

Prevalencia de los trastornos de la conducta alimentaria: consideraciones metodológicas

M. Ángeles Peláez Fernández^{1*}, Francisco J. Labrador Encinas²,

Rosa M. Raich Escursell³

¹University of Toronto, Canada; ²Universidad Complutense de Madrid, España;

³Universidad Autónoma de Barcelona, España

RESUMEN

En este artículo se revisan los diseños metodológicos e instrumentos diagnósticos utilizados para estimar la prevalencia de los trastornos de la conducta alimentaria en población general. También se presentan las ventajas, inconvenientes y sesgos asociados al empleo de cada uno de estos diseños metodológicos. Asimismo, se señalan las dificultades inherentes vinculadas al estudio de la prevalencia de los trastornos alimentarios. Se apuntan una serie de propuestas metodológicas a seguir para estimar de forma precisa el alcance de estas patologías en población general.

Palabras clave: Trastornos de la conducta alimentaria, anorexia, bulimia, prevalencia.

ABSTRACT

Prevalence of eating disorders: Methodological considerations. In the present study, the methodological designs and diagnostic instruments used in the estimation of the prevalence of eating disorders among general population are revised. The advantages, inconveniences and biases associated with the use of each methodological procedure are shown. The inherent difficulties attached to the estimation of eating disorders prevalence are also presented. Lastly, we propose some methodological advises in order to study accurately the prevalence of eating disorders among general population.

Keywords: Eating disorders, Anorexia, Bulimia, Prevalence.

Los Trastornos de la Conducta Alimentaria (TCA) constituyen en la actualidad cuadros de gran relevancia social, especialmente en población de más alto riesgo: mujeres jóvenes y adolescentes de países desarrollados o en vías de desarrollo. En esta población se ha producido un incremento de dos a cinco veces en las cifras de prevalencia de TCA en las últimas tres décadas, aceptándose un porcentaje de 0,5-1% de anorexia nerviosa (AN), de 1-3% de bulimia nerviosa (BN) y de aproximadamente 3% de Trastornos de la Conducta Alimentaria no Especificados (TCANE), según el Protocolo de los TCA del Instituto Nacional de la Salud (Insalud, 1995), la Asociación Americana de Psiquiatría (DSM-IV-TR; APA, 2000) y revisiones recientes de estudios epidemiológicos de TCA (Hoek & van Hoeken, 2003). Dichas fuentes establecen una ratio de prevalencia de TCA de uno a nueve en varones respecto a mujeres.

*La correspondencia sobre este artículo puede ser dirigida a la primera autora: Department of Psychology, University of Toronto at Mississauga, Ontario, Canada. Email: marianpf@cop.es.

Estas patologías se caracterizan por la gravedad de la sintomatología asociada, elevada resistencia al tratamiento y riesgo de recaídas (Kaye et al., 2000; Mizes & Palermo, 1997; Vitello & Lederhendler, 2000) así como alto grado de comorbilidad o presencia de otros trastornos como abuso de drogas, depresión, falta de habilidades sociales y baja autoestima social y corporal (APA, 2000; Gila, Castro, Gomez & Toro, 2005; Lindberg & Hjern, 2003). En el caso de la AN el trastorno es asociado además a un elevado riesgo de mortalidad y reducido porcentaje de recuperación (Kaye et al., 2000; Mizes & Palermo, 1997; Vitello & Lederhendler, 2000). Debido a estos factores entre otros, los TCA constituyen hoy la tercera enfermedad crónica entre la población femenina adolescente y juvenil en las sociedades desarrolladas y occidentalizadas.

En las últimas tres décadas y, como consecuencia de la creciente preocupación por estos cuadros, se han realizado numerosos estudios epidemiológicos tratando de estimar tanto el alcance de los TCA como los factores de riesgo asociados. Algunos estudios (Gila, Castro, Toro & Salamero, 2004; Fairburn, Welch, Doll, Davies & O'Connor, 1997; Fairburn, Cooper, Doll & Welch, 1999; Lindberg & Hjern, 2003) muestran ciertos factores psicosociales de riesgo como la presencia de trastorno psiquiátrico o abuso de drogas en los padres, así como obesidad premórbida, críticas recibidas por el aspecto físico y alteración de la imagen corporal en mujeres adolescentes.

Con el objetivo de estimar el alcance de estas patologías, se han realizado numerosos estudios de prevalencia, especialmente en Norteamérica y Europa Occidental, desde aproximaciones metodológicas diferentes (i.e., utilizando diversos diseños experimentales, así como instrumentos diferentes de preselección e identificación de casos).

Debido a que la programación de recursos sanitarios (i.e. unidades asistenciales, personal sanitario) se apoya sobre los datos de los estudios epidemiológicos en población general, resulta imprescindible que estos estudios sean precisos (i.e., que reflejen con la mayor exactitud posible el alcance de estos trastornos en la población), para garantizar una planificación y provisión racional y eficaz de estos servicios de asistencia.

La precisión en epidemiología depende de una serie de factores metodológicos que es preciso conocer y controlar como: tipo de trastorno, método de selección de las muestras, tipo de diseños e instrumentos utilizados para la preselección e identificación de casos, etc.

Antes de abordar estos factores de forma individual y detallada, y de indicar las dificultades y sesgos asociados al tipo de trastorno y a las diferentes instrumentos y diseños epidemiológicos, se presentan de forma descriptiva los diferentes tipos de estudios de prevalencia de TCA que se recogen en la literatura, con objeto de facilitar la comprensión al lector.

TIPOS DE ESTUDIOS DE PREVALENCIA

En función de su metodología e instrumentos de detección de casos, los estudios de prevalencia pueden ser englobados en tres grupos generales entre los que cabe señalar aquellos estudios apoyados en cuestionarios de autoinforme para establecer un diagnóstico de TCA; estudios de doble fase que utilizan entrevista diagnóstica, pero no

evalúan la adecuación del instrumento de screening; y, por último, estudios de doble fase que utilizan entrevista diagnóstica y evalúan la eficacia del instrumento de screening o que entrevistan a toda la muestra (Fairburn y Beglin, 1990).

La mayoría de los trabajos realizados en la década de los 80 (previos a la publicación de los criterios diagnósticos DSM) constituyen ejemplos del primer tipo (*estudios apoyados en cuestionarios de autoinforme para establecer un diagnóstico de TCA*) en tanto que calculan las tasas de prevalencia de AN y BN únicamente a partir de las respuestas de los sujetos a los cuestionarios de autoinforme. Constituyen, por tanto, una primera aproximación epidemiológica a los TCA, cuando todavía no se disponía de criterios diagnósticos establecidos universalmente. Algunos de estos estudios (menos de la mitad de los publicados de esta época) aportan datos de frecuencia de episodios de sobreingesta o frecuencia de ayuno o dietas estrictas. En las investigaciones que aportan esta información, más de la tercera parte de los sujetos informan episodios recurrentes de sobreingesta y más del 25% indican realizar dietas estrictas o ayuno. No se aportan datos acerca de la preocupación por la forma y peso corporales. Estos trabajos ofrecen una tasa de prevalencia de BN del 10% en mujeres jóvenes y adolescentes.

Aunque este diseño es más propio de la década de los 80, también se han publicado en la década de los 90 algunos estudios epidemiológicos de este tipo, como el de Cotrufo, Barretta, Monteleone y Maj (1998). Estos estudios tienen la ventaja de ser económicos y rápidos ya que pueden aplicarse de forma grupal y en una única fase, sin realización de entrevista posterior; y sencillos, puesto que no requieren de un método de identificación de sujetos para localizarlos en una segunda fase. Sin embargo, conviene ser cautelosos en la interpretación de resultados de este cuerpo de investigaciones por tres cuestiones metodológicas básicas:

Elección de las muestras. Estos estudios emplean mayoritariamente técnicas de muestreo restrictivas: La mitad de ellos toman estudiantes como sujetos, siendo frecuentes los muestreos con estudiantes de Psicología de prestigiosas universidades privadas. También se han utilizado muestras clínicas de pacientes ingresados o en consulta. Estas poblaciones no son representativas de la población general. (Whitaker et al, 1990). P. ej., los sujetos con un significativo bajo peso estarían poco representados en los estudios realizados con usuarios de servicios de planificación familiar; mientras que las mujeres con BN estarían sobre-representadas en los estudios que incluyen muestras clínicas, debido a los problemas físicos y psicológicos que usualmente acompañan a este trastorno.

Tasas de respuesta. La tasa-media de respuestas obtenidas en estos trabajos (i.e., sujetos que completan los cuestionarios) es de 74,4% (en un rango del 34%-100%). En los estudios en los que la tasa de respuesta es muy baja existe un alto riesgo de que los sujetos que respondan a los cuestionarios no sean representativos de la población. Los sujetos con AN o BN pueden mostrarse más o menos renuentes que otros a devolver cuestionarios acerca de actitudes y comportamientos alimentarios, que, usualmente, tienden a ocultar. Las investigaciones que emplean poblaciones "cautivas" (e.g., pacientes ingresados) obtienen una tasa-media de respuesta del 90%, pero adolecen de validez externa.

Método de detección de "casos". Los cuestionarios de autoinforme no constituyen un método adecuado para la detección de casos de TCA. Las características principales de estos trastornos (e.g., episodios de sobreingesta) no tienen un significado específico y aceptado universalmente, por lo que los enunciados de los cuestionarios que hacen referencia a frecuencia de "atracones" presentan habitualmente dificultades de definición e interpretación. (Probablemente sea esta una de las razones por las que las frecuencias de atracones —obtenidas mediante cuestionario— sean tan elevadas). También aparecen dificultades de precisión con ítems como "sentirse gordo" o "haber estado preocupado por el peso". No puede establecerse un perfil diagnóstico (e.g., presencia de ideas sobrevaloradas de la delgadez, o preocupación por la forma y peso corporales) basándose en las respuestas a estos enunciados.

Debido a estas dificultades, asociadas al empleo de cuestionarios de autoinforme para establecer diagnóstico de TCA, existe un acuerdo en la literatura científica acerca de la conveniencia del uso de entrevista diagnóstica clínica "cara-a-cara" con el sujeto para establecer un diagnóstico fiable. (Fairburn y Beglin, 1990; Steinhausen, Winkler y Meier, 1997; Sullivan, Bulik y kindler, 1998).

Los estudios de doble fase que utilizan entrevista diagnóstica, pero no evalúan la adecuación del instrumento de screening constituyen un ejemplo sesgado de lo que se denomina en epidemiología 'metodología en doble fase' en tanto que aplican un cuestionario de barrido a todos los sujetos de la muestra en la primera fase, para detectar la población en riesgo a padecer TCA; y en la segunda fase entrevistan únicamente a aquellos sujetos que han igualado o superado el punto de corte establecido en el instrumento de barrido aplicado en la primera fase. (e.g., Nadaoka et al., 1996)

Estos estudios cuentan con la ventaja de ser económicos, al entrevistar sólo a un número reducido de sujetos: aquellos que constituyen población de riesgo por obtener puntuaciones iguales o superiores al punto de corte del instrumento de barrido en la primera fase; y sencillos, puesto que no incorporan una selección aleatoria de sujetos que no constituyen a priori población de riesgo. Sin embargo, también deben considerarse con precaución, puesto que, al no realizar entrevista diagnóstica a ningún sujeto que puntúa por debajo del punto de corte, este diseño no permite estimar la especificidad (i.e., capacidad para detectar los no-casos o controles) ni, lo que es más importante en un estudio epidemiológico, la sensibilidad (i.e., capacidad de discriminar todos los casos) del instrumento de barrido. Sin datos de sensibilidad, no es posible calcular el porcentaje de falsos negativos en la muestra ni, por tanto, estimar de forma precisa la prevalencia del trastorno.

Entre los *estudios de doble fase que utilizan entrevista diagnóstica y evalúan la eficacia del instrumento de screening o que entrevistan a toda la muestra* se engloba tanto a aquellos *estudios que realizan entrevista diagnóstica a todos los sujetos de la muestra*, como a los *estudios de doble fase o identificación de caso en dos etapas*.

Los trabajos de Drewnowski, Hopkins y Kessler (1988); Bushnell, Wells, Hornblow, Oakley-Browne y Joyce (1990); Hoek (1991); Garfinkel et al. (1995) constituyen ejemplos de estudios epidemiológicos de TCA en que se entrevistó a todos los sujetos de la muestra.

Los estudios de doble fase utilizan el clásicamente aceptado diseño de *identifi-*

cación de "caso" en dos etapas o metodología en doble fase caracterizado por la aplicación de un cuestionario de barrido en la primera fase a todos los sujetos de la muestra, para detectar los posibles casos de TCA (i.e., aquellos sujetos que igualan o superan el punto de corte establecido en el instrumento de barrido) y los controles (i.e., aquellos que puntúan por debajo de dicho punto de corte); y entrevista, en la segunda fase, a todos los sujetos preseleccionados como posibles casos, junto a un número equivalente de presuntos controles, elegidos al azar.

En algunos trabajos como el de Johnson-Sabine, Wood y Patton (1988) además de entrevistar a un grupo de sujetos que no constituyeron población de riesgo según el cuestionario de barrido, se investigó, a aquellos sujetos que no respondieron al cuestionario de barrido o que respondieron de forma incompleta, encontrando entre ellos los únicos casos de AN presentes en todo el estudio. Otros autores, como Rathner y Messner (1995) incorporaron algunas variaciones a este diseño típico, como la inclusión de un test para evaluar la sensibilidad y especificidad del instrumento de barrido; en vez de entrevistar en la segunda fase a un grupo aleatorio de sujetos que no superaron el punto de corte del instrumento.

Este último tipo de diseño es el más complejo de los tres e implica un mayor coste y posibilidad de rechazo por parte de los sujetos, ya que supone entrevistar a toda la muestra; o bien, entrevistar a una muestra aleatoria de sujetos que no superaron el punto de corte, además de aquellos preseleccionados como posibles casos en la primera fase. No obstante, ambos diseños están considerados como los más fiables para estimar la prevalencia de los TCA, ya que aportan una cifra exacta de casos y no-casos (diseño de entrevista a todos los sujetos de la muestra); ó bien, un porcentaje aproximado de falsos negativos entre los sujetos no entrevistados (diseño de doble fase), lo que permite acotar con exactitud la tasa de prevalencia.

CONSIDERACIONES METODOLÓGICAS ASOCIADAS A LOS ESTUDIOS DE PREVALENCIA DE TCA

A la hora de establecer la prevalencia de los TCA, tanto si se sigue un procedimiento de una fase como de dos, se han de tener en cuenta algunas dificultades inherentes a los estudios de prevalencia en este tipo de trastornos y que será preciso manejar:

Dificultades inherentes a los estudios de prevalencia de TCA

Entre las dificultades inherentes al estudio de prevalencia de TCA se encuentran la baja prevalencia de los TCA, las dificultades asociadas a la definición diagnóstica de "caso" y las dificultades asociadas a la evaluación de la sintomatología específica de los TCA.

Debido a la *baja prevalencia de los TCA* (alrededor del 4,5% en mujeres adolescentes y jóvenes y del 1% entre los varones de esta misma edad), se hace necesario tomar muestras muy amplias de sujetos entre la población general para hallar casos, especialmente cuando se realizan estudios con varones. Requieren, por ello, un importante esfuerzo personal y económico, sobre todo en los procedimientos de una fase en

los que se entrevista a toda la muestra.

Es preciso asimismo hacer especial mención a las *dificultades asociadas a la definición diagnóstica de "caso TCA"* ya que prácticamente en todos los estudios recientes se han adoptado los criterios establecidos por el DSM-IV o CIE-10 para el diagnóstico de AN o BN. Algunos de estos criterios, como la cantidad de pérdida de peso, duración de amenorrea y frecuencia de episodios de sobreingesta o conductas compensatorias inadecuadas, son criterios estrictos y arbitrarios (Hsu, 1996) y llevan a la exclusión como caso de AN o BN a los sujetos en los que no se presenten alguno de los criterios.

En los manuales DSM-IV (APA, 1994) y DSM-IV-TR (APA, 2000) se incluyen como trastornos de la conducta alimentaria no especificados (TCANE) seis cuadros de síntomas en los que falta una característica para completar el cuadro de caso de AN o BN. De estos seis cuadros de síntomas, dos corresponden a sujetos que cumplen todos los criterios de AN exceptuando la pérdida de menstruación durante tres ciclos consecutivos (TCANE de Tipo 1) y exceptuando un peso inferior al 25% del peso esperable según altura, edad y constitución (TCANE de Tipo 2). Los otros cuatro son cuadros que corresponden a síntomas parciales de BN. El establecimiento de estos cuadros incompletos limita el diagnóstico a la presencia de esos conglomerados de síntomas. En el caso de los cuadros restrictivos, sólo se contemplan dos tipos, por lo que quedan sin diagnóstico aquellos cuadros en los que falta alguna característica de AN que no sea la amenorrea o el infrapeso.

Por último, en cuanto a las *dificultades asociadas a la evaluación de la sintomatología específica de los TCA*, cabe indicar que los episodios de sobreingesta (denominados comúnmente "atracones") y las conductas compensatorias inadecuadas (e.g., vómitos autoinducidos) por su naturaleza privada, suponen un desafío para su evaluación. No es factible su observación en situaciones naturales por otras personas; y su evaluación, tanto a través de cuestionarios como de entrevistas, ha de fiarse del informe de los propios sujetos. Esta limitación compromete la fiabilidad de la evaluación, dada la tendencia de los sujetos con TCA a la ocultación o encubrimiento de su sintomatología.

El concepto de "atracon" resulta ambiguo y susceptible de diversas interpretaciones (Hsu, 1996; Wilfley, Schwartz, Spurrell, y Fairburn, 1997). Se caracteriza por la ingesta de una cantidad de comida excesiva, dadas las circunstancias; y por una sensación de pérdida de control en la misma. Puesto que los episodios de sobreingesta cursan frecuentemente con una sensación de falta de control; se hace difícil para el sujeto tanto la conciencia de la emisión de la conducta como el recuerdo de su cantidad y frecuencia. Este deterioro del recuerdo dificulta la evaluación de los TCA, especialmente cuando se realizan mediante cuestionarios.

Ventajas e inconvenientes asociados a los estudios de una fase y de dos fases

Junto con las dificultades inherentes a la realización de investigaciones epidemiológicas de TCA, independientemente del procedimiento utilizado, se han de considerar los problemas específicos asociados a los estudios de una y doble fase a la

hora de decidir utilizar un diseño u otro.

Los estudios de una fase establecen la prevalencia de un determinado trastorno mediante un cuestionario o entrevista diagnóstica que se administra a todos los sujetos de la muestra. Los estudios de doble fase utilizan un cuestionario de barrido en la primera fase. Los sujetos que igualan o superan el punto de corte establecido en dicho cuestionario son preseleccionados para establecer su diagnóstico mediante entrevista en una segunda fase. Para evitar los sesgos asociados al instrumento de barrido, una muestra aleatoria de los que no superen el punto de corte, también son entrevistados. Los estudios de doble fase resultan “económicos” al no entrevistar a todos los sujetos de la muestra. Sin embargo, presentan varios inconvenientes:

Dificultad de garantizar el anonimato de los sujetos. Si se pretende “rescatar” a parte de la muestra que participa en la primera fase (i.e., los sujetos que puntúen igual por encima del punto de corte establecido en el instrumento de barrido; y una muestra aleatoria de los que puntúen por debajo), para ser entrevistados en la segunda fase, éstos deben indicar algún tipo de dato personal en dicho instrumento. Los datos solicitados a los sujetos en la primera fase pueden ser personales (e.g., nombre y apellidos); o bien, otro tipo de códigos o claves identificadoras (e.g., nombre ficticio, cifras o dígitos). Esta última opción es susceptible de tener mejor acogida entre los participantes, pero se dificulta al acceso posterior a los mismos, ya que el investigador debe llegar a los sujetos seleccionados a través de sus claves o códigos.

En ambos casos, tanto si el investigador solicita a los sujetos indicar su nombre personal o bien, una clave identificadora en la primera fase, puede guardarse la confidencialidad de los sujetos (siempre que se comunique el diagnóstico por vía telefónica o privada, y no de forma pública –p. ej. a través de los profesores y/o en las aulas-); sin embargo, los diseños de dos fases tienen el riesgo de potenciar o agravar los sesgos producidos por los rechazos de los sujetos a participar o a responder de forma sincera a las preguntas. Este sesgo es esperable en el caso de los TCA, especialmente caracterizados por la ocultación de la sintomatología. Cuando este procedimiento se realiza en colegios con menores de edad, los sesgos por ocultación / negación de sintomatología se incrementan. En este sentido, Johnson-Sabine, Wood y Patton (1988) en un estudio de prevalencia de AN entre alumnas matriculadas en colegios públicos de Londres, investigaron a las 17 adolescentes que no respondieron a los cuestionarios, encontrando que dos de ellas estaban recibiendo tratamiento para AN. Éstas fueron los únicos casos de AN presentes en todo su estudio. Según los cálculos de Williams y MacDonald (1986), este tipo de sesgo maximiza la probabilidad de falsos negativos con una infravaloración del 5% en las cifras de prevalencia.

El hecho de solicitar a los sujetos un nombre o clave identificatoria (condición de no-anonimato) también puede suponer la negativa de los directores, padres y/o profesores a la colaboración, en caso de realizar el estudio epidemiológico en Centros de Enseñanza.

Posible pérdida de sujetos durante el transcurso temporal entre la primera y la segunda fase. Se espera que los casos de TCA estén sobre-representados entre los que rehúsen ser entrevistados en una segunda oleada, por esta tendencia a la ocultación de síntomas. Así, King (1989) encontró que las 3 mujeres que se negaron a realizar la

entrevista diagnóstica habían consultado en varias ocasiones a su médico de cabecera por preocupaciones acerca de su peso.

Cuando se realiza el muestreo en "poblaciones cautivas" (e.g., sujetos escolarizados o pacientes ingresados en hospitales) es más difícil que se produzca esta "pérdida" de sujetos entre las dos fases; aunque se incrementa la probabilidad de ocultación de síntomas durante la entrevista, especialmente en el caso de sujetos escolarizados.

Sesgos debidos al instrumento de barrido utilizado. En función de la sensibilidad y especificidad para un punto de corte establecido, se pueden "escapar" casos o considerar como casos a sujetos sin trastorno.

Debido a los sesgos que plantean los estudios de prevalencia en dos fases, especialmente cuando se realizan en el ámbito de los TCA, algunos trabajos epidemiológicos de TCA han entrevistado a toda la muestra en una sola fase (Drewnowski, Hopkins y Kessler, 1988; Bushnell, Wells, Hornblow, Oakley-Browne y Joyce, 1990; Hoek, 1991). Con esta opción se minimizan los sesgos asociados al no-anonimato (i.e., al hecho de incluir claves identificatorias), al transcurso temporal entre las dos fases y al propio instrumento de barrido. Además, la aplicación de la entrevista cara-a-cara permite clarificar conceptos y solicitar respuestas completas y lógicas. Pero implica también una serie de inconvenientes:

Tener que entrevistar uno a uno a todos los sujetos de la muestra. En el caso de que se opte por el método de entrevista para establecer diagnóstico TCA en los diseños de una fase, se deberá entrevistar a todos los sujetos, lo que implica un importante gasto de tiempo y esfuerzo de tipo personal y económico; así como el riesgo de la negativa y/o imposibilidad por parte de los centros escolares a facilitar los recursos necesarios para realizar las entrevistas (e.g., aulas, flexibilidad de horarios, acuerdo con los profesores para que salgan uno a uno los alumnos de sus clases, etc.).

Relación establecida entre el entrevistador y la persona entrevistada. Las entrevistas ejecutadas oralmente por un investigador tienen la desventaja de estar influidas por la relación establecida entre el entrevistador y la persona entrevistada. Este riesgo resulta muy difícil de controlar, aún en el caso de que el mismo entrevistador realice todas las entrevistas. (Cummings, Strull, Nevitt y Hulley, 1993).

Efectuar la entrevista de modo distinto en cada ocasión. Resulta inevitable que las entrevistas se realicen de un modo, como mínimo, ligeramente distinto en cada ocasión, aunque ambos tipos de instrumentos (cuestionario y entrevista diagnóstica) puedan estar estandarizados.

Poder de observación limitado y deseabilidad social. Los dos tipos de instrumentos de evaluación son susceptibles de errores ocasionados por el poder de observación limitado y la deseabilidad social (i.e., deseo del sujeto en estudio de dar respuestas socialmente aceptables), especialmente en el caso de las entrevistas.

Ventajas e inconvenientes asociados al uso de entrevista diagnóstica vs. cuestionario

Vinculado a la utilización de una vs. dos fases en los procedimientos epidemiológicos de TCA, se plantea en la literatura la utilización de entrevistas diagnósticas vs. cuestionarios diagnósticos a la hora de estimar la prevalencia de los TCA. Las

entrevistas cara-a-cara permiten que el investigador pueda definir términos poco precisos como “atracones”, clarificar cuestiones, ayudar al sujeto a recordar la frecuencia y la intensidad de sus conductas y solicitar respuestas complejas y lógicas; pero resultan poco económicas y corren el riesgo de estar influidas por la relación entrevistador-entrevistado y por la deseabilidad social. Los cuestionarios diagnósticos pueden administrarse en grupo y de forma anónima y no están sujetos a la influencia de la relación con el entrevistado; pero no es posible asegurar con su aplicación la correcta interpretación de las preguntas por parte de los sujetos.

Otra cuestión que se plantea en la literatura es la entrevista diagnóstica a utilizar en la estimación de la prevalencia de los TCA. Las entrevistas más utilizadas para el diagnóstico de TCA son las siguientes: a) La sección H (para TCA) de la entrevista diagnóstica *Composite International Diagnostic Interview* (CIDI), desarrollada por la Organización Mundial de la Salud (WHO, 1990) con base en el *Diagnostic Interview Schedule* (DIS) y la *Present State Examination* (PSE), y b) la *Eating Disorders Examination* (EDE) de Fairburn y Cooper (1993) (Wilson, 1993; Smith, Marcus y Eldredge, 1994; Wilfley, Schwartz, Spurrell y Fairburn, 1997). Aunque las dos entrevistas se adecuan a criterios diagnósticos definidos por el DSM-IV – DSM-IV-TR y CIE-10 y ambas fueron diseñadas para aplicarse por entrevistadores entrenados no-clínicos; existe un acuerdo entre diferentes autores en que la entrevista mejor validada para el diagnóstico de TCA es la EDE. Por su alta fiabilidad y validez, Williamson, Anderson, Jackman y Jackson (1995), recomiendan su uso en investigación y evaluación de eficacia de tratamientos; y Wilson (1993) y Smith, Marcus y Eldredge (1994) la consideran como el método de elección para la evaluación de la psicopatología específica de los TCA.

La versión número 12 de la EDE (Fairburn y Cooper, 1993) consta de 62 preguntas que evalúan 4 subescalas generales: Restricción alimentaria, Preocupación por la comida, Preocupación por el peso y Preocupación por la figura. Cada pregunta cuenta con una escala de 7 puntos. Todas las preguntas se refieren a las últimas cuatro semanas. En la EDE se incluyen una serie de indicaciones para el entrevistador puntúe cada respuesta de 0 a 7, obtenga puntuaciones medias para cada subescala, y obtenga una puntuación global de la severidad de la patología. Asimismo, se aclara mediante definición el significado de términos como: “pérdida de control”, “gran cantidad de comida”, etc.

A pesar de las ventajas de la EDE, su utilización implica una serie de inconvenientes (Wilfley et al., 1997): para ser fiable, requiere un extenso entrenamiento previo por parte del entrevistador y la administración de la entrevista completa supone unos 60 minutos. Estos inconvenientes (extenso entrenamiento del entrevistador y duración de las entrevistas) son comunes al empleo de cualquier entrevista diagnóstica. No obstante y, a pesar de ello, el uso de cuestionarios no supone una adecuada alternativa como método de identificación de casos, ya que no aseguran la correcta interpretación de los ítems diagnósticos por parte de los sujetos.

Esta limitación de los cuestionarios como método diagnóstico ha sido demostrada por Fairburn y Beglin (1994), Black y Wilson (1996) y Wilfley et al. (1997). Estos autores estudiaron la correlación entre las evaluaciones y diagnósticos obtenidos por la entrevista EDE y el cuestionario *Eating Disorder Examination-Q* (EDE-Q) de Fairburn

y Beglin (1994), que es una versión escrita y reducida de la entrevista *Eating Disorder Examination* (EDE) de Fairburn y Cooper (1993). Las tres investigaciones muestran de forma consistente una correlación positiva, aunque moderada, entre los resultados obtenidos mediante la entrevista EDE y el cuestionario de autoinforme EDE-Q. La correlación fue superior en características que no presentan problemas de definición (i.e., frecuencia de vómitos autoinducidos o promedio -en días por semana- de abuso de laxantes); encontrándose los mayores índices de discrepancia en la evaluación de días por semana en los que tuvieron lugar los episodios de sobreingesta. De forma consistente, estos valores fueron superiores en el cuestionario.

Como alternativa al uso de una entrevista diagnóstica completa, se plantea incluir únicamente aquellos ítems que se refieren a los criterios diagnósticos DSM-IV-TR (APA, 2000) de TCA. Esta medida, que ha sido empleada de forma frecuente en la literatura (Hay, 1998; Colton, Woodside y Kaplan, 1999), facilita el proceso de identificación de casos así como la participación de los sujetos, sin que ello suponga una disminución de la eficacia diagnóstica. Con esta restricción, la duración de las entrevistas se reduce a unos diez o quince minutos por persona.

Otras consideraciones metodológicas

Junto a las cuestiones metodológicas señaladas hasta el momento al hilo de los diferentes tipos de diseños e instrumentos, es preciso tener en cuenta, además, una serie de consideraciones metodológicas básicas a la hora de realizar un estudio epidemiológico de TCA, como edad y sexo de los sujetos, procedimiento de selección de la muestra y adaptación de las pruebas al entorno sociocultural.

Edad de los sujetos que participan en el estudio. La media de edad de establecimiento de la BN se acota en los 18-20 años. Algunos estudios epidemiológicos de BN se realizan con sujetos que no han llegado a esta edad, o que sólo la han pasado parcialmente, con lo que se subestima la prevalencia real del trastorno en población femenina adolescente y juvenil. Otros estudios abarcan una franja de edad muy corta (habitualmente adolescentes), excluyendo casos de aparición temprana o de desarrollo tardío. El estudio de Wlodarczyk-Bisaga y Dolan (1996), riguroso en los demás aspectos metodológicos, utiliza una muestra de mujeres escolares de 14-16 años para estimar la prevalencia de los TCA.

La media de edad de aparición de la AN se sitúa en la adolescencia. No obstante, es preciso utilizar un intervalo de edad amplio, que cubra apariciones precoces y tardías del trastorno. Por ejemplo, Lucas et al. (1991) emplean un intervalo de edad (10-19 años), que podría considerarse como adecuado para estudiar la incidencia de la AN en mujeres americanas.

Procedimiento de selección de la muestra. Algunos estudios estiman la prevalencia de los TCA basándose en los registros de los servicios médicos o psiquiátricos. De esta manera se subestima la prevalencia real del trastorno, que también afecta a población no-ingresada. Sobre todo se subestima la prevalencia de la BN y de los TCANE que, por lo general, no precisan ingreso hospitalario.

En otros casos se emplean procedimientos de muestreo no-probabilístico (i.e.,

selección no-aleatoria de los sujetos), como sujetos voluntarios o accesibles (p. ej. tomar como muestra sólo los alumnos de un centro de enseñanza). Este procedimiento dificulta la generalización de los resultados.

Existe un acuerdo en la necesidad de emplear métodos probabilísticos de selección muestral. De hecho, en casi la totalidad de las investigaciones epidemiológicas realizadas en esta década se utiliza un muestreo aleatorio. En su mayoría se emplea un muestreo multietápico por conglomerados (i.e., seleccionar al azar primero los municipios, después los centros y, al final, los individuos). Este muestreo, aunque no asegura totalmente la representatividad de la muestra (se corre el riesgo, p. ej. de seleccionar municipios o centros –conglomerados naturales– no representativos de la población general), es aceptado en este tipo de estudios.

Sólo se ha empleado un muestreo aleatorio estratificado en el estudio epidemiológico de Bijl, Zessen, Rijk y Langendoen (1998) realizado en Holanda con varios trastornos entre los que se incluyeron los TCA. El primer paso en la selección de la muestra fue la elección de los municipios en función de 5 categorías de urbanización clasificadas por el *Netherlands Central Bureau of Statistics*. El segundo paso fue la elección de una muestra aleatoria de hogares de los registros de correos. El número de hogares elegidos en cada municipio se determinó en función de su población. El tercer paso fue la elección de los individuos. Se utilizó un amplio intervalo de edad (18-64 años). Los diagnósticos fueron basados en el *Composite International Diagnostic Interview* (CIDI), que recoge criterios del DSM-III-R. Se obtuvieron unas tasas de prevalencia vital de 2,0% en AN y 1,1% de BN en mujeres; y de 0,0% de AN y 0,2% de BN en varones.

Sexo de los sujetos. En todos los estudios epidemiológicos en el ámbito de los TCA que incluyen ambos sexos se encuentran unas tasas de prevalencia significativamente superiores en mujeres. La proporción es de 9 a 1 (i.e., 90% de los casos son mujeres). Por esta razón, muchos estudios utilizan muestras de población femenina exclusivamente. No existe un acuerdo, por el momento, acerca de la conveniencia de incluir varones.

Pruebas adaptadas al entorno sociocultural. Un aspecto metodológico al que no se ha concedido en la literatura la importancia que merece es la adecuación del instrumento de barrido y de la entrevista diagnóstica al entorno sociocultural de la población a la que se pretenden generalizar los resultados. En este sentido, es preciso elegir cuestionarios y formatos de entrevista que hayan sido traducidos y adaptados al país al que pertenece la población-diana.

CONCLUSIONES

Como resumen de las observaciones realizadas en los distintos apartados pueden avanzarse una serie de pautas a tener en cuenta en la realización de un estudio epidemiológico de TCA. Para identificar los casos de TCA entre los sujetos de la muestra se hace necesario incluir un *método fiable de detección y diagnóstico de casos*, que siga criterios aceptados universalmente (e.g., criterios diagnósticos DSM-IV-TR; APA, 2000). En este sentido, no pueden considerarse adecuados para la identificación

de casos ni aquellos instrumentos que sólo informan de conductas y/o actitudes de riesgo, ni los cuestionarios de autoinforme (aunque incluyan ítems diagnósticos) ya que no garantizan la correcta interpretación de sintomatología diagnóstica por parte de los sujetos, especialmente en el caso de ítems ambiguos. Sólo las entrevistas que sigan criterios diagnósticos universales (i.e. CIDI, EDE) pueden ser válidas para este fin.

En el caso de que *no* sean diagnosticados todos los sujetos de la muestra (p. ej. en los procedimientos de dos fases), será preciso incluir en el diseño: a) *un método fiable de selección de población de riesgo*, que actúe de barrido en una primera oleada. Este instrumento de barrido debe administrarse a todos los sujetos de la muestra; y b) *un método que permita calcular la sensibilidad y especificidad del instrumento de barrido*, lo cual permitirá hacer una estimación del porcentaje esperable de falsos negativos y, por tanto, acotar de forma más fiable las cifras de prevalencia. Ello se consigue en la práctica incluyendo un *muestreo aleatorio* de los sujetos que hayan puntuado por *debajo* del punto de corte establecido en el cuestionario de barrido; realizando entrevista diagnóstica a estos sujetos además de a aquellos que han igualado o sobrepasado tal puntuación.

Tanto las pruebas de diagnóstico como las de barrido, si las hubiera, deben estar *adaptadas al entorno sociocultural de la muestra seleccionada*, es decir, deben estar traducidas al lenguaje de los sujetos de la muestra y baremadas con poblaciones afines (e.g., población del mismo país o entorno de los sujetos del muestreo).

Se deben utilizar *muestras representativas de la población* a la que se pretende generalizar el estudio. Para lo cual, será conveniente: a) seleccionar muestras suficientemente *amplias* de la Comunidad; y b) incluir en el muestreo, sujetos de *diferentes estratos socioeconómicos y culturales* dentro de la Comunidad. Para conseguir este segundo objetivo en la práctica, se podría seleccionar sujetos de distintos municipios y/o áreas de la Comunidad, así como de diferentes Centros de referencia (p. ej. Centros de Enseñanza públicos, privados concertados, privados no concertados). La técnica de muestreo a emplear ha de ser de tipo *probabilístico* (e.g., muestreo aleatorio estratificado o muestreo aleatorio por conglomerados) para garantizar la participación de todos los sujetos, no sólo de los voluntarios. Se deben, asimismo, utilizar muestras con *intervalos de edad suficientemente amplios* (e.g., de 10-12 años a 21-25), para recoger casos de aparición temprana y desarrollo tardío del trastorno.

Ante la escasez de datos acerca de la relación de variables demográficas generales con la presencia de TCA en población de riesgo, se propone la realización de estudios de prevalencia que evalúen la contribución relativa de estos factores con los TCA.

REFERENCIAS

- American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders: DSM-IV*. Washington, APA.
- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders: DSM-IV-TR*. Washington, APA.
- Bijl, RV., Zessen, G., Ravelli, A., de Rijk, C. & Langendoen, Y. (1998). The Netherlands mental health

- survey and incidence study (NEMESIS): objectives and design. *Social Psychiatry Epidemiology*, 33, 581-586.
- Black, CMD., & Wilson, GT. (1996). Assessment of eating disorders: Interview versus questionnaire. *International Journal of Eating disorders*, 20, 43-50.
- Bushnell, JA., Wells, JE., Hornblow, AR., Oakley-Browne, MA., & Joyce, P. (1990). Prevalence of three bulimia syndromes in the general population. *Psychological Medicine*, 20, 671-680.
- Cotrufo, P., Barretta, V., Monteleone, P., & Maj, M. (1998). Full-syndrome, partial-syndrome and subclinical eating disorders: an epidemiological study of female students in Southern Italy. *Acta Psychiatrica Scandinavica*, 98, 112-115.
- Cummings, SR., Strull, W., Nevitt, MC., & Hulley, SB. (1993). Planificación de las mediciones: cuestionarios. En SB. Hulley & SR. Cummings (Eds.), *Diseño de la investigación clínica: Un enfoque epidemiológico*. Barcelona: Doyma.
- Drewnowski, A., Hopkins, SA. & Kessler, RC. (1988). The prevalence of bulimia nervosa in the U.S. college student population. *American Journal of Public Health*, 78, 1322-1325.
- Fairburn, CG., & Beglin, SJ. (1990). Studies of the epidemiology of bulimia nervosa. *American Journal of Psychiatry*, 147, 401-408.
- Fairburn, CG. & Cooper, Z. (1993). The Eating Disorder Examination. En CG. Fairburn & GT. Wilson (Eds.), *Binge Eating: Nature, Assessment and Treatment*. New York: Guilford Press.
- Fairburn, CG. & Beglin, SJ. (1994). Assessment of eating disorders: Interview or self-report questionnaire? *International Journal of Eating Disorders*, 16, 363-370.
- Fairburn, CG., Welch, SL., Doll, HA., Davies, BA., & O'Connor, ME. (1997) Risk Factors for Bulimia Nervosa. *Archives of General Psychiatry*, 54, 507-517.
- Fairburn, CG., Welch, SL., Doll, HA., Davies, BA. & O'Connor, ME. (1999). Risk factors for bulimia nervosa: a community-based case-control study. *Archives of General Psychiatry*, 56, 468-476.
- Garfinkel, PE., Lin, E., Goering, P., Spegg, C., Goldbloom, DS., Kennedy, S., Kaplan, AS. & Woodside, DB. (1995). Bulimia Nervosa in a Canadian Community Sample: Prevalence and Comparison of Subgroups. *The American Journal of Psychiatry*, 152, 1.052-1.058.
- Gila, A., Castro, J., Toro, J. & Salamero, M. (2004). Subjective Body Image Dimensions in Normal Female Population: Evolution through Adolescence and Early Adulthood. *International Journal of Psychology and Psychological Therapy*, 4, 1-10.
- Gila, A., Castro, J., Gómez, MJ. & Toro, J. (2005). Social and Body Self-Esteem in Adolescents with Eating Disorders. *International Journal of Psychology and Psychological Therapy*, 5, 63-71.
- Hoek, HW. (1991) The incidence and prevalence of anorexia nervosa and bulimia nervosa in primary care. *Psychological Medicine*, 21, 455-460.
- Hoek, HW. (2003). Review of the Prevalence and Incidence of Eating Disorders. *International Journal of Eating Disorders*, 34, 383-396.
- Hsu, LKG. (1996). Epidemiology of the Eating Disorders. *Psychiatry Clinic of North America*, 19, 681-700.
- Instituto Nacional de la Salud (1995). *Protocolo de trastornos del comportamiento alimentario*. Madrid: INSALUD. Secretaría General.
- Johnson-Sabine, E., Wood, W., Patton, G. (1988). Abnormal eating attitudes in London schoolgirls – a prospective epidemiological study: factors associated with abnormal response on screening questionnaires. *Psychological Medicine*, 18, 615-622.
- Kaye, WH., Klump, KL., Frank, GKW. & Strober, M. (2000). Anorexia and Bulimia Nervosa. *Annual*

Review of Medicine, 51, 299-313.

- King, MB. (1989). Eating disorders in a general practice population: prevalence, characteristics and follow-up at 12 to 18 months. *Psychological Medicine Monographic Supplement*, 14, 1-14.
- Lindberg, L. & Hjern, A. (2003) Risk Factors for Anorexia Nervosa: A National Cohort Study. *International Journal of Eating Disorders*, 34, 397-408.
- Lucas, AR., Beard, M., O'Fallon, WM. & Kurland, LT. (1991). 50-year trends in the incidence of anorexia nervosa in Rochester, Minn.: A population-based study. *American Journal of Psychiatry*, 148, 917-922.
- Mizes, JS. & Palermo, TM. (1997). Eating Disorders. In RT. Ammerman & M. Hersen (Eds.), *Handbook of prevention and treatment with children and adolescents* (pp. 572-603). New York: Wiley.
- Nadaoka, T., Oiji, A., Takahashi, S., Morioka, Y., Kashiwakura, M., & Totsuka, S. (1996). An epidemiological study of eating disorders in a northern area of Japan. *Acta Psychiatrica Scandinavica*, 93, 305-310.
- Rathner, G., & Messner, K. (1995). Detection of eating disorders in a small rural town: an epidemiological study. *Psychological Medicine*, 23, 175-184.
- Smith, DE., Marcus, MD., & Eldredge, KL. (1994). Binge eating syndromes: A review of assessment and treatment with an emphasis on clinical application. *Behavior Therapy*, 25, 635-658.
- Steinhausen, HC, Winkler, C., & Meier, M. (1997). Eating Disorders in Adolescence in a Swiss Epidemiological Study. *International Journal of Eating Disorders*, 22, 147-151.
- Sullivan, PF., Bulik, CM., Kindler, KS. (1998). The epidemiology and classification of bulimia nervosa. *Psychological Medicine*, 38, 599-610.
- Vitello, B. & Lederhendler, I. (2000). Research on eating disorders: Current status and future prospects. *Biological Psychiatry*, 47, 777-786.
- Whitaker, A., Jonson, J., Shaffer, D., Rapoport, JL., Kalikow, K., Walsh, BT., Davies, M., Braiman, S. & Dolinsky, A. (1990). Uncommon troubles in young people: Prevalence estimates of selected psychiatric disorders in a nonreferred adolescent population. *Archives of General Psychiatry*, 47, 487-496.
- Wilfrey, DE., Schwartz, MB., Spurrell, EB. & Fairburn, CG. (1997). Assessing the specific psychopathology of binge eating disorder patients: Interview or self-report? *Behavior Research and Therapy*, 35, 1151-1159.
- Wilson, GT. (1993). Assessment of binge eating. En CG. Fairburn & GT. Wilson (Eds.), *Binge eating: Nature, assessment and treatment* (pp. 227-249). New York: Guilford Press.
- Williams, P & MacDonald, A. (1986). The effect of non response bias on the results of two-stage screening surveys. *Society of Psychiatry*, 21, 1.982-1.986.
- Williamson, DA., Anderson, DA., Jackman, LP. & Jackson, SJ. (1995). Assessment of eating disordered thoughts, feelings, and behaviors. En DB. Allison (Ed.), *Handbook of assessment methods for eating behaviors and weight-related problems. Measures, theory, and research*. London: SAGE.
- Włodarczyk-Bisaga, K. & Dolan, B. (1996). A two-stage epidemiological study of abnormal eating attitudes and their prospective risk factors in Polish schoolgirls. *Psychological Medicine*, 26, 1021-1032.
- World Health Organization (1990). *Composite International Diagnostic Interview (CIDI), Version 1.0*. World Health Organization, Geneva

Recibido, 4 marzo, 2004
Aceptado, 15 junio, 2005