ELSEVIER

Journal of Psychosomatic Research 53 (2002) 1043–1047

Journal of Psychosomatic Research

Assessment of eating disorders Comparison of interview and questionnaire data from a long-term follow-up study of bulimia nervosa

Pamela K. Keel*, Scott Crow, Traci L. Davis, James E. Mitchell

Departments of Psychology and Psychiatry, Harvard University, Cambridge, University of Minnesota, Minneapolis, and University of North Dakota and Neuropsychiatric Research Institute, Fargo, USA

Abstract

Objective: This paper examines diagnostic agreement between interview and questionnaire assessments of women participating in a long-term follow-up study of bulimia nervosa. **Methods:** Women ($N = 162$) completed follow-up evaluations comprising questionnaires and either face-to-face or telephone interviews. **Results:** Consistent with previous research, rates of eating disorders were higher when assessed by questionnaire than when assessed by interview; however, rates of full bulimia nervosa were similar. Overall diagnostic agreement was adequate for eating disorders (κ=.64) but poor for bulimia nervosa (κ=.49), with greater agreement between questionnaires and telephone interviews (κ's range: .67–.71) than between questionnaires and face-to-face interviews (κ's range: .35–.58). **Conclusion:** Findings support the possibility that increased rates of eating pathology on questionnaire assessments may be due, in part, to increased candor when participants feel more anonymous. Questionnaire assessments may not be inferior to interview assessments; they may reveal different aspects of disordered eating.

Keywords: Eating disorders; Bulimia nervosa; Assessment

Introduction

Because eating disorders are associated with lifetime prevalence estimates between 0.5% and 3.0% of adolescent and young adult women [1], community-based studies of eating disorders require large samples in order to ensure adequate power for analyses. This is particularly true for studies wishing to evaluate eating disorders in males as well as females, as only 10% of anorexia and bulimia nervosa occur in men [1]. However, it is cumbersome and expensive to conduct individual clinical interviews with large samples. Thus, many large (e.g., > 1000 participants) community-based studies of eating pathology in females and males have relied on questionnaire assessments [2–7], although there are some exceptions [8].

Questionnaire assessments of eating pathology have been designed to reflect the DSM diagnostic criteria for eating disorders [9–13]. Comparison of such surveys and interview data suggests that these surveys demonstrate high sensitivity (few false negatives) but lower specificity (greater false positives) [10,12]. French and colleagues [10] utilized a survey of disordered eating and dieting in a school-based sample of adolescents and then conducted a semistructured expert interview of eating pathology in a subsample of 43 students. They used the percentage of all cases that were identified by interviews and detected by surveys (detection fraction) as a measure of survey sensitivity and they used the percentage of all cases identified by survey and confirmed by interviews (confirmation fraction) as a measure of survey specificity. The range for detection fractions was 69.2–100% across the following behaviors: dieting, vomiting, diet pills, laxatives, fasting and binge eating, indicating that a very high percentage of cases identified by interview were detected by the self-report survey. Conversely, the range for confirmation fractions was 13.6–66.6% for the following behaviors: dieting, vomiting, diet pills, fasting and binge eating, indicating only moderate confirmation of self-reported behaviors during the interview. In a study utilizing a similar survey

* Corresponding author. 1320 William James Hall, 33 Kirkland Street, Cambridge, MA 02138, USA. Tel.: +1-617-495-5592; fax: +1-617-495-4990.

E-mail address: pkeel@wjh.harvard.edu (P.K. Keel).

0022-3999/02/$ – see front matter
PII: S0022-3999(02)00491-9

measure, Leon et al. [12] found that self-report surveys utilized to diagnose eating disorders (anorexia nervosa, bulimia nervosa and EDNOS) demonstrated a 95% detection fraction and a 53% confirmation fraction when compared to semistructured interviews. A combination of high sensitivity and lower specificity would explain the increased prevalence of eating disorders when assessed by questionnaire vs. interview [14].

Factors that may increase false positives on surveys include failure to ask questions concerning necessary diagnostic criteria, confusing questions or lay definitions that differ from clinical definitions and are over-inclusive. For example, women from the community described 42% of their subjective binge episodes (assessed by the Eating Disorders Examination clinical interview) as "binge episodes" even though these would not meet DSM-IV criteria for a binge episode [15]. Unlike questionnaires, interviews provide additional opportunities to clarify questions and ensure the application of clinical definitions. However, explanations of low questionnaire specificity are predicated on the assumption that clinical interviews are superior to questionnaire assessments. French et al. [10] posited that participants may provide more candid responses on questionnaires because they feel greater anonymity. This could explain the increased proportion of individuals reporting disordered eating on surveys than in interviews. If this were true, then data from previous studies [10,12] could be interpreted as representing poor sensitivity in interviews rather than poor specificity of questionnaires. French et al. [10] concluded that more research is required to understand whether respondents over-report disordered eating on questionnaires or whether they are "less willing to disclose such practices in a private interview" (p. 45).

The purpose of this study was to compare diagnoses of bulimia nervosa and eating disorders based on questionnaire and interview assessments in a large sample of women diagnosed with bulimia nervosa more than ten years previously. This sample provided a useful group to study because rates of eating pathology would be higher than in a community-based sample and greater balance in the distribution of the dependent variable increases statistical power. Further, interview assessments were conducted both in person and over the telephone for this sample. Mitchell et al. [16] suggested that participants may be more willing to report eating problems over the phone than in a face-to-face interviews. Thus, two levels of perceived anonymity (face-to-face vs. telephone) were present for interview assessments. All questionnaires were completed in private within the same month of interview assessment. If feelings of embarrassment or shame reduce reporting of eating disorder symptoms in an interview compared to a questionnaire, then we would predict a higher level of agreement between questionnaires and telephone interviews and than between questionnaires and face-to-face interviews.

Methods

Subjects

Subjects for the present study were initially evaluated in the University of Minnesota's Eating Disorders Clinic and diagnosed with bulimia nervosa between 1981 and 1987 as part of one of two studies [16,17]. Follow-up assessments occurred more than a decade after presentation (mean [S.D.] length of follow-up = 11.5 [1.9] years). At baseline, all subjects were required to meet DSM-III criteria for bulimia, with the additional criterion of binge eating coupled with self-induced vomiting or laxative abuse at a minimum frequency of three times each week for 6 months before evaluation. Additional baseline inclusion and exclusion criteria for these subjects are presented in the original papers [16,17]. Of the 222 subjects sought for participation, 22 (9.9%) could not be located, 1 (0.5%) was deceased, 1 (0.5%) was severely disabled and 21 (9.5%) either declined participation or did not complete participation before data collection ended. Thus, 177 women participated in follow-up assessments representing 80.5% of the total sample excepting those individuals unable to participate due to death or disability. The ascertainment rate was 90.1%. Subjects who participated did not differ from subjects who did not participate on any baseline variables (specific results are reported elsewhere [18]). Mean age was 35.33 years (S.D. = 5.14). The sample was predominantly Caucasian (99%) with only two non-Caucasian participants. Due to missing questionnaire data, diagnoses could be generated from both interview and questionnaire data for only 162 women. Thus, for the purposes of comparison, all results are reported for these 162 women.

Procedure

Subjects completed questionnaires at home and participated in interviews that were conducted either in person or by telephone. Written informed consent was obtained from each subject prior to the interview at the time questionnaires were received. All interviews were conducted in private by the first author or research assistant trained using the DSM-IV SCID training tapes and supervision was available from a licensed clinical psychologist. Interviews were audiotaped to determine reliability. Among the 162 participants for the current study, face-to-face interviews were conducted with 92 (57%) subjects and 70 (43%) completed telephone interviews. Among individuals selecting telephone interviews, 74% lived either out of state or more than a 3-h drive from the research office. Analyses of eating disorder diagnostic status (EDDS) at follow-up, as determined from interviews, revealed no significant difference between face-to-face and telephone interviews ($\chi^2 = 1.51$, $df = 1$, $P = .22$). Additionally, no significant differences were found between face-to-face