タイトル	Development, reliability, and validity of a tool for the assessment of ethical decision-making skills among health-related researchers in Japan
別タイトル	日本の医学系研究者のための倫理的意思決定スキル評価ツールの開発:信頼性と妥当性の検討
作成者（著者）	中田, 亜希子
公開者	東邦大学
発行日	2018.03.14
掲載情報	東邦大学大学院医学研究科 博士論文. 12.
資料種別	学位論文
内容記述	主査 : 澁谷和俊 / タイトル : Development, reliability, and validity of a tool for the assessment of ethical decision-making skills among health-related researchers in Japan /著者 : Akiko Nakada, Tomohide Ibuki, Taichi Kishi, Kenji Matsui, Atsushi Namiki, Shimon Tashiro /掲載誌 : Toho Journal of Medicine /巻号・発行年等 : 4(1):25-34, 2018
著者版フラグ	ETD
報告番号	32661甲第885号
学位記番号	甲第600号
学位授与年月日	2018.03.14
学位授与機関	東邦大学
DOI	info:doi/10.14994/tohojmed.2017-012
メタデータのＵＲＬ	https://mylibrary.toho-u.ac.jp/webopac/TD53271004

Original Article

Development, Reliability, and Validity of a Tool for the Assessment of Ethical Decision-making Skills among Health-related Researchers in Japan

Akiko Nakada[1,2)*] Tomohide Ibuki[3)] Taichi Kishi[4)]
Kenji Matsui[5)] Atsushi Namiki[1,4)] and Shimon Tashiro[2)]

[1)]Department of Medical Education, Toho University Graduate School of Medicine
[2)]Center for Research Administration and Support, National Cancer Center
[3)]Liberal Arts Department, Faculty of Science and Technology, Tokyo University of Science
[4)]Center for Medical Education, Faculty of Medicine, Toho University
[5)]Department of Research Ethics and Bioethics, National Cerebral and Cardiovascular Center

ABSTRACT

Background: To support research ethics education in Japan, we attempted to develop a tool with which to evaluate ethical decision-making skills among Japanese researchers.

Methods: We developed a questionnaire that was answered by 164 health-related researchers, who participated through an Internet research company, and graduate students admitted to the Toho University Graduate School of Medicine. The questionnaire consisted of four scenarios (32 questions total; 16 pre-test questions and 16 post-test questions), pre-existing scales used to evaluate construct validity, and questions about participant characteristics including previous experience in research ethics education. Each scenario contains eight short scenes, each of which is accompanied by a question asking participants to select two actions that they would be most likely to take. To evaluate reliability, we analyzed the Cronbach's alpha value and the Spearman's rank correlation coefficient ρ for two tests that were separated after the study.

Results: The pre-test's and post-test's Cronbach's alpha values are .861 and .873, respectively, and the Wilcoxon signed-rank test showed no significant difference between the two tests ($p=.085$, N.S.). There was a significant correlation between ethical decision-making skills and research ethics education (number of times; $p=.417$, $p<.0001$). Existing scales were used to analyze the construct validity and four of the six scales showed a significant correlation.

Conclusions: Research ethics education increases the scores on the scale constructed in this study, demonstrating its validity. Although its construct validity may require further verification, the constructed scale is highly reliable and has equivalent pre-test and post-test values.

Toho J Med 4 (1): 25–34, 2018

1, 4) 5–21–16 Omorinishi, Ota, Tokyo 143–8540, Japan
2) 5–1–1 Tsukiji, Chuo, Tokyo 104–0045, Japan
3) 2641 Yamazaki, Noda, Chiba 278–8510, Japan
5) 5–7–1 Fujishiro, Suita, Osaka 565–8565, Japan
*Corresponding Author: tel: 03–3762–4151
e-mail: nakada.toho@gmail.com
DOI: 10.14994/tohojmed.2017-012
Received Oct. 28, 2017: Accepted Dec. 15, 2017
Toho Journal of Medicine 4 (1), Mar. 1, 2018.
ISSN 2189–1990, CODEN: TJMOA2

KEYWORDS: ethical decision-making, reliability, validity, measure, Japanese researchers

Introduction

Since the 1980s, research misconduct has become a social problem in developed countries. To address this problem, each country has sought to improve its ethics education system for researchers. For example, in the United States, researchers applying for grants from funding agencies and organizations are required to take short-term courses in research ethics, and the National Institutes of Health (NIH) recommends nine areas of ethical concern as topics for instruction in the responsible conduct of research (2009 edition)[1, 2]: a) conflict of interest—personal, professional, and financial—; b) policies regarding human subjects, live vertebrate animal subjects in research, and safe laboratory practices; c) mentor/mentee responsibilities and relationships; d) collaborative research including collaborations with industry; e) peer review; f) data acquisition and laboratory tools, management, sharing and ownership; g) research misconduct and policies for handling misconduct; h) responsible authorship and publication; i) the scientist as a responsible member of the society, contemporary ethical issues in biomedical research, and the environmental and societal impacts of scientific research. Similarly, in Japan, major funding agencies and organizations have developed standard textbooks and e-learning systems and require their grant applicants to receive research ethics education.[3–5]

However, no study has carefully investigated the effect of ethics education programs. Thus, whether research ethics education can prevent research misconduct is unclear. In Japan, the effect of ethical education is evaluated using simple knowledge tests, but it is unknown whether researchers can make proper ethical decisions based on knowledge acquired in ethical education programs. This is partly because, beyond simple knowledge tests, no scale has been developed to measure the effect of research ethics education.

In contrast, funding agencies and organizations in the United States have emphasized ethics skills and attitudes, in addition to knowledge, and have actively developed rating scales.[6] When testing the effect of research ethics education, ethical problem solving is the first area to be evaluated because researchers require ethical problem-solving skills to apply their knowledge in their fields. In fact, recent studies have developed two measures: the Ethical Decision-Making Measure[7] and its improved version, the Professional Decision-making in Research (PDR) measure.[8]

However, these measures were developed with American researchers in mind and may be unsuitable for Japanese researchers, with different environmental and educational backgrounds. For example, when the United States government was about to regulate conflicts of interest, universities opposed the regulation, successfully putting it on hold. At the same time, academic societies voluntarily developed guidelines for regulating conflicts of interest and urged universities to establish regulations and policies in line with those guidelines. However, this type of aggressive exchange between authorities and universities does not occur in Japan.[9] In addition, the United States embraces diversity even for universities, allowing them to establish their own policies.[9] In contrast, Japanese universities tend to seek standard policies. A similar trend is observed with respect to university faculty positions. While tenured positions are common in the United States,[9] non-tenure-track positions, including those with term limits or for specified projects, have been increasing in Japan.[10] Accordingly, it is not practical to adapt decision-making measures developed in the United States to Japan because of the marked differences in the research environments and university authorities/policies in the two cultures.

Therefore, in this study, to support research ethics education in Japan, we used the PDR measure proposed by DuBois et al. (2016) as a reference with which to develop a tool to evaluate ethical decision-making skills among Japanese researchers.

Methods

Development of a questionnaire form

We developed a questionnaire form consisting of four scenarios, existing scales used to evaluate construct validity, and questions about participant characteristics, including experience in research ethics education. When using the form as a rating scale, scenarios 1 and 2 (16 questions in total) and scenarios 3 and 4 (16 questions in total) are combined and used as a pretest and post-test, respectively. Accordingly, when developing the form, scenarios were developed to cover all nine areas of ethical concern proposed by the NIH in each test, each consisting of two sce-

Questionnaire item 1. Read the scenario and answer the corresponding question.

Scenario 1
You are an associate professor in the college of pharmacy and currently conduct research in collaboration with clinicians. Some antipsychotic drugs are known to increase body weight as a side effect, and you are interested in famotidine* because you think the drug suppresses weight gain. You will conduct an investigator-initiated clinical study to determine the novel actions of famotidine.
Note*: Famotidine is covered by the National Health Insurance for treatment of stomach ulcers and acute gastritis. The drug is commercially available.

Scene 1
Your research team has developed a single-blind study* to compare treatment outcomes of patients treated with the antipsychotic drug olanzapine alone and those treated with olanzapine plus famotidine. In the study, the principal investigator is Professor A, who is a clinician, and you are a co-investigator. Patients treated by collaborating physicians are actively recruited for participation in the study, but patients with renal and hepatic damage are excluded from the study due to potentially serious adverse events.
Note*: in a single-blind study, it is not possible to identify which group contains patients.

Question 1
The study began after the approval by the Ethics Committee of the college. However, in a laboratory meeting, you are informed that the number of patients enrolled as subjects is insufficient. After hearing this report, what type of proposal will you make? From the answers provided below, choose two answers that best describe your actions and check the box (□) to the left.

□a. Investigate whether the number of patients is not sufficient or only a small number of patients have consented to participate in the study.

□b. After learning that the number of patients is insufficient, you propose to change current exclusion criteria for renal and hepatic damage and to enroll patients with mild impairment as study subjects.

□c. After learning that only a small number of patients have submitted consent forms, even though the number of potential subjects is large enough, you propose to increase the number of collaborative facilities, despite the complicated process involved, to enroll more patients.

□d. Or you propose to double the reward for participation in the study.

□e. Or you instruct attending physicians to emphasize the effect of famotidine and explain to potential subjects that the drug is expected to lower the body weight.

□f. You propose that clinicians who are research collaborators gather together and review the current method of informed consent.

Fig. 1 Sample questionnaire items

narios. Each scenario comprises eight short scenes, each of which is accompanied by a question asking participants to select two actions that they are most likely to take (Fig. 1).

In accordance with the PDR measure of DuBois et al. (2016), each question has six actions, three of them ethical and three of them unethical, to choose from. When both actions selected by a participant are ethical, the individual receives one point/question, with a possible total of 16 points/test. This point allocation is thought to be appropriate not only because the format complies with the PDR format but also because of the difficulty of deciding whether participants have decision-making skills in cases where at least one of the actions they choose is unethical.

The scenarios of ethical decision-making used in the research were developed based on the nine areas of ethical concerns taught in the Responsible Conduct of Research (RCR) program of the NIH, which are also covered in a standard textbook on research ethics education entitled "For the Sound Development of Science: The Attitude of a Conscientious Scientist," compiled by the Japan Society for the Promotion of Science (JSPS) (JSPS standard textbook). This is because, unlike in the United States, Japanese funding agencies and organizations requesting research ethics education, such as JSPS and the Ministry of Education, Culture, Sports, Science and Technology, do not currently recommend the RCR education format.

Three researchers who specialize in research ethics (AN, TI, and ST) created multiple-choice questions and answers and subsequently developed a prototype ethics skills rating scale consisting of 32 questions and existing scales after exchanging opinions several times. With the help of several health-related researchers, we formulated succinct multiple-choice questions and answers, minimizing the number of questions on existing scales for the evaluation of construct validity and experience in research ethics education. The selection of existing scales will be described below.

Due to the nature of the ethics skills rating scale currently being developed, we anticipate an increase in scores before and after research ethics education, but because this is a cross-sectional study, scores may vary widely, depending on years of experience in research ethics education. Therefore, we created questions about the participants' experience/participation in (i) research, (ii) research ethics education offered in seminars or in the classroom (yes/no and number of times), (iii) research ethics e-learning programs (in Japanese) offered by the Collaborative Institutional Training Initiative (CITI), which is globally well-known, (iv) JSPS standard textbooks (three stages), and (v) The Lab: Avoiding Research Misconduct, an educational DVD developed by the Office of Research Integrity, United States Department of Health and Human Service.

To improve the prototype questionnaire form, opinions were exchanged at Expert Group Meetings (15 experts) on health-related research ethics (organized by co-author KM) on several occasions. After additional mail exchanges between several experts, feedback was incorporated into the questionnaire form for revision and correction, eventually leading to the development of a final version of the questionnaire form with proper description/expression of actions in the questions and answers.

Selection of existing scales for the evaluation of construct validity

DuBois et al. (2016) used five existing scales in the PDR measure: the Narcissistic Personality Inventory-16 (NPI-16), the Global Cynicism Scale, the Propensity to Moral Disengagement Scale, the How I Think about Research Scale for evaluating compliance disengagement, and the Crowne-Marlowe Social Desirability Scale. However, because the validity of its Japanese version has not been reported, we selected the following existing scales for the reasons stated below:

NPI: DuBois et al. used NPI-16, a shorter version of the NPI scale; however, the Japanese version was developed from the longer version, NPI-35. Therefore, to reduce the burden of participants while maintaining subordinate concept questions with proven reliability and validity, we kept "need for attention," "sense of grandeur," and "self-conviction," but excluded "leadership" and "praise for the body" from this study.[11]

Cynicism: Because we were unable to obtain the Japanese version of the Global Cynicism Scale used by DuBois et al., we used the Cynicism Questionnaire developed by Izawa et al.,[12] which is thought to define the overall negative thoughts about others, with emphasis on others' negative aspects such as callousness, dishonesty, and selfishness.

Belief about society scale: Individuals' awareness of the role of laws and regulations in society, their cooperation and collaboration with others, and the meaning of selfish behaviors and individual rights are defined as social recognition.[13] In this study, as alternatives to "moral disengagement" and "compliance disengagement" used in the PDR measure, we employed three scales to evaluate regulative belief, symbiotic belief, and selfish belief about society,[13] as these three types of belief are thought to be closely associated with actions that researchers may undertake.

A critical thinking disposition scale: We felt the five existing scales mentioned above were not sufficient to fully evaluate convergent validity. In a study by Plemmons and Kalichman[14] that investigates ethics skills in research and the goals of RCR instructors, RCR instructors selected ethical decision-making and critical thinking as the most and second most important ethics skills, respectively. Therefore, to evaluate convergent validity, we included the Critical Thinking Disposition Scale[15] in the questionnaire form.

Marlowe-Crowne social desirability was included in the PDR study's questionnaire to show subjects whether their answers are socially undesirable. However, we chose not to include the Marlowe-Crowne social desirability scale, as it is not directly involved in the evaluation of construct validity.

Subjects of questionnaire survey

Subjects were 150 health-related researchers who participated in the study through an Internet research company (an outsourcing company, Rakuten Research, Inc.) and graduate students admitted to the Toho University Graduate School of Medicine in 2016. The Internet re-

search company was instructed to build a website and collect data from website registrants, who were composed of the three groups. The first group is graduate students in the department of medicine, nursing, and pharmacy. The second group is faculty members and researchers in the department of medicine, nursing, and pharmacy who are engaging in medical research. The third group is healthcare professionals (physicians, nurses, pharmacists, and dentists) who were involved in medical research. After providing consent, subjects answered questions presented online (the online questionnaire survey began on January 18, 2017 and ended on January 20, 2017 as the expected number of responses had been reached).

A questionnaire survey of graduate students at the Toho University began on January 13, 2017, the mandatory day when courses began. Survey responses were collected on January 20, 2017, and the submission of the questionnaire form was considered after subjects' consent to participate in the study.

Analytical method

It has been reported that a single concept or construct underlies the PDR measure. Therefore, in this study, we assumed ethical decision-making skills to be one construct and analyzed it one-dimensionally by verifying the factor loadings based on significant scree plot and polychoric correlation data. A varimax rotation was also used in the analysis.

To evaluate reliability, we analyzed the overall Cronbach's alpha value as well as Cronbach's alpha and Spearman's rank correlation coefficient ρ for two tests that were separated after the study. In addition, the Wilcoxon signed-rank test, a nonparametric test for paired samples, was performed to test for a significant difference ($p<.05$) between the median values of the two tests. The mean and standard deviation values of each scenario (four scenarios in total) were calculated.

Data from our ethics skills rating scale and subject characteristics were analyzed using descriptive statistics. The correlations between scores from the ethics skills rating scale, research experience, and experience in research ethics education (number of times) were analyzed using the Spearman's rank correlation coefficient ρ, and the correlation between ethics skill scores and experience in research ethics education (yes/no) was analyzed using the Mann-Whitney U test. To analyze the correlation between ethics skill scores and study using the standard textbook, the Mann-Whitney U test was used to compare mean ethics skill scores between subjects who read the entire textbook and those who did not read the entire textbook.

Furthermore, to show whether research is supported by theory, the Spearman's rank correlation coefficient ρ was used to analyze the construct validity obtained using existing scales.

The correlation coefficient and inter-subject correlation were evaluated using the criteria of Oshio (2004): ±.00-.20: very little correlation; ±.20-.40: low correlation; ±.40-.70: moderate correlation; and ±.70-1.00: high correlation.[16] Because some values were missing from the survey responses returned from the graduate students at the Toho University, n analysis was performed by excluding the missing values. A statistical analysis was performed using SPSS software (version 24) and EasyEstimation item response theory analysis software (version 1.8.0).[17]

Ethics considerations

Subjects who participated in the study through the Internet research company and graduate students at the Toho University were informed about the voluntary nature of the survey. This study was approved by the Ethics Committee of the Toho University School of Medicine (approval No. A16085).

Results

Subject characteristics

Subject characteristics are shown in Table 1. Over 56% of the subjects were in their 40s and 50s. Healthcare professionals working in a clinical setting accounted for approximately 60% of subjects who answered research-related questions. Because there was a large difference in the lab viewing experience (four subjects answering Yes, 160 subjects answering No), this question was excluded from the analysis of research ethics education.

Evaluation of reliability and equivalence

A mean score was calculated for each question (32 in total), scenario (four in total), and test (two in total), as well as for the entire survey (Table 2). We also calculated Cronbach's alpha for the scenarios, tests, and survey and correlation coefficients for the scenarios and tests.

The Wilcoxon signed-rank test of the two tests (mean scores, 12.04 and 11.59) showed no significant difference between the tests ($p=.085$, N.S.).

Verification of the one-dimensionality of ethical decision-making skills based on polychoric correlation coefficients

The scree plot in Fig. 2 shows eigenvalues obtained

Table 1 Subject characteristics (n = 164)

Sex	n (%)
Men	112 (68.3)
Women	50 (30.5)
Not specified	2 (1.2)
Age group	**n (%)**
20s	25 (15.2)
30s	26 (15.9)
40s	41 (25.0)
50s	51 (31.1)
60s	17 (10.4)
70s	3 (1.8)
Not specified	1 (0.6)
Experience in research ethics committee	**n (%)**
Yes	32 (19.4)
No	129 (78.7)
Not specified	3 (1.8)
(1) Workshop or seminar related to ethics committee	**n (%)**
No	83 (50.6)
Yes	81 (49.4)
(2) Learned research ethics in undergraduate or graduate course, or FD	**n (%)**
Yes§	95 (57.9)
No	69 (42.1)
Neither (1) nor (2)	51 (31.1)

Research (credit: years)	14.03 ± 12.10 [0.75, 50] *
Field of specialization	**n (%)**
Medicine	85 (51.8)
Dentistry	11 (6.7)
Nursing	22 (13.4)
Pharmacy	29 (17.7)
Other	17 (10.4)
Research field	**n (%)**
General education	4 (2.4)
Basic	53 (32.3)
Clinical	98 (59.8)
Other	8 (4.9)
Not specified	1 (0.6)
CITI Japan course	**n (%)**
Yes	50 (30.5)
No	113 (68.9)
Not specified	1 (0.6)
Standard textbook-based learning	**n (%)**
Read completely	23 (14.0)
Read partly	53 (32.3)
Unread	88 (53.7)

§, One academic hour is counted as 1 hour. *, Mean ± SD, [minimum, maximum]

from all perspectives to verify the one-dimensionality of ethical decision-making skills. Because the first factor explained ≥50% of all perspectives, we concluded that a single-factor structure was appropriate and, thus, the requirement for one-dimensionality was satisfied.

Correlation between existing scales and research ethics education

Existing scales were used to analyze the construct validity, and four of the six scales showed a significant correlation (Table 3). We investigated factors that influence ethical decision-making skills to verify our hypothesis that ethical decision-making skills can be improved through research experience or through research ethics education seminars and courses (number of times or yes/no). Although no significant correlation was observed with research experience (ρ=.002, p=.981, N.S.), there was a significant correlation between ethical decision-making skills and research ethics education (number of times; ρ=.417, $p<.0001$) (Table 4).

Discussion

In this study, we developed an ethics skills rating scale consisting of one test with 16 questions about decision-making skills among researchers. The Cronbach's alpha values obtained in the pre- and post-tests suggested that the reliability and internal consistency of the scale are both high. In addition, there was no significant difference between the mean scores of the two tests, and the results of the two tests had a high Spearman's rank correlation coefficient, suggesting that the two tests are equivalent.

In addition, the scores varied depending on the participants' backgrounds in research ethics education. Therefore, it is quite reasonable to assume that the ethics skills ranking scale measures the ethical decision-making skills of researchers. Because the pretest had a higher mean score in this cross-sectional study, it may be necessary to investigate whether research ethics education actually increases scores for ethical decision-making skills, with the aim of improving the validity of our ethics skills rating

Table 2 Descriptive statistics for individual and combined questions, scenarios, and tests.

	No.	Mean ± SD	Analysis by scenario Mean ± SD [minimum, maximum] Cronbach's alpha	Analysis by test Mean ± SD [minimum, maximum] Cronbach's alpha	Analysis of all factors Mean ± SD [minimum, maximum] Cronbach's alpha
Scenario 1	1	.79 ± .411			
	2	.55 ± .499			
	3	.66 ± .474			
	4	.83 ± .377	5.76 ± 2.11 [0, 8] .746		
	5	.66 ± .474			
	6	.65 ± .478			
	7	.80 ± .402		12.04 ± 3.87 [1, 16] .861	
	8	.82 ± .388			
Scenario 2	9	.82 ± .388			
	10	.81 ± .393			
	11	.80 ± .398			
	12	.85 ± .355	6.27 ± 2.07 [0, 8] .787		
	13	.83 ± .377		Wilcoxon signed-rank test of paired samples, p = .085 Spearman's rank correlation coefficient, p = .734 **	23.61 ± 7.58 [4, 32] .925
	14	.76 ± .427			
	15	.66 ± .476			
	16	.74 ± .441			
Scenario 3	17	.82 ± .383			
	18	.63 ± .483			
	19	.62 ± .488			
	20	.59 ± .493	5.71 ± 2.26 [0, 8] .789		
	21	.77 ± .419			
	22	.70 ± .462			
	23	.82 ± .388			
	24	.75 ± .435		11.59 ± 4.11 [1, 16] .873	
Scenario 4	25	.80 ± .403			
	26	.77 ± .420			
	27	.71 ± .456			
	28	.57 ± .496	5.88 ± 2.11 [0, 8] .763		
	29	.88 ± .321			
	30	.82 ± .388			
	31	.76 ± .431			
	32	.57 ± .497			

Scenarios 1 and 2, n = 164; scenarios 3 and 4, n = 163

**, p < .01.

In order to avoid making some model answer arbitrarily, we will not present the whole scenario and questions in this paper but disclose them only to applicants who wish to use them. If you would like to use this measure, please contact the Correspondence Author.

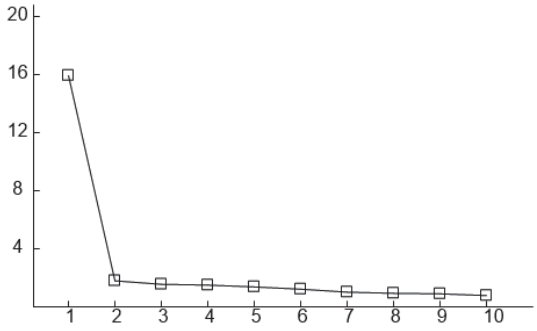

Fig. 2 Scree plot

scale further.

On the other hand, it is of significant interest that experience in research or on an ethics committee is unlikely to affect scores on the ethics skills rating scale. In Japan, funding agencies and organizations began to require research ethics education for grant applications in 2015. This suggests that we may not observe any correlation between research experience and ethical decision-making skills in this study. Indeed, the need to provide research ethics education not only to young researchers but also to

Table 3 Association with existing scales (evaluation of construct validity)

Existing scales	Spearman's rank correlation coefficient	p-value
Narcissistic Personality Inventory (need for attention, sense of grandeur, and self-conviction)	− .235 **	.003
Global cynicism scale	− .051	.518
Regulative belief about society scale	− .065	.409
Symbiotic belief about society scale	.288 **	< .0001
Selfish belief about society scale	− .175 *	.025
Critical Thinking Disposition Scale	.322 **	< .0001

*, p< .05: **, p< .01; n = 163

Table 4 Association between scores from a skill checklist and experiences in research ethics education

Item	Correlation with scores from a skill checklist (Spearman's rank correlation coefficient)	p-value
Research experience	.002	.981
How many times did you enroll in a research ethics education program for ethics committee members, undergraduate/graduate course, or FD?	.417 **	< .0001

Item		n (%)	Mean of the sum of skill scores	p-value
Have you studied research ethics in the education program for ethics committee members, undergraduate/graduate course, or FD?	Yes	112 (68.7)	25.19	< .0001
	No	51 (31.3)	20.13	
CITI Japan course§	Yes	49 (30.2)	26.37	.001
	No	113 (69.8)	22.42	
Textbook-based learning	Read	23 (20.9)	25.65	.089
	Unread	87 (79.1)	22.80	

n = 163; §, n = 162.

senior researchers who have had no previous official research ethics education either as students or research fellows has been advocated in recent years,[18,19] supporting the findings of this study.

In Japan, it is pointed out that ethics committee members may not have had adequate research ethics education.[20,21] In addition, as ethical decision-making skills required by ethics committee members may not match ethical decision-making skills required for conducting responsible research, it is not surprising that they may lack ethical decision-making skills. The newly developed measure appears to be useful for the evaluation of ethics training at the Ethical Review Board.

Not all of the existing scales used in this study were correlated with construct validity. Among existing scales exhibiting a correlation with the present ethics skills rating scale, NPI-16 had a low negative correlation, suggesting discriminant validity. Although the association between our ethics skills rating scale and the Selfish belief about society scale was judged loose, the value was closed to 0.2, meaning to have "low correlation." In addition, the Symbiotic belief about society scale and the Critical Thinking Disposition Scale both had positive correlations, suggesting convergent validity. The relatively low correlation between our scale and these four scales obtained in the research are thought to be appropriate because the concept that these four scales measure does not match the concept of ethical decision-making skills. We believe that, combined, these results support the validity of the construct of our scale.

Unlike the PDR measure, the ethics skills rating scale developed in this study is not correlated with the Global

Cynicism Scale. Izawa et al. developed the Global Cynicism Scale, aiming to measure cynicism associated with hostility, and it is the only Global Cynicism Scale in Japanese. Additionally, based on the mean score of each sample showing opposite results from those in previous studies, Izawa et al. stated that hostility may not be apparent among the members of the Japanese society and that it is important to consider the cultural differences between Japan and other countries.[13] Moreover, multiple regression analysis revealed that the PDR measure had small loading values and did not remain as a factor, suggesting that the construct validity of our ethics skills rating scale is supported.

The regulative belief about society scale was built upon the concept that strict regulations and laws are needed to make society a nice place to live, and subjects with high scores on this scale are thought to be positive about regulations. However, although we used the scale to obtain correlations based on the assumption that feeling positive about regulations is the same as being aware of compliance issues, no correlation was observed in this study. Accordingly, we believe this scale was not appropriate for assessing the construct validity of our ethics skills rating scale. In contrast, there was a strong correlation with the Critical Thinking Disposition Scale, suggesting that having skills for ethical decision-making is different from simply being aware of research compliance issues.

On the other hand, actual research ethics education offered at research institutions and universities in Japan is largely provided in the form of lectures, teaching laws, and guidelines about research over a short time and does not foster the ability to think independently as researchers. In addition, the ethical standards set by Japanese research institutions are often related to compliance with the rules regarding proper use of research funding or compliance with the laws and related regulations.[22] According to Nakagomi et al. (2013), students and faculty members in the graduate school of nursing receive approximately two academic hours of research ethics education during admission orientation or in a lecture.[20] Indeed, skill ratings had a strong correlation with critical thinking disposition but not compliance awareness, suggesting that it may be time to review the current education system to establish a more practical system in the future.

Study limitations

The survey respondents in this study are believed to include not only professional researchers but also healthcare providers who are involved in research. Therefore, one of the limitations of this study is an inability to reveal associations with research experience that show deep involvement in research activity as a job. Additionally, this study did not compare the effect of face-to-face research ethics education in improving ethical decision-making skills with that of the CITI Japan courses. Therefore, further study is needed to establish effective educational approaches, taking into account that this scale will take about 30 min per test.

Conclusion

The evaluation tool developed in this study is a Japanese scale for assessing skills related to behavioral choices among health-related researchers, i.e., ethical decision-making in research. Although it may be necessary to verify its construct validity, the scale is highly reliable and has equivalent pre-test and post-test statistics. Research ethics education increases scores on the scale, showing the scale's validity.

Because the scale developed in this study is for use by health-related researchers, future challenges include the development of a highly versatile evaluation tool that can be used in studies with human as well as non-human subjects.

Acknowledgements: The authors acknowledge Ms. Kayo Takashima (the Institute of Medical Science, University of Tokyo) and Ms. Kuniko Aizawa (National Cerebral and Cardiovascular Center) for their advice on our questionnaire.

Funding: Financial support for this study was provided by the National Cancer Center and the Department of Medical Education, Toho University Graduate School of Medicine, in 2016.

Conflicts of interest: The authors have no conflicts of interest to disclose.

References

1) NIH homepage [Internet]. Update on the requirement for instruction in the responsible conduct of research. Notice Number: NOT-OD-10-019. 2009 [cited 2016 Sept 23]. Available from: https://grants.nih.gov/grants/guide/notice-files/NOT-OD-10-019.html
2) Isobe T. A comparative study of the regulatory and legal systems of clinical research in Western countries and Japan. Comprehensive research report for fiscal year 2013-2014 2015. p. 29. Japanese.
3) Japan Society for the Promotion of Science [Internet]. JSPS elec-

tronic application system for grants-in-aid for scientific research: points for verification and pledge regarding the maintenance of propriety in research activities and appropriate use of research grants [cited 2017 Aug 30]. Available from: http://www.jsps.go.jp/j-grantsinaid/15_hand/data/h28/kakenhi_checklist_e.pdf
4) Japan Science and Technology Agency [Internet]. Enrolling in educational programs on research integrity [cited 2017 Aug 31]. Available from: http://senryaku.jst.go.jp/teian/en/top/juko.html
5) Japan Agency for Medical Research and Development [Internet]. Research ethics education program [cited 2017 May 11]. Available from: http://www.amed.go.jp/kenkyu_kousei/kyoiku_program.html. Japanese.
6) Antes A, Dubois J. Aligning objectives and assessment in responsible conduct of research instruction. Elliott SL, Fischer BA, Grinnell F, Zigmond MJ (Eds) Perspectives on research integrity Washington, DC: ASM Press; 2015. p. 34-48.
7) Mumford MD, Devenport LD, Brown RP, Connelly S, Murphy ST, Hill JH, et al. Validation of ethical decision making measures: evidence for a new set of measures. Ethics Behav. 2006; 16: 319-45.
8) DuBois JM, Chibnall JT, Tait RC, Vander Wal JS, Baldwin KA, Antes AL, et al. Professional decision-making in research (PDR): the validity of a new measure. Sci Eng Ethics. 2016; 22: 391-416.
9) Miyata Y. [University-Industry collaboration and academic integrity in the United States] Tokyo: Tamagawa University Press; 2013. Japanese.
10) Ishii M. [Chapter 4: University faculty: professionals without training plans]. Hashimoto K (Ed) [Japanese Model in Professional Education] Tokyo: Tamagawa University Press; 2009. p. 64-83. Japanese.
11) Konishi M, Okawa M, Hashimoto T. [Development of narcissistic personality inventory (NPI-35)]. Japan Society of Personality Psychology. 2006; 14: 214-26. Japanese.
12) Izawa S, Nomura S. [Development and validation of the cynicism questionnaire]. Japanese Journal of Behavioral Medicine. 2004; 10: 66-72. Japanese.
13) Yoshida T, Ando N, Motoyoshi T, Fujita T, Hirooka S, Saito K, et al. [Study of Social Annoyance] (1). Bulletin of the School of Education. Nagoya University. 1999; 46: 53-73. Japanese.
14) Plemmons DK, Kalichman MW. Reported goals of instructors of responsible conduct of research for teaching of skills. J Empir Res Hum Res Ethics. 2013; 8: 95-103.
15) Hirayama R, Kusumi T. [Effect of critical thinking disposition on interpretation of controversial issues: evaluating evidence and drawing conclusions]. Japanese Journal of Educational Psychology. 2004; 52: 186-98. Japanese.
16) Oshio A. [Psychometrics and survey data analysis using SPSS and Amos: up to factor analysis and covariance structure analysis] Tokyo: Tokyo Tosho; 2004. Japanese.
17) Kumagai R [Internet]. [Page for item response theory and easy estimation] [cited 2017 April 25]. Available from: http://irtanalysis.main.jp/index.php. Japanese.
18) Yamazaki S. [It's time to think about authorship]. Journal of Information Processing and Management [Internet]. 2013 [cited 2017 Aug 30]; 56: 636-9. Available from: https://www.jstage.jst.go.jp/article/johokanri/56/9/56_636/_pdf. Japanese.
19) Kuroki T. [Research misconduct: fabrication, falsification, and plagiarism by scientists] Tokyo: Chuokoron Shinsha; 2016. Japanese.
20) Nakagomi S, Komatsu H, Nawa H, Yamada S, Katada N, Ota K, et al. [Survey report on nursing research ethics review systems in research institutes]. Journal of Japan Academy of Nursing Science. 2013; 32: 45-52. Japanese.
21) Muto K, Sato K, Shirai Y. [Seven suggestions for reforming research ethics review committees]. Journal of the Japan Association for Bioethics. 2005; 15: 28-34. Japanese.
22) Hada T, Tateishi S. [Chapter 8: Academic honesty viewed from national survey]. In: [Promoting research integrity in cross-national trends and challenges]. Institute for excellence in Higher education, Tohoku University. Sendai: Tohoku University Press; 2015. p. 153-75. Japanese.