A two-item version of the Japanese Consultation and Relational Empathy measure

Family Practice, 2022, **39**, 1169–1175
https://doi.org/10.1093/fampra/cmac034
Advance access publication 26 April 2022
Research Methods

A 2-item version of the Japanese Consultation and Relational Empathy measure: a pilot study using secondary analysis of a cross-sectional survey in primary care

Noriyuki Takahashi[1,2,*], Takaharu Matsuhisa[2], Kunihiko Takahashi[3], Muneyoshi Aomatsu[4], Stewart W. Mercer[5], Nobutaro Ban[6,7]

[1]Department of Education for Community-Oriented Medicine, Nagoya University Graduate School of Medicine, Nagoya, Japan
[2]Department of General Medicine/Family & Community Medicine, Nagoya University Hospital, Nagoya, Japan
[3]Department of Biostatistics, M&D Data Science Center, Tokyo Medical and Dental University, Tokyo, Japan
[4]Department of Medical Education, Saku Central Hospital, Saku, Japan
[5]Centre for Population Health Sciences, Usher Institute, University of Edinburgh, Edinburgh, United Kingdom
[6]Medical Education Center, Aichi Medical University School of Medicine, Nagakute, Japan
[7]Nakatsugawa Community Healthcare Center, Nakatsugawa, Japan

*Corresponding author: Department of Education for Community-Oriented Medicine, Nagoya University Graduate School of Medicine, 65 Tsurumai-cho, Showa-ku, Nagoya 466-8550, Japan. Email: nori1007@med.nagoya-u.ac.jp

Background: The Consultation and Relational Empathy (CARE) measure is a patient-reported measure of physician empathy which is widely used internationally. The Japanese version of the CARE measure has very high internal reliability, suggesting that a shorter version may have adequate validity and reliability.

Objective: To investigate a valid shorter version of the Japanese CARE measure.

Methods: We conducted a pilot study using secondary analysis of previous data obtained from 9 general practitioners and 252 patients and used to develop the Japanese CARE measure. All 1,023 possible combinations of the Japanese CARE items (n = 1–10) were candidates for the short measure. The internal consistency (Cronbach's alpha) and the correlations between candidate short questionnaires and the original questionnaire were calculated. After selecting the most valid short questionnaire, inter-rater reliability was determined using generalizability theory, and construct validity (Spearman's rho) was determined using patient satisfaction.

Results: Two items were selected for a pilot shorter version: item 6 "Showing care and compassion" and item 9 "Helping you to take control." These showed high internal consistency and correlations with the 10-item measure (Cronbach's alpha = 0.920, correlation = 0.979). Forty-five questionnaires per doctor allowed us to reliably differentiate between practitioners. The construct validity for the pilot short measure was high (Spearman's rho 0.706, $P < 0.001$).

Conclusion: We generated a pilot 2-item version of the Japanese CARE measure. This pilot 2-item version provides a basis for future validation studies of short CARE measures in other languages.

Key words: empathy, general practice, Japan, medical education, patients, surveys and questionnaires

Introduction

Physician empathy is the foundation of patient–physician communication and affects both patient satisfaction and treatment effectiveness.[1–3] There are several definitions of empathy, including "the ability to think about and understand another person's feelings and thoughts as if you were that person" and "an emotional response (affective), dependent upon the interaction between trait capacities and state influences."[4,5] Patient involvement is an important aspect of empathy education, because patients can help to develop empathic attitudes among health professionals.[6] General practitioners (GPs) tend to underestimate their own empathy compared with patient-perceived physician empathy,[7] suggesting that patient support is needed to help physicians recognize their own empathic responses.

The Consultation and Relational Empathy (CARE) measure is a widely used measure of patient-perceived physician empathy.[2] The original English version was developed by Mercer et al. in 2004.[8] The CARE measure was designed for use in research and clinical practice as a patient-rated process measure of physician empathy.[8,9] The CARE measure has been translated into 9 other languages: Chinese, Japanese, Croatian, Dutch, Swedish, Portuguese, Kannada, Spanish, and Italian.[10–18] The Japanese version of the CARE measure was developed by Aomatsu et al. in 2014 for outpatient GP visits and has shown preliminary validity and internal consistency.[11] Matsuhisa et al. examined the inter-rater reliability of the Japanese CARE measure, and found that it reliably discriminates between doctors.[19] The Japanese CARE measure has been used in several studies. One such study,

Key messages

- The 10-item CARE measure assesses patient-perceived physician empathy.
- A short measure may be useful in daily clinical work in ageing society.
- We generated a pilot 2-item version based on the 10-item Japanese CARE measure.
- This pilot 2-item version showed good discrimination between doctors.
- This pilot version provides a basis for future validation of short versions.

a quasi-randomized controlled trial, showed that patient-perceived empathy was not affected by physician attire.[20]

Although a 5-item version of the CARE measure was developed for children,[21] there is no validated shorter version for adults. A shorter version of the CARE measure would be useful in many countries for 2 reasons. First, because of global ageing, many potential evaluators are older patients.[22] Older people may find it easier to complete a shorter questionnaire.[23] Second, more accessible measures of physician empathy are needed for use in daily clinical settings.[24] A potential barrier for patient participation in healthcare provider education is that such involvement can be time-consuming.[25] Furthermore, a shorter version of the CARE measure is particularly needed in Japan. This is because the Japanese CARE measure assesses a single construct,[11] and shows the highest internal consistency of all translations of the measure (Cronbach's alpha = 0.984).[8,10–18] This indicates item homogeneity; however, very high alpha values can indicate item redundancy and the potential to shorten the measure.[26] Thus, this study aims to investigate a valid shorter version of the 10-item Japanese CARE measure.

Method

Research design

We conducted a secondary analysis of a cross-sectional survey in 2011.[11]

Setting

Data on the Japanese CARE measure and patient background were collected from 317 patients as potential participants who had consulted 1 of 9 GPs in the general medicine department of the same university hospital between July and December 2011. Of these, 252 patients with no missing or unevaluable data were included in the present analysis. The GPs were all men, 3 were senior residents, 2 were medical staff, and 4 were faculty members.[11,19] Patient characteristics, such as age, sex, and educational background, and trends in consultation time and consultation satisfaction, have been described in detail in previous studies.[11,19]

Exploratory selection of items for short measure

The CARE measure, including the Japanese version, is a 10-item questionnaire with 5 response options ranging from poor to excellent (scores 1–5) and a "not applicable" option[8,11] (Supplementary Fig. S1). We selected valid questions for the short version. We considered it unnecessary to change any of the existing questions, and the clinical weightings of all 10 questionnaire items were considered equal. Therefore, all combinations (1,023 possible combinations) of n items (n = 1–10) selected from the 10 items were considered. These combinations were used as candidate short questionnaires. A scatterplot was created to visually examine all candidate items. In the scatterplot, the correlations between the candidate questionnaire scores and total scores on the 10-item Japanese CARE measure were plotted on the vertical axis and values of Cronbach's alpha plotted on the horizontal axis. These axes indicators were selected for the following reasons.

On the vertical axis, correlations between scores on each candidate questionnaire and the total score on the 10-item Japanese CARE measure were examined to assess criterion validity.[26] We considered candidates that showed a strong correlation with the 10-item measure (even after item reduction) to be valid candidates for the short version. Pearson's correlation coefficient was used and a 2-sided test conducted (significance level: 1%). On the horizontal axis, Cronbach's alpha was calculated for the candidate questionnaires to assess internal consistency. Internal consistency indicates the extent to which questionnaire items measure the same construct (homogeneity) and is generally assessed using Cronbach's alpha.[26] In medical research, alpha values ≥0.7 indicate that a questionnaire is reliable, and values ≥0.9 are desirable; however, values that are very high suggest redundancy.[26] We used this value as an indicator of item redundancy. In this study, 1,013 possible combinations of 2–10 items were examined (1 item was excluded because its internal consistency could not be measured).

Based on the scatterplot, favourable combinations were examined through researcher discussions. The following factors were considered: (i) higher correlations between a candidate questionnaire and total scores on the 10-item Japanese CARE measure were preferable, (ii) Cronbach's alpha should be ≥0.90 when selecting 2 or more items, (iii) a smaller number of questions in the short version was preferable, to differentiate it from the 10-item version, and (iv) the question content should be clinically valid. The authors NT, TM, and MA were mainly responsible for selection of candidate items for the short version. The item selection was confirmed by SWM and NB.

Reliability study: inter-rater reliability

We used generalizability theory (G theory) to determine the number of reliable questionnaires needed to differentiate short CARE measure scores for each GP for the questions selected for the short measure. G theory was developed from classical test theory, which uses all the data to simultaneously measure the sources of error variance, and to find strategies to minimize the effects of error variance.[26] The G coefficients (i.e. effect sizes) obtained in the generalizability study were intracluster correlation coefficients (ICCs). Using GP as a group cluster, ICCs were obtained as follows, in accordance with previous studies.[19,26,27]

$$\mathrm{ICC} = \frac{\sigma^2_{\mathrm{GP}}}{\sigma^2_{\mathrm{GP}} + \sigma^2_{\mathrm{P}}}$$

σ^2_{GP} is the variance of the total Japanese CARE measure score of the target GPs and others, and σ^2_{P} is the variance of the

random error generated by patients. If the sample size is n, then

$$\sigma_P^2 = \frac{\sigma^2}{n}$$

σ^2 is the variance of the Japanese CARE measure score for each patient. σ_{GP}^2 and σ^2 were obtained using analysis of variance. In this study, GPs were measured and patients were nested in GPs; that is, each patient was involved with only 1 GP. A Decision study was then used to predict changes in ICCs.[26] Using the smallest number of questionnaires that satisfied ICC = 0.8, we estimated the mean and 95% confidence interval of the predicted scores for each GP (interval estimation).[19,27] From the results, we obtained scores to identify the top 2 scorers and the bottom 2 scorers, which we used as the reference range. The criteria for setting the reference range followed a previous study of the 10-item Japanese CARE measure.[19]

Validity study: construct validity

Because physician empathy affects patient satisfaction, previous studies have assessed construct validity by measuring the relationship between physician empathy and patient satisfaction.[11,17,28] Therefore, we analysed patient satisfaction in this study. To examine construct validity, Spearman's rho was used in a 2-sided test with a 1% level of significance.

Data analysis

The statistical analysis was conducted using SPSS version 26 (IBM Corporation, Armonk, NY, USA) to estimate Cronbach's alpha and correlation coefficients. Scatterplots were generated using R and RStudio[29] using the package *ggplot2*.[30] G theory analysis was performed using G string IV.[31]

Results

Figure 1 shows the scatterplots of Cronbach's alpha values and the correlations between candidate questionnaires and the 10-item total scores for combinations of 10Cn (n = 2–10) questions (1,013 items; 10C1 10 items are not shown in the scatterplot). The internal consistency and correlations with the 10 items remained high even when the number of questions was reduced. Most Cronbach's alpha values were >0.90, except for a few cases in which 2 items were selected. The correlations with the total Japanese CARE score were >0.90 in all cases, indicating a strong correlation. To ensure internal consistency (Cronbach's alpha), we chose 10C2 as the smallest group that exceeded 0.90, which is desirable for questionnaire use. Among the 10C2s, questions 6 and 9 were selected because their internal consistency was >0.90 and they showed the highest correlation with the 10-item Japanese CARE measure (questions 6 and 9, Cronbach's alpha = 0.920, correlation coefficient = 0.979). Question 6 is "Showing care and compassion" and question 9 is "Helping you to take control." The total score was 10 points; the descriptive statistics for each group, including 10C1, are shown in Supplementary Table S1.

Table 1 shows the means and SDs of the total scores on the 10-item measure for each GP and scores on questions 6 and 9. The mean values for the short measure were slightly higher than for the 10-item measure, but (as in the 10-item version) there was no ceiling effect. Table 2 shows the score distributions for questions 6 and 9. The G string IV analysis produced

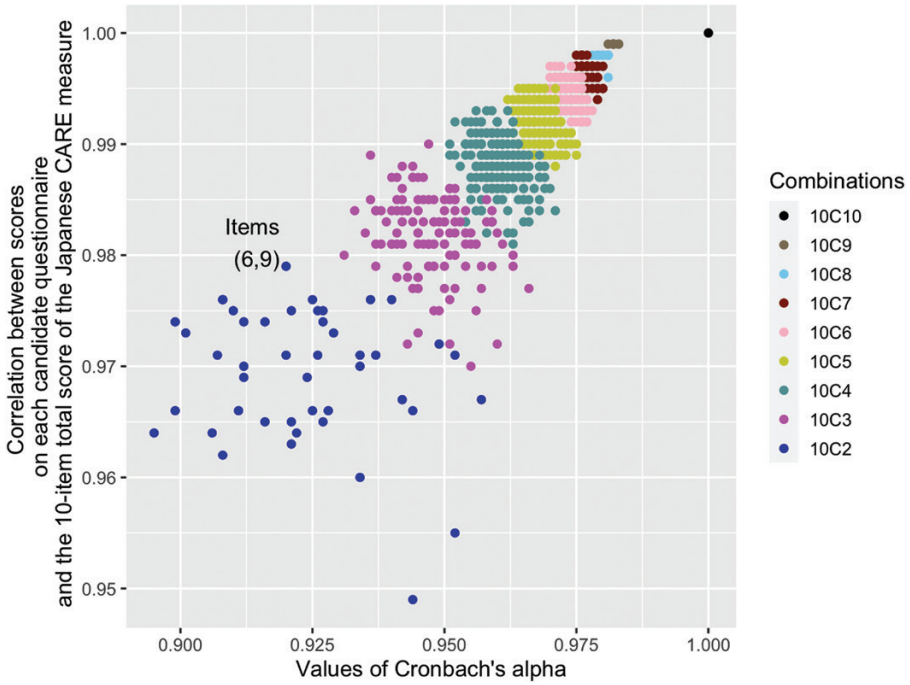

Fig. 1. Scatterplot showing Cronbach's alpha for each candidate questionnaire and correlations between candidate questionnaire scores and 10-item measure total scores. The scatterplot is colour coded and grouped by 10Cn (n = 2–10) combinations selected from the 10 items. The scatterplots were generated using data from the 10-item Japanese version of the Consultation and Relational Empathy (Japanese CARE) measure completed by 252 patients (2011). The horizontal axis shows Cronbach's alpha values for each candidate questionnaire. The vertical axis shows the correlations between scores on each candidate questionnaire and 10-item total scores.

a σ^2_{GP} of 0.277, σ^2 of 3.064, and G coefficient (i.e. effect size) of 0.648. The Decision study showed that the ICC exceeded 0.8 for n = 45 (Table 3). In other words, 45 questionnaires were needed to reliably differentiate the empathy scores for each GP. In the interval estimation of the mean score, the standard range was between 7 and 8 points. The mean score was 7.68 and the SD was 1.82 (Fig. 2).

The correlation between physician empathy and patient satisfaction was high, indicating good construct validity (Spearman's rho 0.706, $P < 0.001$).

Discussion

The aim of this pilot study was to investigate a valid shorter version of the Japanese CARE measure using secondary analysis. We selected 2 of 10 questions using the scatterplot data, which comprised Cronbach's alpha values and correlations between the short and 10-item versions. In the 10C2 group, questions 6 and 9 were selected for the short version because they showed the highest correlations. Cronbach's alpha for questions 6 and 9 was 0.920, lower than that for the original Japanese version (0.984), but equal to the original English version (0.92).[8,11] The construct validity indicated that the pilot short Japanese CARE measure is a valid measure of physician empathy.[11] Although it is essential to validate the 2-item version using a new sample, these findings suggest the feasibility of a short version of the CARE measure.

Question 6 focuses on care and compassion towards patients, whereas question 9 focuses on patient autonomy. Mercer et al. describe in detail the exploratory process for the content and wording of the 10 items during the development of the original CARE measure.[8] Of the 10 questions, questions 1–6 assess an emotive component, and the remaining questions assess cognitive/behavioural components. Thus, the 2 questions we selected represent both components. The addition of another item may allow us to create a short measure that is highly correlated with the 10-item measure. In particular, questions 3, 6, and 9 had higher correlations (0.990) than questions 6 and 9 (0.979). However, the alpha value increased from 0.920 to 0.947, suggesting that questions 3, 6, and 9 would have greater redundancy than questions 6 and 9. This is not surprising because questions 3 and 6 both assess emotive components.

The 2-item measure required 45 patient questionnaires to ensure inter-rater reliability, which is more than the 38 required for the 10-item Japanese CARE measure.[19] This may be because there was less interphysician variability in the scores on the 2 items compared with scores on the 10-item measure. In the Chinese version of the CARE measure, inter-rater reliability was assessed with fewer than 30 respondents, although the comparison was between family physicians and organ specialists.[32] However, the original English CARE measure required approximately 50 raters,[27] so the number of raters required is within the acceptable range for a short questionnaire. For questions 6 and 9, the standard range can be set to between 7 and 8 points, indicating a range of excellent/standard/inferior. This produces a practical screening tool that is particularly useful to identify learners who need a more detailed assessment of empathy expression.

Study strengths

This investigation suggested that a shorter CARE measure may be useful for rapid, easy assessment of physician empathy. The original 10-item questionnaire takes approximately 5–10 min to complete.[12] Reducing the questionnaire to one-fifth of its original length is likely to reduce completion time and thus reduce barriers to patient involvement. The use of fewer questions may reduce the burden of item comprehension, especially for older patients. However, this measure may be insufficient if detailed patient feedback on physician empathy is required. In such cases, other

Table 1. Means and SDs of the 10-item Japanese measure total scores for each GP and scores on items 6 and 9 (2011).

	Sample size	Mean Japanese CARE measure score (SD)[a]	Mean item (6, 9) score (SD)[b]
GP1	14	34.8 (6.9)	6.7 (1.5)
GP2	32	35.3 (8.6)	6.9 (1.9)
GP3	15	36.7 (8.0)	7.2 (1.7)
GP4	24	37.0 (9.4)	7.5 (1.8)
GP5	18	37.7 (6.9)	7.5 (1.5)
GP6	47	38.4 (9.1)	7.6 (2.1)
GP7	43	39.2 (8.4)	7.8 (1.8)
GP8	50	42.9 (7.1)	8.5 (1.5)
GP9	9	45.2 (3.8)	9.0 (0.9)
Total	252	38.8 (8.5)	7.7 (1.8)

Japanese CARE measure, the Japanese version of the Consultation and Relational Empathy measure.
[a]The total score is on a 50-point scale.
[b]Scores on questions 6 and 9 are on a 10-point scale.

Table 2. Score distributions for the 2-item version for 252 patients with no "not applicable" or missing values (2011).

Two-item version of the Japanese CARE measure	Poor (%)	Fair (%)	Good (%)	Very good (%)	Excellent (%)	Mean score
6. Showing care and compassion (*seeming genuinely concerned, connecting with you on a human level; not being indifferent or "detached"*)	2 (0.8)	18 (7.1)	62 (24.6)	99 (39.3)	71 (28.2)	3.87 ± 0.93
9. Helping you to take control (*exploring with you what you can do to improve your health yourself; encouraging rather than "lecturing" you*)	3 (1.2)	20 (7.9)	65 (25.8)	98 (38.9)	66 (26.2)	3.81 ± 0.96

Japanese CARE measure, the Japanese version of the Consultation and Relational Empathy measure.

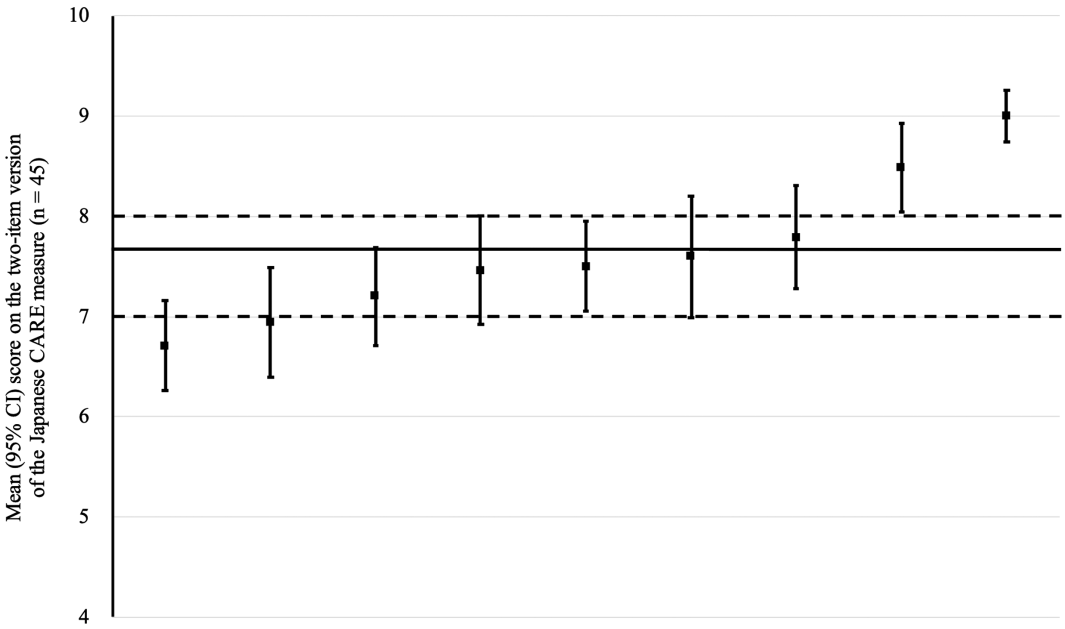

Fig. 2. Interval estimates (means and 95% confidence intervals [CIs], *n* = 45) for each GP's score on the short version of the questionnaire (questions 6 and 9) (2011). The vertical axis shows each GP's scores on the Japanese version of the Consultation and Relational Empathy (Japanese CARE) measure for questions 6 and 9. The total score is 10. The horizontal axis represents each GP. Plots are interval estimates of the mean at *n* = 45 and whiskers indicate 95% CIs. *N* = 45 was calculated using the generalizability theory Decision study. The solid line shows the mean score (7.68). The broken lines show the standard range (7–8 points).

Table 3. Generalizability theory Decision study results for the 2-item version (items 6 and 9; 2011).

Number of patients per GP	ICC[a]
1	0.083
10	0.475
20	0.644
30	0.731
40	0.783
44	0.799
45	0.803
50	0.819

[a]The analysis examined the number of patients (*n*) above ICC 0.8 (45).

additional educational strategies, such as use of the 10-item measure or video reviews, should be used. Therefore, the 2 items should be used as a screening tool to determine whether additional assessment is necessary. The 10-item questionnaire is robust and has been used in many studies since its publication.[2] The 10-item questionnaire can collect accurate data on patient-perceived physician empathy, such as the effect of empathy on treatment effectiveness. The 2-item and original versions of the Japanese CARE measure therefore have clearly different roles.

The pilot 2-item version provides a basis for future validation studies of a short CARE measure in Japanese and other languages. Care must be taken in applying these findings to short CARE measures in other languages, because Cronbach's alpha values for the translated versions of the CARE measure suggest that item homogeneity varies across languages.[8,11,12,15,18] However, the present findings suggest the utility of developing short versions of the CARE measure in other languages.

Study limitations

There are various ways of developing short versions of existing scales.[33–35] A limitation of secondary analysis is the potential discrepancy between the study objectives and the data.[35,36] Responses to a questionnaire item are affected by other items[26] (i.e. responses to 2 of 10 items would differ from responses to the same 2 items presented on their own). Further research is required using randomized studies to compare the results of independent measurement using the 2-item CARE measure with the gold standard of the 10-item version.[26,37] Item response theory is often used in clinical studies, but it is best used in cases where it is difficult to pool all questions, such as in educational examinations.[26,35] Thus, in this study, we selected items using traditional correlations and internal consistency analysis to address the issue of item redundancy.

We analysed data for 252 complete 10-item Japanese CARE questionnaires. Therefore, the number of "not applicable" or missing values if responses to only 2 items had been collected is unclear; this information is important in questionnaire development.[26] The assessment of the original English CARE measure found less than 3.2% not applicable responses for question 6, but 10.8% for question 9, although the number of missing values was not disclosed.[27] For the Japanese CARE measure, the number of not applicable responses was also very low (0.0% for question 6 and 0.6% for question 9), but the missing values tended to be high (8.2% for question 6 and 8.8% for question 9).[11] Conversely, for the Spanish version, there were 0.8% and 2.2% not applicable responses for questions 6 and 9, respectively, and no missing values.[17] Although the estimated number of not applicable responses and missing

values was not high, further research is needed to confirm this.

The present data retain previously identified selection bias.[11,19] Patients were selected consistently rather than randomly, and were recruited by participating physicians. Patients with specific diseases (anxiety disorders and dementia) were excluded. Furthermore, the sample of only 9 GPs at a single centre was relatively small compared with previous studies.[10–18,27] Multicentre settings with more GPs and outpatients may provide more valid findings.

Despite these limitations, we believe that the study has substantial value. The findings demonstrated the potential of a short version of the CARE measure. The purpose of the measure is to evaluate physician empathy from the patient's perspective in clinical settings, and to use the findings in physician education.[8] The 2-item version fits the original purpose of the CARE measure and could provide a more practical way of assessing physician empathy.

Conclusion

This research suggests that physician empathy can be assessed using just 2 items from the 10-item Japanese CARE measure. This pilot 2-item version provides a basis for future validation studies of a short version of the CARE measure for other languages.

Acknowledgement

We thank Diane Williams, PhD, from Edanz (https://jp.edanz.com/ac) for editing a draft of this manuscript.

Supplementary material

Supplementary material is available at *Family Practice* online.

Funding

This work was supported by a Japan Society for the Promotion of Science KAKENHI grant (JP20K10375). The funder had no role in the design of the study, data collection, management, analysis, interpretation of the data, or writing of the manuscript.

Ethical approval

The secondary data analysis was conducted with the approval of the institutional review board to which the first author belongs (approval number: 2021-0199). Patients and physicians who participated in the previous study provided their informed consent; however, no contact information was collected for these participants. Therefore, at the request of the ethical review committee, we informed the patients of the present study by information disclosure on the university website. We ensured that physicians had the opportunity to refuse participation by providing information about the study on the university website.

Conflict of interest

NT reports that his current affiliated institution was established by donations from Aichi Prefecture and Nagoya City, Japan, and that he has received grants and personal fees outside the submitted work from Novartis Pharma K.K. The remaining authors declare no conflicts of interest.

Data availability

The data underlying this article cannot be shared publicly owing to the confidentiality of the data of individuals who participated in the previous study. However, the data will be shared on reasonable request to the corresponding author.

References

1. Kim SS, Kaplowitz S, Johnston MV. The effects of physician empathy on patient satisfaction and compliance. *Eval Health Prof*. 2004;27(3):237–251.
2. Howick J, Steinkopf L, Ulyte A, Roberts N, Meissner K. How empathic is your healthcare practitioner? A systematic review and meta-analysis of patient surveys. *BMC Med Educ*. 2017;17(1):136.
3. Dambha-Miller H, Feldman AL, Kinmonth AL, Griffin SJ. Association between primary care practitioner empathy and risk of cardiovascular events and all-cause mortality among patients with type 2 diabetes: a population-based prospective cohort study. *Ann Fam Med*. 2019;17(4):311–318.
4. Rogers CR. A theory of therapy, personality, and interpersonal relationships, as developed in the client-centered framework. In: Koch S, editor. *Psychology: a study of a science. Vol. 3. Formulations of the person and the social context*. New York (NY): McGraw Hill; 1959. p. 184–256.
5. Cuff BM, Brown SJ, Taylor L, Howat DJ. Empathy: a review of the concept. *Emot Rev*. 2016;8(2):144–153.
6. Gordon M, Gupta S, Thornton D, Reid M, Mallen E, Melling A. Patient/service user involvement in medical education: a best evidence medical education (BEME) systematic review: BEME Guide No. 58. *Med Teach*. 2020;42(1):4–16.
7. Hermans L, Olde Hartman TC, Dielissen PW. Differences between GP perception of delivered empathy and patient-perceived empathy: a cross-sectional study in primary care. *Br J Gen Pract*. 2018;68(674):e621–e626.
8. Mercer SW, Maxwell M, Heaney D, Watt G. The consultation and relational empathy (CARE) measure: development and preliminary validation and reliability of an empathy-based consultation process measure. *Fam Pract*. 2004;21(6):699–705.
9. Mercer SW, Reynolds WJ. Empathy and quality of care. *Br J Gen Pract*. 2002;52(suppl):S9–S12.
10. Fung CS, Hua A, Tam L, Mercer SW. Reliability and validity of the Chinese version of the CARE measure in a primary care setting in Hong Kong. *Fam Pract*. 2009;26(5):398–406.
11. Aomatsu M, Abe H, Abe K, Yasui H, Suzuki T, Sato J, Ban N, Mercer SW. Validity and reliability of the Japanese version of the CARE measure in a general medicine outpatient setting. *Fam Pract*. 2014;31(1):118–126.
12. Hanzevacki M, Jakovina T, Bajic Z, Tomac A, Mercer S. Reliability and validity of the Croatian version of Consultation and Relational Empathy (CARE) Measure in primary care setting. *Croat Med J*. 2015;56(1):50–56.
13. van Dijk I, Scholten Meilink Lenferink N, Lucassen PL, Mercer SW, van Weel C, Olde Hartman TC, Speckens AE. Reliability and validity of the Dutch version of the Consultation and Relational Empathy Measure in primary care. *Fam Pract*. 2017;34(1):119–124.
14. Crosta Ahlforn K, Bojner Horwitz E, Osika W. A Swedish version of the Consultation and Relational Empathy (CARE) measure. *Scand J Prim Health Care*. 2017;35(3):286–292.
15. Manzoni ACT, Fagundes FRC, Fuhro FF, Cabral CMN. Translation, cross-cultural adaptation to Brazilian Portuguese, and analysis of measurement properties of the Consultation and Relational Empathy Measure. *J Chiropr Med*. 2019;18(2):106–114.
16. Rajput S, Kumar A, Puranik MP, Shanbhag N. Exploring the educational opportunity and implementation of CARE among dental

students in India. *J Educ Health Promot*. 2020;9:249. doi:10.4103/jehp.jehp_22_20
17. García del Barrio L, Rodríguez-Díez C, Martín-Lanas R, Costa P, Costa MJ, Díez N. Reliability and validity of the Spanish (Spain) version of the consultation and relational empathy measure in primary care. *Fam Pract*. 2020;38(3):353–359.
18. Natali F, Corradini L, Sconza C, Taylor P, Furlan R, Mercer SW, Gatti R. Development of the Italian version of the Consultation and Relational Empathy (CARE) measure: translation, internal reliability, and construct validity in patients undergoing rehabilitation after total hip and knee arthroplasty. *Disabil Rehabil*. 2022. Online ahead of print. doi:10.1080/09638288.2022.2037742
19. Matsuhisa T, Takahashi N, Aomatsu M, Takahashi K, Nishino J, Ban N, Mercer SW. How many patients are required to provide a high level of reliability in the Japanese version of the CARE Measure? A secondary analysis. *BMC Fam Pract*. 2018;19(1):138.
20. Matsuhisa T, Takahashi N, Takahashi K, Yoshikawa Y, Aomatsu M, Sato J, Mercer SW, Ban N. Effect of physician attire on patient perceptions of empathy in Japan: a quasi-randomized controlled trial in primary care. *BMC Fam Pract*. 2021;22(1):59.
21. Arigliani M, Castriotta L, Pusiol A, Titolo A, Petoello E, Brun Peressut A, Miorin E, Elkina I, Marzona F, Cucchiaro D, et al. Measuring empathy in pediatrics: validation of the Visual CARE measure. *BMC Pediatr*. 2018;18:57.
22. Kanasi E, Ayilavarapu S, Jones J. The aging population: demographics and the biology of aging. *Periodontol 2000*. 2016;72(1):13–18.
23. Broadbent E, Petrie KJ, Main J, Weinman J. The brief illness perception questionnaire. *J Psychosom Res*. 2006;60(6):631–637.
24. Buckman R, Tulsky JA, Rodin G. Empathic responses in clinical practice: intuition or tuition? *CMAJ*. 2011;183(5):569–571.
25. Bombard Y, Baker GR, Orlando E, Fancott C, Bhatia P, Casalino S, Onate K, Denis JL, Pomey MP. Engaging patients to improve quality of care: a systematic review. *Implement Sci*. 2018;13(1):98.
26. Streiner DL, Norman GR, Cairney J. *Health measurement scales: a practical guide to their development and use*. 5th ed. New York (NY): Oxford University Press; 2015.
27. Mercer SW, McConnachie A, Maxwell M, Heaney D, Watt GC. Relevance and practical use of the Consultation and Relational Empathy (CARE) Measure in general practice. *Fam Pract*. 2005;22(3):328–334.
28. Zachariae R, Pedersen CG, Jensen AB, Ehrnrooth E, Rossen PB, von der Maase H. Association of perceived physician communication style with patient satisfaction, distress, cancer-related self-efficacy, and perceived control over the disease. *Br J Cancer*. 2003;88(5):658–665.
29. R_Core_Team. *R: a language and environment for statistical computing*. Vienna (Austria): R Foundation for Statistical Computing; 2021.
30. Wickham H, Chang W, Henry L, Pedersen TL, Takahashi K, Wilke C, Woo K, Yutani H, Dunnington D. ggplot2: create elegant data visualisations using the grammar of graphics [software]. 2021 [accessed 2021 Nov 21]. https://cran.r-project.org/web/packages/ggplot2/index.html
31. MacMaster University. MacMaster Education Research, Innovation and Theory (MERIT) program. Generalizability theory tool [software]. 2021 [accessed 2021 Nov 21]. https://healthsci.mcmaster.ca/merit/research/g_string_vi
32. Mercer SW, Fung CS, Chan FW, Wong FY, Wong SY, Murphy D. The Chinese-version of the CARE measure reliably differentiates between doctors in primary care: a cross-sectional study in Hong Kong. *BMC Fam Pract*. 2011;12:43.
33. Osse BH, Vernooij-Dassen MJ, Schadé E, Grol RP. A practical instrument to explore patients' needs in palliative care: the Problems and Needs in Palliative Care questionnaire—short version. *Palliat Med*. 2007;21(5):391–399.
34. Osman A, Freedenthal S, Gutierrez PM, Wong JL, Emmerich A, Lozano G. The Anxiety Depression Distress Inventory-27 (ADDI-27): a short version of the Mood and Anxiety Symptom Questionnaire-90. *J Clin Psychol*. 2011;67(6):591–608.
35. Cliffe T, Beinart H, Cooper M. Development and validation of a short version of the supervisory relationship questionnaire. *Clin Psychol Psychother*. 2016;23(1):77–86.
36. Windle PE. Secondary data analysis: is it useful and valid? *J Perianesth Nurs*. 2010;25(5):322–324.
37. Smith GT, McCarthy DM, Anderson KG. On the sins of short-form development. *Psychol Assess*. 2000;12(1):102–111.