

tolerable to extremely severe, leading to fear of death. He had concomitant symptoms of palpitation, sensation of shortness of breath, feeling of choking, and simultaneous fear of losing control, and his episodes met the DSM-IV-TR criteria for a panic attack. When such an attack occurred, he was unable to remain still because of extreme fear, and psychiatric symptom management was urgently required.

We informed the patient that his vital signs were normal and that there was no threat to his life. We explained that an autonomic imbalance arising from anxiety was the cause of his attacks, promised relief from his symptoms, and showed him how to perform diaphragmatic breathing. After the intravenous bolus injection of diazepam (5 mg), his panic attack temporarily improved, but he was unable to rest. Paroxetine (20 mg/day) and clonazepam (0.5 mg/day) were administered through a gastrostoma, and the intensity and frequency of his attacks gradually decreased. Although his attacks ceased completely 14 days after surgery, his fear of additional attacks persisted, with no symptoms of agoraphobia; thus, the administration of paroxetine and clonazepam was continued. His physical condition recovered, and he was discharged with a tolerably mild degree of neck stiffness.

#### Comment

We have described two representative cases of first panic attack episodes in head and neck cancer patients that appeared to have been triggered by neck stiffness after radical surgery. In general, physical symptoms resulting from anxiety, such as palpitations or breathlessness, tend to be the main component of panic attack symptoms.

In the present cases, the features of their panic attack symptoms were distinct in that their most terrifying bodily sensation was neck stiffness, as if they were being strangled.

The onset of these panic attacks can be partially explained by the cognitive model of panic (Fig. 1).<sup>7</sup> Cervical discomfort, similar to tightening, probably increased the patients' apprehension—particularly under postoperative circumstances where the patients were forced to stay in bed with many drains and other lines inserted. Apprehension induces physical sensations associated with anxiety, such as palpitations and shortness of breath, feelings of

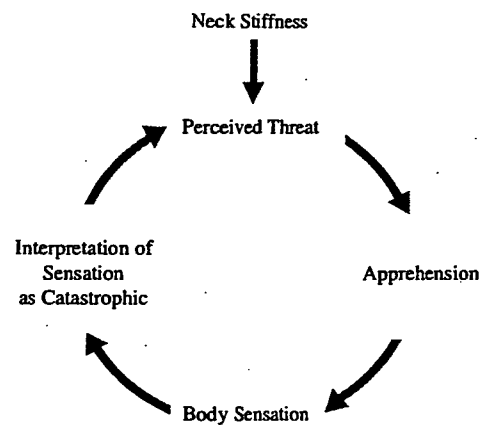


Fig. 1. The panic attack process, as illustrated by a cognitive model of panic.

choking, and so on. The patients interpreted both these apprehension-induced sensations and their neck stiffness as being catastrophic and were extremely fearful of death. Once this vicious cycle of symptoms had started, the patients were unable to cope by themselves. Their fear and physical sensations continued to grow, ultimately resulting in a panic attack.

Neck stiffness and pain sometimes appear after radical neck surgery and may be an aspect of the postradical neck pain syndrome.<sup>4-6</sup> Injury to the superficial cervical plexus is suspected to be one of the causes of this pain. Sist et al. reported the characteristics of 25 cases of post-radical neck pain syndrome. In their series, most of the patients reported that the intensity of their pain varied throughout the day, and many psychological or autonomic symptoms—including fear, punishment, and feelings of suffocation—emerged concomitant with the pain.<sup>6</sup> Although standardized psychiatric evaluations were not performed in this previous study, panic attacks or limited-symptom attacks might have occurred in some of the patients.

In the present cases, the panic attacks were alleviated after using a psychoeducational approach and commonly used psychopharmacological treatments. In general, delirium is first suspected when patients express restlessness after operations, but clinicians should be aware of the possible development of panic attacks after radical neck surgery. Patients undergoing radical head and neck surgery may be at risk for a panic attack; further research regarding this hypothesis is necessary.

Ken Shimizu, MD, PhD  
Hiroya Kinoshita, MD  
Tatsuo Akechi, MD, PhD  
Yosuke Uchitomi, MD, PhD  
Psycho-Oncology Division, Research  
Center for Innovative Oncology  
National Cancer Center Hospital East  
Kashiwa, Chiba, Japan

Mizuo Andoh, MD  
Head and Neck Surgery Division  
National Cancer Center Hospital  
Tokyo, Japan

doi:10.1016/j.jpainsymman.2007.08.007

### *References*

1. Derogatis LR, Morrow GR, Fetting J, et al. The prevalence of psychiatric disorders among cancer patients. *JAMA* 1983;249:751–757.
2. Slaughter JR, Jain A, Holmes S, et al. Panic disorder in hospitalized cancer patients. *Psychooncology* 2000;9:253–258.
3. American Psychiatric Association. Diagnostic and statistical manual of mental disorders, 4th ed., text revision. Washington, DC: American Psychiatric Association, 2000.
4. Portenoy RK. Cancer pain: pathophysiology and syndromes. *Lancet* 1992;339:1026–1031.
5. Foley KM. Pain syndromes in patients with cancer. *Med Clin North Am* 1987;71:169–184.
6. Sist T, Miner M, Lema M. Characteristics of post-radical neck pain syndrome: a report of 25 cases. *J Pain Symptom Manage* 1999;18:95–102.
7. Clark DM. A cognitive approach to panic. *Behav Res Ther* 1986;24:461–470.



## The new GRID Hamilton Rating Scale for Depression demonstrates excellent inter-rater reliability for inexperienced and experienced raters before and after training

Hideaki Tabuse <sup>a,b</sup>, Amir Kalali <sup>c</sup>, Hideki Azuma <sup>b</sup>, Norio Ozaki <sup>d</sup>, Nakao Iwata <sup>e</sup>, Hiroshi Naitoh <sup>e</sup>, Teruhiko Higuchi <sup>f</sup>, Shigenobu Kanba <sup>g</sup>, Kunihiro Shioe <sup>h</sup>, Tatsuo Akechi <sup>b</sup>, Toshi A. Furukawa <sup>b,\*</sup>

<sup>a</sup> Holy Cross Hospital, Kujiri 2431-160, Izumi-cho, Toki, 509-5142 Japan

<sup>b</sup> Department of Psychiatry and Cognitive-Behavioral Medicine, Nagoya City University Graduate School of Medical Sciences, Mizuho-cho, Mizuho-ku, Nagoya, 467-8601 Japan

<sup>c</sup> International Society for CNS Drug Development, San Diego, CA, USA

<sup>d</sup> Department of Psychiatry, Nagoya University Graduate School of Medicine, Nagoya, Japan

<sup>e</sup> Department of Psychiatry, Fujita Health University, Toyoake, Japan

<sup>f</sup> Musashi Hospital, National Center of Neurology and Psychiatry, Kodaira, Japan

<sup>g</sup> Department of Neuropsychiatry, Kyushu University Graduate School of Medical Sciences, Fukuoka, Japan

<sup>h</sup> Department of Neuropsychiatry, Interdisciplinary Graduate School of Medicine and Engineering, University of Yamanashi, Yamanashi, Japan

Received 16 January 2006; received in revised form 15 May 2006; accepted 23 July 2006

### Abstract

The Hamilton Rating Scale for Depression (HAMD) is the *de facto* international gold standard for the assessment of depression. There are some criticisms, however, especially with regard to its inter-rater reliability, due to the lack of standardized questions or explicit scoring procedures. The GRID-HAMD was developed to provide standardized explicit scoring conventions and a structured interview guide for administration and scoring of the HAMD. We developed the Japanese version of the GRID-HAMD and examined its inter-rater reliability among experienced and inexperienced clinicians ( $n=70$ ), how rater characteristics may affect it, and how training can improve it in the course of a model training program using videotaped interviews. The results showed that the inter-rater reliability of the GRID-HAMD total score was excellent to almost perfect and those of most individual items were also satisfactory to excellent, both with experienced and inexperienced raters, and both before and after the training. With its standardized definitions, questions and detailed scoring conventions, the GRID-HAMD appears to be the best achievable set of interview guides for the HAMD and can provide a solid tool for highly reliable assessment of depression severity.  
© 2006 Elsevier Ireland Ltd. All rights reserved.

**Keywords:** Depressive disorder; Psychiatric status rating scales; Mental status schedule; Reproducibility of results

### 1. Introduction

The Japanese Society of Clinical Psychopharmacology has long realized the need to standardize the administration of the Hamilton Rating Scale for Depression

\* Corresponding author. Tel.: +81 52 853 8271; fax: +81 52 852 0837.

E-mail address: [furukawa@med.nagoya-cu.ac.jp](mailto:furukawa@med.nagoya-cu.ac.jp) (T.A. Furukawa).

(HAMD) (Hamilton, 1960), the *de facto* international standard for the assessment of depression (Furukawa et al., 2005), within Japan and appointed a team headed by Dr. Higuchi to develop a model training program in 2000. In the course of these efforts, we learned that a group of researchers had met in the USA in 1999 and proposed to establish a common set of standards for scoring and administering the HAMD that would be acceptable to the Food and Drug Administration and be used by pharmaceutical, academic and clinical researchers. This proposal led to the formation of the Depression Rating Scale Standardization Team (DRSST), a group of individuals representing clinicians, academia, government and the pharmaceutical industry. The goal of this group was to standardize the administration and scoring of the HAMD without significantly altering the original intent of Hamilton's items or the scoring profile rather than to develop a new instrument (Kalali et al., 2002; Bech et al., 2005).

The product of their efforts is the GRID-HAMD, which has three components: the GRID scoring system, (scoring intensity and frequency separately to obtain the severity score), the manual of scoring conventions with detailed anchor descriptions and more behavioral exemplars, and a semi-structured interview guide. The DRSST clarified and operationalized ambiguous anchor descriptions and incorporated the new definitions into the individual items. The GRID-HAMD can be downloaded free of charge at the International Society for CNS Drug Development homepage (<http://www.iscdd.org>). Given the many versions of the scale in use, the DRSST concluded that standardization would improve the current scale and lay the groundwork for development of a new scale.

The Japanese team felt that the GRID-HAMD would set a new standard in depression rating and decided to develop the Japanese training program around it. We developed the Japanese version of the GRID-HAMD (see Section 2) and then conducted a model training course for the GRID-HAMD in March 2004. The primary purpose of the study is to examine the inter-rater reliability of the Japanese version of the GRID-HAMD among experienced and inexperienced Japanese psychiatrists and psychologists, how rater characteristics may affect it, and how training can enhance it.

## 2. Methods

### 2.1. Participants

Psychiatrists ( $n=52$ ), clinical psychologists ( $n=12$ ) and medical students ( $n=6$ ) from three university medical schools in Japan (Nagoya City University, Nagoya Uni-

versity and Fujita Health University) took part in a full day training course for the newly developed Japanese version of the GRID-HAMD. Of the 70 participants, 20 had no previous experience with any version of the HAMD, whereas 17 had administered it between one and five times and 33 had administered it six or more times. However, only 16 of the last group had ever received formal training in the administration of the instrument. The mean (S.D.) of clinical experience was 6.3 (6.1) years for the psychiatrists and 3.5 (3.0) years for the clinical psychologists.

### 2.2. Instrument

The Japanese version of the GRID-HAMD was developed in collaboration with the DRSST. The original English version of the GRID-HAMD was translated into Japanese by TAF. A team of seven psychiatrists, all of them experts in depression treatment and research, checked the translation and amended it where necessary, based upon the consensus of the team. Two research assistants, both proficient in English and one with a Bachelor's degree in psychology, and both blind to the original English version, then back-translated the Japanese translation of the probe questions into English. AK checked the backtranslation and pointed out possible discrepancies, based upon which TAF retranslated the questioned sentences into Japanese. This process was repeated three times, until AK was able to ascertain semantic equivalence between the original and back-translated versions.

### 2.3. Procedure

We used three pairs of videotapes of pre- and post-treatment administration of the HAMD. Two pairs used simulated Japanese patients (one man and one woman) and the other pair used a simulated English patient. The Japanese man, woman and their interviewers were played by professional actors and actresses, based on rough scenarios but including a substantial amount of ad lib interactions. The participants' general impression was that the patients were very well played and appeared natural, but that the interviewers appeared rather stiff. The English patient's interviews had Japanese subtitles. Each interview lasted between 15 and 40 min. The experts' consensus total scores for the six videotapes were 26 for the Japanese man pre-treatment, 10 post-treatment, 37 for the Japanese woman pre-treatment, 19 post-treatment, 21 for the English woman pre-treatment and 0 post-treatment. These videotapes were prepared independently of and before our training workshop for the GRID-

HAMD. The interviewers in these videotapes by and large followed the conventions of the Structured Interview Guide for the Hamilton Depression Rating Scale (SIGH-D) (Williams, 1988), which sometimes did not probe specifically enough into the frequency of the symptoms during the last week.

The participants in the workshop used GRID-HAMD to rate each interview. When the videotape failed to ask for the frequency, the participants were instructed to assume that the frequency was 50% of the time. This was the case for items 2, 3, 7, 10, 11, 12, and 13 of the pre-treatment videotape of the Japanese man, for items 2, 11, and 13 of the post-treatment videotape of the Japanese man, for items 2, 3, 6, 7, 10, 12, 13, and 15 of the pre-treatment video of the Japanese woman, and for items 2, 5, 7, 10, 12, and 13 of the post-treatment video of the Japanese woman. In other words, 24 out of the 68 items (35%) required participants to rely on this rating convention.

Because the rating difficulty might differ between the videotapes of the Japanese man and woman, the participants were randomly divided into two groups, and each group saw either the man's videos or the woman's videos first. There was no discussion immediately following the

two videos. The videos therefore served as pre-training and post-training assessments of the raters' reliability. After this pre-training assessment in the reliability of the GRID-HAMD, their training began with a lecture on the history of the Hamilton Rating Scale for Depression and a general discussion of assessment in psychiatry. The training of the GRID-HAMD formed the core of the workshop and used the English woman's videotapes. After scoring each English woman's videotapes, possible discrepancies and questions were discussed among the participants and the trainers. The three pairs of videotapes were therefore presented during this 1-day course as shown in Fig. 1.

#### 2.4. Analyses

The inter-rater reliability for each item of the GRID-HAMD and for its total score was estimated by way of the ANOVA intraclass correlation coefficient (ICC) (one-way random effects model, single rater) of the SPSS (SPSS Inc., 2002). Because of its intrinsic paradoxical characteristic whereby we obtain low ICC despite high agreement (Feinstein and Cicchetti, 1990), we did not calculate ICC

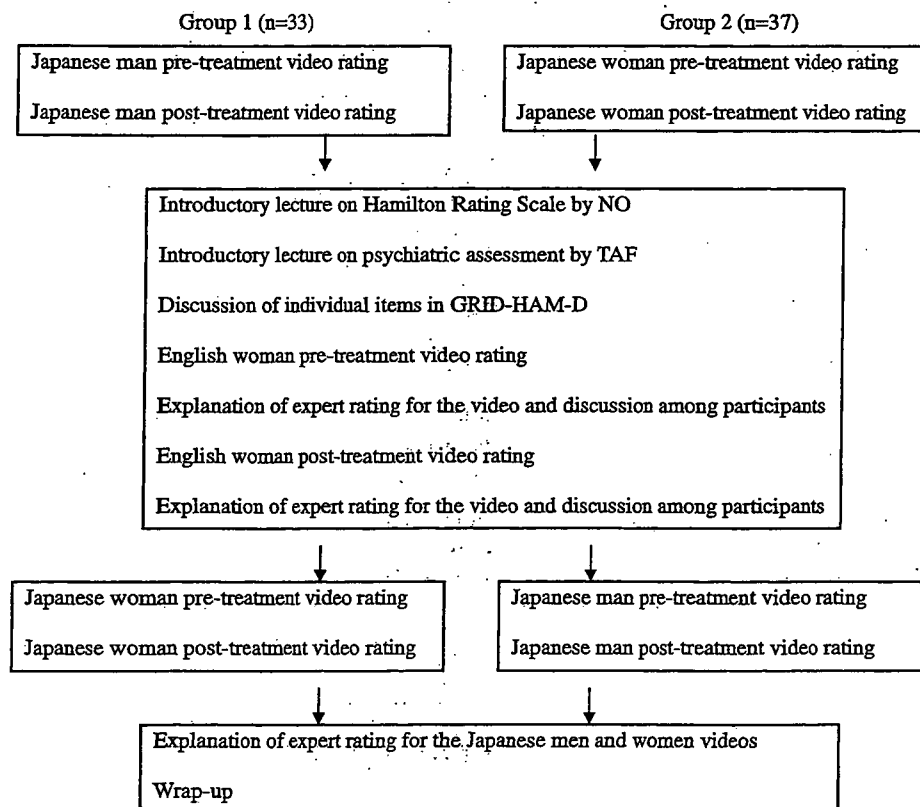


Fig. 1. Procedure of the model training program.

when one rating predominated ( $\geq 0.90$  of all the ratings) for a particular item. It has been suggested that ICCs between 0.61 and 0.80 are “substantial” and those greater than 0.80 are “almost perfect” (Landis and Koch, 1977).

Because we were interested in the effects of experience and training, we subgrouped our participants based on their previous experience with the HAMD as follows.

Group A ( $n=20$ ) No previous experience with the HAMD.  
Group B ( $n=17$ ) Have administered the HAMD between one and five times

Group C ( $n=17$ ) Have administered the HAMD six or more times, but have never had formal training in its administration

Group D ( $n=16$ ) Have administered the HAMD six or more times, and have received formal training in its administration.

To examine the influence of the rating convention of assigning 50% frequency to such items as where the interviewer failed to ask for frequency in the videotape, we ran a supplementary sensitivity analysis by comparing

Table 1  
ANOVA ICC for each item and the total score of the GRID-HAMD for the four subgroups of participants before and after training

Item	Group A		Group B		Group C		Group D	
	Before training	After training	Before training	After training	Before training	After training	Before training	After training
1 Depressed mood	0.83 (0.64–0.95)	0.89 (0.76–0.97)	0.78 (0.50–0.96)	0.84 (0.61–0.97)	0.91 (0.74–0.98)	0.84 (0.62–0.97)	0.84 (0.57–0.97)	0.92 (0.55–0.99)
2 Guilt	0.69 (0.41–0.89)	0.60 (0.30–0.86)	0.41 (0.07–0.84)	0.71 (0.40–0.94)	0.67 (0.34–0.93)	0.59 (0.27–0.91)	0.70 (0.34–0.94)	0.58 (0.19–0.91)
3 Suicide	0.89 (0.75–0.97)	0.87 (0.70–0.96)	0.92 (0.79–0.99)	0.90 (0.73–0.98)	0.92 (0.78–0.99)	0.89 (0.72–0.98)	0.94 (0.83–0.99)	0.97 (0.91–1.00)
4 Insomnia, early	0.90 (0.76–0.97)	0.90 (0.77–0.97)	1.00 (1.00–1.00)	1.00 (1.00–1.00)	0.72 (0.41–0.95)	0.85 (0.63–0.97)	0.86 (0.62–0.98)	1.00 (1.00–1.00)
5 Insomnia, middle	0.78 (0.55–0.93)	0.75 (0.50–0.92)	0.66 (0.32–0.93)	0.79 (0.51–0.96)	0.78 (0.50–0.96)	0.70 (0.39–0.94)	0.87 (0.64–0.98)	0.78 (0.47–0.96)
6 Insomnia, late	0.86 (0.69–0.96)	0.93 (0.83–0.98)	0.91 (0.74–0.98)	0.97 (0.91–1.00)	0.87 (0.66–0.98)	0.95 (0.86–0.99)	0.92 (0.77–0.99)	0.96 (0.8–0.99)
7 Work and activities	0.63 (0.34–0.87)	0.69 (0.42–0.90)	0.70 (0.37–0.94)	0.72 (0.40–0.94)	0.73 (0.42–0.95)	0.75 (0.47–0.95)	0.73 (0.39–0.95)	0.71 (0.35–0.94)
8 Psychomotor retardation	0.67 (0.39–0.89)	0.80 (0.68–0.94)	0.67 (0.34–0.93)	0.78 (0.50–0.96)	0.85 (0.63–0.97)	0.83 (0.61–0.97)	0.59 (0.20–0.91)	0.61 (0.22–0.92)
9 Psychomotor agitation	na	na	-0.05 (-0.18–0.43)	-0.05 (-0.18–0.43)	0.13 (-0.10–0.67)	na	na	0.00 (-0.21–0.57)
10 Anxiety, psychic	0.75 (0.50–0.92)	0.38 (0.08–0.74)	0.64 (0.30–0.92)	0.75 (0.45–0.95)	0.73 (0.41–0.95)	0.74 (0.45–0.95)	0.82 (0.54–0.97)	0.96 (0.88–0.99)
11 Anxiety, somatic	0.88 (0.73–0.96)	0.82 (0.61–0.94)	0.80 (0.53–0.96)	0.87 (0.66–0.98)	0.89 (0.70–0.98)	0.86 (0.65–0.97)	0.93 (0.80–0.99)	0.89 (0.69–0.98)
12 Loss of appetite	0.87 (0.72–0.96)	0.91 (0.78–0.97)	0.89 (0.71–0.98)	0.95 (0.84–0.99)	0.88 (0.69–0.98)	0.82 (0.58–0.97)	0.91 (0.74–0.99)	0.87 (0.64–0.98)
13 Somatic symptoms, general	0.53 (0.21–0.82)	0.50 (0.20–0.81)	0.64 (0.30–0.93)	0.48 (0.13–0.87)	0.57 (0.21–0.90)	0.36 (0.07–0.81)	0.63 (0.25–0.93)	0.44 (0.05–0.86)
14 Sexual interest	na	na	na	na	na	na	na	na
15 Hypochondriasis	0.60 (0.30–0.86)	0.62 (0.33–0.87)	0.53 (0.18–0.89)	0.73 (0.42–0.95)	0.69 (0.37–0.94)	0.83 (0.60–0.97)	0.85 (0.59–0.97)	0.73 (0.39–0.95)
16 Loss of weight	0.63 (0.34–0.87)	0.65 (0.36–0.88)	0.64 (0.30–0.92)	0.79 (0.51–0.96)	0.71 (0.39–0.94)	0.84 (0.73–0.98)	0.77 (0.46–0.96)	0.71 (0.35–0.94)
17 Loss of insight	na	na	na	na	0.03 (-0.15–0.55)	na	na	na
Total	0.95 (0.87–0.98)	0.95 (0.87–0.91)	0.93 (0.82–0.99)	0.95 (0.86–0.99)	0.97 (0.91–1.00)	0.95 (0.85–0.99)	0.97 (0.90–1.00)	0.99 (0.96–1.00)

Figures in parentheses indicate the 95% confidence intervals.

na = not applicable due to too little variation because the particular score predominated and more than 90% of the obtained ratings were the same.

Group A ( $n=20$ ): No previous experience with the HAMD.

Group B ( $n=17$ ): Have administered the HAMD between one to five times.

Group C ( $n=17$ ): Have administered the HAMD six or more times, but have never had formal training in its administration.

Group D ( $n=16$ ): Have administered the HAMD six or more times, and have received formal training in its administration.

the average ANOVA ICCs between items for which the interviewers did not ask about frequency in more than half of the videotapes (items 2, 7, 10, 12, and 13) and those for which the interviewers asked (items 1, 4, 5, 6, and 15).

### 3. Results

Table 1 shows the ANOVA ICCs and their 95% confidence intervals for each item and the total score of the GRID-HAMD as applied to the Japanese man and woman's videotapes, for Groups A through D, both before and after training with lectures and practice with the English woman's videotapes. Excluding items 9, 14, and 17 (Psychomotor agitation, Sexual interest, and Loss of insight), which showed too little variation among raters to calculate meaningful chance-corrected agreement coefficients, and item 13 (Somatic symptoms, general), which often had ANOVA ICCs below 0.60, the inter-rater reliability of individual items was already largely in the substantial to excellent range before the training and did not show much increase after the training. Thus the respective averages of the ANOVA ICCs for individual items were 0.75 and 0.74 for Group A before and after training, 0.73 and 0.81 for Group B, 0.78 and 0.79 for Group C, and 0.81 and 0.79 for Group D. The ANOVA ICCs for the total score were almost perfect for all groups both before and after the training (range: 0.93 to 0.99). The average ICC for the items where the interviewers asked for frequency was 0.83 (range: 0.70 to 0.92) and that for the items where they failed to ask and where therefore the subjects were instructed to assume 50% frequency was 0.69 (range: 0.52 to 0.89).

### 4. Discussion

Our results suggest that when we relied on the GRID-HAMD scoring conventions, the inter-rater reliability of the total score was excellent to almost perfect and that satisfactory inter-rater reliability for individual items was also achievable, even with inexperienced raters and even without training. These findings are at variance with some previous studies on inter-rater reliability for HAMD items, which often reported poor reliability at the individual item levels. Cicchetti and Prusoff (1983) assessed reliability before treatment initiation and 16 weeks later at trial end. Before treatment, only one item was sufficiently reliable and 13 items had coefficients below 0.50. After treatment, again only one item was sufficiently reliable and 11 items had coefficients below 0.50. Craig et al. (1985) also found that only one item had adequate inter-rater reliability. On the other hand, Moberg et al. (2001) reported that nine items

showed adequate reliability when the standard HAMD depression scale was administered, but all items showed adequate reliability when the scale was administered with the SIGH-D interview guidelines of Williams (1988). Our findings appear to extend theirs. Narita et al. (2002) pointed out specific weaknesses/ambiguities in the rating instructions in the SIGH-D, especially with regard to items for middle insomnia, somatic anxiety, loss of weight, depersonalization/derealization, and loss of insight; all of these are well anchored in the GRID-HAMD.

On the other hand, our results suggested that inter-rater reliability for general somatic symptoms may be low. However, we suspect that this was due to the difference in emphasis between SIGH-D item 13 and GRID-HAMD item 13, the former following the traditional HAM-D interpretation and focusing on heaviness and aches and the latter emphasizing fatigue and anergia in accordance with DSM-IV criterion symptoms.

With regard to the total score of the HAMD, most of the previous studies reported substantial to satisfactory inter-rater reliability, with ICCs ranging from 0.46 to 0.99 (Bagby et al., 2004). Some investigators provided evidence that the skill level or expertise of the interviewer and the provision of structured queries and scoring guidelines affect reliability (O'Hara and Rehm, 1983; Hooijer et al., 1991). Our findings suggest that with the use of explicit scoring conventions as outlined in the GRID-HAMD, even inexperienced raters can achieve satisfactory inter-rater reliability. We failed to show a significant effect of expertise or training; possibly because of the ceiling effect of these already high baseline reliability coefficients, although the raw scores do hint at even higher reliability coefficients after training and for more experienced users.

Weaknesses of the present study may be as follows: Firstly, the present study is based on videotaped interviews with simulated patients. Although the actor and actress played their roles naturally, with much ad lib interaction, the generalizability of the present findings to bona fide patients cannot be taken for granted and warrant another study. However, it should be pointed out that experienced physicians have been reported to be unable to differentiate standardized patients from real patients when they were sent unannounced into a physician's office, even when the physician was told in advance that this would be occurring (Kobak et al., 2003). The videotaped reliability study with simulated patients may also have inflated reliability estimates in comparison with test-retest design with real patients, which would more accurately reproduce clinical realities. Secondly, we used videotapes that had been made prior to and independently of our workshop for the GRID-HAMD.

The interviewers in the videotapes therefore did not abide by the GRID-HAMD conventions but roughly followed the SIGH-D questions. They therefore did not probe specifically enough about the frequency of some symptoms. The rating convention of assigning a 50% frequency to such items may have inflated the reliability estimates, but our sensitivity analysis did not support this possibility. Had the interviewers in the videotapes followed the GRID-HAMD interview guides, it is safe to assume that reliability could have been even higher. Thirdly, the videotaped interviews in the present study were such that there was little variation for three out of 17 items of the GRID-HAMD. We could therefore not ascertain satisfactory reliability for these items. In future studies we need to prepare videotapes that allow more variation in ratings for these items. Fourthly, although the ICCs did not change materially before and after the training, it must be pointed out that the present findings do not obviate the need for clinical expertise in depression assessment, as almost all the participants had substantial clinical experience already. In order to assure satisfactory rater performance, the raters' ability to conduct assessments on real patients is important in itself, in addition to the reliability of the instrument (Lipsitz et al., 2004). Lastly, the present study was conducted in Japanese with the Japanese version of the GRID-HAMD. The Japanese version was developed in strict adherence to the established back-translation procedure to ensure its linguistic equivalence with the English original, and we believe the present findings can be replicated with the original version as well, as it is thanks to the well-structured, adequately explained nature of the GRID-HAMD and not to any particularities of its Japanese version that we could achieve satisfactory reliability. Strictly speaking, however, the cross-cultural generalizability of the present findings must await independent replication studies in English and other languages and cultures.

Recently, a comprehensive review of the HAMD by Bagby et al. (2004, 2005) concurred that the GRID-HAMD is a major improvement over the previous versions in developing clear structured interview prompts and scoring guidelines, and in standardizing the scoring system. However, the retention of "loss of insight" that makes neither a conceptual nor an empirical contribution to the severity of depression or the lack of such DSM-IV criterion symptoms as "loss of concentration" remain major difficulties with the GRID-HAMD. Also the report from a 2002 National Institute of Health sponsored conference in the US on the assessment of depression and anxiety in clinical trials recommended the GRID-HAMD as the optimal way to administer the HAMD. A recent National Institutes of Health sponsored conference on

assessment of suicidality also recommended the GRID-HAMD as the preferred version of the HAMD for assessing suicidality.

We feel that the GRID-HAMD is the best achievable set of semi-structured guides for the HAMD, the *de facto* standard in depression rating for over four decades, and this fact was corroborated in the present study by its robust reliability findings. In conclusion, the GRID-HAMD appears to provide a solid tool for highly reliable assessment of depression severity for both experienced and inexperienced mental health professionals.

## References

- Bagby, R.M., Ryder, A.G., Schuller, D.R., Marshall, M.B., 2004. The Hamilton Depression Rating Scale: has the gold standard become a lead weight? *American Journal of Psychiatry* 161, 2163–2177.
- Bagby, R.M., Schuller, D.R., Ryder, A.G., Marshall, M.B., 2005. Dr. Bagby and colleagues reply. *American Journal of Psychiatry* 162, 2397–2398.
- Bech, P., Engelhardt, N., Evans, K.R., Gibertini, M., Kalali, A.H., Kobak, K.A., Lipsitz, J.D., Williams, J.B., Pearson, J.D., Rothman, M., 2005. Why the Hamilton Depression Rating Scale endures. *American Journal of Psychiatry* 162, 2396.
- Cicchetti, D.V., Prusoff, B.A., 1983. Reliability of depression and associated clinical symptoms. *Archives of General Psychiatry* 40, 987–990.
- Craig, T.J., Richardson, M.A., Pass, R., Bregman, Z., 1985. Measurement of mood and affect in schizophrenic inpatients. *American Journal of Psychiatry* 142, 1272–1277.
- Feinstein, A.R., Cicchetti, D.V., 1990. High agreement but low kappa: I. The problems of two paradoxes. *Journal of Clinical Epidemiology* 43, 543–549.
- Furukawa, T.A., Streiner, D.L., Azuma, H., Higuchi, T., Kamijima, K., Kanba, S., Ozaki, N., Aoba, A., Murasaki, M., Miura, S., 2005. Cross-cultural equivalence in depression assessment: Japan–Europe–North American study. *Acta Psychiatrica Scandinavica* 112, 279–285.
- Hamilton, M., 1960. A rating scale for depression. *Journal of Neurology, Neurosurgery and Psychiatry* 23, 56–62.
- Hooijer, C., Zitman, F.G., Griez, E., van Tilburg, W., Willemsse, A., Dinkgreve, M.A., 1991. The Hamilton Depression Rating Scale (HDRS): changes in scores as a function of training and version used. *Journal of Affective Disorders* 22, 21–29.
- Kalali, A., Williams, J., Kobak, K.A., Lipsitz, J., Engelhardt, N., Evans, K., Olin, J., Pearson, J., Rothman, M.L., Bech, P., 2002. The new GRID HAM-D: pilot testing and international field trials. *International Journal of Neuropsychopharmacology* 5, S147–S148.
- Kobak, K.A., Lipsitz, J.D., Feiger, A., 2003. Development of a standardized training program for the Hamilton Depression Scale using internet-based technologies: results from a pilot study. *Journal of Psychiatric Research* 37, 509–515.
- Landis, J.R., Koch, G.G., 1977. The measurement of observer agreement for categorical data. *Biometrics* 33, 159–174.
- Lipsitz, J., Kobak, K., Feiger, A., Sikich, D., Moroz, G., Engelhardt, A., 2004. The Rater Applied Performance Scale: development and reliability. *Psychiatry Research* 127, 147–155.
- Moberg, P.J., Lazarus, L.W., Mesholam, R.I., Bilker, W., Chuy, I.L., Neyman, I., Markvart, V., 2001. Comparison of the standard and structured interview guide for the Hamilton Depression Rating Scale



- in depressed geriatric inpatients. *American Journal of Geriatric Psychiatry* 9, 35–40.
- Narita, T., Kim, N., Nakane, Y., Ozaki, N., 2002. Reliability and validity of the Structured Interview Guide of the Hamilton Depression Rating Scale (SIGH-D). *Japanese Journal of Clinical Psychopharmacology* 6, 77–82 [in Japanese].
- O'Hara, M.W., Rehm, L.P., 1983. Hamilton Rating Scale for Depression: reliability and validity of judgments of novice raters. *Journal of Consulting and Clinical Psychology* 51, 318–319.
- SPSS Inc., 2002. *SPSS for Windows Version 11.5*. SPSS Inc., Chicago.
- Williams, J.B., 1988. A structured interview guide for the Hamilton Depression Rating Scale. *Archives of General Psychiatry* 45, 742–747.

## Regular Article

# Emotional distress and its correlates among parents of children with pervasive developmental disorders

ATSUROU YAMADA, MD,<sup>1,2</sup> MIYOSHI SUZUKI,<sup>2</sup> MISUZU KATO, MSc,<sup>2</sup> MIE SUZUKI,<sup>2</sup> SAYUMI TANAKA, MD,<sup>2,3</sup> TAKUO SHINDO, MD,<sup>2,4</sup> KAZUO TAKETANI, MD,<sup>2</sup> TATSUO AKECHI, MD, PhD<sup>1</sup> AND TOSHI A. FURUKAWA, MD, PhD<sup>1</sup>

<sup>1</sup>Department of Psychiatry and Cognitive-Behavioral Medicine, Nagoya City University Graduate School of Medical Sciences, <sup>2</sup>Yagoto Hospital, Nagoya, <sup>3</sup>Department of Neuropsychiatry, Toyohashi Municipal Hospital and <sup>4</sup>Toyota West Hospital, Aichi, Japan

### Abstract

A number of studies have reported that parents of autistic children face higher levels of stress, but few studies examined the stress associated with the home care of children with pervasive developmental disorders (PDD) other than autistic disorder. The aims of the present study were therefore to (i) evaluate the emotional stress level of parents caring for their children with PDD; and (ii) explore the correlates of their emotional stress. Participants were 147 families (147 mothers and 122 fathers) of 158 children with PDD (42 with autistic disorder, 35 with Asperger's disorder and 81 with PDD not otherwise specified). K6 was used to measure the stress level of the parents. Marital relationships and personality were assessed with the Intimate Bond Measure and the NEO Five-Factor Inventory, respectively. The parents also rated the characteristics of their children with PDD through the Pervasive Developmental Disorder–Autism Society Japan Rating Scale (PARS). The mean K6 score of the mothers was significantly higher than that of the women in the general population in Japan. Stepwise multiple regression indicated that the emotional stress of the mothers was correlated with the personality traits of Neuroticism and Agreeableness, perceived Control by the husband, and the children's PARS score. Clinicians can deliver better service by paying appropriate attention to the emotional distress of mothers of children with not only autistic disorder but also other PDD.

### Key words

Asperger's disorder, autistic disorder, emotional stress, parents, pervasive development disorder.

## INTRODUCTION

A number of studies have established that the parents of autistic children face higher levels of stress than parents of children with other chronic diseases or developmental disorders.<sup>1–8</sup> Studying distress among parents of children with autistic and related disorders is very important because understanding the situations leading to high stress should assist treatment providers

to help these parents and, through this, to promote the children's welfare as well. However, few studies have examined the stress associated with the home care of children with pervasive developmental disorders (PDD) other than autistic disorder.

For example, only a few reports have addressed the problems associated with the care of children with Asperger's disorders. According to several reports, providing home care to children with Asperger's disorders or high-functioning autistic disorders is also associated with high levels of stress, as compared to the stress levels associated with the home care of normal children. Mothers often face, among other problems, challenges to their mental health when caring for such children,<sup>9,10</sup> and the stress levels seem to be higher in mothers than in fathers.<sup>11</sup> Furthermore, this kind of

Correspondence address: Atsurou Yamada, MD, Department of Psychiatry and Cognitive-Behavioral Medicine, Nagoya City University Graduate School of Medical Sciences, Mizuho-Cho, Mizuho-ku, Nagoya 467-8601, Japan. Email: atsurou.yamada@mbr.nifty.com

Received 26 March 2007; revised 23 July 2007; accepted 30 July 2007.

stress seems to show some correlation with the behavioral problems exhibited by the children.<sup>9</sup> There are fewer reports on the stress associated with providing home care to children with pervasive developmental disorders not otherwise specified (PDD-NOS), even though this condition is more prevalent than autistic disorder.

These circumstances pertaining to the study of PDD other than autistic disorder vis-à-vis those of autistic disorder may be attributable to the following factors: (i) Asperger's disorder and PDD-NOS began to attract attention as an individual diagnostic entity only recently, unlike autistic disorder; (ii) children with Asperger's disorders or some children with PDD-NOS often have less severe symptoms and higher levels of cognitive functioning than children with autistic disorders, which might lead to lower levels of stress associated with the home care of these children and explain the lesser attention paid by investigators to this condition; (iii) specific criteria have been defined for the diagnosis of autistic disorder, but no such clear criteria have been established for that of PDD-NOS; and (iv) children with Asperger's disorders or PDD-NOS are unlikely to receive care at expert facilities, making it difficult for an adequate number of subjects to be recruited for studies on this condition.

We therefore undertook the present study in order to (i) evaluate the emotional stress level of parents caring for children with autistic disorders or other PDD including Asperger's disorders and PDD-NOS; and (ii) explore the correlates of emotional stress among parental characteristics such as personality and marital relationships as well as among the children's intelligence, clinical characteristics and behaviors.

## METHODS

### Measures

#### *Parental measures*

**Brief demographic questionnaire and familial profiles.** Questions were asked about the age, occupation, educational career and history of psychiatric treatment of the parents and the presence/absence of additional children with PDD in the family.

**K6.** K6 is a self-report questionnaire containing six items that tap general psychological distress in the past 30 days.<sup>12</sup> This scale has been developed using the modern psychometric theory and has been shown to outperform some existing scales.<sup>13,14</sup> Its Japanese version has been developed recently using the standard back-translation method and has been validated in a

large general population psychiatric epidemiological survey ( $n = 2436$ ).<sup>15</sup> The internal consistency reliability was satisfactory with Cronbach's alpha coefficient of 0.77, and the questionnaire showed an excellent efficiency at screening anxiety and mood disorders in the general population, with area under the curve of receiver operating characteristic of 0.94 (95% confidence interval: 0.88–0.99).<sup>15</sup>

**Intimate Bond Measure.** This measure is designed to evaluate the nature of the relationship with the marital partner. It is a self-report measure for evaluation of two dimensions of marital relationships assessing partner's Care and Control, and is composed of 24 questions. The Care dimension reflects care expressed emotionally as well as physically, with constructs of warmth, consideration, affection and companionships. The Control dimension suggests domination, instructiveness, criticism, authoritarian attitudes and behaviors. The reliability and validity of this measure have been confirmed.<sup>16,17</sup>

**NEO Five-Factor Inventory.** This is a concise measure of the five domains of personality according to the five-factor model. Twelve items are provided for each of the five dimensions of Neuroticism, Extraversion, Openness, Agreeableness, and Conscientiousness. Each item is answered on a 5-point Likert scale (strongly disagree, disagree, neutral, agree, strongly agree). The Japanese version has shown good reliability and validity in the general population of Japan.<sup>18</sup>

#### *Child measures*

**Intelligence and other clinical characteristics.** Data pertaining to the characteristics of individual children, such as the age, sex, diagnosis, IQ or developmental quotient (DQ), and the time of first visit to the clinic, were collected at the outpatient clinic. The IQ of 63 children was measured with the Wechsler Intelligence Scale for Children-Third Edition (WISC-III),<sup>19,20</sup> and the DQ of 88 children was measured with the Kyoto Scale of Psychological Development 2001 (KSPD). The KSPD is a standardized scale and is used in Japan to measure the developmental age from infant to adult in Postural-Motor, Cognitive-Adaptive and Language-Social areas.<sup>21</sup>

**Pervasive Developmental Disorder-Autism Society Japan Rating Scale.** This scale is composed of 57 questions in eight domains of the characteristics of children with PDD. Of the 57, 34 questions pertain to infancy, 33 to childhood and 33 to the behaviors

observed during adolescence and adulthood. Twenty-nine questions pertaining to behaviors observed during both childhood and adolescence/adulthood were selected for the present study.

In a previous study it was shown that each of the infantile retrospective evaluation items and adolescence/adulthood current evaluation items had good internal consistency and discriminant validity.<sup>22</sup>

**Questions about school.** The following three questions about school were asked. The frequency of school attendance was rated on a four-grade scale: (i) very frequent (4–5 days per week); (ii) relatively frequent (2–3 days per week); (iii) rare (1 day per week); and (iv) not at all. Troubles with other children were rated on a four-grade scale: (i) absent; (ii) occasional; (iii) frequent; and (iv) very frequent. Participation in school lessons was rated on a four-grade scale: (i) able to participate at most times; (ii) unable to participate sometimes; (iii) often unable to participate; and (iv) unable to participate ever.

### Participants

The study was explained to the consecutive sample of 198 families of the children with PDD who visited the Outpatient Pediatric Neuropsychiatry Clinic of the Toyohashi Municipal Hospital, Japan, during the 3-month period September–November 2006. The diagnosis of the mental disorder was confirmed by the lead author (a pediatric psychiatrist), based on the DSM-IV diagnostic criteria.<sup>23</sup>

Parents of children satisfying the following criteria were considered eligible to the study: (i) children were of elementary school or junior high school age; and (ii) children were diagnosed as having autistic disorder, Asperger's disorder or PDD-NOS according to the DSM-IV diagnostic criteria<sup>23</sup> at the outpatient clinic. Siblings of the children who satisfied (i) and (ii) were also included even if they themselves did not visit the clinic during the aforementioned period.

We excluded: (i) children for whom parents did not serve as the main care providers at home; (ii) children or parents whose mother tongue was not Japanese; and (iii) children with tuberous sclerosis.

Children with the severest form of mental retardation were not excluded so long as they satisfied the diagnostic criteria of PDD.

After giving a brief account at the outpatient clinic, we handed a questionnaire to those consenting to the study, requesting them to mail it after filling out the necessary information. In the case of parents with two or more children with PDD, child measures were performed for each child.

**Table 1.** Subject data

	Mother	Father
<i>n</i>	147	122
Age (years)	38.3 ± 4.6	41.0 ± 5.7
Parent-child relationship		
Real	147	121
Foster	0	1
Marital status		
Married	138	122
Divorced	9	0
Employment		
Housework or out of work	74	1
<40 h per week	64	5
>40 h per week	9	116
Education		
Up to junior high school	3	6
High school	77	66
Junior college or beyond	67	50
No. children with PDD		
One	124	102
Two or more	23	20
History of psychiatric treatment		
Positive	27	6
Negative	120	116

PDD, pervasive developmental disorder.

Responses were collected from 147 families (74.2%), covering 158 children. The data of 145 families were obtained between September and December 2006. Table 1 lists parents' demographic data. The present study included 158 children including 42 with autistic disorders (35 boys), 35 with Asperger's disorder (29 boys) and 81 with PDD-NOS (65 boys). The mean ages of the children were 9 years 3 months for the autistic disorder group (SD: 2 years 6 months), 10 years 3 months for the Asperger's disorder group (SD: 2 years 5 months), and 9 years 7 months for the PDD-NOS group (SD: 2 years 7 months). The duration of outpatient treatment was 2 ± 1 years for the autistic disorder group, 1 year 9 months ± 1 year for the Asperger's disorder group, and 1 year 9 months ± 1 year for the PDD-NOS group. The distribution of the IQ/DQ in the three groups of children was as follows: IQ/DQ >70: autistic disorder group, 4; Asperger's disorder group, 34; PDD-NOS group, 60; IQ/DQ 50–69: autistic disorder group, 10; PDD-NOS group, 16; IQ/DQ 35–49: autistic disorder group, 14; PDD-NOS group, 4; IQ/DQ 20–34: autistic disorder group, 7; IQ/DQ <20: autistic disorder group, 5; PDD-NOS group, 1; IQ/DQ unknown: autistic disorder group, 2; Asperger's disorder group, 1.

The study protocol was approved by the Ethics Committee of Toyohashi Municipal Hospital. Written

consent was obtained from each of the parents who filled out and returned the questionnaire.

**Data analysis**

First, the K6 scores of the mothers and fathers were compared with those of the general population in Japan,<sup>24</sup> while controlling for the effect of age because K6 scores tend to decrease with advancing age.

Next, we analyzed variables to explore the characteristics that might serve as correlates of the emotional stress in the parents having one child with PDD using Pearson's correlation coefficient. Regarding the influence of the husband-wife relationship and the personalities of the mother and father, we explored the correlation between K6 scores of the mother and father using Pearson's correlation coefficient, in cases of one affected child and of two or more affected children with PDD.

Finally, stepwise multiple regression of the K6 scores of the mother and father was conducted, with the child characteristics, husband-wife relationship and personality traits serving as the independent variables. The parents were divided into two groups: one group with only one affected child and the other with two or more children affected by PDD. In order to avoid false-positive findings due to multiple comparisons,  $P < 0.01$  was regarded as denoting statistical significance. We used SPSS for Windows 11.5J (SPSS, Chicago, IL, USA) for all data analyses.

**RESULTS**

The mean K6 scores of the mothers and fathers were  $5.3 \pm 5.1$  and  $2.9 \pm 3.4$ , respectively. The mean K6 score of the mothers was significantly higher than that of the women of the corresponding age in the general population (mean =  $2.2 \pm 3.1$ ;  $t = -6.9$ , d.f. = 196,  $P < 0.001$ ), while the mean K6 score for the fathers corrected for age did not differ significantly from that of the men in the general population (mean,  $2.5 \pm 3.5$ ;  $t = -1.0$ , d.f. = 320,  $P = 0.32$ ).

None of the child's age, duration of outpatient care, subtype of PDD, or IQ/DQ was correlated with the K6 scores of the mothers or the fathers. When the relationship of the mothers' and fathers' K6 scores to the child's situation at school was analyzed, a significant correlation was noted between the mother's K6 score and the child's troubles with schoolmates ( $r = 0.31$ ,  $P < 0.001$ ). The Pervasive Developmental Disorder-Autism Society Japan Rating Scale (PARS) scores were significantly associated with K6 scores for mothers ( $r = 0.43$ ,  $P < 0.001$ ) and fathers ( $r = 0.26$ ,  $P = 0.009$ ).

Table 2 shows the correlations between K6 scores and Intimate Bond Measure (IBM) scores, the personality

**Table 2.** Correlations between K6 scores and IBM scores or NEO-FFI scores

	IBM		NEO-FFI					
	Care by spouse	Control by spouse	Neuroticism	Extraversion	Openness	Agreeableness	Conscientiousness	
	<i>n</i>	<i>r</i>	<i>n</i>	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>	
Mothers having one child with PDD	117	-0.23	124	0.56*	-0.36*	-0.05	0.43*	0.12
Fathers having one child with PDD	102	0.01	102	0.60*	-0.11	-0.12	0.20	0.28*
Mothers having two or more children with PDD	21	0.29	23	0.67*	-0.44	-0.36	-0.15	0.05
Fathers having two or more children with PDD	20	0.04	20	0.63*	-0.24	-0.09	0.37	0.52

\*  $P < 0.01$ .

IBM, Intimate Bond Measure; NEO-FFI, NEO Five-Factor Inventory; PDD, pervasive developmental disorder.

Table 3. Stepwise multiple regression analysis for K6 scores

		Standardized coefficients $\beta$	<i>t</i>	adjusted $R^2$
Mothers having one child with PDD ( <i>n</i> = 112)	NEO-FFI Neuroticism	0.39**	5.1	0.48
	NEO-FFI Agreeableness	0.23**	3.0	
	PARS	0.25**	3.6	
	IBM Control	0.22**	3.1	
Mothers having two or more children with PDD ( <i>n</i> = 20)	NEO-FFI Neuroticism	0.73**	4.6	0.51
Fathers having one child with PDD ( <i>n</i> = 97)	NEO-FFI Neuroticism	0.61**	7.4	0.36
Fathers having two or more children with PDD ( <i>n</i> = 19)	NEO-FFI Neuroticism	0.63**	3.3	0.36

\*\*  $P < 0.01$ . IBM, Intimate Bond Measure; NEO-FFI, NEO Five-Factor Inventory; PARS, Pervasive Developmental Disorder-Autism Society Japan Rating Scale; PDD, pervasive developmental disorder.

traits as measured on NEO Five-Factor Inventory (NEO-FFI) for mothers and fathers. The K6 scores of the mothers having one child with PDD were directly proportional to the father's degree of Control. No such relationship was noted in the K6 scores of mothers having two or more children with PDD. Among parents having only one child with PDD, the K6 scores of the mothers were positively correlated with their Neuroticism and Agreeableness scores and negatively with their Extraversion scores, while the K6 scores of the fathers were positively correlated with the Neuroticism and Conscientiousness scores. In contrast, when mothers and fathers having two or more children with PDD were analyzed, the K6 scores of both the mothers and fathers were correlated with only the Neuroticism scores.

Finally, we explored the influences of all of the aforementioned variables on the K6 scores (Table 3). In families having only one child with PDD, the K6 scores of the mothers were significantly correlated with Neuroticism, Agreeableness, PARS and IBM Control scores. In families having two or more children with PDD, the mothers' K6 scores were correlated with only their Neuroticism scores. In the case of fathers, the K6 scores were correlated with only the Neuroticism scores, irrespective of whether they had one or more affected child(ren) with PDD.

## DISCUSSION

We found that higher emotional stress levels are seen among mothers taking care of children with all kinds of PDD. Numerous reports regarding the stress levels in the parents of children with autistic disorders have suggested that the mothers of children with autistic disorders show higher stress levels than the fathers,<sup>2,8,25</sup> that depressive symptoms are often seen,<sup>4,11,26-28</sup> and that symptoms of anxiety are noted frequently.<sup>29</sup> A similar

trend has also been reported for parents of children with Asperger's disorders or high-functioning autistic disorders.<sup>9-11</sup> The present results that the emotional stress levels were higher among the mothers than the fathers of children with PDD is consistent with these previous reports. One of the reasons for this finding may be that mothers tend to be more involved in the care of their children than fathers.<sup>26,30,31</sup>

What characteristics of children with PDD increase the stress level of mothers? Hastings *et al.* evaluated the problematic behaviors of children with autistic disorders using the Parent Report version of the Developmental Behavior Checklist and reported that the mother's stress level can be predicted from the number of problematic behaviors of the child but not adaptive behavior or autistic symptoms.<sup>28</sup> Konstantareas and Homatidis reported that the stress level of the parents of children with autistic disorders became higher if the child commits self-injury.<sup>25</sup> Hastings and Freeman *et al.* reported that mothers were more stressed by the behaviors of children with autistic disorders than fathers.<sup>29,32</sup> In cases of Asperger's disorder and high-functioning autistic disorders also, an association of the behavioral problems (e.g. hyperactivity) with the mother's health has been reported.<sup>9</sup> In the present study the children's troubles with schoolmates were also associated with the mother's emotional stress level, and the PARS scores were associated with the emotional stress level of parents having one child with PDD. PARS is a screening test consisting of items that tap both behavior problems and symptoms of PDD.<sup>22</sup> In cases of children with PDD it is difficult to distinguish their behavior problems from their symptoms. Thus the results of the present study endorse the previous findings.

Are there any other correlates associated with increased emotional stress levels? Surprisingly, irrespective of the number of children with PDD in the

family and the age or IQ/DQ of the child with PDD, the personality tendency of the parents was the most closely associated with the emotional stress levels. In this connection, previous reports have found that hardiness or coping styles may serve as a predictor of satisfactory adaptation of the mother to children with PDD.<sup>7,33,34</sup> Although no previous report has suggested a direct association of the stress level with the personality of the parents, the personality trait of Neuroticism might be the most important correlate of the high emotional stress levels in parents taking care of children with PDD.

Many previous reports have shown that the relationship of a mother to the father is a source of stress for the mother.<sup>28,29</sup> Furthermore, the father's support, in terms of verbal expression, has been shown to serve as a predictor of the course of care.<sup>26</sup> And perceived availability of support appears to be more important than actual support.<sup>7,31</sup> We may therefore say that cooperation by the father and support from the father are probably necessary to alleviate the stress in the mother. Especially the father should avoid dominative and intrusive attitudes and behaviors, and support the mother emotionally as well as physically. This knowledge may be useful for devising new means of intervention.

There are some limitations to the present study. First, only parents of children receiving outpatient care were included. Therefore, it is possible that these parents faced higher levels of stress (higher severity of symptoms, more problematic behaviors etc.) than the parents of children with PDD who do not require outpatient care. Another limitation is related to the fact that we analyzed children with PDD all together, without differentiating their subtypes. There are probably more ambiguities in the differential diagnosis when PDD rather than autistic disorder is selected as the target disorder.

Despite these limitations, responses were collected from a relatively large number of subjects in the present study, and the study analyzed factors that had not been adequately evaluated before (e.g. personality tendency, husband-wife relationship, school etc.). Further studies with a similar design to replicate this finding in the community-dwelling subjects and their families are warranted.

## REFERENCES

- Dumas JE, Wolf LC, Fisman SN, Culligan A. Parenting stress, child behavior problems, and dysphoria in parents of children with autism, Down syndrome, behavior disorders, and normal development. *Exceptionality* 1991; **2**: 97-110.
- Holroyd J. The Questionnaire on Resources and Stress: An instrument to measure family response to a handicapped family member. *J. Community Psychol.* 1974; **2**: 92-94.
- Koegel RL, Schreibman L, Loos LM *et al.* Consistent stress profiles in mothers of children with autism. *J. Autism Dev. Disord.* 1992; **22**: 205-216.
- Olsson MB, Hwang CP. Depression in mothers and fathers of children with intellectual disability. *J. Intellectual Disabil. Res.* 2001; **45**: 535-543.
- Rodrigue JR, Morgan SB, Geffken GR. Psychosocial adaptation of fathers of children with autism, Down syndrome, and normal development. *J. Autism Dev. Disord.* 1992; **22**: 249-263.
- Sanders JL, Morgan SB. Family stress and adjustment as perceived by parents of children with autism or Down syndrome: Implications for intervention. *Child Family Behav. Ther.* 1997; **19**: 15-32.
- Weiss MJ. Hardiness and social support as predictors of stress in mothers of typical children, children with autism, and children with mental retardation. *Autism* 2002; **6**: 115-130.
- Wolf LC, Noh S, Fisman SN, Speechley M. Psychological effects of parenting stress on parents of autistic children. *J. Autism Dev. Disord.* 1989; **19**: 157-166.
- Allik H, Larsson JO, Smedje H. Health-related quality of life in parents of school-age children with Asperger syndrome or high-functioning autism. *Health Qual. Life Outcomes* 2006; **4**: 1.
- Little L. Differences in stress and coping for mothers and fathers of children with Asperger's syndrome and non-verbal learning disorders. *Pediatr. Nurs.* 2002; **28**: 565-570.
- Gray DE. Gender and coping: The parents of children with high functioning autism. *Soc. Sci. Med.* 2003; **56**: 631-642.
- Kessler RC, Andrews G, Colpe LJ *et al.* Short screening scales to monitor population prevalences and trends in non-specific psychological distress. *Psychol. Med.* 2002; **32**: 959-976.
- Kessler RC, Barker PR, Colpe LJ *et al.* Screening for serious mental illness in the general population. *Arch. Gen. Psychiatry* 2003; **60**: 184-189.
- Furukawa TA, Kessler RC, Slade T, Andrews G. The performance of the K6 and K10 screening scales for psychological distress in the Australian National Survey of Mental Health and Well-Being. *Psychol. Med.* 2003; **33**: 357-362.
- Furukawa TA, Kawakami N, Saitoh M *et al.* The performance of the Japanese version of the K6 and K10 in the World Mental Health Survey Japan. *Int. J. Methods Psychiatr. Res.* (in press).
- Wilhelm K, Parker G. The development of a measure of intimate bonds. *Psychol. Med.* 1988; **18**: 225-234.
- Furukawa TA, Hori S, Azuma H *et al.* Parents, personality or partner? Correlates of marital relationships. *Soc. Psychiatry Psychiatr. Epidemiol.* 2002; **37**: 164-168.

18. Yoshimura K, Ono Y, Nakamura K, Nathan JH, Suzuki K. Validation of the Japanese version of the NEO Five-Factor Inventory in a large community sample. *Psychol. Rep.* 2001; **88**: 443–449.
19. Wechsler D. *Manual for the Wechsler Intelligence Scale for Children*, 3rd edn. Psychological Corporation, San Antonio, TX, 1991.
20. Wechsler D. *The Japanese Version of Manual for the Wechsler Intelligence Scale for Children*, 3rd edn. Nihon Bunka Kagakusha, Tokyo, 1998 (in Japanese).
21. Ikuzawa M, Matsushita H, Nakase A. *The Manual for Kyoto Scale Psychological Development 2001*. Kyoto International Social Welfare Exchange Centre, Kyoto, 2002 (in Japanese).
22. Kamio Y, Yukihiro R, Adachi J *et al.* Reliability and Validity of the Pervasive Developmental Disorder (PDD)–Autism Society Japan Rating Scale (PARS): A behavior checklist for adolescent and adults with PDDs. *Clin. Psychiatry (Seishin-Igaku)* 2006; **48**: 495–505 (in Japanese).
23. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*, 4th edn, DSM-IV. American Psychiatric Association, Washington, DC, 1994.
24. Kawakami N, Takeshima T, Ono Y *et al.* Twelve-month prevalence, severity, and treatment of common mental disorders in communities in Japan: Preliminary finding from the World Mental Health Japan Survey 2002–2003. *Psychiatry Clin. Neurosci.* 2005; **59**: 441–452.
25. Konstantareas MM, Homatidis S. Assessing child symptom severity and stress in parents of autistic children. *J. Child Psychol. Psychiatry* 1989; **30**: 459–470.
26. Bristol MM, Gallagher JJ, Schopler E. Mothers and fathers of young developmentally disabled and non-disabled boys: Adaptation and spousal support. *Dev. Psychol.* 1988; **24**: 441–451.
27. Moes D, Koegel RL, Schreibman L, Loos LM. Stress profiles for mothers and fathers of children with autism. *Psychol. Rep.* 1992; **71**: 1272–1274.
28. Hastings RP, Kovshoff H, Ward NJ, degli Espinosa F, Brown T, Remington B. Systems analysis of stress and positive perceptions in mothers and fathers of pre-school children with autism. *J. Autism Dev. Disord.* 2005; **35**: 635–644.
29. Hastings RP. Child behaviour problems and partner mental health as correlates of stress in mothers and fathers of children with autism. *J. Intellect. Disabil. Res.* 2003; **47**: 231–237.
30. Konstantareas MM, Homatidis S. Mothers' and fathers' self-report of involvement with autistic, mentally delayed, and normal children. *J. Marriage Fam.* 1992; **54**: 153–164.
31. Milgram NA, Atzil M. Parenting stress in raising autistic children. *J. Autism Dev. Disord.* 1988; **18**: 415–424.
32. Freeman NL, Perry A, Factor DC. Child behaviours as stressors: Replicating and extending the use of the CARS as a measure of stress: A research note. *J. Child. Psychol. Psychiatry* 1991; **32**: 1025–1030.
33. Gill MJ, Harris SL. Hardiness and social support as predictors of psychological discomfort in mothers of children with autism. *J. Autism. Dev. Disord.* 1991; **21**: 407–416.
34. Dunn ME, Burbine T, Bowers CA, Tantleff-Dunn S. Moderators of stress in parents of children with autism. *Community Ment. Health J.* 2001; **37**: 39–52.



## Japanese cancer patients' communication style preferences when receiving bad news

Maiko Fujimori<sup>1</sup>, Patricia A. Parker<sup>2</sup>, Tatsuo Akechi<sup>1,3</sup>, Yuji Sakano<sup>4</sup>, Walter F. Baile<sup>5</sup> and Yosuke Uchitomi<sup>1\*</sup>

<sup>1</sup> Psycho-Oncology Division, Research Center for Innovative Oncology, National Cancer Center Hospital East, Kashiwanoha 6-5-1, Kashiwa, Chiba 277-8577, Japan

<sup>2</sup> Department of Behavioral Science, The University of Texas M.D. Anderson Cancer Center, 1515 Holcombe Blvd, Houston, TX 77030, USA

<sup>3</sup> Department of Psychiatry, the Nagoya City University Medical School, Kawasumi 1, Mizuho-cho, Mizuho-ku, Nagoya, Aichi 467-8601, Japan

<sup>4</sup> School of Psychological Science, the Health Sciences University of Hokkaido, Ainosato 2-5, Kita-ku, Sapporo, Hokkaido, 002-8072 Japan

<sup>5</sup> Section of Psychiatry, Department of Neuro-oncology, The University of Texas M.D. Anderson Cancer Center, 1515 Holcombe Blvd, Houston, TX 77030, USA

\*Correspondence to: Psycho-Oncology Division, Research Center for Innovative Oncology, National Cancer Center Hospital East, Kashiwanoha 6-5-1, Kashiwa, Chiba 277-8577, Japan.  
E-mail: yuchitomi@eastncc.go.jp

### Abstract

This study describes the communication style preferences of Japanese patients when receiving bad news, examines the factor structure of the measure for patients' preferences (MPP) in a Japanese population, and explores variables that may be associated with patients' communication style preferences. Five hundred twenty-nine cancer outpatients completed several psychosocial measures including the Japanese version of the MPP (MPP-J), the Mental Adjustment to Cancer Scale (MAC), and the Hospital Anxiety and Depression Scale (HADS). The patients desired detailed information and a supportive environment when receiving bad news. The MPP-J demonstrated a 5-factor structure: support, facilitation, medical information, clear explanation, and encouraging question-asking. Regression analyses indicated that a female gender, the fighting spirit and anxious preoccupation dimensions of the MAC were positively associated with all 5 MPP-J factors. In conclusion, Japanese cancer patients' preferences for communication when receiving bad news differ somewhat from those of American patients. Japanese physicians should encourage patients to ask questions and should consider the demographic (e.g. gender), medical (disease status) and psychosocial characteristics (fighting spirit and anxious preoccupation) of patients when delivering bad news.

Copyright © 2006 John Wiley & Sons, Ltd.

Keywords: cancer patients' preferences; bad news; communication; patient–physician relationship

Received: 15 February 2006

Revised: 17 August 2006

Accepted: 18 August 2006

### Introduction

Medical specialists must be able to discuss complex medical information with patients. Oncologists, in particular, must be able to disclose bad news like a life-threatening diagnosis, recurrence, or the discontinuation of anticancer treatments. A physician's communication skills when disclosing bad news about cancer can reportedly influence a patient's anxiety levels and subsequent psychological adjustment [1–4]. Therefore, several guidelines and interventions have been developed to enhance physicians' skills when breaking bad news [5–8]. Communication skills training programs may also improve health care providers' behaviors, beliefs, and confidence in their ability to communicate with patients [9–11].

Most guidelines and interventions for breaking bad news have been developed for patients from western countries, and few studies have examined the preferences of Japanese cancer patients. Japanese patients are traditionally more passive and receptive to a 'paternalistic' approach in their

relationship with their oncologists [12,13]. Typically, the physician and the family decide how much information should be given to the patient [14]. Traditionally, Japanese physicians are less likely to disclose cancer diagnoses to patients than in the West. However, Japanese culture is changing, and patients now want more information and expect to participate more actively in their care [15]. Therefore, it is particularly important to consider and understand a patient's communication style preferences in light of the changing cultural values regarding medical care and the patient–physician relationship.

The importance of communicating with patients on an individual basis has also been recognized, and several associations between medical or psychosocial variables and patients' communication style preferences have been examined. Some evidence suggests that sex, age, education level, and physical status may be associated with a patient's communication style preferences [16–18]. As far as we know, only one validated measure has been developed to examine the communication style

preferences of cancer patients when receiving bad news: the measure of patients' preferences (MPP). In the US, Parker *et al.* [18] conducted a cross-sectional survey by administering the MPP to cancer patients. They found that patients placed the greatest importance on the physician's expertise and the delivery of information regarding their physical condition and treatment options. The patients' communication style preferences could be categorized into three dimensions: the content of the message, the setting, and supportive elements [18]. These results were reproduced in another study examining men with prostate cancer in Canada [19].

The purpose of the present study was to describe the communication style preferences of Japanese cancer patients when receiving bad news, to examine whether the factor structure of the MPP was similar to that found in a US population, and to explore the associations between the dimensions of patients' communication style preferences and demographic, medical and psychological adjustment variables.

## Methods

### Patients

The study participants were outpatients at the National Cancer Center Hospital East in Japan who had attended follow-up medical appointments with their oncologist. The eligibility criteria were as follows: an age of 20 years or older, physician approval regarding the capacity of the patient to complete the survey, the ability to speak and write in Japanese, and the delivery of bad news (i.e. cancer diagnosis, recurrence, treatment failure, or discontinuation of active anticancer treatment) at least 3 months prior to the start of the study. This time frame was chosen to allow patients time to adjust to the idea of their diagnosis and to reflect on their experience. The dates on which the patients were informed of bad news were obtained from the patients' medical charts, and the patients were also asked when they had received bad news.

### Procedure

This study was a cross-sectional consecutive survey. The Institutional Review Board and the Ethics Committee of the National Cancer Center of Japan approved this study. Each participant signed a written informed consent.

All the patients who were being treated by oncologists in our outpatient clinic and who met the eligibility criteria were approached in person and invited to participate in the study. After giving their informed consent, the patients completed a series of questionnaires and were asked to return them by mail. A single attempt was made by

telephone or post to obtain any missing information.

## Measures

### Measure of patients' preferences, Japanese version

The original MPP is a 46-item self-rating scale developed to assess what items are important to American patients with cancer when receiving bad news [18]. All the items are worded as statements, and the responses are recorded using a five-point Likert scale. The response options for each item are as follows: (1) not at all important; (2) optional, can take it or leave it; (3) important; (4) very important; and (5) essential, every doctor should do it. Participants were asked to think about the time when they were first told they had cancer or that their cancer had recurred and to respond to the questions as they would have liked to have been told.

We collaborated with the original author of the MPP (P.P.) to create a Japanese version of the MPP (MPP-J). To do this, we first translated the instructions and items of the MPP into Japanese; the instructions and items were then back translated into English by a native English speaker who was also fluent in Japanese. To examine the readability of the MPP-J, the instrument was piloted among 42 cancer outpatients attending the NCCHE. None of the patients reported problems completing the questionnaire. The results of this pilot study suggested that the Japanese version was easily understood and appropriate for use with Japanese cancer patients.

### Mental adjustment to cancer scale

The Japanese version [20] of the Mental Adjustment to Cancer (MAC) Scale [21] is a 40-item measure developed to evaluate cancer patients' mental adjustment to their cancer after diagnosis. The scale, which has been shown to have adequate validity and reliability, consists of five subscales: fighting spirit, anxious preoccupation, fatalism, helplessness/hopelessness, and avoidance.

### Hospital anxiety and depression scale

The Japanese version of the Hospital Anxiety and Depression Scale (HADS; [22]) assesses anxiety and depressive symptoms and has been shown to be a reliable and valid measure. The scale consists of 14 items [23]. In the present study, the total HADS score was used to assess psychological distress.

### Demographic and medical variables

Demographic information, including age, education level, occupation, and marital status, was collected in the survey. Medical information

regarding the type of cancer, disease stage, presence of recurrence or metastasis, and the time since the patients had received their initial diagnosis was obtained from the patients' medical charts.

### Statistical analysis

To examine the Japanese cancer patients' communication style preferences, we examined the distribution of the patients' responses to the 46 items of the MPP-J. We then performed a factor analysis using the maximum-likelihood method with a varimax rotation to examine the reproducibility of the original three-factor structure of the MPP. Items with a communality of 0.25 or greater and factor loadings of 0.30 or greater have a high degree of correlation and are generally considered reliable. Thus, items meeting these criteria were retained. Because the factor structure of the MPP-J did not reproduce the original factor structure of the MPP, we performed an exploratory factor analysis using the maximum-likelihood method with a varimax rotation. We applied a number of factors based on the results of a scree plot (the 3-factor, 4-factor, 5-factor, 6-factor and 7-factor eigenvalues were 1.85, 1.79, 1.52, 1.11, and 1.08, respectively). We then selected the best rotation based on the goodness of fit index (GFI) and the comparative fit index (CFI) of each factor. The CFI and GFI can range from 0 to 1, and a value of 0.90 or more indicates a good model fit. Cronbach's alpha was calculated to evaluate the internal consistency reliability of each factor. We then performed a series of stepwise multiple regression analyses to examine the unique contribution of demographic (age, sex, marital status, employment status, and education level), medical (cancer site, presence of recurrence or metastasis, current treatment, and the time since the patients had received their initial diagnosis), and psychological status (each MAC subscale and the total HADS score) variables on each subscale of the MPP-J. Three participants were excluded from this statistical analysis because of missing data. The statistical analyses were performed using SPSS 12.0 statistical software. A *P*-value of less than 0.05 was considered significant.

## Results

### Patients

One thousand fifty-six consecutive outpatients who were being treated by 36 oncologists were screened for inclusion. Of the 656 patients who were eligible, 34 were not interested in learning about the study, and 47 could not be contacted. Of the remaining 575 patients who were approached in person, nine refused to participate. Thus, 566 outpatients were given questionnaires; 529 questionnaires (93.5%)

**Table 1.** Demographic characteristics of the subjects (*N* = 529)

		<i>N</i>	%
Age (years; mean, SD range)		62, 11, 26–97	
Sex	Male	274	51.8
Employment status	Employed	190	35.9
Marital status	Married	452	85.4
Household size	Living alone	19	3.6
Education	9 or less years	98	18.5
Cancer site	Digestive	185	34.7
	Breast	125	23.5
	Head and neck	112	21.2
Recurrence or metastasis	Lung	107	20.2
	Yes	299	56.5
	Absence	229	43.3
Treatment received	Surgery	426	80.5
	Chemotherapy	221	41.8
	Radiation therapy	153	28.9
	Hormone therapy	45	8.5
Current anti-cancer treatment	Other	16	3.0
	Yes	134	25.5
Bad news received regarding	Diagnosis	529	100
	Recurrence	164	31
	Disease progression	38	7
	Absence of active anticancer treatment	1	0.2

were returned. Overall, 80.6% (529/656) of the eligible patients participated in the study. The demographic characteristics of the participants are listed in Table 1. The mean time since the patients had received their initial diagnosis was 3.3 years (S.D. = 2.7 years; range, 0.3–11.9 years). Three patients had missing data regarding demographic or medical variables (education [2 patients] and recurrence or metastasis [1 patient]).

### Ratings for important aspects of delivering bad news

The 46 MPP-J items are listed in descending order based on the perceived importance of the items, and the mean score and standard deviation of the original MPP items are shown in Table 2. The mean scores of the MPP-J items were generally about 0.5 points lower than the scores in the US study. The highest rated item was 'Being told in person rather than over the phone (item 1),' and the lowest rated item was 'The doctor holding my hand or touching my arm while telling me the news (item 7).'

We then examined the rank order of each item in the present study and those of the items in the original MPP study to identify the perceived importance to patients in Japan and the US. For example, item 43 ('Having the doctor inform my family members about my diagnosis') was ranked 19th in the present study and 39th in the US study. Meanwhile, item 5 ('Having my doctor give me his/her full attention') was ranked 27th in the present

**Table 2.** The rating scores of the Japanese cancer patients' preferences of communication

No. Item	The present study		Parker et al. [18]		
	Mean	SD	Mean	SD	Ranking
1 Being told in person	4.23	0.81	4.31	1.05	24
16 Telling me best treatment option	4.14	0.77	4.70	0.57	2
40 Telling me he/she will do everything to cure my cancer	4.14	0.81	4.42	0.82	19
24 Feeling confident about doctor's skill	4.03	0.70	4.59	0.55	5
12 Giving news in clear, simple language	4.02	0.77	4.56	0.70	7
15 Being given detailed information about test results	4.00	0.77	4.45	0.74	15
25 Being up to date on research on my type of cancer	3.99	0.82	4.72	0.49	1
22 Being honest about my condition	3.90	0.75	4.61	0.64	4
38 Doctor really listens to me	3.84	0.84	4.44	0.71	16
34 Making me feel comfortable to ask question	3.79	0.84	4.35	0.74	22
10 Telling news directly	3.77	0.73	4.56	0.62	8
23 Telling me how cancer may affect my daily functioning	3.75	0.83	4.39	0.76	20
26 Taking the time to answer my questions completely	3.74	0.89	4.66	0.56	3
4 Setting time aside	3.72	0.84	4.43	0.76	18
28 Being given enough time to ask my questions	3.67	0.89	4.57	0.63	6
39 Offering hope about my condition	3.64	0.87	4.11	0.97	28
27 Stopping to ask if I have questions	3.64	0.91	4.46	0.56	13
43 Informing family about my diagnosis	3.62	0.89	3.38	1.31	39
17 Describing treatment options in detail	3.62	0.97	4.48	0.72	12
21 Giving a lot of information about my cancer	3.60	0.91	4.22	0.89	26
14 Asking how much I want to know about my cancer	3.59	0.87	4.36	0.97	21
13 Asking how much I want to know about my treatment	3.57	0.83	4.43	0.97	17
44 Informing family about my prognosis	3.54	0.90	3.42	1.30	36
46 Giving written summary to take home	3.50	1.01	4.06	1.06	29
33 Doing things to show his/her concern for me	3.47	0.88	3.91	0.94	31
3 Being told in private setting	3.47	0.97	3.86	1.14	32
5 Giving me full attention	3.46	0.89	4.53	0.68	10
9 Waiting until all tests in before giving news	3.45	0.94	4.15	1.03	27
2 Being told by a doctor who knows me well	3.43	0.97	3.33	1.25	41
18 Letting me know all of the treatment options	3.43	1.04	4.55	0.65	9
6 Maintaining eye contact	3.40	0.91	4.05	0.95	30
20 Telling me the prognosis	3.40	0.96	4.46	0.87	14
32 Telling how to contact him/her	3.34	0.95	4.22	0.87	25
41 Making me feel ok to show emotional reactions	3.34	0.99	3.76	1.02	34
19 Telling me about new experimental therapies	3.33	0.99	4.33	0.77	23
37 Encouraging me to talk about feelings	3.29	0.92	3.42	1.14	37
31 Arranging for another meeting to discuss treatment	3.25	0.96	3.82	1.02	33
29 Telling me about support services available	3.22	0.92	3.60	1.03	35
36 Telling me it's ok if I become upset	3.11	0.94	3.40	1.14	38
8 Telling me as soon as possible	3.10	0.92	4.51	0.72	11
35 Comforting me if emotional	2.91	0.92	3.29	1.10	42
30 Telling me about resources in the community	2.80	0.94	3.35	1.04	40
11 Warning me there will be unfavorable news	2.74	1.01	3.01	1.32	44
45 Helping me figure out how to tell family about the cancer	2.73	1.05	2.62	1.25	45
42 Having another healthcare provider present	2.58	0.92	3.14	1.23	43
7 Holding hand/touching arm	1.68	0.76	2.23	1.17	46

study and 10th in the US study, and item 18 ('My doctor letting me know all of the different treatment options') was ranked 30th in the present study and 9th in the US study.

#### Factor structure of the MPP-J

Factor analysis was used to examine whether the three-factor structure found in the original MPP study could be replicated in our sample. We found that the MPP-J did not reproduce the original MPP 3-factor structure.

The results of an exploratory factor analysis yielded the following five factors from among the 42 items of the MPP-J (Table 3).

**Factor 1: Emotional support.** Nine items loaded on this factor, accounting for 14.5% of the total variance. These items covered supportive aspects of communication styles and included offering comfort and support to the patient, and helping to communicate information to family members and friends of the patients.

**Factor 2: Medical information.** Ten items loaded on this factor, accounting for 11.8% of the total variance. These items emphasized medical