

Características de los Instrumentos de Tamizaje del Trastorno por Uso de Alcohol en Jóvenes Universitarios de Argentina

Karina Conde, Paula Victoria Gimenez y Mariana Cremonte

*Instituto de Psicología Básica, Aplicada y Tecnología, Universidad Nacional de Mar del Plata, Argentina
Consejo Nacional de Investigaciones Científicas y Técnicas, Mar del Plata, Argentina*

INFORMACIÓN DEL ARTÍCULO

Historia del artículo:
Recibido el 28 de octubre de 2017
Aceptado el 4 de diciembre de 2017

Palabras clave:

Alcohol
AUDIT
Tamizaje
Universidad
Argentina

Keywords:

Alcohol use
AUDIT
Screening
University
Argentina

RESUMEN

Los trastornos por uso de alcohol son preocupantemente altos en la población universitaria. Contar con instrumentos de tamizaje es fundamental para la medición epidemiológica y la práctica clínica. El objetivo de este trabajo es comparar el desempeño de los instrumentos de tamizaje más usados a nivel internacional (AUDIT, AUDIT-C, TWEAK, RAPS-QF y CAGE) en estudiantes universitarios de Argentina. Los instrumentos de tamizaje, junto a la Entrevista Diagnóstica Compuesta Internacional, que permite obtener diagnósticos compatibles con el DSM 5, fueron administrados a una muestra probabilística ($N = 2,430$) de estudiantes de una universidad pública nacional. Entre otros, se evaluaron los puntos de corte óptimos, la sensibilidad, la especificidad, la consistencia interna y la estabilidad temporal, así como la presencia y la capacidad de discriminación de los ítems. La apreciación global de los resultados sugiere que tanto el AUDIT como el RAPS-QF serían los instrumentos preferibles para detectar la presencia de trastorno por uso de alcohol en esta población.

Characteristics of alcohol use disorder screening tests in young university students in Argentina

ABSTRACT

The prevalence of alcohol use disorders in university students is alarmingly high. Screening instruments are fundamental for both, epidemiological research and clinical practice. The aim of this article is to compare the performance of the most widely used screening instruments (AUDIT, AUDIT-C, TWEAK, RAPS-QF, and CAGE) in Argentinean university students. Screening tests, along with the Alcohol Section of the Composite International Diagnostic Interview that yields DSM 5 compatible diagnoses were administered to a probabilistic sample ($N = 2,430$) of university students from a national public university. The following characteristics were evaluated among others: optimal cut-off points, sensitivity, specificity, internal consistency, temporal stability, and item presence and discrimination capacity. The global appraisal of results suggests that the AUDIT and RAPS-QF would be preferable for evaluating AUD in this population.

Una gran proporción de jóvenes padecen o están en riesgo de presentar trastorno por uso de alcohol, especialmente los universitarios, que consumen más que otras poblaciones (Slutske, 2005). Sin embargo, se estima que la mayoría de las personas que padecen un trastorno por uso de alcohol no son identificadas. Una manera de mejorar la detección, y el tratamiento posterior, es desarrollando y evaluando instrumentos de tamizaje. El tamizaje se puede definir como el uso adecuado de procedimientos para detectar y diferenciar aquellos con un posible trastorno por uso de alcohol de aquellos con una baja probabilidad de tenerlo fundamentados en evidencia empírica (Connors y Volk, 2003). En la precisión de estos métodos se fundamentarán luego posteriores evaluaciones y decisiones clínicas (Allen, 2003).

A lo largo de las últimas décadas se han construido distintos instrumentos destinados a tal fin. Uno de los más investigados es el Test de Identificación de Trastornos por uso de Alcohol (AUDIT), creado por la Organización Mundial de la Salud (OMS) para ser utilizado en atención primaria (Babor, Higgins-Biddle, Saunders y Monteiro, 2001). Las propiedades psicométricas del AUDIT han sido estudiadas en diferentes países y poblaciones, mostrándose como una herramienta válida para detectar trastornos por uso de alcohol (Lundin, Hallgren, Balliu y Forsell, 2015). En muestras de estudiantes universitarios este es el test más estudiado, obteniendo resultados que indican una buena sensibilidad y especificidad (Fleming, Barry y MacDonald, 1991), a la vez que presenta una adecuada fiabilidad

Para citar este artículo: Conde, K., Victoria Gimenez, P. y Cremonte, M. (2018). Características de los instrumentos de tamizaje del trastorno por uso de alcohol en jóvenes universitarios de Argentina. *Clínica y Salud*, 29, 27-33. <https://doi.org/10.5093/clysa2018a5>

Financiación: El Consejo Nacional de Investigaciones Científicas y Técnicas y la Universidad Nacional de Mar del Plata han contribuido parcialmente a la financiación de este proyecto. Correspondencia: kconde@mdp.edu.ar (K. Conde).

ISSN: 1130-5274/© 2018 Colegio Oficial de Psicólogos de Madrid. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

(García Carretero, Novalbos Ruiz, Martínez Delgado y O'Ferrall González, 2016).

Si bien el AUDIT es un cuestionario de corta duración (2-4 minutos), existe una versión breve, creada con el objetivo de favorecer la detección rápida: el AUDIT-C (Bush, Kivlahan, McDonell, Fihn y Bradley, 1998). Incluye sólo 3 preguntas sobre frecuencia de consumo, cantidad y frecuencia de consumo excesivo episódico. Aunque cumple con el objetivo de ser fácil y rápido de administrar, el riesgo de detectar falsos positivos es alto (Nordqvist, Johansson y Bendetsen, 2004). El AUDIT-C ha sido poco estudiado en universitarios, pero algunos investigadores sugieren que podría ser una herramienta con validez y fiabilidad para esta población, por lo que recomiendan continuar estudiándolo (García Carretero et al., 2016).

Otro instrumento de tamizaje en esta área es el TWEAK (acrónimo de Tolerance, Worried, Eye-opener, Amnesia, Kut-down), construido para ser aplicado a mujeres gestantes (Russell, 1994) y ampliamente evaluado en esta población (Moraes, Viellas y Reichenheim, 2005; López, Lichtenberger, Conde y Cremonte, 2017). Si bien se ha extendido su estudio a otras poblaciones encontrando resultados favorables (Jones, 2011), en ciertos contextos, como por ejemplo en las salas de emergencia, el desempeño de otros instrumentos sería mejor (Cremonte, Ledesma, Cherpitel y Borges, 2010). Un cuestionario que ha sido construido para su uso en salas de emergencias es el Rapid Alcohol Problems Screen (Cherpitel, 1995), que en su versión más reciente (RAPS-QF) agrega preguntas sobre frecuencia de consumo con el objetivo de aumentar la sensibilidad; detectaría el 90% de los casos de abuso de alcohol (Cherpitel, 2002).

El CAGE (acrónimo de Cut-down, Annoyed, Guilty, Eye-opener) es otro instrumento de tamizaje ampliamente utilizado, construido para detectar personas que han sufrido dependencia a lo largo de su vida (Ewing, 1984). Por lo tanto, no sería adecuado para detectar el consumo excesivo episódico de alcohol, sino el consumo continuado durante largos periodos (Rigler, 2000). Este instrumento presentaría buenas propiedades psicométricas, pero, al igual que otros, su desempeño varía en función de la población a la que se aplica (Dhalla y Kopec, 2007).

Además de las diferencias según la población en la que se utilicen, los instrumentos de tamizaje presentarían variaciones en su funcionamiento en hombres y mujeres. Por ejemplo, el RAPS-QF tendría una sensibilidad adecuada para ambos géneros, mientras que el CAGE tendría una sensibilidad más baja en mujeres (Cherpitel, 2002). Sin embargo, se podría considerar que, en general, los instrumentos de tamizaje presentan una sensibilidad más baja cuando se aplican a mujeres que a hombres, por lo cual sería apropiado utilizar diferentes puntos de corte, siendo inferiores para la población femenina (Cherpitel, 1999).

Por otro lado, a pesar de que estos instrumentos hayan sido evaluados según los criterios del DSM-IV e ICD-10, los estudios sobre su adecuación al sistema diagnóstico del DSM 5 son mucho menos abundantes. Se registran pocas investigaciones con instrumentos como el AUDIT-C (Hagman, 2015) y el RAPS-QF (Cherpitel y Ye, 2015), que informan que la actuación de los instrumentos mejoraría con el nuevo sistema diagnóstico.

En resumen, se han evaluado distintos instrumentos de tamizaje del consumo de alcohol y los problemas relacionados en distintas poblaciones, con diferencias en el desempeño según el género y el sistema de diagnóstico utilizado. Se ha observado que son escasos los estudios sobre las propiedades psicométricas de los instrumentos de tamizaje más utilizados en otras poblaciones (AUDIT, AUDIT-C, TWEAK, RAPS-QF, CAGE) con muestras de universitarios, una población de riesgo en lo que a consumo de alcohol se refiere. En esta, como en otras poblaciones universitarias, el consumo de alcohol y los problemas relacionados suelen ser mucho más prevalentes que en la población general y se ha buscado implementar estrategias de detección rápida y a gran escala para llevar a cabo intervenciones que reduzcan el consumo de alcohol. Adicionalmente, algunos ins-

trumentos (como el TWEAK y el RAPS-QF) no han sido evaluados en esta población. Por lo tanto, el objetivo de este estudio será evaluar las propiedades psicométricas de distintos instrumentos de tamizaje (AUDIT, AUDIT-C, TWEAK, RAPS-QF, CAGE) del trastorno por uso de alcohol de acuerdo al DSM 5 en estudiantes universitarios de la ciudad de Mar del Plata. Se espera así contribuir a la detección precoz y a intervenciones efectivas en esta población de riesgo.

Método

Participantes

Participaron en este estudio 2,430 estudiantes, de una muestra probabilística de todas las Facultades de una universidad pública de Mar del Plata, institución que cuenta con más de 20,000 estudiantes de toda la región. El procedimiento de muestreo fue aleatorio por conglomerados, seleccionando una asignatura por año, por Facultad, de primer a tercer año, mediante un generador de números aleatorios. Para el retest se seleccionaron aleatoriamente 122 estudiantes de la muestra original que fueron reevaluados transcurrida una semana de la evaluación original. Los datos se recolectaron entre los años 2009 y 2014.

Instrumentos

Se utilizó un cuestionario auto-administrado que evaluó:

Trastorno por uso de alcohol. Se estableció la presencia-absencia de un trastorno por uso de alcohol según el DSM versión 5 (variable dicotómica sí [al menos dos criterios cumplidos]/no [menos de dos criterios cumplidos]), con la Sección de Alcohol de la Entrevista Internacional de Diagnóstico Compuesta (Tacchini, Coppola, Musazzi, Altamura e Invernizzi, 1994). La Entrevista Internacional de Diagnóstico Compuesta es una evaluación clínica que permite el diagnóstico de un trastorno por uso de alcohol según los sistemas de diagnóstico más utilizados. A pesar de su utilidad epidemiológica, resulta mucho más extensa que un instrumento de tamizaje, más apropiado para contextos en que se requiere una rápida detección (Kessler y Üstün, 2004).

Alcohol Use Disorder Identification Test (AUDIT). Este cuestionario consta de 10 ítems que evalúan para los últimos doce meses, 1) frecuencia usual de consumo de alcohol y 2) cantidad típica consumida (en tragos o unidades estándar, cualquier bebida con 11 g de alcohol puro), junto con la frecuencia de las siguientes conductas: 3) consumo de 6 o más tragos por ocasión, 4) no poder parar de consumir, 5) que la persona no pudiera realizar lo que se esperaba de ella por el consumo, 6) el consumo matutino después de una ocasión de consumo en exceso, 7) culpa o remordimiento por el consumo, 8) episodios de amnesia anterógrada inducidos por el alcohol, 9) lesiones propias o a terceros mientras estaba intoxicado y 10) preocupación del entorno por su manera de beber. Se puede obtener un puntaje de entre 0 y 40, teniendo cada ítem un puntaje de 0 a 4 (Saunders, Aasland, Babor, de la Fuente y Grant, 1993).

AUDIT-C (AUDIT-Consumption Questions). Este instrumento incluye únicamente los primeros 3 ítems del AUDIT (aquellos referidos a la conducta de consumo de alcohol) y su puntaje es de 0 a 12 (Bush et al., 1998).

Tolerance, Worried, Eye-Opener, Amnesia, Kut-down (TWEAK). Consta de 5 ítems que evalúan si en el último año la persona presentó o no: 1) una tolerancia mayor a 5 tragos, 2) quejas por parte de personas cercanas respecto de su consumo, 3) consumo en la mañana o en ayunas, 4) necesidad de reducir la ingesta de alcohol, 5) episodios de amnesia anterógrada inducidos por el alcohol en el último año. En esta escala, con un puntaje de entre 0 y 7, una respuesta positiva en los dos primeros ítems equivale a 2 puntos y los restantes 1 (Russell, 1994).

Rapid Alcohol Problems Screen - Quantity Frequency (RAPS-QF). Esta escala de 5 ítems tiene un valor de un punto por respuesta positiva en el último año a cada uno. Estos son: 1) sentimiento de culpa o remordimiento (remorse), 2) episodios de amnesia anterógrada (amnesia), 3) imposibilidad para realizar lo que se esperaba de ellos (perform), 4) consumo en la mañana o en ayunas (starter), 5) consumo de 5 o más tragos en una ocasión y frecuencia de consumo de al menos una vez por mes (Cherpitel, 2002). El periodo evaluado fueron los últimos doce meses.

Cut-down, Annoyed, Guilty, Eye-Opener (CAGE). Este instrumento tiene 4 ítems y un puntaje positivo por cada uno de ellos. Indaga si la persona tuvo alguna vez durante los últimos doce meses: 1) necesidad de reducir la ingesta de alcohol, 2) quejas por parte de personas cercanas respecto de su consumo, 3) sentimiento de culpa o remordimiento, 4) consumo en la mañana o en ayunas (Ewing, 1984).

Procedimiento

Los estudiantes presentes en el día de la administración completaron un cuestionario de auto-reporte en lápiz y papel de una duración aproximada de 20 minutos. Las respuestas fueron anónimas y confidenciales y entregadas bajo consentimiento informado. Los investigadores con formación en el área, presentes durante la administración, entregaron hojas con impresos de contacto e información general sobre el estudio. Esta investigación contó con el aval del Comité de Ética del Instituto Nacional de Epidemiología de Mar del Plata.

Análisis de Datos

En primer lugar se realizaron análisis sobre los puntos de corte más adecuados para cada instrumento según la presencia/ausencia de un trastorno por uso de alcohol. Se utilizaron dos índices de estimación, el índice de Youden y el de igualdad entre sensibilidad y especificidad. La diferencia entre ambos es que mientras que el primero aspira a un punto de corte con valores máximos de sensibilidad y especificidad, a la vez que un acuerdo máximo con el criterio, el segundo busca un punto de corte que implique la menor distancia entre sensibilidad y especificidad. Para cada uno de los puntos de corte se estimó la sensibilidad, especificidad, predicción de valores positivos, predicción

de valores negativos, índice de probabilidad de diagnóstico positivo e índice de probabilidad de diagnóstico negativo para el total de la muestra y luego utilizando el género como covariable. Posteriormente se evaluaron las propiedades psicométricas de los instrumentos, utilizando como indicadores el área bajo la curva ROC [Receiver Operating Characteristic] como medida para estimar la capacidad de discriminación entre sanos y enfermos del puntaje total de cada test, el alfa de Cronbach (ítems ordinales) o el índice Kuder-Richardson 20 (ítems dicotómicos) para medir la consistencia interna, el coeficiente K2 de Livingston para corregir los coeficientes consistencia interna según el punto de corte de cada instrumento, la medida de acuerdo no paramétrica W de Kendall para evaluar la confiabilidad test-retest y el índice kappa de Cohen, junto con su valor verdadero según las frecuencias marginales observadas, para estimar el acuerdo entre un diagnóstico positivo en cada test y un diagnóstico positivo de trastorno por uso de alcohol. Por último, se analizó el índice de discriminación de cada ítem con la correlación ítem-total y punto biserial corregida (según fueran ítems ordinales o dicotómicos respectivamente) y la proporción de cada respuesta positiva en los ítems dicotómicos. Para realizar estos análisis se utilizaron los paquetes OptimalCutpoint (López-Ratón, Rodríguez-Álvarez, Suarez y Sampedro, 2014), irr (Gamer, Lemons, Gamer, Robinson y Kedall's, 2012), psych (Revelle y Revelle, 2015) del software R y el software ICalfa (Merino, 2015).

Resultados

Datos Sociodemográficos

El porcentaje de mujeres fue 66% [IC 95% 64-68] y de varones 34% [IC 95% 32-36], de edades entre los 17 y 77 años, *M* = 22.32 [IC 95% 22.07-22.56], *DE* = 6.13; la prevalencia de trastorno por uso de alcohol según DSM 5 fue del 22% [IC 95% 20-23].

Puntos de Corte Óptimos

El índice Youden arrojó valores de punto de corte más bajos para el AUDIT y AUDIT-C y, dado que la sensibilidad fue más alta, se prefirió este índice al de igualdad entre sensibilidad y especificidad. El punto de corte para el AUDIT fue de 7 y para el AUDIT-C de 4. Los puntos de corte óptimos para el TWEAK, RAPS-QF y CAGE fueron de 3, 2, y 1

Tabla 1. Análisis de puntos de corte de instrumentos de tamizaje según trastorno por uso de alcohol, estudiantes universitarios, *N* = 2,430, Mar del Plata

	Índice de Youden							Igualdad entre sensibilidad y especificidad						
	PC	Se	Sp	PPV	NPV	DLR+	DLR-	PC	Se	Sp	PPV	NPV	DLR+	DLR-
<i>Total</i>														
AUDIT	7	.83	.74	.48	.94	3.25	.22	8	.75	.81	.52	.92	3.89	.31
AUDIT-C	4	.87	.57	.36	.94	2.06	.22	5	.72	.72	.41	.90	2.53	.39
TWEAK	3	.64	.85	.55	.89	4.41	.42	3	.64	.85	.55	.89	4.41	.42
RAPS-QF	2	.80	.78	.50	.93	3.64	.26	2	.80	.78	.50	.93	3.64	.26
CAGE	1	.68	.78	.47	.90	3.13	.41	1	.68	.78	.47	.90	3.13	.41
<i>Mujeres</i>														
AUDIT	7	.81	.80	.49	.95	4.10	.24	7	.81	.80	.49	.95	4.10	.24
AUDIT-C	4	.84	.64	.35	.94	2.34	.25	5	.65	.78	.41	.91	3.02	.44
TWEAK	1	.92	.62	.36	.97	2.41	.12	2	.77	.69	.36	.93	2.47	.33
RAPS-QF	2	.77	.81	.48	.93	4.06	.28	2	.77	.81	.48	.93	4.06	.28
CAGE	1	.65	.81	.45	.91	3.53	.42	1	.65	.81	.45	.91	3.53	.42
<i>Varones</i>														
AUDIT	9	.74	.78	.56	.89	3.36	.33	9	.74	.78	.56	.89	3.36	.33
AUDIT-C	5	.81	.58	.42	.89	1.91	.33	6	.68	.70	.47	.85	2.29	.46
TWEAK	3	.68	.79	.55	.86	3.22	.41	3	.68	.79	.55	.86	3.22	.41
RAPS-QF	2	.82	.72	.53	.91	2.97	.25	2	.82	.72	.53	.91	2.97	.25
CAGE	1	.71	.72	.49	.86	2.49	.40	1	.71	.72	.49	.86	2.49	.40

Nota. PC = punto de corte; Se = sensibilidad; Sp = especificidad; PPV = predicción de valores positivos; NPV = predicción de valores negativos; DLR+ = índice de probabilidad de diagnóstico positivo; DLR- = índice de probabilidad de diagnóstico negativo.

Tabla 2. Características psicométricas de los instrumentos de tamizaje para trastorno por uso de alcohol, estudiantes universitarios, N = 2,430, Mar del Plata

	AUC (IC 95%)	PC	M (DE)	Fiabilidad (IC 95%)	Fiabilidad según punto de corte ³	Fiabilidad temporal ⁴	Medida de acuerdo ⁵
AUDIT	.87 (.85-.89)	7	5.88 (5.11)	.78 (.77-.79) ¹	.79	.95**	.45 (.73)
AUDIT-C	.79 (.77-.81)	4	3.86 (6.28)	.64 (.61-.66) ¹	.64	.94**	.30 (.74)
TWEAK	.82 (.80-.84)	3	1.60 (1.63)	.48 (.45-.51) ²	.70	.88**	.47 (.51)
RAPS-QF	.85 (.83-.86)	2	1.27 (1.36)	.66 (.64-.68) ²	.74	.86**	.48 (.69)
CAGE	.75 (.73-.77)	1	1.48 (0.82)	.52 (.49-.55) ²	.66	.74**	.40 (.53)

Nota. IC = intervalo de confianza; AUC = área bajo la curva ROC (Receiver Operating Characteristic).

¹Alfa de Cronbach.

²Índice Kuder-Richardson 20.

³Coefficiente K2 de Livingston.

⁴W de Kendall.

⁵Kappa de Cohen (valor observado según las frecuencias marginales) para trastorno por uso de alcohol.

** $p < .001$

Tabla 3. Índice de discriminación de los ítems¹ para el Test de Identificación de Trastornos por Uso de Alcohol (AUDIT) y el AUDIT-Consumption Questions (AUDIT-C), estudiantes universitarios, N = 2,430, Mar del Plata

	Ítem 1	Ítem 2	Ítem 3	Ítem 4	Ítem 5	Ítem 6	Ítem 7	Ítem 8	Ítem 9	Ítem 10
AUDIT	.44	.39	.67	.48	.56	.30	.40	.65	.41	.40
AUDIT C	.41	.41	.60							

Nota. ¹Correlación ítem-total corregida.

Tabla 4. Índice de discriminación de los ítems para instrumentos breves de tamizaje de trastornos por uso de alcohol, estudiantes universitarios, N = 2,430, Mar del Plata

	Proporción de la presencia del indicador					Índice de Discriminación ¹				
	Ítem 1	Ítem 2	Ítem 3	Ítem 4	Ítem 5	Ítem 1	Ítem 2	Ítem 3	Ítem 4	Ítem 5
TWEAK	.50	.07	.07	.26	.11	.32	.31	.24	.39	.29
RAPS-QF	.22	.26	.25	.07	.48	.42	.49	.51	.27	.40
CAGE	.11	.07	.22	.07	---	.37	.34	.37	.19	---

Nota. ¹Correlación ítem-total punto-biserial corregida.

respectivamente. La sensibilidad fue alta para el AUDIT, AUDIT-C y el RAPS-QF y menor para el TWEAK y el CAGE que, en cambio, tuvieron una alta especificidad. El instrumento con mejor sensibilidad/especificidad fue el AUDIT. La predicción de valores positivos fue menor en el AUDIT-C y la predicción de valores negativos fue similar en todos los instrumentos (Tabla 1).

Se hallaron diferencias en los puntos de corte para mujeres y varones en el AUDIT, AUDIT -C y TWEAK (Tabla 1). El desempeño del TWEAK fue mejor en mujeres que en varones en torno a la sensibilidad y la predicción de valores negativos.

Con algunas diferencias, cualquier positivo en los instrumentos incrementó la probabilidad de tener un trastorno por uso de alcohol, mientras que un negativo la disminuyó.

Características Psicométricas

El área bajo la curva ROC fue satisfactoria en todos los casos, pero baja en el instrumento CAGE. La fiabilidad fue buena para el AUDIT, pero baja en el resto de los instrumentos. Corregida por el punto de corte, la fiabilidad fue buena para AUDIT, TWEAK y RAPS-QF. La estabilidad test-retest fue alta en todos los instrumentos, mientras que el acuerdo con el trastorno por uso de alcohol según el DSM 5 fue moderado. Aun así, controlando por sesgos de distribución, fue alto para el AUDIT, AUDIT-C y RAPS-QF y moderado para TWEAK y CAGE (Tabla 2).

Análisis a Nivel del Ítem

Para el AUDIT y AUDIT-C la discriminación de los ítems fue satisfactoria, aunque bastante menor para el ítem 6 del AUDIT, que evalúa consumo matutino, y particularmente alta para los ítems de consumo de 6 o más tragos por ocasión, y de episodios de amnesia anterógrada inducidos por el alcohol (Tabla 3).

En el TWEAK, RAPS-QF y CAGE el ítem sobre consumo matutino tuvo una discriminación baja y también una prevalencia baja. En cambio, el ítem sobre amnesia anterógrada pareció tener un buen funcionamiento en el TWEAK y RAPS-QF (Tabla 4).

Discusión

El objetivo de este estudio fue evaluar las propiedades psicométricas de instrumentos de tamizaje de trastorno por uso de alcohol en estudiantes universitarios. Algunos de estos instrumentos fueron evaluados en otros contextos en poblaciones similares, pero con criterios distintos al DSM 5. Para el AUDIT los puntos de corte establecidos en la literatura son de entre 5 y 11 (Adewuya, 2005; Fleming et al., 1991; García Carretero et al., 2016; Kokotailo et al., 2004), mientras que el más apropiado en la población estudiada fue 7. Los puntos de corte del AUDIT-C, TWEAK y RAPS-QF obtenidos aquí fueron más altos que los propuestos inicialmente por quienes los diseñaron (Bush et al., 1998; Cherpitel, 2002; Russell, 1994), probablemente debido a las características de consumo de esta población (consumo alto con consecuencias negativas, e.g., amnesia) pero sin cumplir criterios de trastorno por uso de alcohol. Similarmente a lo hallado en una investigación en estudiantes universitarios para el AUDIT-C y para otras poblaciones en general, los puntos de corte del AUDIT, AUDIT-C y TWEAK fueron menores para las mujeres (DeMartini y Carey, 2012). Si bien hay indicadores de que el consumo de alcohol y problemas asociados están aumentando entre las mujeres argentinas, siguen siendo menores que los de los varones, lo que podría explicar este resultado.

Considerados globalmente, los instrumentos con mejor desempeño respecto de sensibilidad/especificidad, propiedades psicométricas y desempeño de ítems fueron el AUDIT y el RAPS-QF. El AUDIT-C fue el instrumento más sensible, pero el menos específico, y su consistencia

interna no fue alta; lo anterior podría deberse a que evalúa sólo el consumo en una población con un consumo alto. El TWEAK tuvo un desempeño global bajo, pero fue muy sensible en mujeres, lo que no resulta sorprendente ya que este instrumento fue desarrollado para su uso en mujeres (Russell, 1994). El CAGE tuvo el desempeño más pobre para discriminar personas con un trastorno por uso de alcohol. Otros estudios en universitarios también notaron que el desempeño del AUDIT sería mejor que el del CAGE (Fleming et al., 1991). Para mejorar sus características psicométricas se han propuesto modificaciones como cambiar la pregunta sobre las discusiones con familiares por una pregunta sobre la conducción bajo los efectos del alcohol (Aertgeerts et al., 2000) o incluso añadirle otras como la influencia del alcohol en los afectos, en la pérdida de recursos y oportunidades y el vandalismo (Taylor, El-Sabawi y Cangin, 2016).

Un posible indicio del motivo del menor desempeño del CAGE en esta población puede encontrarse en el análisis a nivel del ítem. En primer lugar, se trata del único instrumento que no incluye preguntas sobre la conducta de consumo de alcohol en ninguna de sus formas. Otros estudios han encontrado que los patrones de consumo de alcohol tienen un importante valor para predecir los problemas por consumo de alcohol (Cremonte, Cherpitel Borges, Peltzer y Santangelo, 2010). Aquí, los ítems con mejores índices de discriminación tuvieron contenido referido a estas preguntas. Asimismo, tampoco incluye preguntas sobre los episodios de amnesia anterógrada. Esta consecuencia del consumo de alcohol también ha demostrado una estrecha relación con el trastorno por uso de alcohol en los jóvenes universitarios argentinos (Conde, Remaggi y Cremonte, 2014) y altos índices de discriminación. Por otro lado, sí incluye preguntas como si la persona se sintió culpable por el consumo, tuvo la necesidad de reducir el consumo o si las personas de su entorno se preocuparon por su consumo, en un contexto donde el consumo de alcohol es ampliamente tolerado y en una población donde la intoxicación es aceptada (Ministerio de Salud de la Nación, 2011). Además, incluye una pregunta sobre el consumo matutino, ítem que tuvo una de las menores prevalencias y capacidad de discriminación, probablemente debido a que evalúa una de las conductas relacionadas con el espectro más severo del trastorno por uso de alcohol, que no es frecuente en esta población de jóvenes.

Entre otras limitaciones generales (por ejemplo las que afectan a los estudios basados en autoinformes), la principal que presenta este estudio se relaciona con el orden de administración de los reactivos de los instrumentos (preguntas de consumo y preguntas evaluando consecuencias), ya que algunos estudios han mostrado sesgos en las respuestas según se administren primero unas u otras (Harford, 1994); sin embargo este tema no ha sido suficientemente explorado como para que haya recomendaciones en uno u otro sentido.

Conclusiones

Hasta donde sabemos este es el primer estudio que compara el desempeño de los instrumentos de tamizaje más usados a nivel internacional en estudiantes universitarios de un país hispanohablante del Cono Sur. Los resultados sugieren que tanto el AUDIT como el RAPS-QF serían los instrumentos preferibles para detectar la presencia de trastorno por uso de alcohol en esta población. Detectar a aquellos estudiantes con trastorno por uso de alcohol es fundamental para orientar las intervenciones que se realicen, tanto a nivel de la clínica individual como de las políticas sanitarias, sin desestimar la necesidad de medidas orientadas a reducir el consumo en esta población.

Extended Summary

Most people who suffer from an alcohol use disorder are misidentified, even in at-risk populations such as young university students.

To improve detection, evaluation, and clinical decisions, we need well developed and assessed screening instruments (procedures to detect those with a high probability of having an alcohol use disorder, and to differentiate them from those with a low probability of having it, based on empirical evidence).

Throughout the last decades, researchers have developed different instruments for this purpose. One of the most studied is the Alcohol Use Disorders Identification Test (AUDIT), created by the World Health Organization to be used in primary care. The AUDIT was assessed in its psychometric properties in different countries and populations. In university students' samples, this test obtained results that indicated a good sensitivity and specificity. At the same time, it has adequate reliability. Although the AUDIT is a short questionnaire, a shorter version was created to favor rapid detection: the AUDIT-C. It includes three questions about usual frequency and quantity of consumption and binge drinking. This AUDIT version has been limitedly studied in university students. Other screening tests in the field are the TWEAK (Tolerance, Worried, Eye-Opener, Amnesia, Kut-down), aimed to pregnant women and later to other populations, the Rapid Alcohol Problems Screen (RAPS-QF), which detects 90% of cases of alcohol abuse in primary care, and the CAGE (Cut-down, Annoyed, Guilty, Eye-Opener), developed to detect people who have suffered dependence some time in their lives. These tests' performance varies according to the population where they are used and when applied to men or women. For example, RAPS-QF has adequate sensitivity for both genders, whereas CAGE has lower sensitivity in women. Although these instruments were evaluated according to DSM-IV and ICD-10 criteria, the studies on their suitability to the DSM 5 system of diagnostic are much less common.

The aim of this study is, therefore, to evaluate the psychometric properties of alcohol use disorders screening tests (AUDIT, AUDIT-C, TWEAK, RAPS-QF, and CAGE), according to DSM 5 criteria and by gender, in university students in the city of Mar del Plata, Argentina.

Method

A total of 2,430 students participated in this study, from a probabilistic sample of all the Faculties of a public university in Mar del Plata. For re-test, we randomly selected 122 students from the original sample and re-evaluated them after one week of the original evaluation. We collected data between 2009 and 2014.

The students present on the day of administration completed the questionnaire. The answers were anonymous and confidential. We delivered the questions only when informed consent was given. This study was authorized by the Ethics Committee of the National Institute of Epidemiology of Mar del Plata. We used a self-administered questionnaire to examine: a) the presence-absence of an alcohol use disorder according to DSM 5, using the Alcohol Section of the Composite International Diagnostic Interview; b) AUDIT, with ten items about the usual frequency and quantity of alcohol consumption, and the frequency of several problematic behaviors (consumption of six or more drinks per occasion, not being able to stop consuming, not performing as expected, morning consumption, guilt or remorse, blackouts, injuries while intoxicated, and concern of significant others) in the last year; c) AUDIT-C, which has the three AUDIT's questions about alcohol consumption; d) TWEAK, a five-item test that evaluates if the person presented tolerance, complaints of significant others, morning consumption, the need to reduce alcohol intake and blackouts in the past year; e) RAPS-QF, a test with five questions about feelings of guilt or remorse, blackouts, inability to perform as expected, morning consumption, and binge drinking in the last 12 months; f) CAGE, an instrument which asks if the person ever had the need to reduce alcohol intake, received complaints from people about their consumption, had feelings of guilt or remorse, and the presence of morning consumption.

We performed analyses to estimate the most appropriate cutoff points for each instrument, based on the presence/absence of an alcohol use disorder (Youden index and equality between sensitivity and specificity). Sensitivity, specificity, prediction of positive and negative values, and index of the probability of positive and negative diagnosis, were estimated for each of the cut-off points for the total of the sample, and later using the gender as a covariate. Psychometric properties of the instruments were evaluated, using as indicators the area under the ROC for discrimination capacity, the Cronbach's alpha or the Kuder-Richardson index 20 for internal consistency, Livingston coefficient K2 to correct the internal consistency coefficients by the cutoff point, Kendall's non-parametric agreement W for test-retest reliability, Cohen's Kappa to estimate the agreement between a positive diagnosis in each test and a positive diagnosis of an alcohol use disorder, and item-total correlation and point-biserial corrected as discrimination index of each item. Analyses were performed with R and the ICalfa software.

Results

The percentage of women was 66% and men 34%, and the mean age was 22.32. The prevalence of alcohol use disorders according to DSM 5 was 22%.

According to the Youden index (the one with the higher sensitivity), the cut-off point for the AUDIT was 7, and for the AUDIT-C 4. The optimal cutoff points for the TWEAK, RAPS-QF, and CAGE were 3, 2, and 1 respectively. Sensitivity was high for the AUDIT, AUDIT-C and RAPS-QF (>.80), and lower for the TWEAK and CAGE (<.70), which, on the other hand, had a high specificity (.78). The instrument with the best sensitivity/specificity was the AUDIT (.83/.74). The prediction of positive values was lower in the AUDIT-C (.36), and the prediction of negative values was similar in all instruments. Differences were found in cut-off points for women and men in the AUDIT (7 versus 9), AUDIT -C (4 versus 5), and TWEAK (1 versus 3). When considering sensitivity and prediction of negative values, the performance of the TWEAK was better in females than males. Any positive outcome in all the instruments increased the probability of having an alcohol use disorder, while a negative one decreased it.

The area under the ROC curve was satisfactory in most cases, but low in the CAGE (.75). The reliability was good for the AUDIT and, as expected, lower in the rest of the instruments (which are shorter). Reliability for the cutoff point of each instrument was highest for the AUDIT (.79) and RAPS-QF (.74). The test-retest stability was high on all instruments (>.70), whereas the agreement with an alcohol use disorder diagnosis according to DSM 5 was high for the AUDIT, AUDIT-C, and RAPS-QF (>.70), and moderate for the TWEAK and CAGE (>.50), when considering the observed values adjusting by the sample's distribution.

Regarding item level analysis, for the AUDIT and AUDIT-C, the item discrimination was mostly satisfactory, yet very low for the morning consumption item (.30), and particularly high for consumption of six or more drinks per occasion (.67) and blackouts (.65) items.

In the TWEAK, RAPS-QF, and CAGE, the morning consumption item also had low discrimination (<.30), and a low prevalence (7%). In contrast, the item about blackouts appeared to work well in the TWEAK and RAPS-QF (>.30).

Discussion

Some of the instruments evaluated here were assessed in other contexts and in similar populations, but with criteria other than DSM 5. For the AUDIT, the cutoff points found in the literature were between 5 and 11, while the most appropriate in this study was 7. The cutoff points of the AUDIT-C, TWEAK, and RAPS-QF found here were higher than those initially proposed, probably due to the consumption

characteristics of the population, namely, a high consumption with many negative consequences, but without meeting criteria for an alcohol use disorder. Similar to what was found in university students and other populations, the AUDIT, AUDIT-C, and TWEAK cutoff points were lower for women, probably due to differences in drinking practices and their consequences.

Considered globally, the instruments with the best performance were AUDIT and RAPS-QF. The AUDIT-C was the most sensitive instrument, but the least specific. This could be due to the fact that it only evaluates consumption, and alcohol intake is high in this population. The TWEAK had a low overall performance but was very sensitive to detect alcohol use disorders in women, which is not surprising since this instrument was developed to be used in women. The CAGE had the poorest performance to discriminate people with an alcohol use disorder. Other studies in universities also noted that the AUDIT's performance may be better than that of the CAGE. We can find some possible reasons of the lower performance of the CAGE in this population at the item level analysis. The CAGE is the only instrument that does not include questions about alcohol consumption and blackouts. Drinking patterns are important predictors of alcohol-related problems, and alcohol blackouts have shown to be related to alcohol use disorders in university students. On the other hand, it does include questions about feelings of guilt, the need to reduce alcohol consumption and the concern of significant others, in a context where alcohol consumption is tolerated, and in a population where intoxication is accepted. In addition, the CAGE has an item on morning consumption, which is not frequent in youth.

Conclusion

To our knowledge, this is the first study to compare the performance of internationally used screening instruments in university students from a Spanish-speaking country of South America. The results suggest that the AUDIT and the RAPS-QF would be the preferred instruments for detecting the presence of an alcohol use disorder in this population and for guiding interventions at an individual level, without underemphasizing the need for comprehensive health policies.

Conflicto de Intereses

Los autores de este artículo declaran que no tienen ningún conflicto de intereses.

Referencias

- Adewuya, A. O. (2005). Validation of the alcohol use disorders identification test (audit) as a screening tool for alcohol-related problems among Nigerian university students. *Alcohol and Alcoholism*, 40, 575-577. <https://doi.org/10.1093/alcalc/agh197>
- Aertgeerts, B., Buntinx, F., Bande-Knops, J., Vanderrneulen, C., Roelants, M., Ansoms, S. y Fevery, J. (2000). The value of CAGE, CUGE, and AUDIT in screening for alcohol abuse and dependence among college freshmen. *Alcoholism: Clinical and Experimental Research*, 24, 53-57. <https://doi.org/10.1111/j.1530-0277.2000.tb04553.x>
- Allen, J. P. (2003). Assessment of alcohol problems: An overview. En J. P. Allen & V. B. Wilson (Eds.), *Assessing Alcohol Problems: A Guide for Clinicians and Researchers*, 2 (pp. 1-11). Bethesda, MD: National Institute on Alcohol Abuse and Alcoholism.
- Babor, T. F., Higgins-Biddle, J. C., Saunders, J. B. y Monteiro, M. G. (2001). *Cuestionario de Identificación de los Transtornos debidos al Consumo de Alcohol. Pautas para su utilización en atención primaria*. Ginebra: Organización Mundial de la Salud. Recuperado de http://www.who.int/substance_abuse/activities/en/AUDITmanualSpanish.pdf
- Bush, K., Kivlahan, D. R., McDonell, M. B., Fihn, S. D. y Bradley, K. A. (1998). The AUDIT alcohol consumption questions (AUDIT-C): an effective brief screening test for problem drinking. *Archives of internal medicine*, 158, 1789-1795. <https://doi.org/10.1001/archinte.158.16.1789>
- Cherpitel, C. J. (1995). Screening for alcohol problems in the emergency room: a rapid alcohol problems screen. *Drug and alcohol dependence*, 40, 133-137. [https://doi.org/10.1016/0376-8716\(95\)01199-4](https://doi.org/10.1016/0376-8716(95)01199-4)

- Cherpitel, C. J. (1999). Gender, injury status and acculturation differences in performance of screening instruments for alcohol problems among US Hispanic emergency department patients. *Drug and alcohol dependence*, 53, 147-157. [https://doi.org/10.1016/S0376-8716\(98\)00122-7](https://doi.org/10.1016/S0376-8716(98)00122-7)
- Cherpitel, C. J. (2002). Screening for Alcohol Problems in the US General Population: Comparison of the CAGE, RAPS4, and RAPS4-QF by Gender, Ethnicity, and Service Utilization. *Alcoholism: Clinical and Experimental Research*, 26, 1686-1691. <https://doi.org/10.1111/j.15300277.2002.tb02471.x>
- Cherpitel, C. J. y Ye, Y. (2015). Performance of the RAPS4/RAPS4-QF for DSM-5 compared to DSM-IV alcohol use disorders in the general population: Data from the 2000-2010 National Alcohol Surveys. *Drug and alcohol dependence*, 151, 258-261. <https://doi.org/10.1016/j.drugalcdep.2015.02.025>
- Conde, K., Remaggi, M. L. y Cremonte, M. (2014). Alcohol y amnesia en universitarios de Argentina y España: prevalencia y predicción. *Salud y drogas*, 14, 37-46.
- Connors, G. J. y Volk, R. J. (2003). *Self-report screening for alcohol problems among adults. Assessing alcohol problems: A guide for clinicians and researchers*, 2 (pp. 21-35). Bethesda, MD: National Institute on Alcohol Abuse and Alcoholism. Recuperado de <https://pubs.niaaa.nih.gov/publications/arh28-1/5-16.htm>
- Cremonte, M., Cherpitel, C. J., Borges, G., Peltzer, R. I. y Santángelo, P. R. (2010). Drinking Patterns and DSM-IV Alcohol Use Disorders' Criteria in Argentinean Emergency Department Patients. En J. Maars (Ed.), *Alcohol Dependence and Addiction*. Hauppauge, NY: Nova Publishers.
- Cremonte, M., Ledesma, R. D., Cherpitel, C. J. y Borges, G. (2010). Psychometric properties of alcohol screening tests in the emergency department in Argentina, Mexico and the United States. *Addictive Behaviors*, 35, 818-825.
- DeMartini, K. S. y Carey, K. B. (2012). Optimizing the use of the AUDIT for alcohol screening in college students. *Psychological assessment*, 24, 954-963. <https://doi.org/10.1037/a0028519>
- Dhalla, S. y Kopec, J. A. (2007). The CAGE questionnaire for alcohol misuse: a review of reliability and validity studies. *Clinical & Investigative Medicine*, 30, 33-41. Recuperado de <http://cimonline.ca/index.php/cim/article/viewFile/447/554>
- Ewing, J. A. (1984). Detecting alcoholism: the CAGE questionnaire. *Jama*, 252, 1905-1907. <https://doi.org/10.1001/jama.1984.03350140051025>
- Fleming, M. F., Barry, K. L. y MacDonald, R. (1991). The alcohol use disorders identification test (AUDIT) in a college sample. *International Journal of the Addictions*, 26, 1173-1185. <https://doi.org/10.3109/10826089109062153>
- Gamer, M., Lemon, J., Gamer, M. M., Robinson, A. y Kendall's, W. (2012). Package 'irr'. *Various coefficients of interrater reliability and agreement*. Recuperado de <https://cran.r-project.org/web/packages/irr/irr.pdf>
- García Carretero, M. Á., Novalbos Ruiz, J. P., Martínez Delgado, J. M. y O'Ferrall González, C. (2016). Validación del test para la identificación de trastornos por uso de alcohol en población universitaria: AUDIT y AUDIT-C. *Adicciones*, 28, 194-204. <https://doi.org/10.20882/adicciones.28.4>
- Hagman, B. T. (2015). Toward efficient screening for DSM-5 alcohol use disorders in college students: performance of the AUDIT-C. *Addictive Disorders & Their Treatment*, 14, 1-15. <https://doi.org/10.1097/ADT.0000000000000003>
- Harford, T. C. (1994). The effects of order of questions on reported alcohol consumption. *Addiction*, 89, 421-424. <https://doi.org/10.1111/j.1360-0443.1994.tb00916.x>
- Jones, L. A. (2011). Systematic review of alcohol screening tools for use in the emergency department. *Emergency Medicine Journal*, 28, 182-191. <https://doi.org/10.1136/emj.2009.085324>
- Kessler, R. C. y Üstün, T. B. (2004). The world mental health (WMH) survey initiative version of the world health organization (WHO) composite international diagnostic interview (CIDI). *International Journal of Methods in Psychiatric Research*, 13, 93-121. <https://doi.org/10.1002/mpr.168>
- Kokotailo, P. K., Egan, J., Gangnon, R., Brown, D., Mundt, M. y Fleming, M. (2004). Validity of the alcohol use disorders identification test in college students. *Alcoholism: Clinical and Experimental Research*, 28, 914-920. <https://doi.org/10.1097/01.ALC.0000128239.87611.F5>
- López, M. B., Lichtenberger, A., Conde, K. y Cremonte, M. (2017). Psychometric properties of brief screening tests for alcohol use disorders during pregnancy in Argentina. *Revista Brasileira de Ginecología y Obstetricia*, 39, 322-329.
- López-Ratón, M., Rodríguez-Álvarez, M. X., Suarez, C. C. y Sampedro, F. G. (2014). OptimalCutpoints: an R package for selecting optimal cutpoints in diagnostic tests. *Journal of Statistical Software*, 61(8), 1-36. <https://doi.org/10.18637/jss.v061.i08>
- Lundin, A., Hallgren, M., Balliu, N. y Forsell, Y. (2015). The use of alcohol use disorders identification test (AUDIT) in detecting alcohol use disorder and risk drinking in the general population: validation of AUDIT using schedules for clinical assessment in neuropsychiatry. *Alcoholism: Clinical and Experimental Research*, 39, 158-165. <https://doi.org/10.1111/acer.12593>
- Merino, C. (2015). *ICalfa: módulo Excel para estimar intervalos de confianza para alfa de Cronbach*. Programa no publicado.
- Ministerio de Salud de la Nación. (2011). *Algunos datos sobre el consumo de Alcohol en Argentina*. Recuperado de <http://www.msal.gov.ar/saludmental/images/stories/info-equipos/pdf/4-algunos-datos-sobre-el-consumo-de-alcohol.pdf>
- Moraes, C. L., Viellas, E. F. y Reichenheim, M. E. (2005). Assessing alcohol misuse during pregnancy: evaluating psychometric properties of the CAGE, T-ACE and TWEAK in a Brazilian setting. *Journal of Studies on Alcohol*, 66, 165-173. <https://doi.org/10.15288/jsa.2005.66.165>
- Nordqvist, C., Johansson, K. y Bendtsen, P. (2004). Routine screening for risky alcohol consumption at an emergency department using the AUDIT-C questionnaire. *Drug and Alcohol Dependence*, 74, 71-75. <https://doi.org/10.1016/j.drugalcdep.2003.11.010>
- Revelle, W. y Revelle, M. W. (2015). Package 'psych'. The Comprehensive R Archive Network.
- Rigler, S. K. (2000). Alcoholism in the elderly. *American Family Physician*, 61, 1710-1716. <https://doi.org/10.1.1.516.2517&rep=rep1&type=pdf>
- Russell, M. (1994). New assessment tools for risk drinking during pregnancy: T-ACE, TWEAK, and others. *Alcohol Research and Health*, 18, 55-61. Recuperado de <https://pubs.niaaa.nih.gov/publications/ahrw18-1/55-61.pdf>
- Saunders, J. B., Aasland, O. G., Babor, T. F., De la Fuente, J. R. y Grant, M. (1993). Development of the alcohol use disorders identification test (AUDIT): WHO collaborative project on early detection of persons with harmful alcohol consumption-II. *Addiction*, 88, 791-804. <https://doi.org/10.1111/j.1360-0443.1993.tb02093.x>
- Slutske, W. S. (2005). Alcohol use disorders among US college students and their non-college-attending peers. *Archives of General Psychiatry*, 62, 321-327. <https://doi.org/10.1001/archpsyc.62.3.321>
- Tacchini, G., Coppola, M., Musazzi, A., Altamura, A. y Invernizzi, G. (1994). [Multinational validation of the composite international diagnostic interview (CIDI)]. *Minerva Psichiatrica*, 35, 63-80.
- Taylor Jr, P., El-Sabawi, T. y Cangin, C. (2016). Improving alcohol screening for college students: screening for alcohol misuse amongst college students with a simple modification to the CAGE questionnaire. *Journal of American college health*, 64, 397-403. <https://doi.org/10.1080/07448481.2016.1168429>