



Original Article

Initial development of a patient-reported outcome measure of disability due to Katakori via evaluating patient comprehensibility and comprehensiveness

HIROSHI TAKASAKI, PhD^{1)*}, YUSUKE HANDA, BSc²⁾

¹⁾ Department of Physical Therapy, Saitama Prefectural University: 820 Sannomiya, Koshigaya, Saitama 343-8540, Japan

²⁾ Graduate School of Rehabilitation Science, Saitama Prefectural University, Japan

Abstract. [Purpose] “Katakori” refers to a nonspecific symptom, including discomfort or dull pain, that is experienced around the occiput and that extends through the cervical spine to the acromion and scapular area. This study aimed to develop a patient-reported outcome measure of disability due to Katakori, namely the Katakori Disability Index, via evaluating patient comprehensibility and comprehensiveness. [Participants and Methods] We conducted a semi-structured interview among participants who had experienced Katakori consistently during the past month to examine patient comprehensibility and comprehensiveness; we particularly used the thinking-aloud method and cognitive debriefing to evaluate comprehensibility. [Results] We initially tested a provisional version of the Katakori Disability Index with 24 items using two 11-point numeric rating scales in a subset of 10 participants. Considering the issues identified concerning comprehensibility and comprehensiveness, we created a second draft of the Katakori Disability Index with two 6-point Likert scales, modified items, and four additional items. The second draft was tested in another subset of 10 participants. We eventually developed a 31-item Katakori Disability Index with modified instructions and items, two additional items, and a post-survey checklist; all these features addressed the concerns identified and suggestions obtained in the second round of interviews. [Conclusion] We developed a 31-item Katakori Disability Index with content validity.

Key words: Content validity, Semi-structured interview, Stiff neck

(This article was submitted Aug. 14, 2021, and was accepted Oct. 5, 2021)

INTRODUCTION

Katakori is a unique Japanese term that refers to a nonspecific symptom, including discomfort or dull pain, around the occiput through the cervical spine to the acromion and scapular area¹⁾. It is one of the most frequent symptoms occurring in the Japanese population (11.4% of females and 5.7% of males)²⁾.

In a systematic review, Aoki et al.³⁾ noted the lack of investigation on the content validity of patient-reported outcome measures (PROMs) of disability due to Katakori. Subsequently, Natsume et al.⁴⁾ identified the conceptual structures of disturbed activities of daily living due to Katakori via semi-structured interviews with 50 participants who have chronic Katakori using the patient elicitation technique (PET). The identified activities were not covered by the Neck Disability Index⁵⁾, which is the most common PROM of neck disability, and a specific PROM of Katakori developed in Japan⁶⁾. These studies have justified the development of a validated PROM that evaluates the magnitude of disability due to Katakori.

*Corresponding author. Hiroshi Takasaki (E-mail: physical.therapy.takasaki@gmail.com)
(Supplementary materials: refer to PMC <https://www.ncbi.nlm.nih.gov/pmc/journals/2193/>)

©2022 The Society of Physical Therapy Science. Published by IPEC Inc.



This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial No Derivatives (by-nc-nd) License. (CC-BY-NC-ND 4.0: <https://creativecommons.org/licenses/by-nc-nd/4.0/>)

Content validity, which is considered to be the most important psychometric property of the PROMs⁷⁾, should be investigated when developing a new PROM. The processes of determining the content validity include item selection and evaluations of patient's comprehensibility and comprehensiveness⁷⁾. Item selection was based on the results of the results of the above-described PET-based study⁴⁾, and a provisional version of the Katakori Disability Index (KDI) was developed.

In the provisional version, we included semi-individualized measures reflecting individual differences between patients, wherein patients rated the importance of an item and the magnitude of its bothersomeness. Such a semi-individualized structure has been used in recent studies^{8, 9)} and known to have a more patient-centric viewpoint compared with traditional and structured PROMs³⁾. However, whether patients find such a semi-individualized structure to be understandable remains unknown. Therefore, this study aimed to develop the KDI by evaluating patient's comprehensibility and comprehensiveness.

PARTICIPANTS AND METHODS

The COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) guideline^{10, 11)} was followed for assessing patient's comprehensibility and comprehensiveness of the provisional KDI. We conducted an online survey to identify and understand our target demographics. In addition, we conducted a semi-structured interview with the study participants via Zoom (Zoom Video Communications, Inc., San Jose, CA, USA) a few days after the online survey to examine the patient's comprehensibility and comprehensiveness, wherein the think-out-loud method and cognitive debriefing were particularly used to assess comprehensibility. This study was approved by an institutional research ethics committee (Saitama Prefectural University, Japan, No. 20025), and the participants provided their written informed consent before start of the data collection.

In December 2020, recruitment was undertaken online through the corresponding author's (HT) website, which was introduced in multiple social networking services. The eligibility criteria included presence of consistent Katakori during the past month, age between 18 and 65 years, knowledge of Japanese as the native language, and willingness and ability to undertake an online interview. The exclusion criteria included a diagnosis of cognitive, neurological, or respiratory disease as well as hospital confinement or administration to a care facility for older adults.

The semi-structured interviews were conducted by one of authors (HT), who had experience with interviewing, using an interview guide. The interview and think-out-loud responses to the provisional KDI were recorded and transcribed by another author (YH). The interview guide contained questions about 1) the participant's understanding of the instructions, recall period, intended meaning and relevance of the items, and response options, 2) general questions about the overall instrument, and 3) missing concepts.

The authors independently reviewed the transcriptions and extracted descriptions that were relevant to the analysis of the patient's comprehensibility and comprehensiveness. A consensus meeting was held, and issues involving patient's comprehensibility and comprehensiveness were identified.

The provisional KDI (Supplementary materials 1 and 2) included 24 items that were determined in a previous study⁴⁾. The same structure of semi-individualized measures as that used in the Satisfaction Recovery Index⁸⁾ was used in the provisional KDI, wherein the participants rated the importance (0=not important at all to me; 10=extremely important to me) and bothersomeness (0=not bothersome at all; 10=extremely bothersome) of each item on an 11-point scale. One item identifying whether responses were reliable or not was included.

The following data were collected via the online survey: 1) demographics (age, gender, and body mass index), 2) duration of symptoms since the last symptom-free month¹²⁾, 3) symptom location¹³⁾, 4) pain intensity as measured using the 4-item Pain Intensity Measure (P4)¹⁴⁾, and 5) quality of life as measured using the EuroQol 5 Dimensions questionnaire (EQ-5D). The P4 is an 11-point numerical rating scale (NRS) for pain intensity, with a higher sum score indicating greater pain intensity (0–40). The EQ-5D includes five quality of life-related items¹⁵⁾, with a higher score indicating better quality of life (score range, 0–1)¹⁶⁾.

The sample size estimation was based on the COSMIN rating of “very good” for a study design with qualitative analysis conducted by interviewing 10 participants. Thus, a subset of 10 participants was initially recruited (round 1). When a substantial modification of the provisional KDI became necessary, we recruited another subset of 10 participants (round 2).

Descriptive analyses were employed to understand the characteristics of the participants. The data were expressed as mean and standard deviation values.

RESULTS

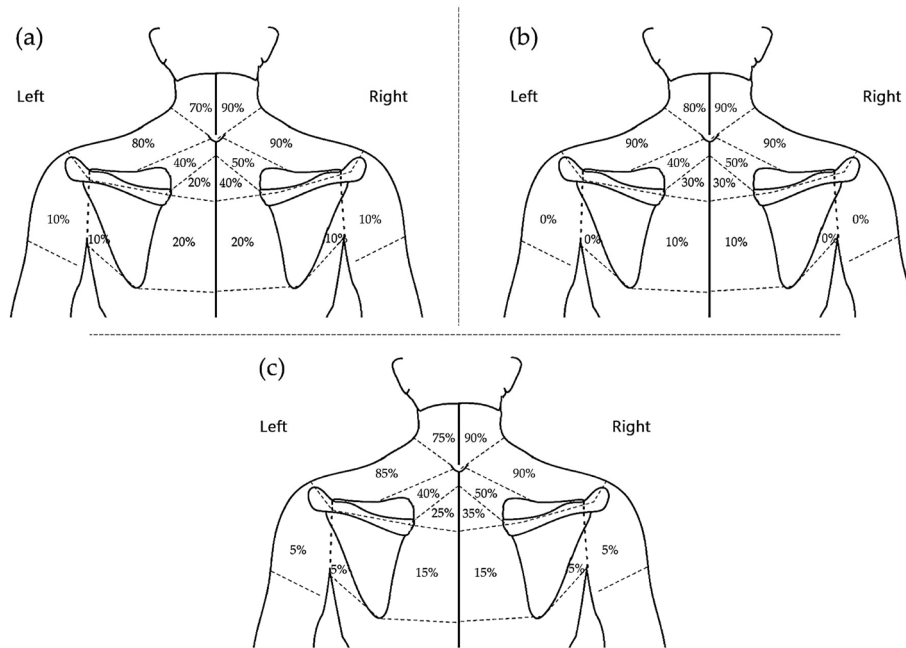
Each author extracted the same concerns about comprehensibility and comprehensiveness from each subset of 10 participants (Table 1). The participants' symptom locations are presented in Fig. 1. Missing data were not detected.

With regard to comprehensibility, participant A (a 32 year-old female) did not have a child and had not been “taking care of children” during the past week but was initially unsure about rating the item's importance to her with a score of 0 or with another score; she ultimately used her common sense in rating the item. Participants B (a 35 year-old female) and C (a 37 year-old female) also wondered about rating items that did not apply to them; both participants eventually gave those items a score of 0. Participant D (a 25 year-old male) had difficulty interpreting “lying on the side” and “holding a strap”,

Table 1. Demographics of the participants

Variables	1st subset of 10 participants	2nd subset of 10 participants
Age (years)	32.1 ± 11.2	26.3 ± 6.6
Gender (males : females)	4: 6	2: 8
Body mass index (kg/m ²)	21.5 ± 3.3	19.4 ± 1.2
Duration of Katakori symptoms (n of those with ≥1 and <3 months: n of those with ≥3 and <6 months: n of those with ≥6 months)	2: 2: 6	1: 1: 8
4-item pain intensity measure (0–40)	14.1 ± 8.6	14.3 ± 7.4
EuroQol 5 Dimension questionnaire (0–1)	0.77 ± 0.10	0.71 ± 0.07

Values are presented with mean ± standard deviation, unless specified.

**Fig. 1.** Location of Katakori symptoms.

(a) The 1st subset of the 10 participants; (b) The 2nd subset of the 10 participants; and (c) All 20 participants.

but cognitive debriefing indicated that he understood the items as “lying down and resting” and “holding on to railings or straps (e.g., in trains and buses)”, respectively, which were correct. Participant E (a 56 year-old female) wondered if “doing calligraphy and handicraft” was included in “doing hobbies with non-intense physical activities (e.g., playing an instrument or singing)”.

With regard to the recall period, participants A and D interpreted it as 1 month. Participant A commented that the highlighted instructions for the recall period were included in the highlighted instructions for scoring.

With regard to the response options, participant D expressed difficulty in selecting a score out of 10 for importance and bothersomeness as well as in visualizing extremely bothersome conditions. Participant F (an 18 year-old female) also described selecting a score out of 10 for importance as difficult. Participant D suggested the use of examples of scoring to make the KDI easier for everyone—for example, a score of 2 would indicate difficulty of typing while using a computer. Participant G (a 23 year-old male) suggested adding “cannot do” as an option under extremely bothersome conditions. Participant C suggested presenting the KDI in two pages and adding “have not done” as an option with a corresponding score of 0 under the items being rated for importance.

With regard to the instructions, participant B commented that it would be important to scan all items first to evaluate the relativity of the score for an item to the scores for all the other items and to consider the modification of the internal scale if necessary. Participant H (a 33 year-old male) suggested that the internal scale be changed as the answer progressed. Participant E noted the difficulty of interpreting the instructions and suggested the use of simpler instructions and layouts.

With regard to comprehensiveness, participant I (an 18 year-old female) cited “disturbed adequate sleep due to Katakori” as a missing concept. Participant C identified “turning the neck to check my surroundings due to Katakori” as a missing

concept. Participant E reported “being frustrated due to Katakori” (e.g., “doing each thing carefully” and “accepting the words and actions of others with a margin in my heart”) as a missing concept.

These feedback and findings led to the development of the second draft of the KDI (Supplementary materials 3 and 4), which has simpler instructions, comes in two pages, uses two 6-point Likert scales, and includes modified and additional items, namely, “turning the neck to check my surroundings”, “doing each thing carefully”, “accepting the words and actions of others with a margin in my heart”, and “sleeping soundly until morning without being disturbed”. All modifications reflected the comments obtained from the interviews. As these modifications were substantial, another subset of 10 participants (Table 1) was recruited and interviewed.

With regard to comprehensibility, participant J (a 21 year-old female) did not have a child and had not been “taking care of children” during the past week but selected a score that reflected her taking care of her little brother. Participant K (a 21 year-old female) had difficulty in choosing a bothersomeness scale for “carrying something light (e.g., handbag and shopping bag)” and “carrying something heavy” as she interpreted bothersomeness to be dependent on the length of the tasks. Participant K was confused as to whether importance referred to the degree of the cause of Katakori or the importance of the activity itself as the response progressed. Participants L (a 31 year-old male) and M (a 26 year-old female) had difficulty in interpreting “lying down and resting.” Cognitive debriefing for participant L indicated that he understood the item to mean “lying to sleep”, which differed from its actual context of taking a rest from lying. Cognitive debriefing for participant M indicated that she misinterpreted “lying down and resting”; she was not sure of the difference between “falling asleep comfortably” and “sleeping soundly without being disturbed until morning”. Participant M was also confused about whether practicing yoga and muscle strength training could be classified as “doing hobbies with intense physical activities (e.g., sports)”. Furthermore, participant M was not sure of the difference between “concentrating on something” and “doing each thing carefully”; she was also unsure how “doing each thing carefully” and “accepting the words and actions of others with a margin in your heart” related to disability due to Katakori. Similarly, participant N (a 25 year-old female) was not sure as to how “doing each thing carefully” related to disability due to Katakori. Participant M suggested the addition of explanations to connect these items to frustration and irritation due to Katakori, so that everyone can better understand their context. Participant O (a 34 year-old female) expressed difficulty in distinguishing between “not bothersome at all” and “not bothersome” as well as between “not bothersome” and “not very bothersome”; she wondered about rating the bothersomeness of items that she did not perform during the past week. However, cognitive debriefing indicated that she correctly understood the rating scale. Furthermore, the impact of the bothersomeness scale was considered negligible for items that were not performed during the past week as the product of two values was used. Participant P (a 22 year-old female) noted that “doing hobbies with non-intense physical activities (e.g., playing an instrument, singing, calligraphy, and handicraft)” had a wide scope and that there would thus be a tendency for intermediate options to be selected, although cognitive debriefing indicated that she correctly understood the rating scale. Participants M and N wondered whether “doing hobbies with non-intense physical activities (e.g., playing an instrument, singing, calligraphy, and handicraft)” included cooking and reading a book.

With regard to the recall period, Participant J overlooked the instructions and took it to mean 6 months. These were attributed to her excessive mental workload. She suggested that colors or a note at the left or right margin of the questionnaire would be useful.

With regard to the instructions, Participant L noted that the DONO TEIDO (in Japanese) notation in the instructions was unnatural. However, cognitive debriefing indicated that his understanding of the instructions was correct. Participant J wondered what ATARI (in Japanese) meant and suggested replacing it with KENTO (in Japanese) in the instructions.

With regard to comprehensiveness, two participants reported missing concepts. Participant L cited difficulty in looking up due to Katakori. Conversely, participant M suggested the inclusion of the appearance of stiff shoulders (as seen when wearing a dress).

Based on these feedback and findings, we developed the final draft of the KDI (Supplementary materials 5 and 6) by modifying expressions in the instructions and items as well as by adding “looking up” and “wearing what you like without worrying about stiff shoulders” as items. In particular, considering the comments of participant J, we added a checklist (Supplementary materials 5 and 6) below the questionnaire to allow respondents to correct their responses even if they had missed the instructions. The final draft of the KDI satisfied all comments and suggestions obtained from the second round of interviews.

DISCUSSION

Patient’s comprehensibility and comprehensiveness are features establishing the content validity of an assessment instrument. To the best of our knowledge, this is the first study to have investigated the patient’s comprehensibility and comprehensiveness of a new PROM of disability due to Katakori using a robust method. Our 31-item KDI is a patient-centric PROM with a semi-individualized structure, and the results of this study suggest that its semi-individualized structure was well received by the respondents.

Our analysis of the patient’s comprehensibility and comprehensiveness revealed unexpected concerns, which we immediately addressed to substantially improve the instrument. The results of this study highlight the value of evaluating a patient’s comprehensibility and comprehensiveness during its development phase. In particular, overlooking and forgetting

instructions were unforeseen issues. As such human errors cannot always be avoided, we added a post-survey checklist in the KDI to allow respondents to correct their scores with proper understanding of the instructions. After the semi-structured interviews with the first subset of 10 participants, the response scale was changed from an 11-point numeric rating scale to a 6-point Likert scale. The modification was based on a suggestion from Simms¹⁷⁾ that a 6-point Likert scale is optimal for a PROM.

In this study, 6 items were added to the 25 items previously identified by a PET-based study⁴⁾. Future research should investigate the structural validity of the KDI and limit the items to not only allow for calculation of sum scores but also minimize patients' effort in completing the questionnaire. Moreover, although *Katakori* is a Japanese term, nonspecific symptoms, including discomfort or dull pain, around the occiput through the cervical spine to the acromion and scapular area are not exclusive to the Japanese population^{18, 19)}. Therefore, future research should also investigate the cross-cultural validity of the KDI.

The limitations of our research are the convenient sampling and the use of biased samples of participants in their 20s and 40s. Further studies on the KDI's structural validity should investigate whether the instrument is easily understood by and applicable to older adults by evaluating missing items with participants from different age groups. Such research should also confirm whether the completion time for the KDI is <10 min, which is considered to be the median value for a feasible web survey²⁰⁾.

In conclusion, by evaluating patient's comprehensiveness and comprehensibility, we developed the 31-item KDI, which will be a foundation for establishing a valid and reliable PROM of disability due to *Katakori*.

Funding and Conflict of interest

None.

REFERENCES

- 1) Yabuki S: Pathogenesis of the neck-shoulder stiffness (*Katakori*). *Rinsho Seikei Geka*, 2007, 42: 413–417.
- 2) Japan Ministry of Health, Labour and Welfare. From Comprehensive Survey of Living Conditions, 2019: Graphical Review of Japanese Household. <https://www.mhlw.go.jp/toukei/list/dl/20-21-h29.pdf>. (Accessed Dec. 22, 2021)
- 3) Aoki K, Hall T, Takasaki H: Reporting on the level of validity and reliability of questionnaires measuring *Katakori* severity: a systematic review. *SAGE Open Med*, 2019, 7: 2050312119836617. [Medline] [CrossRef]
- 4) Natsume K, Handa Y, Takasaki H: Determination of the conceptual structures of the disturbed activities of daily living due to *Katakori* by evaluating patients' values or perceptions using a patient-elicitation technique. *J Phys Ther Sci*, 2021, 33: 683–688. [Medline] [CrossRef]
- 5) Vernon H, Mior S: The Neck Disability Index: a study of reliability and validity. *J Manipulative Physiol Ther*, 1991, 14: 409–415. [Medline]
- 6) Miyazaki J, Kuge H, Sakaguchi S, et al.: Design of a questionnaire on stiff shoulders with activities of daily living as indices: Inter-rater reliabilities of questionnaires on shoulder pain and stiff shoulders, and the repeatability of the latter questionnaire. *Orient Med Pain Clin*, 2014, 44: 2–10.
- 7) Prinsen CA, Vohra S, Rose MR, et al.: How to select outcome measurement instruments for outcomes included in a “Core Outcome Set”—a practical guideline. *Trials*, 2016, 17: 449. [Medline] [CrossRef]
- 8) Walton DM, MacDermid JC, Pulickal M, et al.: Development and initial validation of the Satisfaction and Recovery Index (SRI) for measurement of recovery from musculoskeletal trauma. *Open Orthop J*, 2014, 8: 316–325. [Medline] [CrossRef]
- 9) Hush JM, Kamper SJ, Stanton TR, et al.: Standardized measurement of recovery from nonspecific back pain. *Arch Phys Med Rehabil*, 2012, 93: 849–855. [Medline] [CrossRef]
- 10) Terwee CB, Prinsen CA, Chiarotto A, et al.: COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study. *Qual Life Res*, 2018, 27: 1159–1170. [Medline] [CrossRef]
- 11) Mookkink LB, de Vet HC, Prinsen CA, et al.: COSMIN risk of bias checklist for systematic reviews of patient-reported outcome measures. *Qual Life Res*, 2018, 27: 1171–1179. [Medline] [CrossRef]
- 12) Stanton TR, Latimer J, Maher CG, et al.: Definitions of recurrence of an episode of low back pain: a systematic review. *Spine*, 2009, 34: E316–E322. [Medline] [CrossRef]
- 13) Iizuka Y, Shinozaki T, Kobayashi T, et al.: Characteristics of neck and shoulder pain (called *katakori* in Japanese) among members of the nursing staff. *J Orthop Sci*, 2012, 17: 46–50. [Medline] [CrossRef]
- 14) Spadoni GF, Stratford PW, Solomon PE, et al.: The evaluation of change in pain intensity: a comparison of the P4 and single-item numeric pain rating scales. *J Orthop Sports Phys Ther*, 2004, 34: 187–193. [Medline] [CrossRef]
- 15) Shiroiwa T, Ikeda S, Noto S, et al.: Comparison of value set based on DCE and/or TTO data: scoring for EQ-5D-5L health states in Japan. *Value Health*, 2016, 19: 648–654. [Medline] [CrossRef]
- 16) van Hout B, Janssen MF, Feng YS, et al.: Interim scoring for the EQ-5D-5L: mapping the EQ-5D-5L to EQ-5D-3L value sets. *Value Health*, 2012, 15: 708–715. [Medline] [CrossRef]
- 17) Simms LJ, Zelazny K, Williams TF, et al.: Does the number of response options matter? Psychometric perspectives using personality questionnaire data. *Psychol Assess*, 2019, 31: 557–566. [Medline] [CrossRef]
- 18) Läubli T, Thomas C, Hinnen U, et al.: [Assessment of musculoskeletal disorders using a questionnaire]. *Soz Präventivmed*, 1991, 36: 25–33 (in German). [Medline]
- 19) Hill CL, Lester S, Taylor AW, et al.: Factor structure and validity of the shoulder pain and disability index in a population-based study of people with shoulder symptoms. *BMC Musculoskelet Disord*, 2011, 12: 8. [Medline] [CrossRef]
- 20) Revilla M, Ochoa C: Ideal and maximum length for a web survey. *Int J Mark Res*, 2017, 59: 557–565.