de los que SÍ lo han hecho.

El árbol de clasificación de la Figura 3, predice si una estudiante mujer se ha emborrachado en los últimos 30 días exclusivamente a partir de la cantidad de amigos que toman regularmente alcohol. Esto es, si el número de amigos que toma regularmente alcohol es la mitad o más de la mitad, entonces SÍ se emborrachan. Respecto a la precisión, el árbol clasifica correctamente al 68,6% de los que NO se han emborrachado durante los últimos 30 días y al 71,3% de los que SÍ.

Se puede observar que la mitad de los hombres muestra una base decisional semejante a la de las mujeres, en las cuales sólo influiría si los amigos beben o no regularmente. La otra mitad de varones influyen además otros condicionantes, como la atención de los padres y la edad.

Discusión

Las personas definen los patrones conductuales (estilos de vida) durante la adolescencia. Son estos los que luego se constituirán en factores de riesgo o de protección para las enfermedades que determinan la mayor carga de enfermedad de la población chilena del siglo XXI, los grupos II y III de la CIE (ENS 2009-2010). Los modelos causales del consumo de alcohol muestran que elementos ambientales como el nivel de impuestos y precios del alcohol, el tipo de mercado que se define para su venta (estancos, mercado libre, densidad de puntos de venta, horarios, edades que facultan

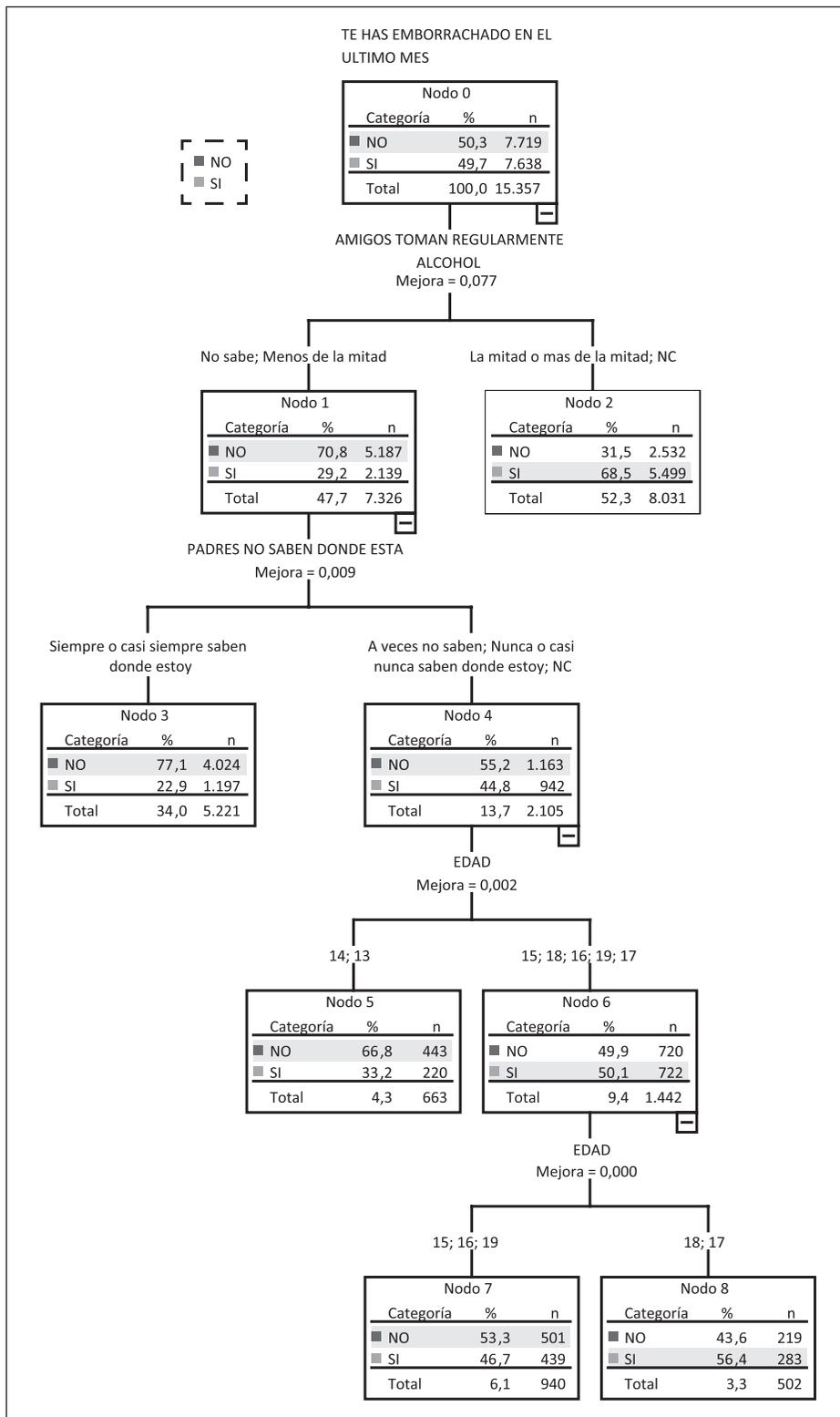


Figura 1. Árbol de Clasificación para Emborrachamiento en los últimos 30 días a través del algoritmo CART.

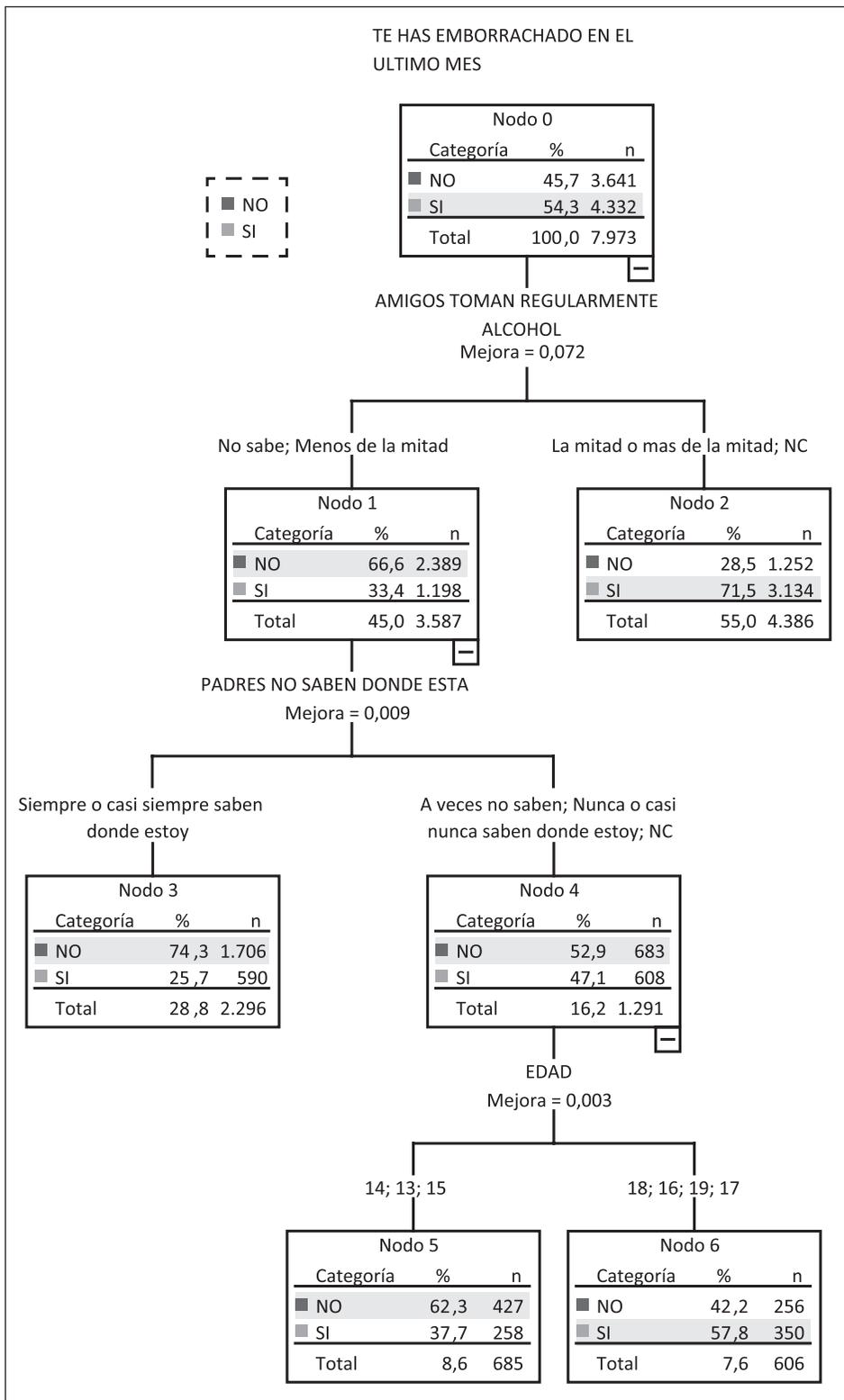


Figura 2. Árbol de Clasificación para emborrachamiento en los últimos 30 días hombres a través del algoritmo CART.

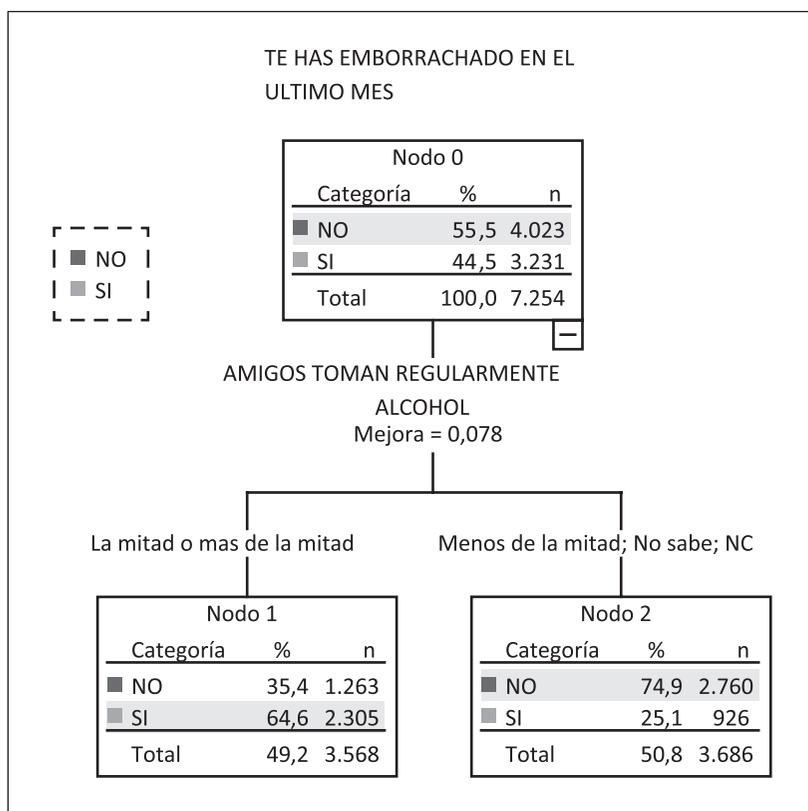


Figura 3. Árbol de Clasificación para Emborrachamiento en los últimos 30 días mujeres a través del algoritmo CART.

para su compra), están asociados y son importantes en la definición del modo de consumir alcohol¹⁸.

En Chile, sin embargo, la experiencia con el proceso de tabaco, problemática sanitaria análoga en muchos aspectos a la del alcohol, está mostrando que pese al cúmulo de evidencia irrefutable científicamente respecto del mayor impacto poblacional de este tipo de variables, las políticas implementadas difícilmente alcanzan variables de este nivel de complejidad. Así ha sucedido, por ejemplo, con la medida de mayor impacto probado en la reducción del consumo de tabaco, el aumento del precio vía impuestos, que se implementó post-terremoto, pero no por motivos sanitarios.

Restaría entonces ocuparse de elementos de segundo orden de impacto poblacional, pero con mayor factibilidad política de implementación. Entre estos, el estilo de crianza de los padres y la conducta de los grupos de pares han mostrado reducir el consumo nocivo de alcohol¹¹. En este ámbito de investigación se ubica el presente

estudio que ha buscado contribuir a identificar perfiles diferenciales de riesgo poblacional para el consumo nocivo de alcohol entre adolescentes con una herramienta, CART, que está teniendo una aplicación exponencial en la investigación en salud en la última década.

A diferencia de una aplicación reciente en el consumo de alcohol entre adolescentes⁹, en nuestro caso la precisión de CART se mostró mejor en el ámbito de la confirmación de la población que sí consume (75,7%; 80,4% y 71,3% para los árboles 1, 2 y 3, respectivamente). Si bien la validez del método aún es materia de discusión¹⁹, las ventajas de la herramienta en el ámbito de la mejor comunicación de los hallazgos científicos y su uso en la implementación de políticas parece relevante para fines de la salud pública.

Las dificultades en el manejo de un concepto que aparece como básico en epidemiología, el riesgo, resume de buen modo una complejidad que contribuye a las dificultades en el establecimiento de una relación más fecunda entre quienes trabaja-

mos en salud pública²⁰. Esto, entre epidemiólogos. Pero las barreras con base en las expectativas, la transferencia, la aceptación y la interpretación de la información generada desde la epidemiología y que debe servir para la acción son aún mayores²¹. Ha sido el reconocimiento de la naturaleza y complejidad de los procesos políticos en salud pública²² el que nos ha motivado a explorar el uso de herramientas que permitan ampliar el espacio de interacción entre quienes declaramos un objetivo común como es la mejoría de las condiciones sanitarias de la población.

Los objetivos sanitarios para la década 2010-2020 se propusieron metas desafiantes que podrán ser conseguidas en la medida que se consideren respuestas (políticas y programas) que reconozcan diferencias al interior de la población general. Esto supone la identificación de subgrupos con especificidades, o perfiles distintivos, respecto de una conducta de riesgo como el consumo nocivo de alcohol. La intervención breve, por ejemplo, pudiera requerir una implementación diferencial para varones y mujeres, si se atienden los distintos caminos de desarrollo psicosocial que estos resultados evidencian.

Los resultados en una materia relevante para la salud pública muestran que es posible una mejor explotación de bases secundarias, más allá de sus debilidades en el tema drogas²³. Trabajadas con métodos más “amigables” que colaboran en acortar brechas epistemológicas²⁴, de bases profundas en nuestras formaciones y quehaceres profesionales y, por tanto, difíciles de suprimir, no debieran ser impedimento para trabajar más y mejor en un modo interdisciplinario y con probabilidades de éxito²⁵.

Como en todo estudio basado en datos secundarios, los resultados se encuentran supeditados a la calidad del registro y al sesgo de deseabilidad social relativo al consumo de sustancias. El CART ha mostrado ventajas respecto de la regresión logística, pero resulta menos accesible; además, tendría un menor poder relativo respecto de la modelación y la interrelación con otras variables al compararlo con el método de redes neuronales²⁶.

Referencias

1. Encuesta Nacional de Salud. Ministerio de Salud de Chile. 2003.
2. Encuesta Nacional de Salud. Ministerio de Salud de Chile. 2009.
3. Objetivos estratégicos de salud década 2011-2020. Ministerio de Salud de Chile. 2011.
4. Ministerio de Salud de Chile. Estrategia Nacional de Alcohol 2010.
5. Ministerio de Salud Chile - Pontificia Universidad Católica de Chile. Segundo estudio de carga de enfermedad y carga atribuible. 2007.
6. Wechsler H, Nelson TF. Binge drinking and the American college student: what's five drinks? *Psychol Addict Behav* 2001; 15 (4): 287-91.
7. Foxcroft DR, Ireland D, Lister-Sharp DJ, Lowe G, Breen R. Primary prevention for alcohol misuse in young people. *Cochrane Database Syst Rev* 2002; (3): CD003024.
8. CONACE. Octavo Estudio Nacional de Drogas en Población Escolar de Chile. 2009.
9. Gervilla García E, Jiménez López R, Montaña Moreno JJ, Sesé Abad A, Cajal Blasco B, Palmer Pol A. CART y OH. La metodología del data mining. Una aplicación al consumo del alcohol en adolescentes. *Adicciones* 2009; 10 (1): 65-80.
10. Iglesias V, Cavada G, Silva C, Cáceres D. Consumo precoz de tabaco y alcohol como factores modificadores del riesgo de uso de marihuana. *Rev Salud Pública* 2007; 41 (4): 517-22.
11. Ryan SM, Jorm AF, Lubman DI. Parenting factors associated with reduced adolescent alcohol use: a systematic review of longitudinal studies. *Aust N Z J Psychiatry*. 2010; 44 (9): 774-83.
12. Breiman L, Friedman JH, Olshen RA, Stone CJ. *Classification and regression trees*. London: Chapman & Hall; 1984.
13. Schiattino I, Silva C. Árboles de clasificación y regresión: modelos CART. *Cienc Trab* 2008; 10 (30): 161-6.
14. The use of classification trees. Cheng X, Wang M, Zhang H. *WIREs Data Mining Knowl Discov* 2011; 155-63.
15. Schiattino I, Silva C. Árboles de Clasificación y Regresión: Modelos Cart. *Cienc Trab* 2008; 10 (30): 161-6.
16. Loh WY. Classification and regression trees. *WIREs Data Mining Knowl Discov* 2011; 1 14-23.
17. Sepúlveda G, Almonte C, Valenzuela C, Avendaño A. Desarrollo moral en adolescentes de 16 a 19 años. Perspectiva psicosocial. En: Devilat, Keith editores. *Sociedad de Psiquiatría y Neurología de la Infancia y Adolescencia. La televisión, el niño y el adolescente*. 1° ed. Chile: Lord Cochrane; 1994. p. 139-54.
18. Babor T, Caetano R, Casswell S, Edwards G, Giesbrecht N, Graham K, et al. *Alcohol, No Ordinary Commodity: Research and Public Policy*. Oxford; 2003.

19. Muller R, Möckel M. Logistic regression and CART in the analysis of multimarker studies. *Clinica Chimica Acta* 2008; 394: 1-6.
20. Almeida Filho N, Castiel LD, Ayres JR. Riesgo: concepto básico de la epidemiología. *Salud colectiva* 2009; 5 (3): 323-44.
21. Marks JS. Epidemiology, Public Health, and Public policy. *Preventing Chronic disease. Public health research, practice and policy* 2009; 6 (4).
22. de Goede J, Putters K, van der Grinten T, van Omers HAM. knowledge in process? Exploring barriers between epidemiological research and local health policy development. *Health research policy and Systems* 2010; 8: 26.
23. Delaney-Black V, Chiodo LM, Hannigan JH, Greenwald MK, Janisse J, Patterson G, et al. Just say "I don't": lack of concordance between teen report and biological measures of drug use. *Pediatrics*. 2010 Nov; 126 (5): 887-93.
24. Garvin T. Analytical paradigms: the epistemological distances between scientists, policy makers, and the public. *Risk Analysis* 2001; 21 (3).
25. Brownson RC, Chiqui JE, Burgeson ChR, Fisher MC, Ness RB. Translating epidemiology into policy to prevent childhood obesity: the case for promoting physical activity in school settings. *Ann Epidemiol* 2010; 20: 436-44.
26. Trujillano J, Sarria-Santamerac A, Esquerdac A, Badia M, Palmac M, March J. Aproximación a la metodología basada en árboles de decisión (CART). Mortalidad hospitalaria del infarto agudo de miocardio. *Gac Sanit* 2008; 22 (1): 65-72.