ACHSM Asia-Pacific Journal of Health Management

RESEARCH ARTICLE

TRANSLATION, ADAPTATION, AND VALIDATION OF SURVEY ON PATIENT SAFETY CULTURE FOR AMBULATORY CARE IN JAPAN

*Khin Zay Yar Myint** [1], *Ikumi Genka* [1], *Junichi Taguchi* [1, 2], *Toshiomi Kusano* [1]

1. Tokyo Midtown Clinic, Minato-ku, Tokyo, Japan
2. Nihonbashi Muromachi Mitsui Tower Midtown Clinic, Chiyoda-ku, Tokyo, Japan

*Correspondence: khinzayarmyint@gmail.com

ABSTRACT

OBJECTIVE

There is no validated questionnaire in Japan to measure the culture of safety in ambulatory care clinics. Therefore, we developed the Japanese version of the Medical Office Survey on Patient Safety Culture (MOSPSC) of the Agency of Healthcare Research and Quality (AHRQ) in the United States with the aim to establish a tool for evaluating and benchmarking the safety culture of outpatient clinics in Japan.

DESIGN AND SETTING

This research uses both qualitative and quantitative approaches to translate, adapt and validate the MOSPSC questionnaire which consists of 62 questions. The process involved seven steps such as translation by two independent bilingual physicians, drafting and reviewing, backtranslation by two separate translation companies, semantic equivalence assessment by AHRQ and revision, pretest, focused discussion, and finalizing the questionnaire after expert review and proofreading. An actual safety culture survey was conducted with mainly online and paper versions at four clinics in Tokyo.

MAIN OUTCOME MEASURES

The survey results were evaluated for patient safety dimensions, reliability and construct validity.

RESULTS

Efforts are made to select appropriate terminology during tool adaptation processes due to different language and medical system between Japan and the United States. The response rate in the actual survey was 66.4% (242/364). Confirmatory factor analysis showed that factor loading and goodness of fit indices were better when 3 items were removed from the original 10-composite model with 38 items. The Cronbach's alpha coefficients of composite measures ranged from 0.62 to 0.78 in the original model and 0.62 to 0.85 in the new model, indicating good internal consistencies.

CONCLUSIONS

Considering the differences in medical systems, culture, and language between the United States and Japan, the instrumented was adapted with a satisfactory content validity and reliability.

KEYWORDS

safety culture, ambulatory care, MOSPSC, validation, adaptation.

BACKGROUND

With increasing complexity and advances in medicine, modern healthcare has become vulnerable to errors. Consequently, fostering a culture of safety has been prioritized in medical institutions in recent years. Safety culture is a sub-facet of the organizational culture that pertains to safety and reflects an organization's commitment to, and effectiveness in, health and safety management [1,2]. According to the Japanese Ministry of Health, Labour, and Welfare, safety culture in healthcare is defined as "A way of thinking or an attitude in which all the staff engaged in medical services give the highest priority to patient safety and strive for its realization, as well as the way organizations work toward that goal"[3].

Dr. James Reason identified five elements of safety culture: an informed culture, a reporting culture, a learning culture, a just culture, and a flexible culture [4]. The outcome of these safety cultures is reflected in medical errors, and it is important to analyze these errors from incident reports, mortality and morbidity conferences, and chart reviews [5-8]. As for a proactive approach to preventing errors, we can evaluate the pre-existing organizational culture in relation to safety, develop action plans for improvements, assess safety interventions and changes over time, and serve as internal and external benchmarking [9-10]. However, measuring safety culture presents several challenges [11]. Measurements include surveys, document reviews, work observations, individual interviews, and focused groups [12]. Among them, surveys are practical and time-efficient and, most importantly, provide ease of benchmarking.

There are different safety culture surveys in healthcare [13]. The most widely adopted international patient safety culture surveys were developed by the Agency of Healthcare Research and Quality (AHRQ). The Japan Council for Quality Health Care has used the validated Japanese version of the Hospital Survey on Patient Safety Culture (HSOPS) by the AHRQ to evaluate the safety culture of hospitals in Japan [14, 15]. Conversely, there is no questionnaire in Japan to evaluate the safety culture of ambulatory medical facilities.

The AHRQ has a safety culture survey called "Medical Office Survey on Patient Safety Culture (MOSPSC)" for ambulatory care. As of September 2022, it has been translated into 15 languages and conducted in 38 countries [16]. Yet, there is no Japanese version of the MOSPSC. Moreover, it is critical to develop a reliable and valid instrument that can be implemented at ambulatory medical institutions in Japan to evaluate safety culture, implement effective measures according to the survey results, and provide benchmarking with other facilities. Therefore, we developed a Japanese version of the AHRQ's MOSPSC to improve medical safety and quality in our group of clinics and expand its use for external benchmarking.

METHODOLOGY

INSTRUMENT

The MOSPSC consists of 62 questions, including free comments, and evaluates the safety culture of the respondents, the team they work for, their immediate supervisor, and the organization in sections A to H [17]. Question items in section A and B are based on a 6-point Likert scale (e.g., 1 = Daily, 2 = Weekly, 3 = Monthly, 4 = Several times in the past 12 months, 5 = Once or twice in the past 12 months, 6 = Not in the past 12 months) . Question items in section C through F are based on 5-point Likert scale (e.g., 1 = Strongly Disagree, 2 = Disagree, 3 = Neither Agree nor Disagree, 4 = Agree, 5 = Strongly Agree). In those sections, score 9 refers to "Does Not Apply or Don't Know". Overall rating in section G was based on 5-point Likert scale (e.g., 1 = Poor, 2 = Fair, 3 = Good, 4 = Very good, 5 = Excellent). The details of the questionnaire items in English were provided as the supplementary data file.

FORWARD TRANSLATION

We developed the Japanese version of the MOSPSC in seven steps: Step 1 was the translation of the original MOSPSC questionnaire into Japanese by two bilingual physicians whose native language is Japanese, one of whom knew the purpose of the translation, and the other who did not. In Step 2, these two translations were reviewed by a review

committee consisting of the above two physicians, the principal investigator, and the chief nurse. The review committee identified noticeable differences in the classification and roles of medical professions between the United States and Japan and question items that were difficult to translate due to the use of double negatives. The review committee also referred to question items already included in the validated Japanese version of HSOPS for appropriate wording and structure [14]. After all these considerations, the two translations were modified and integrated into the first Japanese draft (T1).

BACK-TRANSLATION AND SEMANTIC EQUIVALENCE

Step 3 was the back translation into English by two external translators from different translation companies who were native English speakers and did not know the original English version of the questionnaire. Step 4 involved examining the semantic equivalence of the two back-translated versions with the original English version. They pointed out five points and we responded accordingly. The details are described in Supplementary data file. Based on our responses, the AHRQ determined that the items closely followed the original survey, and we proceeded with the next steps in administering the survey at our medical facilities. In Step 5, based on the feedback from the AHRQ, the review committee modified the integrated version (T1) into a revised version (T2).

PILOT TEST AND FOCUSED DISCUSSION

In phase I of Step 6, we created T2 as a web questionnaire using the Microsoft form application. We then conducted a preliminary test with 34 employees at the Tokyo Midtown Clinic in Roppongi. The preliminary test included twelve medical staff, eight nurses, five administrative staff, three clinical technicians, three radiological technicians, two nurse aids, and one manager. In phase II of Step 6, a focused discussion with five staff members from five different professions (i.e., doctors, nurses, clinical and radiological technicians, and administrative staff) was conducted to evaluate the clarity of the questions and the selection of appropriate words. Details are provided in Supplementary data file. The questionnaire edited in accordance with the focused group discussion results was termed T3.

PROOFREADING BY AN EXPERT AND ACTUAL SURVEY

Finally, in step 7, an expert Japanese language proofreader assessed the grammar, style, clarity, pertinence, and understandability of the contents. He suggested changing all sentences from the polite form to the dictionary form in Japanese. We developed the final version of the questionnaire (T4) to conduct an actual survey at four clinics in the Tokyo area: Tokyo Midtown Roppongi, Nihonbashi, Ariake, and Tokyo Bay, all of which belong to the Medical Corporation Midtown Clinic. A survey link was provided to all staff working at these facilities. For those who did not have email addresses or personal computers at the workplace, paper survey questionnaires were distributed, and patients were asked to drop them into a designated internal post box after completion. The survey was distributed and collected between September 21, 2023, to November 30, 2023. The timeline and flowchart for developing the questionnaire are outlined in Supplementary Figure file.

ETHICAL CONSIDERATIONS

The survey was conducted anonymously, in a self-administered manner. After explaining the purpose and other essential contents of the survey on the first page, we asked for consent. Only those who provided consent on the first page were included in the final analysis. The study was approved by the Institutional Review Board of The Midtown Clinic Medical Corporation (Rin-2021-10).

DATA ANALYSIS

First, we entered the data from the paper survey forms into a web form and saved the data file in Microsoft Excel. There was no missing data. We then coded the responses in accordance with the scales in each section: 1 to 5 in sections A and B, and 1 to 6 in sections C to G, with 9 coded for Does not Apply/Don't know. The next step was to check the data file for possible data entry errors using frequency tables for each item to ensure that all responses were within a valid range. We followed the AHRQ's MOSPSC data analysis guidelines to calculate the frequency percentages of positive responses to the single-item measures in sections A and B. We categorized the 38 items in sections C to F into 10 composite measures of patient safety culture and calculated the percent positive response, means, and composite measure scores.

Negatively worded items were reverse-coded, as recommended by the AHRQ. The results were then compared with those from the AHRQ database.

Cronbach's alpha coefficient and the average inter-item correlation (AIIC) were used to analyze reliability. We considered Cronbach's alpha values >0.6 adequate for each composite, and at least 0.9 for the entire questionnaire. AIIC evaluates the correlation of items within each composite, and a value between 0.15 and 0.5 is considered evidence that the items are measuring the same underlying composite [18]. For structural validity, we performed confirmatory factor analysis (CFA) to determine how well the Japanese version fits the composite structure of the original MOSPSC. We did not perform exploratory factor analysis because we aimed to maintain the same factor structure as the original MOSPSC for benchmarking, and the MOSPSC has already been validated across different languages. First, we conducted Kaiser-Meyer-Olkin (KMO) and Bartlett tests of sphericity to evaluate whether our sample size was appropriate for factor analysis. The cut-off value for factor loadings was set to 0.40. We set the criteria for goodness-of-fit indices based on previous studies [19]. CFA was repeated if the factor loadings or model fit indices were not acceptable.

Finally, we calculated polychoric correlations for the 10 composites of patient safety culture to verify construct validity, as the items were ordinal variables. Figures and tables were generated using Microsoft Excel and PowerPoint. Comparison with AHRQ database was conducted using the Medical Office Survey Excel Tool Version 6.2 July 2024.20 Data analysis and statistical tests were performed using the R software (R version 4.4.2). The CFA was performed using the "*lavaan*" package, and the Cronbach's alpha and AIIC calculations were performed using the "*psych*" package. Polychoric correlations and correlation plot were generated using the "*polycor*" and "*corrplot*" packages.

RESULTS

CHARACTERISTICS OF PARTICIPANTS

A total of 242 employees responded to the survey: 165 participated via the online form, and 77 returned the paper questionnaires. The response rate was 66.4% (242/364). Among 242 responders, 20 people did not give consent to participate in the study nor answer any survey questions. Therefore, we analyzed the data of the 222 participants who provided consent to participate in the study and answered the questionnaire (Figure 1). Detailed characteristics of the participants are presented in Table 1. The rates of participation according to occupation were 61%, 27%, 71%, 74%, and 60% for physicians, leaders and managers, administrative or clerical staff, nurses, and other clinical staff or paramedics, respectively.

FIRGURE 1. PARTICIPANT FLOWCHART

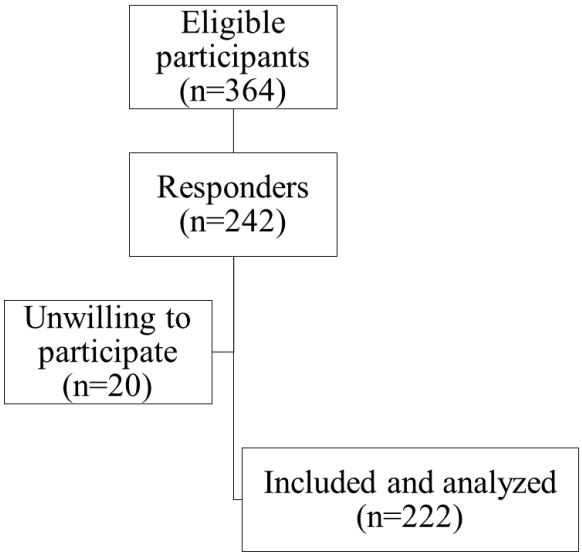


TABLE 1. CHARACTERISTICS OF THE PARTICIPANTS (N=222)

Variable	n	(%)
Years of experience		
Less than 2 months	9	(4%)
2 months to less than 1 year	32	(14%)
1 year to less than 3 years	47	(21%)
3 years to less than 6 years	70	(32%)
6 years to less than 11 years	48	(22%)
11 years or more	16	(7%)
Staff position		
Physician (MD or DO)	22	(10%)
Leaders and managers	15	(7%)
Administrative or Clerical Staff	55	(25%)
Nurse (RN), Licensed Vocational Nurse (LVN), Licensed Practical Nurse (LPN)	69	(31%)
Other Clinical Staff or Clinical Support Staff	60	(27%)
Others	1	(<1%)

SINGLE ITEM MEASURES AND COMPOSITE MEASURES

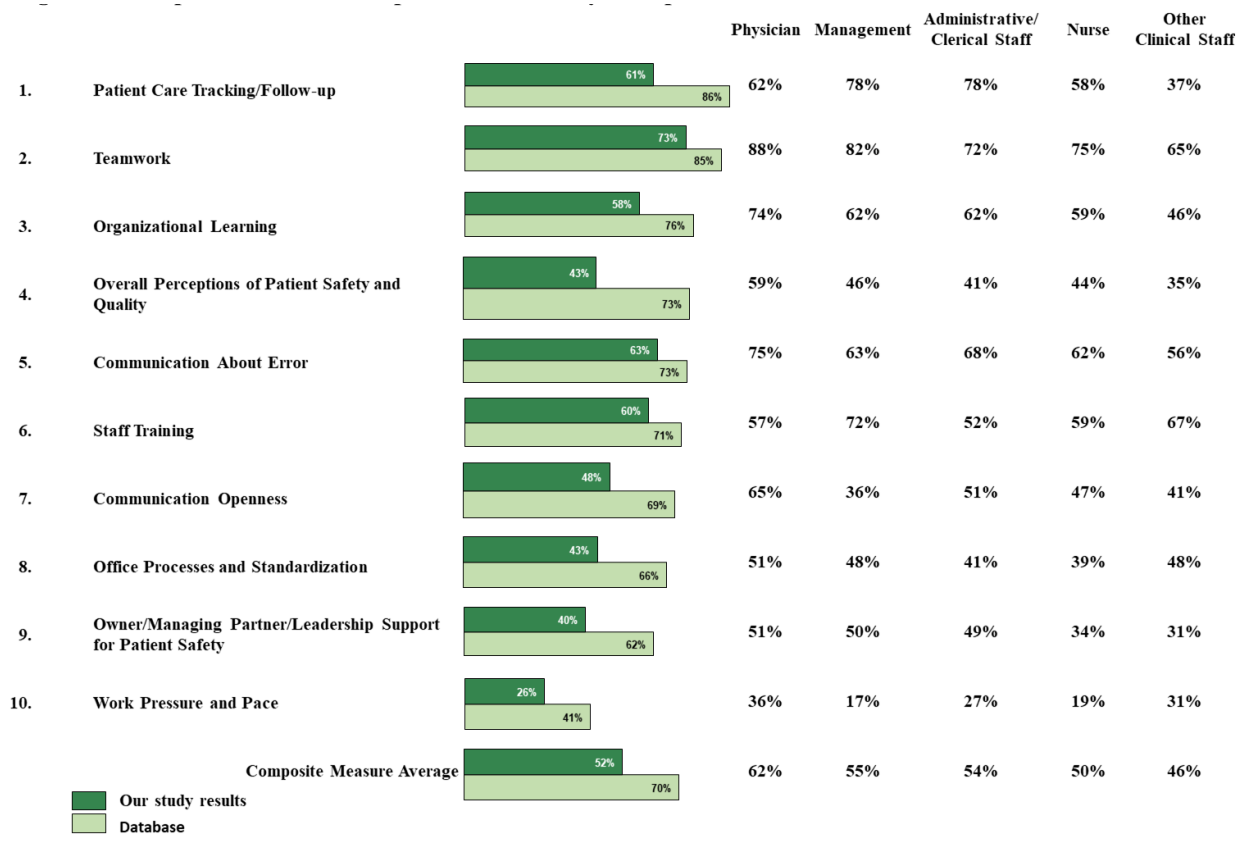
Table 2 shows the positive perception rate of Patient Safety and Quality Issues, information exchange with other settings, and the overall positive rating of quality and patient safety compared to the AHRQ database. The results of the first two measures were superior to those of the AHRQ database, but those of the last two measures were inferior. Figure 2 shows the positive response rates for the composite measures compared to the AHRQ database and among different occupations. The positive response results of the 10 composites were lower than those in the AHRQ database.

TABLE 2. SINGLE ITEM MEASURE COMPARATIVE RESULTS

Note: Database results are based on data from 1,164 medical offices included in the 2024 SOPS Medical Office Database Report; 2) Your medical office's % positive composite measure results are not displayed if 5% or less.		Positive		Neutral		Negative	
		Our study	AHRQ Database	Our study	AHRQ Database	Our study	AHRQ Database
List of Patient Safety and Quality Issues							
A1	A patient was unable to get an appointment within 48 hours for an acute/serious problem.	96%	64%	0%	9%	4%	27%
A2	The wrong chart/medical record was used for a patient.	100%	97%	0%	2%	0%	1%
A3	A patient's chart/medical record was not available when needed.	96%	93%	4%	3%	0%	4%
A4	Medical information was filed, scanned, or entered into the wrong patient's chart/medical record.	97%	97%	1%	2%	2%	1%
A5	Medical equipment was not working properly or was in need of repair or replacement.	94%	88%	6%	6%	0%	6%

A6	A pharmacy contacted our office to clarify or correct a prescription.	64%	61%	10%	13%	26%	26%
A7	A patient's medication list was not updated during his or her visit.	95%	78%	5%	7%	0%	15%
A8	The results from a lab or imaging test were not available when needed.	95%	76%	4%	10%	1%	14%
A9	A critical **abnormal** result from a lab or imaging test was not followed up within 1 business day.	100%	92%	0%	3%	0%	5%
Information Exchange with Other Settings							
B1	Over the past 12 months, how often has your medical office had **problems exchanging accurate, complete, and timely information with:** Outside labs/imaging centers?	97%	73%	2%	12%	1%	15%
B2	Over the past 12 months, how often has your medical office **had problems exchanging accurate, complete, and timely information with**: Other medical offices/outside physicians?	95%	74%	1%	12%	3%	14%
B3	Over the past 12 months, how often has your medical office had **problems exchanging accurate, complete, and timely information with**: Pharmacies?	95%	74%	1%	9%	4%	17%
B4	Over the past 12 months, how often has your medical office **had problems exchanging accurate, complete, and timely information with**: Hospitals?	95%	80%	1%	9%	4%	11%
Overall Ratings on Quality							
G1A	Patient-centered: Is responsive to individual patient preferences, needs and values.	29%	67%	36%	25%	35%	8%
G1B	Effective: Is based on scientific knowledge.	23%	69%	40%	25%	37%	6%
G1C	Timely: Minimizes waits and potentially harmful delays.	24%	54%	32%	28%	43%	18%
G1D	Efficient: Ensures cost-effective care (avoids waste, overuse, and misuse of services).	18%	58%	27%	29%	55%	13%
G1E	Equitable: Provides the same quality of care to all individuals regardless of gender, race, ethnicity, socioeconomic status, language, etc.	33%	81%	33%	15%	35%	4%
Overall Rating on Patient Safety							
G2	Overall, how would you rate the systems and clinical processes your medical office has in place to prevent, catch, and correct problems that have the potential to affect patients?	16%	64%	31%	27%	52%	9%

FIRGUE 2: COMPOSITE MEASURE COMPARATIVE RESULTS BY STAFF POSITION



RELIABILITY

Cronbach's alpha value of each composite ranged from 0.5 to 0.6 in the pilot test and above 0.6 in the actual survey, indicating better internal consistency. Additionally, the actual survey achieved good overall consistency (>0.9). AIIC ranged from 0.31 to 0.71 in the pilot test and 0.33 to 0.49 in the actual survey. Therefore, AIIC improved, indicating that the items measured the same construct. Detailed results are listed in Table 3. It also showed that our adapted version had lower Cronbach's alpha values compared to the AHRQ reliability statistics, which is based on the pilot test data from 202 medical offices and more than 4200 staff [20]. Notably, composite 8, "Staff training," had the lowest internal consistency (0.8 versus 0.62).

TABLE 3. RELIABILITY OF EACH COMPOSITE MEASURE

Composite measures	Number of items	Cronbach's alpha (AHRQ)	Cronbach's alpha (pilot)	Cronbach's alpha (actual)	AIIC (pilot)	AIIC (actual)
1. Communication About Error	4	0.75	0.59	0.63	0.31	0.32
2. Communication Openness	4	0.81	0.71	0.73	0.42	0.41
3. Office Processes and Standardization	4	0.77	0.83	0.69	0.54	0.36
4. Organizational Learning	3	0.82	0.76	0.74	0.52	0.49
5. Overall Perceptions of Patient Safety and Quality	4	0.79	0.83	0.78	0.56	0.48
6. Leadership Support for Patient Safety	4	0.76	0.88	0.75	0.62	0.42

7. Patient Care Tracking/Follow-up	4	0.78	0.87	0.72	0.71	0.41
8. Staff Training	3	0.8	0.58	0.62	0.31	0.37
9. Teamwork	4	0.83	0.69	0.70	0.40	0.38
10. Work Pressure and Pace	4	0.76	0.82	0.68	0.52	0.33
Total	38		0.81	0.94	0.34	0.22

CONSTRUCT VALIDITY

A KMO of 0.74 and the significance of the Bartlett's test of sphericity ($p < 0.05$) suggested that our sample size was appropriate for factor analysis. Standardized factor loadings ranged from 0.09 to 0.9, and some indicators for model fit did not meet the criteria ($X^2 = 1362.6$, df = 620, $p < 0.001$, RMSEA = 0.074 [90% confidence interval: 0.068-0.079], CFI = 0.719, SRMR = 0.108). As some factors (D7, E2, and C3) showed inadequate factor loadings (0.17, 0.09, and 0.26, respectively) and Cronbach's alpha could be improved by dropping these items, we decided to run another CFA after deleting those items.

The new data model of 35 items was also appropriate for factor analysis (KMO=0.81, Bartlett's test of sphericity <0.05). Standardized factor loadings ranged from 0.51 to 0.91, and goodness-of-fit indices were either excellent or acceptable ($X^2 = 1055$, df = 515, $p < 0.001$, RMSEA = 0.069 [90% confidence interval: 0.063-0.075], CFI = 0.814, SRMR = 0.087). Cronbach's alpha values improved, but the average inter-item correlations of the composites marginally increased. Supplementary table shows the CFA results of the 38-item model and the changes in reliability statistics after three items were deleted. The deleted questions are shown in bold.

Figure 3 shows the polychoric correlations of the new 35-item model of the adapted version. Positive correlations were observed among 10 composites of patient safety culture ranging from 0.04 to 0.66. The highest correlation (r=0.66) was found between "Office Processes and Standardization" and "Work Pressure and Pace". Conversely, the lowest correlation (r=0.05) was observed between "Organizational Learning" and "Work Pressure and Pace" as well as "Owner/Managing Partner/Leadership Support for Patient Safety" and "Teamwork".

FIRGURE 3: POLYCHORIC CORRELATIONS OF 35-ITEM JAPANESE MOSPSC



DISCUSSIONS

This is the first study to translate, adapt, and validate the MOSPSC into Japanese. In 2011, Ito et al. culturally adapted the HSOPS for use in Japan. Their instrument had acceptable levels of internal consistency and construct validity, and the composite structures were almost identical to the original HSOPS [21]. Similarly, in our study, both the original 38-item and new 35-item models indicated good internal consistencies with a high overall Cronbach's alpha value (>0.9), and moderate-to-high alpha values in composite measures (>0.6). The AIIC values indicated that the items in the composite measured the same construct, supporting homogeneity and unidimensionality. The model structure with 10 composites was also acceptable when evaluated in the confirmatory factor analysis.

When goodness-of-fit indices were evaluated, the 35-item model had better goodness-of-fit and internal consistency than the original 38-item model. The 35-item model was an adapted version after deleting three items with low factor loadings on composites 1, 6, and 10 from the original model. Two of the deleted items were the only negatively and positively worded items in composites 1 and 6, and the third item in composite 10 was the one we modified the sentence structure as the AHRQ pointed out that it should be tested again as the two back-translations differed. Therefore, it is necessary to investigate the wording or structure of these items in further studies on the MOSPSC in Japan.

Consistent with previous studies on cross-cultural adaptation of MOSPSC in other countries [22-25], our Japanese version of MOSPSC showed a lower internal consistency compared to the AHRQ reliability statistics, especially "Staff training". In this composite, training on new processes, providing on-the-job training, and assigning untrained tasks were evaluated and yielded inconsistent responses, which may be responsible for the relatively low internal consistency. These results also identified areas for improvement in staff training and education.

In our Japanese version, we identified a strong patient safety dimension with regards to access to care, patient identification, availability and completeness of medical records, maintenance of medical equipment, medication updates, and timeliness of test results. Similarly, the positive rating results regarding exchanging accurate, complete, and timely information with labs, imaging centers, other medical offices, outside physicians, pharmacies, hospitals, etc. were high (>95%) and superior to the AHRQ database. However, the overall positive rating of quality was very low (18% to 33%) when compared to the AHRQ database, which ranged from 54% to 81%. The overall positive rating for patient safety was quite low, accounting for only one-fourth of the AHRQ database results (16% versus 64%). These results identified areas for improvement and maintenance of their strengths. They also indicated the need to raise awareness among leaders and staff regarding the results of the MOSPSC survey.

Among the composites of patient safety culture, teamwork was rated highest (73% positive rating), particularly among physicians (88%). "Teamwork" was also identified as the highest safety domain in a patient safety culture survey conducted using the MOSPSC in Yemen (96%), Portugal (76%), and Greece (82%) [22-24]. Conversely, work pressure and pace were poorly rated (26%), especially among leadership (17%) and nurses (19%). This might be due to an imbalance between workload and human resources, particularly among nurses. The shortage of nurses has long been an issue due to the increasing healthcare demand associated with a rapidly aging society. According to the Ministry of Health, Labor, and Welfare (MHLW), there could be a shortage of 60,000–270,000 nurses by 2025 [26]. The healthcare working environment in Japan is also extremely harsh due to free access to the healthcare system. According to 2020 OECD report, the number of outpatient visits per physician per year was 3.5 times that of the United States [27]. "Work Pressure and Pace" was also the lowest scoring dimension on patient safety culture surveys in Portugal (21%), Greece (46%), Slovenia (10.7%), and Yemen (50%) [22-25].

Our study has a few limitations. First, the survey was conducted in only four clinics in the Tokyo area. This could affect the representativeness of the sample and generalizability of the findings to other outpatient clinics. However, since it is the Japanese outpatient version of the medical safety survey questionnaire, whose reliability and validity have been verified for the first time, it will be useful for improving medical safety and quality in our medical group, as well as in other outpatient

medical institutions, serving as a benchmark between facilities. Second, we did not examine other psychometric properties (e.g., convergent validity, discriminant validity, and test-retest reliability) beyond structural validity. As we plan to conduct this survey annually, we can evaluate and improve it in the future.

CONCLUSIONS

We adapted the Japanese version of the MOSPSC and verified that the instrument had good reliability and construct validity. This is the first study to develop an instrument that can be used to measure patient safety culture, benchmark other facilities, and promote patient safety efforts in Japanese ambulatory clinics. Owing to the increasing healthcare demand, shortage of human resources, and sophistication of healthcare in Japan, our tool is very relevant for continuously monitoring safety culture and identifying areas of improvement to focus actions for improvement.

ACKNOWLEDGEMENTS

The authors thank the Agency of Healthcare Research and Quality (AHRQ), Dr. Keisuke Hata, Dr. Ichiro Akagi, Dr. Ken Endo, Dr. Kazuhiro Hara, members of the expert panel, and staff at four clinics in the Tokyo area: Tokyo Midtown Roppongi, Nihonbashi, Ariake and Tokyo Bay.

DECLARATION OF COMPETING INTERESTS

Corresponding authors, on behalf of all the authors of a submission, disclose no conflict of interest regarding this research article.

References

1. Cooper MD. Towards a model of safety culture. Safety Science [Internet]. 2000;36(2):111–36. Available from: http://www.behavioral-safety.com/articles/Towards_a_model_of_safety_culture.pdf
2. Naji GMA, Isha ASN, Mohyaldinn ME, Leka S, Saleem MS, Rahman SMNBSA, et al. Impact of Safety Culture on Safety Performance; Mediating Role of Psychosocial Hazard: An Integrated Modelling Approach. International Journal of Environmental Research and Public Health [Internet]. 2021 Aug 13;18(16):8568. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8394037/
3. Yokoyama R, Hasegawa T. The Survey on Patient Safety Culture; Visualization of Patient Safety Culture. Journal for Japan Society for Safety Engineering. 2021 60(6)[cited 2024 Dec 11];454–61. Available from: https://www.jstage.jst.go.jp/article/safety/60/6/60_454/_article/-char/ja/
4. Reason J. Human Error. Cambridge: Cambridge University Press; 1990.
5. Rufo J. The Differences Between Human Error, At-Risk Behavior, and Reckless Behavior Are Key to a Just Culture [Internet]. ECRI and ISMP. 2020. Available from: https://home.ecri.org/blogs/ismp-alerts-and-articles-library/the-differences-between-human-error-at-risk-behavior-and-reckless-behavior-are-key-to-a-just-culture
6. Thomas EJ, Petersen LA. Measuring Errors and Adverse Events in Health Care. Journal of General Internal Medicine [Internet]. 2003 Jan 1;18(1):61–7. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1494808/
7. Garrouste-Orgeas M, Philippart F, Bruel C, Max A, Lau N, Misset B. Overview of medical errors and adverse events. Annals of Intensive Care. 2012 Feb 16;2(1).
8. Mulac A, Anne Gerd Granås. Detecting Medication Errors and Adverse Drug Events: A Review of Current Methodologies. Duo Research Archive (University of Oslo). 2022 Jan 1;135–60.
9. Nieva VF, Sorra J. Safety culture assessment: a tool for improving patient safety in healthcare organizations. Quality and Safety in Health Care [Internet]. 2003 Dec 1;12(90002):ii17-23. Available from: https://qualitysafety.bmj.com/content/12/suppl_2/ii17
10. Bienassis K de, Klazinga N. Comparative assessment of patient safety culture performance in OECD countries. OECD health working papers [Internet]. 2024 Jul 12 [cited 2024 Aug 8]; Available from: https://www.oecd-ilibrary.org/social-issues-migration-health/comparative-assessment-of-patient-safety-culture-performance-in-oecd-countries_d0552328-en

11. Ellis LA, Falkland EC, Hibbert P, Wiig S, Ree E, Schultz T, et al. Issues and complexities in safety culture assessment in healthcare. Frontiers in Public Health [Internet]. 2023 Jun 15;11. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10309647/
12. International Atomic Energy Agency. Performing safety culture self-assessments. Vienna: International Atomic Energy Agency; 2016.
13. Colla JB. Measuring patient safety climate: a review of surveys. Quality and Safety in Health Care. 2005 Oct 1;14(5):364–6.
14. Kigawa M, Fujita S, Seto K, Matsumoto K, Hirao T, Hasegawa T, et al. Questionnaire on Patient Safety Culture for Hospitals by the U.S. Agency for Healthcare Research and Quality and its applicability in Japan. The Journal of Japan Society for Health Care Management [Internet]. 2010 May [cited 2025 Dec 16];11(1):2–14. Available from: https://www.jstage.jst.go.jp/article/jhm/11/1/11_2/_article/-char/ja/
15. List of hospitals participated in Hospital Survey on Patient Safety Culture. [Internet]. Japan Council for Quality Health Care. Japan Council for Quality Health Care; 2025 [cited 2025 Jun 1]. Available from: https://www.jq-hyouka.jcqhc.or.jp/safe/users/
16. International Use of SOPS [Internet]. Ahrq.gov. 2022 [cited 2025 Dec 16]. Available from: https://www.ahrq.gov/sops/international/index.html
17. Medical Office Survey on Patient Safety Culture [Internet]. www.ahrq.gov. Available from: https://www.ahrq.gov/sops/surveys/medical-office/index.html
18. Clark LA, Watson D. Constructing validity: New Developments in Creating Objective Measuring instruments. Psychological Assessment. 2019 Mar 21;31(12):1412–27.
19. Hooper D, Coughlan J, Mullen MR. Structural Equation Modelling: Guidelines for Determining Model Fit. Electronic Journal of Business Research Methods [Internet]. 2008 Sep 1;6(1):pp53-60–0. Available from: https://academic-publishing.org/index.php/ejbrm/article/view/1224
20. Data Entry and Analysis Tools [Internet]. Agency for Healthcare Research and Quality, Rockville, MD. U.S. Department of Health & Human Services; 2022 [cited 2025 Jun 15]. Available from: https://www.ahrq.gov/sops/resources/data-entry-analysis-tools.html
21. Ito S, Seto K, Kigawa M, Fujita S, Hasegawa T, Hasegawa T. Development and applicability of Hospital Survey on Patient Safety Culture (HSOPS) in Japan. BMC Health Services Research. 2011 Feb 7;11(1).
22. Webair HH, Al-assani SS, Al-haddad RH, Al-Shaeeb WH, Bin Selm MA, Alyamani AS. Assessment of patient safety culture in primary care setting, Al-Mukala, Yemen. BMC Family Practice. 2015 Oct 13;16(1).
23. Eiras M, Escoval A, Silva C. Patient Safety Culture in Portuguese Primary Care: Validation of the Portuguese Version of the Medical Office Survey [Internet]. www.intechopen.com. IntechOpen; 2018. Available from: https://www.intechopen.com/chapters/62934
24. Antonakos I, Souliotis K, Psaltopoulou T, Tountas Y, Papaefstathiou A, Kantzanou M. Psychometric Properties of the Greek Version of the Medical Office on Patient Safety Culture in Primary Care Settings. Medicines. 2021 Jul 26;8(8):42.
25. POTOČNIK TV, Zalika KLEMENC-KETIŠ. Perception of Patient Safety Culture at the Primary Care Level: The Case of the Community Health Centre Ljubljana. Slovenian Journal of Public Health. 2023 Dec 27;63(1):21–9.
26. Nursing-related policies [in Japanese] [Internet]. www.mhlw.go.jp. Ministry of Health, Labor and Welfare; 2023. Available from: https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/0000079675.html
27. OECD iLibrary | OECD Health Working Papers [Internet]. Oecd-ilibrary.org. 2006. Available from: https://www.oecd-ilibrary.org/social-issues-migration-health/oecd-health-working-papers_18152015

SUPPLEMENTARY TABLE (PART 1). CONFIRMATORY FACTOR ANALYSIS RESULTS AND RELIABILITY STATISTICS OF 38-ITEM MODEL BEFORE AND AFTER REMOVING 3 ITEMS WITH LOW FACTOR LOADING

	Item	Mean score	Standard deviation	Standadrized factor loading	α if an item is dropped	AIIC if an item is dropped	α (38-item model)	α (35-item model)	α (AHRQ)
1.	**Communication About Error**						0.63	0.75	0.75
D7.	**Staff feel like their mistakes are held against them. (negatively worded)**	**2.44**	**1.11**	**0.17**	**0.75**	**0.53**			
D8.	Providers and staff talk openly about office problems.	3.23	1.16	-0.52	0.55	0.32			
D11.	In this office, we discuss ways to prevent errors from happening again.	4.12	0.96	-0.84	0.44	0.21			
D12.	Staff are willing to report mistakes they observe in this office.	3.93	0.99	-0.89	0.43	0.21			
2.	**Communication Openness**						0.73	0.73	0.81
D1.	Providers in this office are open to staff ideas about how to improve office processes.	3.30	0.96	0.63	0.74	0.40			
D2.	Staff are encouraged to express alternative viewpoints in this office.	3.10	0.98	0.75	0.63	0.30			
D4.	Staff are afraid to ask questions when something does not seem right. (negatively worded)	2.36	0.99	-0.54	0.68	0.40			
D10.	It is difficult to voice disagreement in this office. (negatively worded)	2.54	1.02	-0.67	0.63	0.30			
	3. Office Processes and Standardization						0.69	0.69	0.77
C8.	This office is more disorganized than it should be. (negatively worded)	3.35	0.99	0.56	0.68	0.43			
C9.	We have good procedures for checking that work in this office was done correctly.	3.55	0.91	-0.62	0.61	0.34			
C12.	We have problems with workflow in this office. (negatively worded)	2.99	0.96	0.64	0.59	0.34			
C15.	Staff in this office follow standardized processes to get tasks done.	3.75	0.76	-0.61	0.62	0.35			
4.	**Organizational Learning**						0.74	0.74	0.82
F1.	When there is a problem in our office, we see if we need to change the way we do things.	4.05	0.65	0.60	0.75	0.61			
F5.	This office is good at changing office processes to make sure the same problems don't happen again.	3.42	0.84	0.77	0.56	0.40			
F7.	After this office makes changes to improve the patient care process, we check to see if the changes worked.	3.36	0.85	0.77	0.62	0.47			

The deleted item in 35-item model was bold.

SUPPLEMENTARY TABLE (PART 2). CONFIRMATORY FACTOR ANALYSIS RESULTS AND RELIABILITY STATISTICS OF 38-ITEM MODEL BEFORE AND AFTER REMOVING 3 ITEMS WITH LOW FACTOR LOADING

	Item	Mean score	Standard deviation	Standardized factor loading	α if an item is dropped	AIIC if an item is dropped	α (38-item model)	α (35-item model)	α (AHRQ)
5.	**Overall Perceptions of Patient Safety and Quality**						0.78	0.78	0.79
F2.	Our office processes are good at preventing mistakes that could affect patients.	3.52	0.83	0.72	0.75	0.51			
F3.	Mistakes happen more than they should in this office. (negatively worded)	2.65	0.99	-0.74	0.68	0.41			
F4.	It is just by chance that we don't make more mistakes that affect our patients. (negatively worded)	2.70	0.95	-0.70	0.74	0.50			
F6.	In this office, getting more work done is more important than quality of care. (negatively worded)	3.09	1.09	-0.64	0.74	0.49			
6.	**Owner/Managing Partner/Leadership Support for Patient Safety**						0.75	0.85	0.76
E1.	They aren't investing enough resources to improve the quality of care in this office. (negatively worded)	2.94	1.21	0.81	0.60	0.33			
E2.	They overlook patient care mistakes that happen over and over. (negatively worded)	2.30	1.06	0.91	0.63	0.35			
E3.	**They place a high priority on improving patient care processes.**	**3.27**	**1.22**	**0.09**	**0.85**	**0.65**			
E4.	They make decisions too often based on what is best for the office rather than what is best for patients. (negatively worded)	2.97	1.28	0.68	0.61	0.34			
7.	**Patient Care Tracking/Followup**						0.72	0.72	0.78
D3.	This office reminds patients when they need to schedule an appointment for preventive or routine care.	4.03	0.90	0.67	0.70	0.46			
D5.	This office documents how well our chronic-care patients follow their treatment plans.	3.63	1.00	0.65	0.62	0.37			
D6.	Our office follows up when we do not receive a report we are expecting from an outside provider.	3.39	1.08	0.58	0.68	0.44			
D9.	This office follows up with patients who need monitoring.	3.69	1.10	0.77	0.63	0.39			

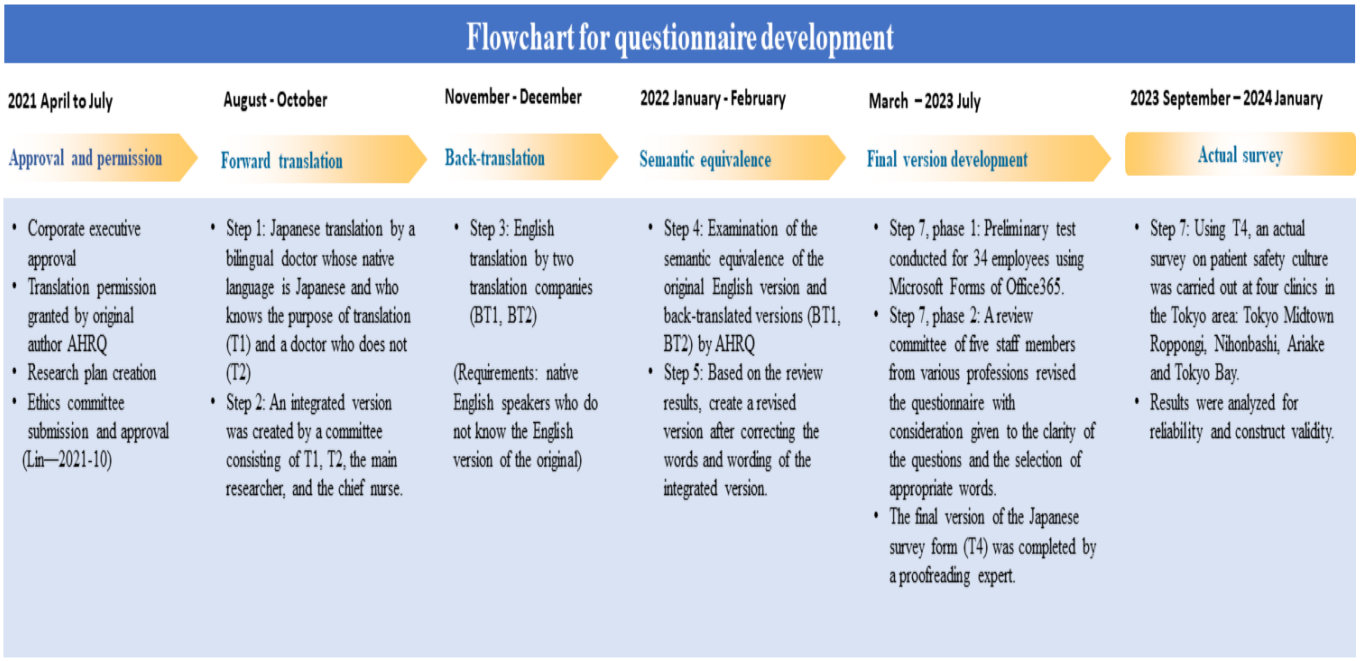
The deleted item in 35-item model was bold.

SUPPLEMENTARY TABLE (PART 3). CONFIRMATORY FACTOR ANALYSIS RESULTS AND RELIABILITY STATISTICS OF 38-ITEM MODEL BEFORE AND AFTER REMOVING 3 ITEMS WITH LOW FACTOR LOADING

	Item	Mean score	Standard deviation	Standardized factor loading	α if an item is dropped	AIIC if an item is dropped	α (38-item model)	α (35-item model)	α (AHRQ)
8.	**Staff Training**						0.62	0.62	0.8
C4.	This office trains staff when new processes are put into place.	3.78	0.83	0.72	0.50	0.00			
C7.	This office makes sure staff get the on-the-job training they need.	3.59	0.91	0.62	0.45	0.00			
C10.	Staff in this office are asked to do tasks they haven't been trained to do. (negatively worded)	2.78	1.10	-0.53	0.63	0.00			
9.	**Teamwork**						0.70	0.70	0.83
C1.	When someone in this office gets really busy, others help out.	4.26	0.72	0.52	0.66	0.40			
C2.	In this office, there is a good working relationship between staff and providers.	3.73	0.75	0.59	0.64	0.38			
C5.	In this office, we treat each other with respect.	3.86	0.78	0.74	0.58	0.32			
C13.	This office emphasizes teamwork in taking care of patients.	3.65	0.94	0.61	0.67	0.41			
10.	**Work Pressure and Pace**						0.68	0.74	0.76
C3.	**In this office, we often feel rushed when taking care of patients. (negatively worded)**	**3.47**	**0.95**	**0.26**	**0.74**	**0.40**			
C6.	We have too many patients for the number of providers in this office. (negatively worded)	3.50	1.08	0.74	0.50	0.20			
C11.	We have enough staff to handle our patient load.	3.06	1.07	-0.49	0.64	0.30			
C14.	This office has too many patients to be able to handle everything effectively. (negatively worded)	3.01	1.07	0.90	0.49	0.20			
					Overall internal consistency		0.94	0.93	

The deleted item in 35-item model was bold.

SUPPLYMENT FURGURE: STEPS INVOLVED IN THE DEVELOPMENT OF MOSPSC JAPANESE VERSION



We developed the Japanese version of MOSPSC through seven steps.