

# Translation, Cross-Cultural Adaptation, And Validation of The Japanese Version of The Patient Education Materials Assessment Tool (PEMAT)

Emi Furukawa (✉ [efurukawa-tho@umin.ac.jp](mailto:efurukawa-tho@umin.ac.jp))

The University of Tokyo Graduate School of Medicine

Tsuyoshi Okuhara

The University of Tokyo

Hiroko Okada

The University of Tokyo

Ritsuko Shirabe

The University of Tokyo Graduate School of Medicine

Rie Yokota

The University of Tokyo Graduate School of Medicine

Reina Iye

The University of Tokyo Graduate School of Medicine

Takahiro Kiuchi

The University of Tokyo

---

## Research Article

**Keywords:** Patient education, education materials, health communication, health literacy, assessment, measurement, readability.

**Posted Date:** September 27th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-909531/v1>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

**Background:** The Patient Education Materials Assessment Tool (PEMAT) systematically evaluates the understandability and actionability of patient education materials. This study aimed to develop a Japanese version of PEMAT and verify its reliability and validity.

**Methods:** After assessing content validation, experts scored healthcare-related leaflets and videos according to PEMAT, to verify inter-rater reliability. In validation testing with laypeople, the high-scoring material group ( $n=800$ ) was presented with materials that received high ratings on PEMAT, and the low-scoring material group ( $n=799$ ) with materials that received low ratings. Both groups responded to understandability and actionability of the materials and perceived self-efficacy for the recommended actions.

**Results:** The Japanese version of PEMAT showed strong inter-rater reliability (PEMAT-P: % agreement=87.3, Gwet's AC1=0.83. PEMAT-A/V: % agreement=85.7%, Gwet's AC1=.80). The high-scoring material group had significantly higher scores for understandability and actionability than the low-scoring material group (PEMAT-P: understandability 6.53 vs. 5.96,  $p<.001$ ; actionability 6.04 vs. 5.49,  $p<.001$ ; PEMAT-A/V: understandability 7.65 vs. 6.76,  $p<.001$ ; actionability 7.40 vs. 6.36,  $p<.001$ ). Perceived self-efficacy increased more in the high-scoring material group than in the low-scoring material group.

**Conclusions:** Our study showed that materials rated highly on PEMAT were also easy for laypeople to understand and action. The Japanese version of PEMAT can be used to assess and improve the usability of patient education materials.

## Background

A variety of patient education materials, including pamphlets, web pages, videos, and smartphone apps, support patients in the medical field, helping them understand their conditions, make decisions, and communicate with their health care providers. However, studies show that patient education materials are often poorly understood by patients, especially those with limited health literacy [1, 2]. Inadequate health literacy is associated with more limited disease control, medical adherence, and patient outcomes [1, 3]. Therefore, patient-friendly materials, regardless of the reader's health literacy, are essential in improving health outcomes for patients.

To address the situation, we focused on the Patient Education Materials Assessment Tool (PEMAT), a reliable and valid instrument developed by the Agency for Healthcare Research and Quality (AHRQ) to evaluate understandability and actionability of patient education materials [4, 5]. Understandability refers to the extent to which consumers of diverse backgrounds can process and explain key messages [4]. Actionability refers to the degree to which consumers of diverse backgrounds and varying levels of health literacy can identify what actions they should take to improve their health, based on the information presented [4]. PEMAT is divided into two parts: PEMAT-P, a scale for printable materials (brochures and PDFs), and PEMAT-A/V, a scale for audiovisual materials (videos and multimedia materials including

smartphone apps). The scores are calculated by taking the sum of the points, dividing by the total possible points, and multiplying by 100 to obtain a percentage. The developers have set the cutoff value for understandability and actionability at 70%.

Studies have used PEMAT to identify issues with patient materials. For example, Yiu et al. evaluated web-based education materials for patients taking non-vitamin K oral anticoagulants. The study revealed a need to include more summaries of information, visual aids, and tangible tools such as checklists [6]. PEMAT is also reported useful to develop or improve patient education materials. Jamil et al. developed an integrated diabetes-periodontitis nutrition and health education module using PEMAT [7].

PEMAT has been translated into Malay [8, 9] and Korean [10], but a Japanese version has not yet been developed. Furthermore, to our knowledge, there are no validated tools to assess whether material is understandable and actionable in Japanese. Therefore, our study aimed to translate and cross-culturally adapt the PEMAT into Japanese and verify its reliability and validity.

## Methods

### Stage 1: Translation of PEMAT into Japanese

We translated the PEMAT questionnaire and user's guide with the permission of AHRQ. The translation and cross-cultural adaptation process were carried out according to the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) Task Force [11] and Guidelines for the Process of Cross-Cultural Adaptation of Self-Report Measures [12]. Two translators whose native language is Japanese independently translated the original PEMAT into Japanese (T1 and T2). A health communication researcher (TO) and a physician (EF) reviewed T1 and T2, integrating them into the Japanese forward-translated version (T-12). Two translators whose native language is English independently back-translated T-12 into English (BT1 and BT2). The back-translators were neither aware of the concept of PEMAT nor involved in the forward translation. After making a comparison and integration of BT1 and BT2, the expert committee, including a health communication researcher (TO) and three medical professionals (physician (EF) and nurses (HU and HO)), created a final Japanese version of PEMAT.

### Stage 2: Assessment of content validity by expert panel

At a panel meeting of twelve experts (medical professionals and health communication researchers), we determined whether the items apply to the Japanese cultural background, and whether the content is appropriate. The questionnaire's original developers examined and provided feedback on the Japanese version.

### Stage 3: Determining the reliability of the instrument

We tested the inter-rater reliability of the PEMAT using patient education materials written in Japanese. The eligibility criteria for the materials were as follows: (1) developed by academic societies, government

offices, or non-profit organizations; (2) including any of the nine topics presented in Health Japan 21 (2<sup>nd</sup> edition) [13] such as nutrition and dietary habits, physical activities and exercise, rest and mental health, smoking, alcohol, dental health, diabetes, cardiovascular disease, and cancer; and (3) materials that could be downloaded for free from the Internet. We searched via the two most popular search engines in Japan, Google Japan [14] and Yahoo! Japan [15]. The search terms in Japanese were 'topic' (where 'topic' was any of the nine topics in Health Japan 21) AND 'pamphlet' OR 'leaflet' OR 'video' OR 'patients' OR 'explanation.' We then selected the first 100 written materials and the first 50 audiovisual materials from the search results.

Four evaluators (EF, RS, RI, and RY) rated the patient education materials according to the Japanese version of PEMAT. EF and RS are physicians; RI is a nutritionist, and RY is a health communication researcher. Two rounds of reliability testing were performed because of the low reliability found in the first round. In each round, the evaluators followed the guidance on the question items and evaluation methods using the Japanese version of the PEMAT User's Guide before evaluating the material. In the first round, two evaluators (EF and RI) independently assessed 50 materials using PEMAT-A/V. EF, RS, and RY each assessed 67 materials with PEMAT-P. Before starting the second round, we revised the explanations and examples in the User's Guide through discussion with the evaluators. Then the evaluators switched the materials to ensure they did not evaluate the same materials as in the first round. After the evaluation, the mean PEMAT score was calculated, and the materials with the highest and lowest scores were selected.

#### Stage 4: Assessment of construct validity by testing with the general public

In this stage, we conducted an online survey to determine whether non-experts found the material with high/low PEMAT scores (from the expert evaluation in stage 3) easy/difficult to understand and take action from. The online survey consisted of two studies to test the validity of the PEMAT-P and PEMAT-A/V, one with a leaflet presentation and the other with a video presentation.

#### *Participants*

Study participants were recruited from registered monitors of an online survey company (Rakuten Insight). Men and women who use Japanese as a native language were eligible to participate in the study. In the PEMAT-P study, participants aged 18 to 69 years were included, and in the PEMAT-A/V part, participants aged 60 to 79 years were included. This is because the age groups targeted by the materials used for intervention in PEMAT-P and A/V were different, as described below. Participants were excluded if they had experience in health care, or were restricted from practicing the action recommended in the materials due to illness or injury. Participants were randomized into two groups using a central computerized random allocation system. One group (intervention group) viewed the material that was highly scored by experts in stage 3, while the other (control group) viewed the low-scoring material.

#### *Materials*

For testing PEMAT-P, participants viewed leaflets that promote healthy eating habits. The PEMAT-P score of the leaflets was 100% for the high-scoring material group and 69.7% for the low-scoring material group. In testing PEMAT-A/V, we used videos on the topic of locomotive syndrome prevention for the elderly. Locomotive syndrome occurs in conditions with a high risk of motor function decline due to locomotive organ impairment [16]. The overall PEMAT-A/V score of each video was 85.4% (intervention group) and 25.0% (control group), respectively.

### *Measures*

The survey company provided participants' gender and age, and participants responded to questions about their educational background, annual family income, occupation, marital status, and self-perceived health. Participants also answered questions about their baseline health literacy and perceived self-efficacy. Health literacy was measured using the 14-item health literacy scale for Japanese adults (HLS-14) [17]. Self-efficacy was measured by the Self-Efficacy Scale for Positive Eating Behavior [18] for PEMAT-P and the Home-Exercise Barrier Self-Efficacy Scale [19] for PEMAT-A/V.

After responding to these questions, participants viewed the relevant materials. They then rated how easy the material was to understand or take action from, on a scale from 1 to 10. They also responded to eight selected items in PEMAT (items 1, 4, 8, 9, 11, 17, 19, and 21) (see **Table 1** and **Table 2**). These items were asked in both the PEMAT-P and PEMAT-A/V studies and were relevant for all the materials presented. At the end of the survey, participants responded about their self-efficacy immediately after the intervention on a scale from 1 to 10. The participants scored the items as 1 if they completely disagreed with the content of the item and 10 if they completely agreed with it.

### *Additional validation with readability scores*

We also used the text readability measurement system, 'jReadability' [20, 21], as the developers of the original PEMAT recommended using readability evaluation tools to evaluate the readability of printed materials along with the PEMAT [22]. To assess readability, we manually retrieved the text from the printable materials and transcribed the audio from the audiovisual materials.

### *Statistical analysis*

In stage 3, the PEMAT scores of the two raters were averaged, and the inter-rater reliability of the PEMAT-A/V tool was calculated. Inter-rater reliability was used to assess the external consistency of the PEMAT, using percentage agreement and Cohen's kappa for two evaluators. We also calculated Gwet's AC1 [23] when low kappas occurred despite a high percentage of agreement [24]. Inter-rater agreement was deemed poor (0), slight (0.01–0.20), fair (0.21–0.40), moderate (0.41–0.60), substantial (0.61–0.80), or almost perfect (0.81–1.0) [25].

In stage 4, the primary analysis was on an intent-to-treat basis. Sample size calculation was performed based on an effect size of 0.2 (Cohen's *d*) [26], a significance level of .05, and a power of 0.8. It was estimated that 394 participants per group were required. Differences between the control and intervention

groups were evaluated using the two-sample t-test for age and the chi-square test or Fisher's exact test for sex, educational background, occupation, annual household income, marital status, and self-perceived health. Welch's t-test was used to compare understandability, actionability, and perceived self-efficacy between the two groups. Pearson's correlation coefficient was used to determine whether there was a correlation between the PEMAT understandability scores and jReadability scores.

All  $p$ -values were two-sided, and  $p < .05$  was considered statistically significant. All analyses were conducted with R version 4.0.3 (2020-10-10).

## Results

### Content validation

At a panel meeting of experts, we deleted item number 5 in the understandability scale. This was because the item refers to use of the active voice, but the distinction between active and passive voice is not clear in Japanese daily conversation and the item was judged to be irrelevant. Consequently, the final Japanese PEMAT consists of 25 items and two scales, including understandability (18 items) and actionability (7 items). Details of the text for each item are presented in **Table 1** and **Table 2**.

### Inter-rater reliability

The median (IQR) of the PEMAT score of the printed materials was 76.7% (62.5–87.5%) for understandability and 70.0% (45.0–91.7%) for actionability. The median (IQR) of the audiovisual materials was 63.6% (45.8–75.0%) for understandability and 66.7% (33.3–83.3%) for actionability.

The Japanese version of PEMAT showed strong inter-rater reliability. For PEMAT-P, agreement was 87.3%,  $\kappa$  was 0.735, and Gwet's AC1 was 0.834. For PEMAT-A/V, agreement was 85.7%,  $\kappa$  was 0.753, and Gwet's AC1 was 0.798.

The kappa range for the understandability items was 0.30–0.84 in PEMAT-P and 0.35–0.84 for PEMAT-A/V. For the actionability items, scores were 0.47–1.00 for PEMAT-P and 0.67–0.81 for PEMAT-A/V. Gwet's AC1 revealed strong agreement for both scales and material types (**Table 1** and **Table 2**).

### Construct validity

#### *Baseline participant characteristics*

For PEMAT-P, out of 1,526 randomized participants, we analyzed 400 in the high-scoring material group and 399 in the low-scoring material group. For PEMAT-A/V, of 1,211 participants randomized, 400 in the high-scoring material group and 400 in the low-scoring material group were analyzed (**Figure 1**). Those who incorrectly answered the screening questions or did not complete the questionnaire were removed from the analysis. Study arm characteristics are described in **Table 3** and **Table 4**. There was no

significant difference in baseline characteristics, including health literacy and self-efficacy scores, between the two groups.

### *Assessment of the material by non-experts*

In both PEMAT-P and PEMAT-AV, the high-scoring material group had significantly higher scores for understandability and actionability than the low-scoring material group (PEMAT-P: overall understandability 6.53 vs. 5.96,  $p < .001$ ; overall actionability 6.04 vs. 5.49,  $p < .001$ ; PEMAT-A/V: overall understandability 7.65 vs. 6.76,  $p < .001$ ; overall actionability 7.40 vs. 6.36,  $p < .001$ ). Similar results were obtained for all the selected PEMAT items (**Table 5**).

### *Self-efficacy*

In PEMAT-AV, perceived self-efficacy increased significantly more in the high-scoring material group than in the low-scoring material group (increase in self-efficacy scores 2.18 vs. 1.46,  $p < .01$ ). In PEMAT-P, the scores increased more in the high-scoring material group than the low-scoring material group; however, the difference did not reach significance (increase in self-efficacy scores 2.22 vs. 1.53,  $p = 0.14$ ). (**Table 6**).

### *Comparison of the PEMAT understandability scores and jReadability*

We calculated an average readability score to examine the correlation with the PEMAT understandability scores. The average readability score was 2.7 (range 1.1–4.4) for PEMAT-P and 2.8 (range 0.8–4.0) for PEMAT-A/V. These scores indicate that the materials were at the ‘upper intermediate’ level, which can be understood by people who comprehend the language of daily life and some technical terms. There was a moderate positive correlation between the understandability scores and the readability score for printable materials (Pearson’s  $r = 0.46$ , 95% CI: 0.27–0.62), and a weak positive correlation for audiovisual materials (Pearson’s  $r = 0.33$ , 95% CI: 0.03–0.57).

## **Discussion**

### **Main findings**

We developed and tested the reliability and validity of the Japanese version of PEMAT, a tool for assessing the understandability and actionability of patient education materials. The inter-rater reliability was moderate when measured by the kappa coefficient but showed a convincingly strong agreement when calculated with Gwet’s AC1. In the development of the original PEMAT, the overall agreement was 69–90%, and Gwet’s AC1 range was 0.56–0.86 (mean 0.74) [4]. In the reliability testing of the Malay version of PEMAT, Wong et al. evaluated 13 leaflets and 13 videos. They found that understandability of PEMAT-P had an agreement of 61.5–91.6%, and Gwet’s AC1 was 0.259–0.965. For actionability, agreement was 69.7–98.3% and Gwet’s AC1 was 0.394–0.980. For PEMAT-A/V, the agreement was 64.1–98.3% and Gwet’s AC1 was 0.394–0.980 for understandability. The agreement was 79.5–91.5%, and the AC1 statistic was 0.397–0.931 for actionability [8]. Our study demonstrated that the reliability of the

Japanese version of PEMAT compares favorably with other language versions. In addition, when testing inter-rater reliability, our study included material on diverse topics: from recommendations for healthy living (having medical checkups, improving dietary and exercise habits) to secondary/tertiary prevention in patients with diabetes and cardiovascular disease. This suggests that the Japanese version of PEMAT can evaluate a wide range of health materials in the real world.

In the validation testing with the non-experts, the high-scoring material group rated higher than the low-scoring material group on the PEMAT-P and PEMAT-AV for all eight selected items. This result suggests that materials that medical professionals rated as easy to understand and act upon using the Japanese version of PEMAT are also easy for the general population to understand and act on. At the time of the development of the original version of PEMAT, there were significant positive correlations between PEMAT-A/V actionability scores and consumer actionability scores [4]. However, there was no clear relationship between understandability as rated by the experts and non-experts' comprehension scores [4]. This may be attributed to the inadequate sample size of  $n = 47$  for consumer testing. In our study, we were able to secure a sufficiently large sample size of nearly 800 participants for each of PEMAT-P and A/V, and overcome the limitations of the original version. In addition, in the Japanese version of PEMAT, the increase in self-efficacy tended to be greater in the high-scoring material group than in the low-scoring material group. As self-efficacy is a predictor of behavior [27], materials that receive high ratings on the PEMAT may encourage individuals' behaviors. Studies have shown that health materials that are easy to understand and act upon may encourage their audiences to adopt healthier behaviors. Arterburn et al. found that understandable decision aids improved the quality of decision-making and reduced uncertainty about the treatment for bariatric surgery in obese patients [28]. Nagle et al. reported pregnant women who viewed a decision aid for prenatal testing of fetal abnormalities were more likely to make an informed decision than those who viewed the less informative material [29].

Nakayama et al. note that 85.4% of Japanese people have inadequate health literacy [30]. However, according to Yamamoto et al., the drug guides for patients written in Japanese are designed to be understandable for patients with at least a high school education level [31]. Thus, it is essential to create and improve materials so they are easy to understand and act on regardless of individuals' health literacy. Evaluating and improving materials using the Japanese version of PEMAT may contribute to supporting behavior change in terms of health literacy.

#### Limitations of this study

There are limitations to our study. First, the validation survey indicated whether the material was easy to understand, but it did not measure whether individuals actually understood the information. This was due to the lack of novelty in the materials on eating behavior and exercise, making it impossible to create a comprehension test that would specifically tap knowledge of the content of the materials. Therefore, it is desirable to have non-healthcare professionals quantitatively evaluate materials on a specific disease to measure comprehension and understandability in PEMAT. Second, we could not assess concurrent validity because we could not find validated tools similar to PEMAT available in Japanese. However, in

the domain of understandability, we observed a moderate positive correlation with jReadability, which has criterion-related validity with the Japanese Language Proficiency Test, supporting the comprehensibility of the Japanese version of PEMAT. Lastly, although we assessed perceived self-efficacy, we could not measure actual behavior. Future research is needed to measure actual outcomes in terms of behavioral change, by following up with participants for some time after the intervention.

## Conclusions

The Japanese version of PEMAT developed in this study is the first reliability-validated tool for assessing the patient-friendliness of patient education materials created in Japanese. This instrument indicated that, as with the original version of PEMAT, the materials that experts rated as easy to understand and act upon using the Japanese version of PEMAT were also easy for laypeople to understand and act on. The Japanese version of PEMAT enables medical professionals to select more understandable and actionable patient education materials. It also allows them to develop and improve patient-friendly materials, ultimately encouraging patients to practice self-management and healthy behaviors.

## Abbreviations

PEMAT: The Patient Education Materials Assessment Tool; AHRQ: the Agency for Healthcare Research and Quality.

## Declarations

### *Ethics approval and consent to participate*

The study was registered as a University Hospital Medical Information Network Clinical Trials Registry (UMIN-CTR) Clinical Trial (Unique trial no. 000042989) on January 13, 2021. The methods of the study adhered to CONSORT guidelines, and the protocol was approved by the Research Ethics Committee of the University of Tokyo Graduate School of Medicine and Faculty of Medicine. All participants provided their informed consent to participate in this study before a screening survey, in accordance with the Declaration of Helsinki.

### *Consent for publication*

Not applicable.

### *Availability of data and materials*

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

### *Competing interests*

The authors declare that there is no competing interests.

### *Funding*

This work was supported by the Japan Society for the Promotion of Science KAKENHI (grant number 19K10615).

### *Authors' contributions*

Emi Furukawa: Conceptualization, Methodology, Data extraction and synthesis, Formal analysis, Investigation, Writing - original draft. Tsuyoshi Okuhara: Writing - review and editing, Funding acquisition. Hiroko Okada: Writing - review and editing. Ritsuko Shirabe: Investigation, Writing - review and editing. Rie Yokota: Investigation, Writing - review and editing. Reina Iye: Investigation, Writing - review and editing. Takahiro Kiuchi: Writing - review and editing, Supervision.

### *Acknowledgments*

The authors are grateful to Ms Cindy Brach from AHRQ for her collaboration and advice during the back-translation of the Japanese version of PEMAT and the validation study. We thank Michelle Pascoe, PhD, from Edanz (<https://jp.edanz.com/ac>) for editing a draft of this manuscript.

## References

1. Berkman ND, Sheridan SL, Donahue KE, Halpern DJ, Crotty K: **Low health literacy and health outcomes: an updated systematic review.** *Ann Intern Med* 2011, **155**(2):97–107.
2. Keller DL, Wright J, Pace HA: **Impact of health literacy on health outcomes in ambulatory care patients: a systematic review.** *Ann Pharmacother* 2008, **42**(9):1272–1281.
3. Dewalt DA, Berkman ND, Sheridan S, Lohr KN, Pignone MP: **Literacy and health outcomes: a systematic review of the literature.** *J Gen Intern Med* 2004, **19**(12):1228–1239.
4. Shoemaker SJ, Wolf MS, Brach C: **Development of the Patient Education Materials Assessment Tool (PEMAT): a new measure of understandability and actionability for print and audiovisual patient information.** *Patient Educ Couns* 2014, **96**(3):395–403.
5. Vishnevetsky J, Walters CB, Tan KS: **Interrater reliability of the Patient Education Materials Assessment Tool (PEMAT).** *Patient Educ Couns* 2018, **101**(3):490–496.
6. Yiu A, Ng KK, Lee VW, Bajorek BV: **Evaluating the Understandability and Actionability of Web-Based Education Materials for Patients Taking Non-vitamin K Oral Anticoagulants.** *Therapeutic Innovation & Regulatory Science* 2019.
7. Jamil NA, Chau SH, Abdul Razak NI, Shamsul K, II, Mohd-Said S, Rani H, Sameeha MJ: **Development and evaluation of an integrated diabetes-periodontitis nutrition and health education module.** *BMC Med Educ* 2021, **21**(1):278.

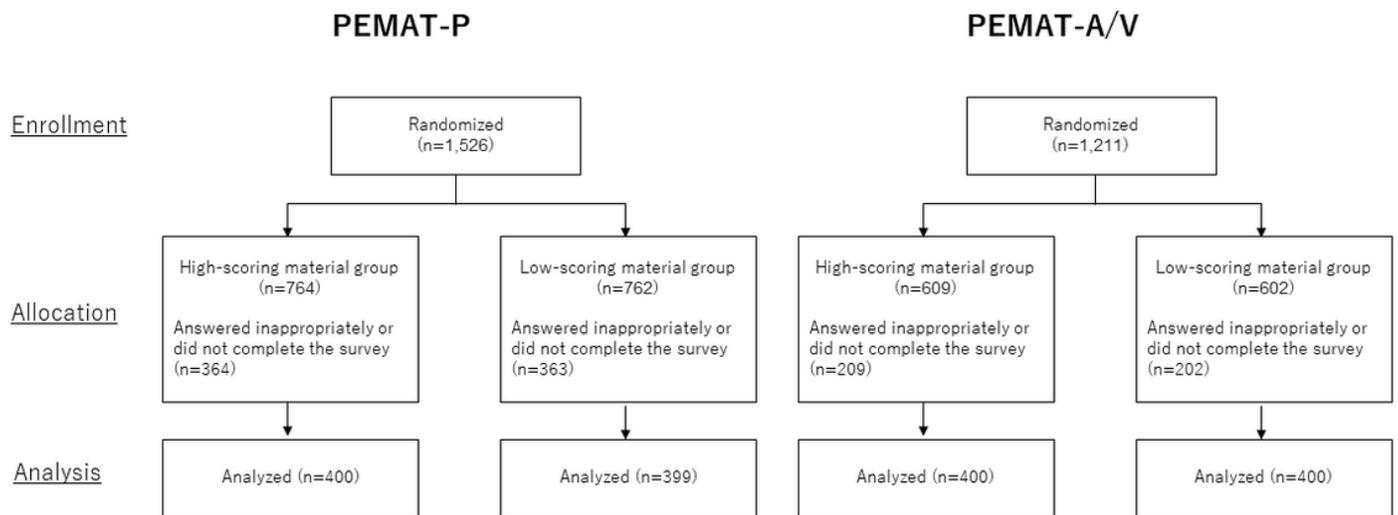
8. Wong ST, Saddki N, Arifin WN: **Inter-Rater Reliability of the Bahasa Malaysia Version of Patient Education Materials Assessment Tool.** *The Medical journal of Malaysia* 2019, **74**:100.
9. Wong ST, Saddki N, Arifin WN: **Validity of the Bahasa Malaysia Version of Patient Education Materials Assessment Tool.** *Malaysian Journal of Public Health Medicine* 2019, **19**:35.
10. Lee H, Kim J, Yoo R, Lee J-y: **Development and Evaluation of Cardiovascular Disease Prevention Education Materials for Middle-aged Korean-Chinese Female Workers: Applying Patient Education Materials Assessment Tool for Printable Materials (PEMAT-P).** *Journal of Korean Academy of Community Health Nursing* 2016, **27**(3):284.
11. Wild D GA, Martin M, Eremenco S, McElroy S, Verjee-Lorenz A, Erikson P: **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**:11.
12. Beaton DE BC, Guillemin F, Ferraz MB.: **Guidelines for the Process of Cross-Cultural Adaptation of Self-Report Measures.** *Spine (Phila Pa 1976)* 2000, **25**(24):6.
13. Ministerial Notification **No. 430 of the Ministry of Health, Labour and Welfare**  
[<https://www.mhlw.go.jp/file/06-Seisakujouhou-10900000-Kenkoukyoku/0000047330.pdf>]
14. **Google Japan** [<https://www.google.com/?hl=ja>]
15. **Yahoo! Japan** [<https://www.yahoo.co.jp/>]
16. Nakamura K, Ogata T: **Locomotive Syndrome: Definition and Management.** *Clin Rev Bone Miner Metab* 2016, **14**:56–67.
17. Suka M, Odajima T, Kasai M, Igarashi A, Ishikawa H, Kusama M, Nakayama T, Sumitani M, Sugimori H: **The 14-item health literacy scale for Japanese adults (HLS-14).** *Environ Health Prev Med* 2013, **18**(5):407–415.
18. Moriya K: **An Examination of the Usefulness of Self Efficacy and Social-Support Scales Providing Support for Positive Health Behavior Change.** *Bulletin of Tenshi College* 2009, **9**:1–20.
19. Arita N, Takenaka K: **Development of a Home-Exercise Barrier Self – Efficacy Scale for Elderly People Requiring Support and Care.** *The Journal of Japanese Physical Therapy Association* 2014, **41**(6):338–346.
20. Hasebe Y, Lee J-h: **Introducing a readability evaluation system for Japanese language education.** In: *The 6th International Conference on Computer Assisted Systems for Teaching and Learning Japanese [CASTEL/J]* University of Hawaii; 2015.
21. **jReadability PORTAL** [<http://jreadability.net/>]
22. Davis TC, Wolf MS, Bass PF, 3rd, Middlebrooks M, Kennen E, Baker DW, Bennett CL, Durazo-Arvizu R, Bocchini A, Savory S *et al.*: **Low literacy impairs comprehension of prescription drug warning labels.** *J Gen Intern Med* 2006, **21**(8):847–851.
23. Gwet KL: **Computing inter-rater reliability and its variance in the presence of high agreement.** *Br J Math Stat Psychol* 2008, **61**(Pt 1):29–48.

24. Feinstein AR, Cicchetti DV: **High agreement but low kappa: I. The problems of two paradoxes.** *J Clin Epidemiol* 1990, **43**(6):543–549.
25. Landis JR, Koch GG: **The measurement of observer agreement for categorical data.** *Biometrics* 1977, **33**(1):159–174.
26. Cohen J: **A power primer.** *Psychol Bull* 1992, **112**(1):155–159.
27. Bandura A: **Self-efficacy: Toward a unifying theory of behavioral change.** *Psychological Review* 1977, **84**(2):191–215.
28. Arterburn DE, Westbrook EO, Bogart TA, Sepucha KR, Bock SN, Weppner WG: **Randomized Trial of a Video-Based Patient Decision Aid for Bariatric Surgery.** *Obesity* 2011, **19**(8):1669–1675.
29. Nagle C, Gunn J, Bell R, Lewis S, Meiser B, Metcalfe S, Ukoumunne OC, Halliday J: **Use of a decision aid for prenatal testing of fetal abnormalities to improve women's informed decision making: a cluster randomised controlled trial [ISRCTN22532458].** *Bjog* 2008, **115**(3):339–347.
30. Nakayama K, Osaka W, Togari T, Ishikawa H, Yonekura Y, Sekido A, Matsumoto M: **Comprehensive health literacy in Japan is lower than in Europe: a validated Japanese-language assessment of health literacy.** *BMC Public Health* 2015, **15**:505.
31. Yamamoto M, Doi H, Yamamoto K, Watanabe K, Sato T, Suka M, Nakayama T, Sugimori H: **Adaptation of the European Commission-recommended user testing method to patient medication information leaflets in Japan.** *Drug Healthc Patient Saf* 2017, **9**:39–63.

## Tables

Due to technical limitations, tables 1 to 6 are only available as a download in the Supplemental Files section.

## Figures



**Figure 1**

Flow diagram of the participants

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [20210915BMCTable1furukawa.xlsx](#)
- [20210915BMCTable2furukawa.xlsx](#)
- [20210915BMCTable3furukawa.xlsx](#)
- [20210915BMCTable4furukawa.xlsx](#)
- [20210915BMCTable5furukawa.xlsx](#)
- [20210915BMCTable6furukawa.xlsx](#)