日本の摂食障害患者に対する Eating Disorder Examination (EDE) の試行：EDE 邦語版の信頼性と妥当性に関する検討

舘 哲朗*1 村上 健*2 鷲塚 輝久*3
生田 憲正*4 西園マーハ文*5 三宅 由子*6

摂食障害診断のための面接評価方法 (EDE) の邦語版を作成し、信頼性と妥当性を検討した。EDE 邦語版を摂食障害患者 61 名に試み、評価者間信頼性、内部一貫性と識別妥当性を検討した。摂食制限、食関心、体重関心、体型関心の 4 つの下位尺度とグローバルスコアの内部一貫性はともに満足な水準を認めた。評価項目、下位尺度、グローバルスコアはともに高い評価者間信頼性を認めた。体型や体重に関する関心の病理性や過食の重症度を測定する尺度として、3 つの下位尺度で神経性無食欲症 (AN) と神経性大食症 (BN) を識別したことから、EDE は摂食障害病理の評価尺度として高い妥当性を持つ、有用な評価方法であることが示された。また、わが国の AN 患者の食事制限、体重関心と体型関心、BN 患者の食事制限は欧米に比べて低値である特徴を認めたことについて考察した。

■ Key words：摂食障害診断面接、神経性無食欲症、神経性大食症、比較文化的問題

2004 年 3 月 8 日受稿、2004 年 9 月 21 日受理
*1 東海大学健康科学部社会福祉学科（連絡先：舘 哲朗、☎ 259-1193 神奈川県伊勢原市望星台）
*2 村上病院 ☎ 東口メンタルクリニック ☎ 国立成育医療センター
*3 東京都精神医学総合研究所 ☎ 国立精神・神経センター精神保健研究所

心身医・2005 年 10 月・第 45 巻 第 10 号 785
Application of the Eating Disorder Examination (EDE) to Japanese Patients with Eating Disorders: Reliability and Validity of the Japanese Version of EDE

Tetsuro Tachi, MD*1 Ken Murakami, MD*2 Teruhisa Washizuka, MD*3
Norimasa Ikuta, MD*4 Aya Maha Nishizono, MD*5
Yuuko Miyake, PhD*6

Abstract

Objective: We produced a Japanese version of the Eating Disorder Examination (EDE) and investigated its reliability and validity.

Methods: EDE Japanese version was trial administered to 69 Japanese patients with eating disorders. The inter-rater reliability, internal consistency, and discriminant validity using eating disorder diagnoses as external standard were examined.

Results: High inter-rater reliability was obtained for evaluation item, subscale and global scores. Satisfactory internal consistency was obtained for 4 subscales and global score. Bulimia nervosa (BN) group showed significantly higher values than anorexia nervosa (AN) group in Weight Concern, Shape Concern and Bulimia subscales.

Discussion: Japanese version of EDE is a measure instrument that can be expected to have high inter-rater reliability and a satisfactory level of internal consistency with respect to evaluation item, subscale and global score. Three EDE subscales discriminated AN from BN, indicating EDE as a highly valid and useful evaluation method for eating disorder psychopathology. Lower scores in Japanese than Western patients for Restraint, Weight Concern and Shape Concern in AN and Restraint in BN suggested cross-cultural issues.

Key words: eating disorder examination, anorexia nervosa, bulimia nervosa, cross-cultural issue

Introduction

Many self-report measures have been developed to evaluate the psychopathology specific to eating disorders. The Eating Disorder Inventory (EDI) is one of the representative measures aiming at comprehensive and multidimensional evaluation of psychological features of abnormal eating behaviors in patients with anorexia nervosa...
(AN) and bulimia nervosa (BN). Since its introduction, the EDI has been proved to possess adequate reliability and validity, and is widely used as a multidimensional scale for evaluating eating disorders. Since self-report instruments can be applied within a short time, they have the advantage of being easy to use in the clinical setting. On the other hand, in research that requires precise evaluation of key psychopathology and symptoms related to eating disorders, definition and identification of psychopathologic trends and symptoms characteristic of eating disorder are difficult to achieve with questionnaires, raising a question on the reliability of evaluation. Take the evaluation of pathological concerns over body shape and weight as an example: identification of the pathological severity may be difficult when the respondents are AN patients, since these concerns are ego-syntonic to them. Moreover, shape concern is commonly found in women at large. Therefore, when making a diagnosis of eating disorders, although evaluation of the degree of obsession over eating and body concerns is important, it is difficult to achieve by questionnaires. Based on these recognitions, Cooper & Fairburn developed the Eating Disorder Examination (EDE).

We have prepared Japanese translations of the 12th edition of EDE and conducted studies to verify whether these instruments are applicable to Japanese patients. The objectives of the present study were to trial administer the Japanese versions of EDE to Japanese patients with eating disorders, and to demonstrate that the EDE can be used in Japanese patients as a diagnostic method by analyzing the inter-rater reliability and internal consistency together with discriminant validity of the EDE. Also, by comparing the EDE results with those reported in Western countries, we shall discuss the major psychopathological symptoms of eating disorders, and furthermore the characteristics of Japanese patients with eating disorders.

Method

1 EDE

The EDE comprises items that evaluate the key psychopathological features of eating disorders, and also diagnostic items that evaluate the frequencies of binge and weight control behaviors. There are 22 items for evaluating the major psychopathology of eating disorders. For each item, the rater questions the state of the subject for the past 4 weeks, and scores on a 7-grade scale. These items are categorized into 4 areas of specific psychopathology: Restraint, Food Concern, Weight Concern, and Shape Concern. For each of the four subscales, all evaluation item scores are totaled and divided by the number of items to contain the subscale score. The four subscale scores were averaged to obtain the global score. In addition to the above four subscale scores, the Bulimia subscale score was included up to EDE 11th version. Prior to this study, we prepared the Japanese version of EDE after obtaining the approval from Professor Fairburn. The Japanese translations were back translated to confirm that the contents of the Japanese versions were almost the same as the original versions.

2 Participants

Among the patients who visited the Psychiatry Outpatient Department of Tokai University Hospital for the main reason of eating disorders, 51 subjects who
gave informed consent to participate in this study were enrolled. Among the subjects of the eating disorder follow-up study conducted at Keio University Hospital in Tokyo, 18 subjects who gave informed consent were also enrolled. The disease types according to the lifetime diagnosis based on structured clinical interview for DSM-IV \(^7\) (SCID) were anorexia nervosa restricting type (AN-R) in 10 cases, anorexia nervosa binge-purging type (AN-BP) in 20 cases, bulimia nervosa purging type (BN-P) in 22 cases, bulimia nervosa non-purging type (BN-NP) in 5 cases, binge eating disorder (BED) in 7 cases, and eating disorder not otherwise specified (ED-NOS) in 5 cases. Table 1 shows characteristics of participants.

### Table 1 Characteristics of participants

<table>
<thead>
<tr>
<th></th>
<th>AN-R (n=10)</th>
<th>AN-BP (n=20)</th>
<th>BN-P (n=22)</th>
<th>BN-NP (n=5)</th>
<th>BED (n=7)</th>
<th>ED-NOS (n=5)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age</strong></td>
<td>20.8 (6.6)</td>
<td>25.5 (5.9)</td>
<td>24.9 (4.8)</td>
<td>26.8 (5.0)</td>
<td>21.9 (3.3)</td>
<td>19.2 (2.6)</td>
</tr>
<tr>
<td><strong>BMI</strong></td>
<td>15.3 (3.2)</td>
<td>15.6 (2.6)</td>
<td>20.3 (3.0)</td>
<td>22.1 (1.9)</td>
<td>23.5 (1.8)</td>
<td>22.5 (11.6)</td>
</tr>
<tr>
<td><strong>Duration of illness</strong>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>less than 1 year</td>
<td>6 (60.0)</td>
<td>4 (20.0)</td>
<td>1 (4.5)</td>
<td>0 (0.0)</td>
<td>1 (14.3)</td>
<td>1 (20.0)</td>
</tr>
<tr>
<td>1-3 years</td>
<td>1 (10.0)</td>
<td>4 (20.0)</td>
<td>4 (18.2)</td>
<td>1 (20.0)</td>
<td>4 (57.1)</td>
<td>3 (60.0)</td>
</tr>
<tr>
<td>4-9 years</td>
<td>2 (20.0)</td>
<td>7 (35.0)</td>
<td>12 (54.5)</td>
<td>2 (40.0)</td>
<td>1 (14.3)</td>
<td>1 (20.0)</td>
</tr>
<tr>
<td>more than 10 years</td>
<td>1 (10.0)</td>
<td>5 (25.0)</td>
<td>5 (22.7)</td>
<td>2 (40.0)</td>
<td>1 (14.3)</td>
<td>0 (0.0)</td>
</tr>
</tbody>
</table>

*Values shown are means (standard deviations) of age

**Values shown are means (standard deviations) of BMI

***Values shown are subjects' numbers (%) of subjects

Both doctors scored independently.

### 4 Statistical analysis

The inter-rater reliability was analyzed by calculating the degree of agreement between the two raters who conducted EDE, using \( \chi \) statistics. The correlation coefficients (Pearson's \( r \)) for the subscale and global scores were also calculated. The reliability of EDE was also assessed from the viewpoint of internal consistency as a scale estimated by the \( \alpha \) reliability coefficients. The validity of EDE was assessed by analyzing the discriminant validity using the diagnoses of current presentation as the external standard.

### Results

#### 1 Reliability of EDE

Table 2 shows the inter-rater reliability of the evaluation item scores expressed in \( \chi \) statistics, and the inter-rater reliability of the four subscale scores and global score expressed in Pearson's \( r \). The \( \alpha \) coefficients for the four subscale scores of Restraint, Eating Concern, Weight Concern, and Shape Concern and global score were 0.75, 0.65, 0.79, 0.79, and 0.81.
Correlation between EDE and specific diagnoses of eating disorders

Table 3 shows the EDE subscale and global scores for 26 cases of AN and 15 cases of BN diagnosed based on current presentation. When the two groups were compared using t-test, BN group showed significantly higher values than AN group in Weight Concern, Shape Concern and Bulimia subscales.

Discussion

Reliability of EDE Japanese version

Inter-rater reliability was assessed step-wise for evaluation item scores, subscale scores and global score. Beglin\(^6\) reported the original EDE showed high inter-rater reliability ($\kappa$ statistics) of 0.70 or above for all the items. In the present study, all 22 items except one showed 0.63 or above, indicating a satisfactory level of inter-rater reliability. The item “discomfort seeing body” had the lowest score. When the inter-rater score distribution was examined, a high correlation (Pearson’s $r$) of 0.87 was observed between the scores of the two raters, indicating that the inter-rater reliability of this item is not necessarily low. The low inter-rater reliability expressed in $\kappa$ statistics thus reflects the fact that the evaluation criteria of the two raters were slightly different, which lowers the rate of complete agreement between the raters. Accordingly, better definition of the evaluation criteria may increase the inter-rater consistency.

Several studies have investigated the inter-rater reliability of the four subscale scores and the global score using the original EDE. Wilson & Smith\(^8\) reported Pearson’s $r$ of 0.97 or above for all the scales, and Rosen et al.\(^9\) reported values ranging from 0.83 to 0.99. The results of the

<table>
<thead>
<tr>
<th>Table 2 Inter-rater reliability of EDE</th>
<th>$\kappa$ statistics</th>
<th>Pearson’s $r$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Restraint subscale</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Restrained over eating</td>
<td>0.65</td>
<td>0.94</td>
</tr>
<tr>
<td>Avoidance of eating</td>
<td>0.87</td>
<td></td>
</tr>
<tr>
<td>Food avoidance</td>
<td>0.90</td>
<td></td>
</tr>
<tr>
<td>Dietary rules</td>
<td>0.92</td>
<td></td>
</tr>
<tr>
<td>Empty stomach</td>
<td>0.90</td>
<td></td>
</tr>
<tr>
<td>Eating Concern subscale</td>
<td>1.00</td>
<td>0.99</td>
</tr>
<tr>
<td>Preoccupation with food and calories</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fear of losing control over eating</td>
<td>0.86</td>
<td></td>
</tr>
<tr>
<td>Social eating</td>
<td>0.76</td>
<td></td>
</tr>
<tr>
<td>Eating in secret</td>
<td>0.69</td>
<td></td>
</tr>
<tr>
<td>Guilt about eating</td>
<td>0.81</td>
<td></td>
</tr>
<tr>
<td>Weight Concern subscale</td>
<td></td>
<td>0.95</td>
</tr>
<tr>
<td>Importance of weight</td>
<td>0.75</td>
<td></td>
</tr>
<tr>
<td>Reaction to prescribed weighing</td>
<td>0.68</td>
<td></td>
</tr>
<tr>
<td>Preoccupation with shape or weight</td>
<td>0.92</td>
<td></td>
</tr>
<tr>
<td>dissatisfaction with weight</td>
<td>0.78</td>
<td></td>
</tr>
<tr>
<td>Desire to lose weight</td>
<td>0.93</td>
<td></td>
</tr>
<tr>
<td>Shape Concern subscale</td>
<td>0.98</td>
<td></td>
</tr>
<tr>
<td>Flat stomach</td>
<td>0.72</td>
<td></td>
</tr>
<tr>
<td>Preoccupation with shape or weight</td>
<td>0.92</td>
<td></td>
</tr>
<tr>
<td>Importance of shape</td>
<td>0.81</td>
<td></td>
</tr>
<tr>
<td>Fear of weight gain</td>
<td>0.80</td>
<td></td>
</tr>
<tr>
<td>Dissatisfaction with shape</td>
<td>0.93</td>
<td></td>
</tr>
<tr>
<td>Discomfort seeing body</td>
<td>0.25</td>
<td></td>
</tr>
<tr>
<td>Avoidance of exposure</td>
<td>0.63</td>
<td></td>
</tr>
<tr>
<td>Feeling of fatness</td>
<td>0.85</td>
<td></td>
</tr>
<tr>
<td>Bulimia subscale</td>
<td></td>
<td>0.88</td>
</tr>
<tr>
<td>Global score</td>
<td></td>
<td>0.99</td>
</tr>
</tbody>
</table>

Discussion

<table>
<thead>
<tr>
<th>Table 3 EDE norms for anorexia nervosa and bulimia nervosa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anorexia nervosa $(n = 26)$</td>
</tr>
<tr>
<td>-------------------------------------------------------------</td>
</tr>
<tr>
<td>Mean</td>
</tr>
<tr>
<td>-------------------------------------------------------------</td>
</tr>
<tr>
<td>Restraint</td>
</tr>
<tr>
<td>Eating Concern</td>
</tr>
<tr>
<td>Weight Concern*</td>
</tr>
<tr>
<td>Shape Concern*</td>
</tr>
<tr>
<td>Bulimia*</td>
</tr>
<tr>
<td>Global score</td>
</tr>
</tbody>
</table>

*Significant difference in EDE score between anorexia nervosa and bulimia nervosa: $t$-test, $p < 0.05$.
Table 4  EDE norms for anorexia nervosa and bulimia nervosa in Western countries

<table>
<thead>
<tr>
<th></th>
<th>Anorexia nervosa</th>
<th>Anorexia nervosa</th>
<th>Bulimia nervosa</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(n=47)</td>
<td>(n=50)</td>
<td>(n=53)</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
</tr>
<tr>
<td>Restraint</td>
<td>3.17</td>
<td>1.47</td>
<td>3.8**</td>
</tr>
<tr>
<td>Eating Concern</td>
<td>2.17</td>
<td>1.62</td>
<td>2.8</td>
</tr>
<tr>
<td>Weight Concern</td>
<td>2.40</td>
<td>1.48</td>
<td>2.9*</td>
</tr>
<tr>
<td>Shape Concern</td>
<td>2.85</td>
<td>1.22</td>
<td>3.5**</td>
</tr>
<tr>
<td>Bulimia</td>
<td>1.58</td>
<td>1.55</td>
<td>1.9</td>
</tr>
</tbody>
</table>

*Significant difference compared with EDE score of Japanese cases (data shown in Table 2) ;
**Significant difference compared with EDE score of Japanese cases (data shown in Table 2) ;
***Significant difference compared with EDE score of Japanese cases (data shown in Table 2) ;

The present study showed a satisfactory level ranging from 0.88 to 0.99, resembling those of Rosen et al. These results indicate that the Japanese version of EDE is a measure instrument that can be expected to have high inter-rater reliability with respect to evaluation item scores, subscale scores and global score.

The α reliability coefficients for the subscale scores of Restraint, Food Concern, Weight Concern, and Shape Concern and the global score were high, ranging from 0.65 to 0.81. In research conducted in Western countries, the α coefficients were in the ranges of 0.68 to 0.90. In the Japanese version also, the four subscales and global subscale can be considered to have attained a satisfactory level of internal consistency.

2 Discriminant validity of EDE and the major psychopathology of eating disorders

The results of the EDE subscale scores for Eating Concern, Weight Concern, and Shape Concern indicated that BN has a higher pathological level than AN. This result is essentially similar to the reports from Western countries (Table 4). Studies that examined the discriminant validity of the original EDE reported that all the subscales except Restraint discriminated BN from AN-R⁶, and that Weight Concern and Shape Concern subscales discriminated BN from AN-R⁹. In the present study also, three EDE subscales, with the exception of Restraint and Eating Concern, discriminated AN from BN. The result that Shape and Weight Concerns were markedly strong in BN supports Russell’s perception⁵⁰ that a morbid fear of becoming fat is the key psychopathology. Evaluation of pathology in this area is indispensable in clinical studies that examine the therapeutic outcome and course of BN. In this context, EDE may be used as a highly valid evaluation method.

It is a well-known clinical observation that restraint is closely related to the development and maintenance of BN, and the pathological level of restraint in BN has been a topic of great interest. However, our present data showed that the Restraint subscale score was not high in BN, rather lower than that of AN. Wilson & Smith⁶ found no significant difference in the EDE Restraint subscale between BN and AN-R, and expressed the view that extreme dietary concern may be a necessary but clearly not a sufficient condition for the development of BN. Our results support their opinion.

Finally, we shall compare the EDE
scores in the present study with reports in Western countries\(^3\)\(^{-6}\) and discuss the differences (Table 4). When the EDE subscale scores obtained in this study are compared with those in Western countries by \(t\)-test, several differences are observed both in AN and BN. For AN, Restraint, Weight Concern and Shape Concern were significantly lower than those reported in Western countries. The same trend was observed for the EDI-2 Drive for Thinness subscale. The Drive for Thinness of AN reported by Garner et al.\(^{11}\) was 15.4, compared to 9.4 in our previous study\(^{11}\), showing a significant difference. Other Japanese studies\(^{12\text{(213)}}\) also reported Drive for Thinness of 4.7 to 8.1 in AN patients, also lower compared to the data of Western countries.

Lee et al.\(^{14}\) studied females in Hong Kong and reported that the fear of fatness was not strong in many of the AN patients, or rather such fear was absent in some cases, stating that in non-Western countries drive for thinness probably does not constitute the central psychopathological feature. The significantly lower score for the EDE Restraint subscale in BN in the present study (1.79) compared to reports in Western countries (3.14 to 3.7) may suggest a similar background. On the other hand, according to the report of Kusano et al.\(^{15}\) on the EDI-2 results of a non-clinical group, Japanese female college students showed significantly stronger Drive for Thinness than students of Western countries. Moreover, Nakai\(^{13}\) reported no significant difference in Drive for Thinness between Japanese AN patients and normal controls. Considering the present results and the above reports together, the following socio-cultural background of eating disorders is hypothesized. Drive for thinness or a dieting trend has become a matter of great concern for young women in Japan, and has developed recently into a social boom over and above the Western society. However, the problem that perpetuates chronicity of AN; in other words, continued restraint and difficulty of recovery from low body weight, may not be a measure of severity of psychopathology related to the drive for thinness.

References

3) Fairburn CG, Cooper Z: The eating examina-

●お知らせ

第5回日本認知療法学会
会期：平成17年12月9日（金）～10日（土）
会場：名古屋銀行協会会館（名古屋市中区丸の内2-4-2）
会長：貝谷久宣（医療法人和楽会）
テーマ：「科学としての認知療法」
参加費：会員3,000円，非会員6,000円（会員・非会員とも）
お問合せ先：第5回日本認知療法学会事務局
〒102-0075 東京都新宿区神楽坂6-42
TEL：03-5206-5281 FAX：03-5206-5285 E-mail：ninchi@ghgroupjp.com
※詳しくは第5回日本認知療法学会のホームページをご覧ください。
http://www.fuanclinic.com/ninti/ninti_top.htm