The Reliability and Validity of the Japanese Version of the Daily Record of Severity of Problems (J-DRSP) and Development of a Short-Form Version (J-DRSP(SF)) to Assess Symptoms of Premenstrual Syndrome Among Japanese Women

Yumie Ikeda (yumieikeda@gmail.com)
Kyoto University https://orcid.org/0000-0002-4686-3093

Miho Egawa
Kyoto University Hospital: Kyoto Daigaku Igakubu Fuzoku Byoin

Kazuya Okamoto
Kyoto University Hospital: Kyoto Daigaku Igakubu Fuzoku Byoin

Masaki Mandai
Kyoto University: Kyoto Daigaku

Yoshimitsu Takahashi
Kyoto University: Kyoto Daigaku

Takeo Nakayama
Kyoto University: Kyoto Daigaku

Research

Keywords: Premenstrual Syndrome, Japanese, DRSP, Reliability, Validity

DOI: https://doi.org/10.21203/rs.3.rs-81000/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License
Abstract

Purpose

To assess the validity and reliability of the Japanese version of the Daily Record of Severity of Problems (J-DRSP, 24 items) for evaluating symptoms of premenstrual syndrome (PMS), and to develop a short form version of the J-DRSP.

Methods

Using the “DRSP-JAPAN” smartphone app, we collected daily J-DRSP records from 7 days before to 10 days after each participant’s menstruation onset date. Factorial validity (exploratory factor analysis: EFA, confirmatory factor analysis: CFA) and criterion validity were examined. We evaluated test-retest reliability (intraclass correlation: ICC) and obtained Cronbach’s alpha coefficients. The short-form version of the J-DRSP was developed using classical test theory.

Results

In total, 304 women participated and 243 recorded symptoms on at least 4 days spanning the week of the luteal phase (CD-6 to CD0) and 4 days spanning the week of the follicular phase (CD4 to CD10), with CD1 representing the menstruation onset date. With CD0 set as one day before menstruation onset date, the EFA revealed a two-factor structure. KMO was 0.992, and Bartlett’s test of sphericity chi-square was 3653.89 (P < 0.001). However, the model fitness of CFA was found to be suboptimal (CFI: 0.83, RMSEA: 0.12). Total scores for J-DRSP and the sum scores for each subscale were higher on CD10 than on CD0 (p < 0.001), suggesting validity for some criteria. ICC values for the total J-DRSP score from CD0 to CD-1, and between CD9 to CD10, were 0.60 (95% CI: 0.48–0.72) and 0.76 (95% CI: 0.69–0.82), respectively. For participants reporting significant interference in their daily life one day out of CD-1 and CD0 or CD9 and CD10, ICCs were 0.79 (0.68–0.86) among 76 participants and 0.88 (0.76–0.84) among 30 participants. Having eliminated some original items after considering factor loading for each item, we developed an 8-item Short-Form J-DRSP (J-DRSP(SF)) comprising 2 factors (S-Psychological and S-Physical, 4 items for each). CFA showed a better model fit (CFI: 0.99, RMSEA: 0.048) with significantly higher scores for the J-DRSP(SF), S-Psychological, and S-Physical in the luteal phase than in the follicular phase (p < 0.001). Cronbach’s alphas for the J-DRSP(SF), S-Psychological, and S-Physical were 0.89, 0.89, and 0.80, respectively. ICC values for the J-DRSP(SF) in the luteal and follicular phases were 0.61 (95% CI: 0.51–0.68) and 0.70 (95% CI: 0.62–0.77), respectively.

Conclusion
The J-DRSP has moderate to good reliability and certain level of validity. The J-DRSP(SF) has a two-factor structure and can be used effectively among Japanese women to assess their PMS symptoms.

**Background**

Premenstrual syndrome (PMS) is the emotional and physical condition that appears in the luteal phase and resolves after menstruation starts. It is estimated that roughly 30% of women of reproductive age suffer from moderate to severe PMS\(^1\). If identified correctly, PMS can be controlled well and treated appropriately through lifestyle modifications and medication. Unfortunately, very few cases are identified and make their symptom controlled, resulting in a large loss of their productivity and major economic burden in Japan\(^2\).

To accurately diagnose PMS, prospective daily recording for at least two consecutive menstrual cycles is critical, as it allows for the relationship between cycles and symptoms to be determined\(^3\). Of the available assessment tools, the Daily Record of Severity of Problems (DRSP) is the most widely used and is recommended by the Royal College of Obstetricians and Gynaecologists (ROCG) guidelines as a valid scale\(^3\).

The Japanese version of the DRSP (J-DRSP) was developed in 2020 in accordance with the necessary translation and cultural adaptation guidelines published by the Patient-Reported Outcomes (PRO) Consortium\(^4\). This process comprised two independent forward translations, two independent back translations, a consensus meeting about content equivalence and cultural adaptation, reviews by other native Japanese healthcare providers, and cognitive interviews with patients with Premenstrual Dysphoric Disorder (PMDD) at Kyoto University. Internal consistency and concurrent validity were confirmed for 119 Japanese women using the Center for Epidemiologic Studies Depression Scale, PMDD Scale, and the Numerical Rating Scale for general health\(^5\).

The next necessary step in this process is to determine the factorial validity and test-retest reliability of the J-DRSP; in addition, development of a short-form of the J-DSRP would be ideal, as a previous study determined that the median time to complete the J-DRSP was 2 minutes, which may be longer than most women wish to spend on assessing their PMS symptoms.

**Aim**

This study aimed to assess the validity and reliability of J-DRSP, and to develop a short-form version of the J-DRSP.

**Methods**

**Instruments**
The original DRSP contains 21 items pertaining to premenstrual symptoms and 3 items to describe dysfunction in daily life caused by these symptoms, and is based on diagnostic criteria for PMDD from the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV). All items are scored daily throughout the menstrual cycle on a scale of 1 (not at all) to 6 (extreme). The J-DRSP was developed according to the formal procedures outlined in the translation guidelines of the latest Patient-Reported Outcomes (PRO) Consortium.

For the present study, we created the “DRSP-JAPAN” smartphone app, which allows users to log their daily symptoms for the J-DRSP. We released the app at the App Store and on Google Play for free in 2018 (Fig. 1). Following user authentication, participants can rate their condition in the DRSP-JAPAN app and log their symptoms on the server every day. The app offers a visual display of daily condition, the severity of which is indicated by various colors and numbers, enabling participants to look back on the various changes in their symptoms over time.

The app has a blank check function, and a reminder mail is sent automatically if data have not been logged for more than three consecutive days.

**Participants**

Participant recruitment flyers were posted between September 2018 and March 2019 at Kyoto University, Doshisha Women’s College of Liberal Arts, the staff lounge rooms of Kyoto University Hospital, and seven gynecology clinics in Japan where many asymptomatic women come for medical check. Women were eligible for study participation if they were over 20 years of age, had self-reported regular menstruation, and used a smartphone or tablet. Those with any current mental disorder or who were taking oral contraceptives or ovulation inhibitors were excluded. All participants installed the DRSP-JAPAN app.

**Measurements**

All data were collected through the DRSP-JAPAN app. At the beginning of the study, we collected demographic data on participant age and self-reported duration of menstruation. In addition to bleeding volume (one item of the J-DSRP), participants were prompted to log the start date of menstruation. Participant J-DSRP data are logged daily until the participants are notified that the research study has come to an end. Data from each cycle comprise those from half of the 7 days before menstruation starts and half of the days spanning day 4 to day 10 after menstruation begins. Once those data are collected, no further data are collected for that cycle.

**Statistical analysis**

In order to assess factorial validity, we conducted exploratory factor analysis (EFA) using data for the 21 symptom items on the day before menstruation began. To measure the factorability of the correlation matrix, we conducted the Kaiser-Meyer-Olkin (KMO) test and Bartlett’s test of sphericity. Eigenvalues and screen plots were used to determine the number of factors, and promax rotation was used. Next, confirmatory factor analysis (CFA) was performed to assess the adoption. We then calculated the comparative fit index (CFI) and root mean square error of approximation (RMSEA) in order to analyze...
model fitness. CFI values range from 0 to 1, with larger values indicating a better fit. A CFI value of 0.95 or higher is regarded as a good fit. RMSEA ranges from 0 to 1, with smaller values indicating a better model fit; values less than 0.06 indicate an acceptable model fit.

To confirm criterion validity, we compared the J-DRSP total score obtained the day before menstruation began to that obtained 10 days after the menstruation start date using the Wilcoxon signed-rank test. P < 0.05 was considered statistically significant.

To examine test-retest reliability, J-DRSP scores from each of the two days of the follicular phase and the luteal phase were compared. Specifically, with cycle day 1 (CD1) set to represent the menstruation onset date, we compared data from CD9 to CD10 and CD-1 to CD0.

We determined the intraclass correlation coefficient (ICC), percent agreement, and weighted kappa coefficient. Percent agreement and kappa coefficients were calculated after the total J-DRSP was categorized into six levels. As a sensitivity analysis, ICC was determined for any participant whose symptoms interfered with daily life. Those were selected using J-DRSP items 22, 23, and 24, which concern dysfunction in productivity, social activities, and relationships by the symptom items from item 1 to 21, and those who scored 1 (none) for all of items 22, 23, and 24 were excluded. To measure internal consistency, we also calculated Cronbach's alphas for the total J-DRSP score as well as for each subscale.

Developing the Short-Form J-DRSP (J-DRSP(SF))

The short-form J-DRSP (J-DRSP(SF)) was created using classical test theory. First, items for which the factor load was less than 0.5 were excluded after conducting EFA for the original J-DRSP with promax rotation. Next, any items demonstrating sufficient model fit were explored using RMSEA and CFI. After developing the J-DRSP(SF), we conducted CFA and determined Cronbach's alphas, ICCs, and criterion validity for the J-DRSP(SF) total score and each subscale. ICC values were calculated in both the follicular and luteal phases, i.e., CD9 to CD10 and CD-1 to CD0, respectively. Sensitivity analysis was conducted again using J-DRSP items 22, 23, and 24. Stata 15.3 was used for all statistical analyses.

Ethical considerations

The Kyoto University Ethics Committee approved the study protocol (R1593). Written informed consent was obtained from each participant.

Results

Participant demographics and characteristics

In total, 304 women (mean age, 21.0 years; range, 20–50 years) gave their informed consent to participate in the study and successfully downloaded the app.
Of these, 243 logged their symptoms on at least 4 days of the week spanning the luteal phase (CD-6 to CD0) and 4 days of the week spanning the follicular phase (CD4 to CD10), with CD1 set as the menstruation onset date. Data from these participants were analyzed.

**Validity and Reliability of the J-DRSP**

Data from the day before menstruation started (CD-1) were collected from 228 participants who logged symptoms on that day; these were used in the EFA. Eigenvalues and the scree plot suggested a two-factorial component, with a KMO of 0.992 and Bartlett's test of sphericity chi-square of 3653.89 (P < 0.001), implying an adequate EFA with a sufficient sample size. Factor 1 and Factor 2 were labeled as the Psychological factor and Physical factor, respectively. The EFA coefficients for each item are shown in Supplementary Table 2. The CFA revealed a CFI of 0.83 and RMSEA of 0.12.

The total J-DRSP score and sum scores for each subscale were higher on CD10 than on CD0 (p < 0.001 for all). ICCs for the total J-DRSP score for the period of CD 0 to CD -1, and CD9 to CD-10, were 0.60 (95% CI: 0.48–0.72) and 0.76 (95% CI: 0.69–0.82), respectively. When all scores were categorized into 6 levels, percent agreement and weighted kappa values for the luteal phase and follicular phase were 98.9% and 0.76, and 97.9% and 0.57, respectively. After excluding participants who reported no daily life interference on at least one day between CD-1 and CD0 or between CD0 and CD10, ICCs were 0.79 (0.68–0.86) among 76 participants, and 0.88 (0.76–0.84) among 30 participants.

**Developing the J-DRSP(SF) and assessing its validity and reliability**

After items 7, 9, 15, and 17 were excluded due to low factor loading, J-DRSP(SF) was examined to ensure sufficient model fit. The final model was created and CFA was conducted (Fig. 3). Two factors were generated and labeled Short-Form Psychological (S-Psych) and Short-Form Physical (S-Phys). RMSEA was 0.048 and CFI was 0.99.

Cronbach's alphas for the J-DRSP(SF), S-Psych, and S-Phys were 0.89, 0.89, and 0.80, respectively. ICCs for the J-DRSP(SF), S-Psych, and S-Phys in the luteal and follicular phases were 0.61 (95%CI: 0.51–0.68) and 0.70 (95%CI: 0.62–0.77), respectively (Fig. 4). Sensitivity analysis revealed ICCs of 0.82 (0.73–0.88) for the J-DRSP(SF) in the luteal phase of 76 symptomatic women, and 0.84 (0.69–0.92) among 30 women. Scores for the J-DRSP(SF), S-Psych, and S-Phys on CD0 were significantly higher than those obtained on CD10 (p < 0.001).

**Discussion**

In this study, daily J-DRSP records were collected from Japanese women using the DRSP-JAPAN smartphone app. According to the app's function, there were no missing data when the participants recorded at least one item, the rate of DRSP completer was higher than the previous study and data management became easier. Although CFA revealed suboptimal factorial validity of the J-DRSP, the
newly developed J-DRSP(SF) showed an acceptable model fit. Test-retest reliability for the J-DRSP and J-DRSP(SF) were good, especially among those with significant PMS symptoms.

The original DRSP has 3 clinically determined subscales that comprise 6 items for Depressive Symptoms, 4 items for Physical Symptoms, and 2 items for Anger/Irritability. There are 9 items left, and didn’t show the result of CFA. First, we conducted EFA, which indicated a two-factor structure. The CFA for a two-factor structure also revealed a suboptimal model fit, mainly because some items contributed to both factors. The total scores for J-DSRP were higher in the luteal phase, demonstrating some validity of the J-DRSP. Cronbach’s alpha was sufficiently high and test-retest reliability was good, especially among symptomatic women, indicating good reliability of the J-DRSP.

The newly developed J-DRSP(SF) contains 8 items in total (4 psychological and 4 physical items). The CFA of the J-DRSP(SF) showed a better model fit than that of the original J-DRSP, suggesting higher validity as a scale to assess PMS symptoms among Japanese women. ICC and Cronbach’s alpha also indicated sufficient reliability of the J-DRSP(SF).

In clinical settings, J-DRSP(SF) in the luteal phase could be used as a self-awareness checklist for women who may not yet realize that their symptoms are related to PMS. In addition, checking the J-DRSP(SF) could help identify whether their main problems are more physical or psychological in nature, which could help with effective clinical management of PMS.

This study has some limitations. First, none of our participants were patients, and those with and without subjective PMS symptoms were included in the analysis. The original DRSP was developed as a PMDD diagnostic tool, and validated as PMDD scale although it is widely used for PMS. Floor effects can explain the lower ICC in this study relative to results from the original DRSP. Second, the sample size might be insufficient to ensure a stable factorial analysis. Third, diagnostic properties of the J-DRSP(SF) were not analyzed in this study. The original DRSP is used as a diagnostic tool to clarify the relationship between PMS-like symptoms and the menstrual cycle. Future studies should compare J-DRSP(SF) to the DRSP as a diagnostic tool.

**Conclusion**

The J-DRSP exhibited moderate to good reliability and certain level of validity. The J-DRSP(SF), with its two-factorial structure, can be used as an optimized form of the J-DRSP to evaluate PMS symptoms among Japanese women.

**Declarations**

*Ethics approval and consent for study participation and publication*

The Kyoto University Ethics Committee approved the study (R1593). Study aims and procedures were explained to all participants, who were ensured that they were permitted to withdraw from the study at
any time. Written informed consent was obtained from each participant.

Consent for publication

Not applicable

Availability of data and materials

The datasets generated and analyzed during the current study are not publicly available, as we did not have participant consent for the secondary use of their data.

Competing interests

The authors have no conflicts of interest to declare.

Funding

This study received financial support from the Kyoto University Center of Innovation Project (100160700002), which is supported by the Japan Science and Technology Agency.

Authors’ contributions

YI designed the original idea and protocol, collected and analyzed data, and wrote the initial draft of the manuscript. ME and YI helped with the study design and conducting of the study. YT and TN guided the study design and statistical analyses, adding their detailed feedback. MM helped with conducting the study and writing the manuscript. All authors read and approved the final manuscript.

Acknowledgements

The authors thank Dr. Tomoki Aoyama and Dr. Haruka Ohta for recruiting participants at Kyoto University and Doshisha Women's College of Liberal Arts. We are grateful to Dr. Haruto Egawa, Dr. Akeo Kawamura, Dr. Yoko Kawaguchi, Dr. Hiroyuki Funatsu, Dr. Chika Shime, Dr. Ruriko Tsushima, and Dr. Kanako Ito for their help in distributing the recruiting flyers at their clinics. We appreciate the dedicated assistance of Ruriko Tsuchiya and thank all of our study participants.

References

1. Borenstein JE, Dean BB, Yonkers KA, Endicott J. Using the daily record of severity of problems as a screening instrument for premenstrual syndrome. Obstet Gynecol. 2007;109(5):1068–75.
2. Survey on Health Promotion and Working Women. 2018. Health and Global Policy Institute March 22, 2018.
3. Green L, O’Brien N, Panay, Craig M. Management of Premenstrual Syndrome: Green-top Guideline No. 48. BJOG. 2016/12/03 edn 2017; e73-e105.
4. Wild D, Grove A, Martin M, et al. Principles of Good Practice for the Translation and Cultural Adaptation Process for Patient-Reported Outcomes (PRO) Measures: report of the ISPOR Task Force for Translation and Cultural Adaptation. Value Health. 2005;8(2):94–104.

5. Yumie Ikeda M, Egawa K, Hiyoshi T, Ueno K, Ueda, Carl B, abeker. Yoshimitsu Takahashi, Takeo Nakayama, Masaki Mandai. Development of a Japanese Version of the Daily Record of Severity of Problems for Diagnosing Premenstrual Syndrome. *Women's Health Reports*. 2020; 1(1).

6. Eremenco S, Pease S, Mann S, Berry P. Patient-Reported Outcome (PRO) Consortium translation process: consensus development of updated best practices. *J Patient Rep Outcomes*. 2017;2(1):12.

7. Lt Hu, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*. 1999;6(1):1–55.

8. Wu L, He Z, Zhao H, et al. Chinese version of Daily Record of Severity of Problems: reliability and validity. *J Adv Nurs*. 2013;69(2):449–56.

9. Endicott J, Nee J, Harrison W. Daily Record of Severity of Problems (DRSP): reliability and validity. *Arch Womens Ment Health*. 2006;9(1):41–9.