Kawasaki Journal of Medical Welfare Vol. 27, No. 1, 2021 13－27

Review Article

Predictors of Success in the National Therapy Licensure Examination in Japan: A Systematic Review

Daichi IIMURA[*1], Shu HARAYAMA[*1], Shinsuke NAGAMI[*1], Rie UETANI[*2], Kaori NAKATA[*2], Noritaka HATTA[*1] and Masashi SHIOMI[*1]

(Accepted April 23, 2021)

Key words: national examination, therapist, predictors of success, systematic review, Japan

Abstract

The purpose of the study is to synthesize the published research articles about the predictors of performance in the national examination of therapy licensure in Japan. A systematic review of the literature was carried out on November 14, 2020. We searched for articles containing the words "national examination" and "point" or "performance" or "scoring" or "factor" or "impact" or "pass/fail" in two Japanese databases and the words "national examination" and "therapist" or "therapy" in three English databases. We screened the literature about the national examinations for therapists-physical, occupational, and speech-in Japan. We examined the statistical relationship between the outcome and independent variables. Twenty-three articles met all our inclusion criteria. Analysis of those articles revealed that the variables that significantly affected the outcome were: scores on tests during college, scores on mock tests, and grade point averages. Most of these papers also identified the variables that did not significantly affect the outcome, that is, scores on tests immediately after the entrance, scores on admission examinations, or admission categories. Some of these inconsistencies were explained by the time at which the test was performed, which subjects were analyzed, or the statistical tests that were used.

1. Introduction

In Japan, the number of licensed physical therapists (PT), occupational therapists (OT), and speech therapists or speech-language-hearing therapists (ST), who are rehabilitation specialists, has been increasing and was estimated at over 300,000 people in 2020[1-3]. To obtain a license, one must pass a national examination held once a year, and currently, the number of higher education institutions that train these students is increasing. Additionally, the students entering these institutions are diversifying due to the decline of the birth rate and the increase in the percentage of students going on to higher education institutions[4]. There is a concern about the increase of prospective students whose academic

[*1] Department of Speech-Language pathology and Audiology, Faculty of Rehabilitation
Kawasaki University of Medical Welfare, 288 Matsushima, Kurashiki, 701-0193, Japan
E-Mail: iimura@mw.kawasaki-m.ac.jp
[*2] Department of Sensory Science, Faculty of Health Science and Technology
Kawasaki University of Medical Welfare

ability is of a lower standard or who are unable to finish the training courses. For students to acquire professional knowledge and skills to become therapists, and to achieve success in the national therapy licensure examination, it could be beneficial to consider what factors contribute to success in the national examination. The purpose of this systematic review is to identify and synthesize the published research articles on the variables that predict performance in the national therapy licensure examination in Japan.

2. Methods

2.1 Protocol

This systematic literature review followed the PRISMA guidelines[5] and included any article that reported on the students' performance in the national therapy licensure examination and related factors that endeavored to describe or analyze this performance. All of the material was published by November 14, 2020.

2.2 Article inclusion

To minimize the article selection bias and to provide a clear determination of those articles to be included in this review, the following criteria were established before the literature search: (a) That it is about the national examination for a therapy license-PT, OT, or ST-in Japan; (b) the outcomes measured are the scoring or pass/fail of the participants in the national examination, or a test that is not a national examination but is designed to simulate the actual national examination, that is, the mock test that students take right before the national examination; (c) that there is a statistical analysis, for example, t-test, analysis of variance (ANOVA), Mann-Whitney U test, or Kruskal-Wallis test to compare the mean or median of two or more groups; chi-square or Fisher's exact test to compare the proportion of participants in some groups, that is, examination-pass or examination-fail groups, correlation analysis of Pearson or Spearman to reveal the relationships between variables, or regression analysis to predict factors that affect the scoring or pass/fail status in the national examination; (d) the original practice is not specified as the independent variable, that is, the article does not aim to clarify the original practice (curriculum or special lectures for the national examination) that lead to success in the national examination; (e) independent variables were measured before the outcome; and (f) the outcome was reported by explaining the relationship of outcome and independent variable. In this review, articles were excluded if they did not mention the national therapy licensure examination in the title or abstract, were a dissertation, or were a conference paper, review, or commentary.

2.3 Information sources

Since no single database is likely to contain every relevant publication, articles were identified using multiple databases. The Japan Medical Abstract Society Database (JAMAS; 1959-present; https://www.jamas.or.jp) and CiNii Articles (1948-present; https://ci.nii.ac.jp) were utilized for searching the Japanese literature. For searching the English literature, PubMed (1966-present; https://pubmed.ncbi.nlm.nih.gov), ERIC (1966-present; https://eric.ed.gov), and Web of Science (1900-present; https://apps.webofknowledge.com) were used to identify all relevant publications. Additionally, reference lists from manual literature searches that met the inclusion criteria were searched for additional citations. The last search was conducted on November 14, 2020.

2.4 Search parameters

We identified relevant articles in the two Japanese databases by searching items that contained the words "national examination (kokkashiken)" and terms referring to predictors of success on the national examination: "point (tensuu)" or "performance (seiseki)" or "scoring (tokuten)" or "factor (youin)" or "impact (eikyou)" or "pass/fail (gouhi)." In the English database, we searched the term "national examination" and "therapist" or "therapy." These search parameters were decided on with the agreement of all authors.

2.5 Article selection

Following the flowchart for the review process by Minds[6] and PRISMA[5], the inclusion of articles was achieved by implementing a two-stage process. Stage 1: Screening of the citation, including the title and the abstract, and Stage 2: Screening of the full-text manuscript. In stage 1, the citations and abstracts obtained during the individual database searches were evaluated independently by the first and second authors and were judged to be potentially eligible to advance to Stage 2 of the inclusion process if they met six inclusion criteria and they did not meet three exclusion criteria. If there was a lack of clarity about the exclusion of an article, it was advanced to Stage 2 for checking of the full-text manuscript. Articles that were advanced to Stage 2, in which the full text of the article was obtained, were evaluated independently by the first and second authors. To be included in the final review, an article had to meet all five of the inclusion criteria. The first and second authors carried out the Stage 1 and Stage 2 processes independently. In the event of a disagreement, or if a question regarding the inclusion of a particular article at either stage arose, these inconsistencies were resolved and modified by discussion. Finally, articles were identified through the systematic review processes in Stage 1 and Stage 2.

2.6 Data extraction

The process of data extraction involved the coding for each article across six primary dimensions: (1) number of participants (2) type of national therapy licensure examination (3) statistical analysis (4) dependent variables (5) independent variables that significantly affected the outcomes and (6) independent variables that did not significantly affect the outcomes. Dependent variables included outcome variables, that is, the scoring or pass/fail result in the national examination or mock test. Independent variables included variables that significantly affect the outcomes and variables that do not significantly affect the outcomes. The coding of all included articles was conducted by the first author, and all authors verified all coding. Any discrepancy in coding was resolved by discussion.

2.7 Data analysis

Findings were reported by tabulating characteristics of the included articles, for example, design, participants, independent variables, and dependent variables. The included articles that met all six criteria were included for analysis.

2.8 Risk of bias

The PRISMA statement recommends the inclusion of systematic assessments for the risk of bias within articles[6]. Hence, the Risk of Bias Assessment Tool for Nonrandomized Studies (RoBANS)[7] was used independently to evaluate the non-randomized studies by the first and second authors. Risk of bias is assessed as a judgment (high, low, or unclear) for an individual element based on the answer to six separate criteria: the selection of participants (selection bias), confounding variables (selection bias), measurement of exposure (performance bias), blinding of outcome assessments (detection bias), incomplete outcome data (attrition bias), and selective outcome reporting (reporting bias)[7]. More rigorous studies are likely to have a lower risk of bias, and yield results that are closer to the truth[8].

3. Results

3.1 Information retrieved for included and excluded studies

Five electronic databases were searched for potential studies to include in this review. The electronic search yielded 541 citations across the following databases: JAMAS (n = 244), CiNii Articles (n = 239), PubMed (n = 38), ERIC (n = 13), and Web of Science (n = 7). After removing articles that duplicate indexing across databases, 426 titles and abstracts were evaluated for advancement to a full-text retrieval stage based on the inclusion criteria described earlier. Additionally, two further articles were identified through a manual bibliographic search, which met the eligibility criteria and were included in the reviews. Of these

428 articles, 391 were excluded because the title and abstract did not match the five inclusion criteria (1st screening). The full texts of the 37 articles were examined and identified following our inclusion and exclusion criteria. Based on these criteria, another 14 articles were excluded leaving a total of 23 articles (2nd screening). Finally, 23 articles[9-31] were identified through Stage 1 and Stage 2. The flowchart in Figure 1 depicts the identification and selection of the articles for this systematic review.

3.2 Summary of included studies

The summary of included studies is shown in Table 1, showing each study's reference, number of participants (N), type of national therapy licensure (PT/OT/ST), statistical analysis, dependent variables (outcomes), variables that significantly affected the outcomes, and variables that did not significantly affect the outcomes.

The number of participants ranged from 18 to 447 among these articles. Of all articles, 21 articles included participants from the PT course/division/department, eight articles included participants from the OT course/division/department, and four articles included participants from the ST course/division/department.

Among these studies, correlation analyses were frequently used; 17 articles used this analysis; seven were Pearson, six were Spearman, and four were unclear. 15 analyses were found that compared the mean or median of two or more groups; seven were t-test, five were Mann-Whiney, two were ANOVA, two were Kruskal-Wallis, and one was a Wilcoxson Signed Rank test. Four analyses were found about the comparison of the proportion of participants in some groups; two were chi-square, and two were Fisher's exact test. Five analyses were also found about the regression analysis: two were multiple regression, two were logistic regression, and one was simple regression analysis. Of these, four articles used receiver operating characteristic (ROC) analysis. In sum, 14 out of 23 studies utilized multiple methods of analysis.

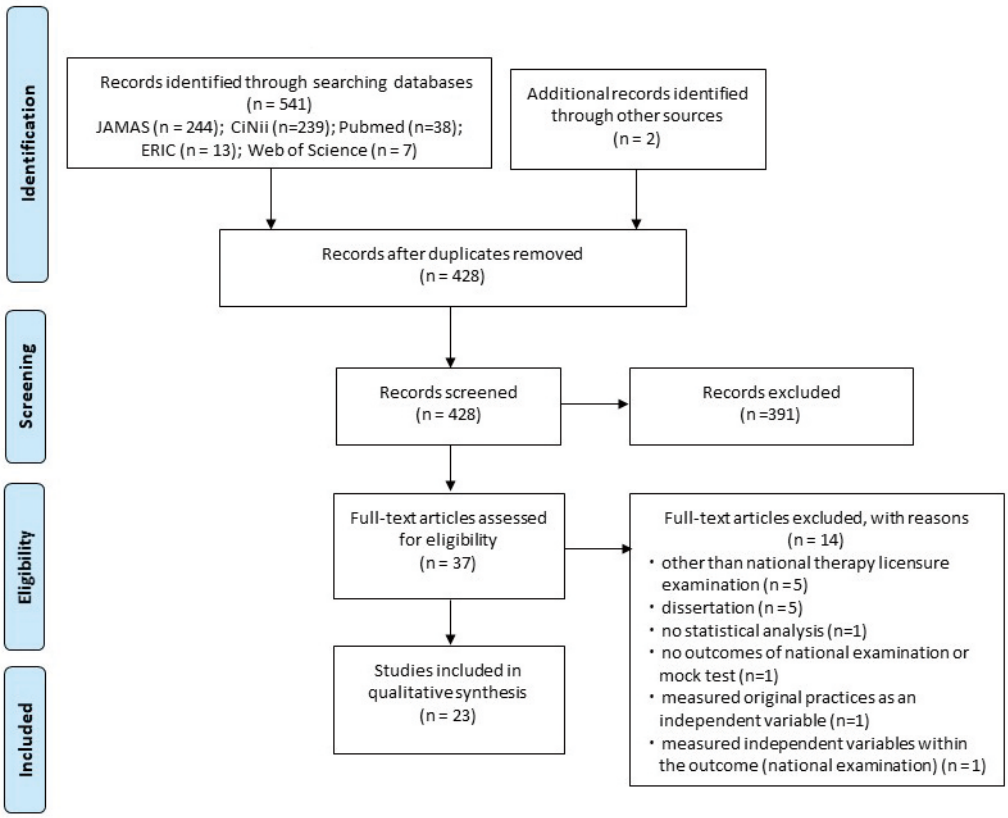

Figure 1 PRISMA flow chart

Table 1 Summary of included studies

Paper	N	PT/OT/ST	Statistical analysis	Outcomes	Variables that significantly affected the outcomes (statistical values) [a]	Variables that did not significantly affect the outcomes (statistical values) [a]
Yanagisawa et al. (2000) [b]	185 (PT), 206 (OT)	PT, OT	C, FX, TT, SR	N-PF, N-S	7 out of 19 grades in subjects (PT course) ($p < .05$); 12 out of 15 grades in subjects (OT course) ($p < .05$)	Admission categories [b] Scores on admission examinations (PT; English: $r = .25$, Science: $r = .21$, Japanese: $r = .08$, Math: $r = .12$) 12 out of 19 grades in subjects (PT course); 3 out of 15 grades in subjects (OT course)
Motooka et al. (2003)	150 (PT), 146 (OT), 151 (ST)	PT, OT, ST	TT	N-PF	Grade in fundamental subjects (PT course) ($p < .05$) Grade in specialized subjects (PT/ST course) ($p < .05$) Grade in all subjects (PT/ST course) ($p < .05$)	Grade in fundamental subjects (ST course)
Tsubota et al. (2011)	59	PT	PC	MT	Scores on periodic tests of Physiology (first year) ($r = .66$, $p < .001$)	
Kochi et al. (2012)	70	PT	PC	MT	Scores on informal memory tests (first year) ($r = .44$, $p < .01$)	Scores on informal calculation tests (first year) ($r = .21$)
Murata et al. (2012)	270	PT, OT, ST	C, CS	N-S	Scores on admission examinations ($r = 0.34$, $p < .05$) Timing of the admission examinations ($\chi^2(2) = 1.21 \sim 7.71$, $ps = .02 \sim .55$) [c]	Admission categories ($\chi^2(5) = 2.44 \sim 8.09$, $ps = .08$-$.66$)
Akagi et al. (2013)	176	PT	SC, FX, TT, MW	N-PF, N-S	GPAs of REE subjects (fundamental subjects, specialized subjects, including first-, second-, third-, and fourth-year students) ($rs = .196 \sim .535$, $p < .05$) GPAs of GEE subjects (fundamental subjects, specialized subjects, including first-, second-, third-, and fourth-year students) ($rs = .279 \sim .523$, $p < .05$) Admission categories to N-S ($p <. 05$)	Admission categories to N-PF ($\chi^2(1) = 0.105$, $p = .746$, $\phi = 0.024$)
Murao (2014)	30	PT	SC	N-S	Grade in specialized subjects (16 out of 61 [c] subjects) ($rs = .38 \sim .57$, $p < .05$)	Grade in specialized subjects (33 out of 61 [c] subjects).
Ashiga et al. (2015)	18	ST	PC	N-S	Scores on tests immediately after entrance ($r = .824$) Grade in subjects ($r = .759$) Grade in clinical practice ($r = .665$)	
Goto et al. (2015)	44	ST	PC, MW	N-PF, N-S	GPAs ($r = .37$, $p = .016$ in PC; $Z = -2.058$, $p = .040$ in MW) Scores on MTs (in Nov and Jan of the fourth year) ($rs = .42 \sim 58$, $ps < .001 \sim .006$ in PC; $Zs = -2.479 \sim -3.879$, $p < .001 \sim .012$ in MW)	Scores on MTs (in Sep, Oct, and Dec of fourth-year students)
Inomoto et al. (2015)	57	PT	PC, MR	N-S	Scores on MTs (21 out of 24 subjects, $rs = .284 \sim .714$, $p < .05$) MR: significantly described by Kinesiology ($p < .001$), Cerebral Vascular Disease ($p = .005$), and Psychology ($p = .044$) ($R^2 = .62$)	Scores on MTs (3 out of 24 subjects, $rs > .05$) MR: not significantly described by 18 out of 21 subjects of MTs

Kubo et al. (2015)	98	PT	MW, WS	N-S	Levels of satisfaction with learning conditions (males) ($p < .05$).	Levels of satisfaction with living conditions (males) Levels of satisfaction with learning and living conditions (females) Levels of satisfaction with learning and living conditions (all participants)
Murao (2015)	53	PT	SC, MW, KW	N-S		Admission categories ($p = .45$) Scores on admission examinations ($rs = .02$〜$.23$, $ps = .50$〜$.98$)
Otaka et al. (2017)	214	PT	C, TT	N-PF	Scores on tests immediately after entrance (Japanese) ($t(212) = 3.422$, $p < .01$) GPAs ($t(212) = 1.92$〜3.07, $p < .01$) MTs immediately before the national exam (fundamental subjects: Anatomy, Kinesiology, and Physiology) ($t(212) = 3.757$, $p < .01$).	Admission categories ($\chi^2(5) = 10.16$, $p > .05$) Deviation values of high school scores($t(212) = 0.765$)
Inomoto et al. (2017)	78	PT	SC, MR	N-S	Average of five MTs (med Oct, med Nov, early Dec, early Jan, and late Jan (Anatomy: $r = .574$, $p < .01$; Physiology: $r = .552$, $p < .01$; Kinesiology: $r = .546$, $p < .01$) MR: significantly described by the MTs in late Jan ($p < .001$, $R^2 = .657$)	MR: not significantly described by MTs in mid-Oct, med Nov, early Dec, and early Jan. ($ps = .131$〜$.774$)
Murao et al. (2017)	93	PT	SC	N-S	GPAs of second- or third-year students ($rs = .217$〜$.465$, $ps < .001$〜$.040$)	GPAs of first- or fourth-year students ($rs = -.085$〜$.192$, $ps = .066$〜$.502$)
Kitamura et al. (2018)	305 (PT), 89 (OT)	PT, OT	PC, TT, LL, ROC	N-PF, N-S	Scores on tests immediately after entrance by PC or TT (PT: $r = .32$, $p < .01$ (N-S) and $p = .016$ (N-PF); OT: $r = .21$, $p < .01$ (N-S) and $p < .001$ (N-PF)) Scores on periodic tests of first- and second-year students by PC or TT (PT: $rs = .53$〜59, $p < .01$ (N-S) and $p < .001$ (N-PF); OT: $rs = .61$〜70, $p < .05$ (N-S) and $p < .001$ (N-PF)) Scores on MTs in early Dec, early Jan, and early Feb in the final year by PC (PT: $rs = .57$〜71, $p < .01$ (N-S) and $p < .001$ (N-PF); OT: $rs = .70$〜78, $p < .01$ (N-S) and $p < .001$ (N-PF)) Scores on MTs in early Dec and early Feb in the final year to N-PF by LL (PT) (OR = 0.92-0.94, 95%CI [0.91-1.00], $ps < .001$〜$.028$) Cutoff: 90.5 (Scores on MTs in early Dec of final-year students) (sensitivity = .638, specificity = .816, AUC = .821) Cutoff: 30.5 (Scores on MTs in early Feb of final-year students) (sensitivity = .700, specificity = .880, AUC = .863)	Scores on tests immediately after entrance by LL (PT) (OR = 1.01, 95%CI [0.99-1.03], $p = .186$) Scores on periodic tests of first- and second-year students by LL (PT) (OR = 0.98-1.01, 95%CI [0.94-1.09], $ps = .158$-691) Scores on MTs in early Jan in the final year to N-PF by LL (PT) (OR = 0.93-0.99, 95%CI [0.94-1.09], $ps = .158$-745)

Kubo et al. (2018)	98	PT	KW	N-S	Scores on CBT in first-year students among four groups divided by CBT ($p < .05$) Scores on CBT in second-year students among four groups divided by CBT ($p < .05$)	
Matsuura et al. (2018)	85	PT, OT	C	N-PF, N-S	GPAs of first-year students (PT: $r = .828$; OT: $r = .594$) GPAs of second-year students (PT: $r = .854$; OT $r = .671$) Scores on graduation examinations (PT: $r = .854$; OT: $r = .641$)	
Kitamura et al. (2019)	297 (PT), 104 (OT)	PT, OT	PC, TT, ROC	N-PF, N-S	Scores on tests immediately after entrance to N-PF (PT: $p < .05$, OT: $p < .01$) Scores on periodic tests of first- and second-year students to N-PF (PT: $p < .05$, OT: $p < .05$) Scores on MTs in early Dec, early Jan, and early Feb to N-PF (PT: $p < .05$, OT; $p < .05$) Scores on MTs in early Feb of third-year students to N-S (PT: $r = .76$, $p < .01$, OT: $r = .78$, $p < .01$) Cutoff: 90.5 (Scores on tests immediately after the entrance) (sensitivity = .72, FP = .44, AUC = .67) Cutoff: 30.5 (Scores on tests of first-year students) (sensitivity = .57, FP = .23, AUC = .73) Cutoff: 42.5 (Scores on foundational tests in the second year of fundamental subjects) (sensitivity = .52, FP = .10, AUC = .77) Cutoff: 68.5～82.5 (Scores on foundational tests in the second year of specialized subjects) (sensitivity = .64～72, FP = .13-29, AUC = .74～75) Cutoff: 41.5～47.5 (Scores on MTs in the third year of fundamental subjects) (sensitivity = .66～.84, FP = .05～.25, AUC = .60～.85) Cutoff: 84.5～103.5 (Scores on MTs in the third year of specialized subjects) (sensitivity = .68～.85, FP = .11～.25, AUC = .62～.83)	
Koshino (2019)	54	PT	ANOVA, ROC	MT	Scores on MTs of first-, second-, and third-year students ($p < .05$) Self-learning quantity of online drills in late Nov, early Dec ($p < .05$) Cutoff: 0.29 (Scores on MTs in the first year of fundamental subjects) (sensitivity = .867, specificity = .629, AUC = .80) Cutoff: 0.31 (Scores on MTs in the second year of fundamental subjects) (sensitivity = .588, specificity = .771, AUC = .73) Cutoff: 0.36 (Scores on MTs in the third year of fundamental subjects) (sensitivity = .688, specificity = .829, AUC = .77)	Self-learning quantity of online drills in early Oct, late Dec, early Jan, late Jan, and early Feb

Serita et al. (2019)	129	PT	SC, MW	N-S	Tension-Anxiety (T-A) measured by POMS second eds.[d] ($p < .01$); Confusion-Bewilderment (C), Depression-Dejection (D), and Fatigue-Inertia (F-I) measured by POMS second eds.[d] ($p < .05$)	Anger-Hostility (A-H), Vigor-Activity (V), and Friendliness (F) measured by POMS second eds[d]
Kitamura et al. (2020)	175	PT, OT	CS, TT, LL, ROC	MT	Scores on periodic tests in the second year (PT: $ps < .01$; OT: $p < .01$ except in the fundamental subjects; $ps < .05$) Scores on MTs in early Dec and early Jan of the final year (PT: $ps < .01$; OT: $ps < .01$) Scores on MTs in early Jan of the final year in fundamental subjects by LL (PT) (OR = 0.85, 95%CI [0.76-0.95], $p = .005$) Cutoff: 49.5 (Scores on MTs in early Jan of the final year of fundamental subjects) (sensitivity = .786, FP = .227, AUC = .842)	Scores on the tests immediately after entrance (PT, OT) Scores on the periodic tests in the first year (PT, OT) Scores on MTs in early Dec of fundamental and specialized subjects and in early Jan in the final year of specialized subjects by LL (PT) (OR = 0.93～1.02, 95%CI [0.84-1.02], $ps = .086$～$.893$)
Yakabi et al. (2020)	74	PT	ANOVA	MT	Main effect of the timing of MTs ($p < .01$)	The main effect of the training type of clinical practice Interaction of the timing of MTs and training type of clinical practice

PT: physical therapist; OT: occupational therapist; ST: speech therapist; PC: Pearson's correlation analysis; SC: Spearman's correlation analysis; C: Correlation analysis (not described which correlation analysis was done); CS: Chi-square test; FX: Fisher's exact test; TT: t-test; ANOVA: Analysis of variance; MW: Mann-Whitney U test; KW: Kruskal-Wallis test; WS: Wilcoxon signed-rank test; SR: Simple regression analysis; MR: Multiple regression analysis; LL: Logistic regression analysis; ROC: receiver operating characteristic analysis; N-PF: Pass/Fail of National Examination; N-S: Score of National Examination; MT: Mock test; GEE: general entrance examination; REE: recommendation entrance examination; GPA: Grade point average; FP = false positive; AUC = Area under curve; CBT: Computer-based testing; n.s.: not significant

[a] if there is no p value defined as a significant, $p < .05$ was considered significant. [b] participants were included other than PT/OT/ST course and total sample size is n = 1105 [c] other 16 out of 61 subjects were regarded as "undeterminable [d] POMS (2nd eds.): Profile of Mood States second edition[32)]

Concerning the outcomes of these studies, 16 articles dealt with the pass/fail status in the national examination (N-PF), eight articles discussed the scores in the national examination (N-S), and five studies examined the mock test (MT), which is a simulation of the national examination. Of these, six articles had set multiple outcomes, that is, N-PF and N-S.

Variables that significantly affected the outcomes were as follows: scores on tests during college (13 articles), for example, periodic tests in each year, periodic tests in foundational work and/or specialized tests, tests in particular subjects, scores on the MT (eight articles), scores on the test immediately after entrance (five articles), and GPAs-grade point averages-(four articles). Concerning admission examinations, the scores for the admission examination are included in two articles, and admission categories are included in one article. Other variables that significantly affected the outcomes were: the levels of satisfaction with learning conditions, the quantity of self-learning assessed by use of online drills, mood, measured by POMS (Profile of Mood States), and the timing of MTs (one article each).

Variables that did not significantly affect the outcomes were: scores on tests during college (seven articles), scores in the MTs (five articles), scores on the tests immediately after entrance (two articles), and the GPA of first- or fourth-year students (one article). Concerning admission examinations, the scores for the admission examinations are included in two articles, and admission categories are included in five articles. Other variables that did not significantly affect the outcomes were: levels of satisfaction with learning or living conditions, the quantity of self-learning, measured by the use of online drills, indication of mood, measured by Poms[32], the training type of clinical practice, and deviation values of high school scores (one article each).

3.3 Quality assessment

The results of the evaluation for the risk of bias are displayed in Figure 2. Risk of bias is assessed as a judgment (high, low, or unclear) for the individual elements.

Most of the studies also were at a high risk of selection bias (selection of participants) because, in our systematic review, the outcome was measured by N-PF, N-S, or MT. There is a high possibility that this could influence the selection effect of participants because participants were limited to those who could take a national examination, that is, students with low grades could not be included in each study. A few studies were judged as unclear or low risk because MTs were set as the outcome and the number of participants in the phase of study design and the actual number of participants were not described in the article. All studies had an unclear risk of selection biases (confounding variables) because of limited reporting of the interpretation of practice effect. Concerning performance bias, we assessed whether exposure, that is, the

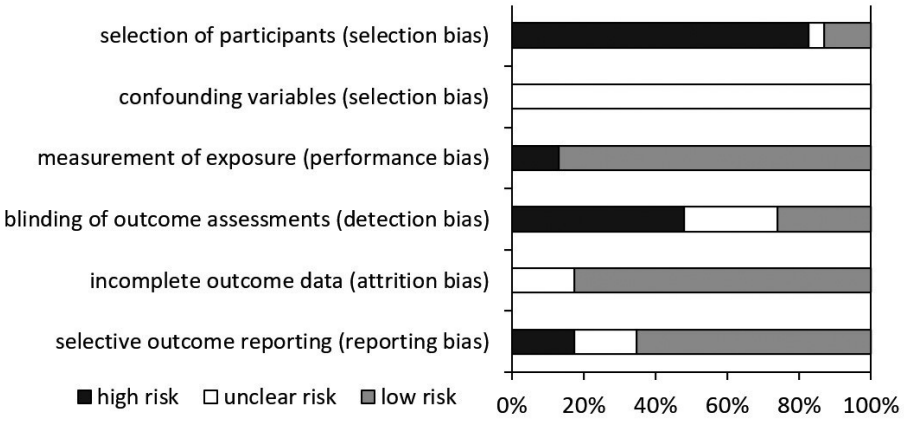

Figure 2 Assessment of risk of bias using the RoBANS

measurement of independent variables, was biased because of self-reporting methods, that is, whether the exposure was blinded. Most studies were judged as low risk because objective measures were used to assess the exposure, but some studies used self-reported measures, which were judged as high risk. Regarding detection bias, we judged the outcome of N-S as high because participant scores were not disclosed in the national examinations, thus the scores must be obtained through the participants' self-scoring, that is, a self-reporting method, and we judged the outcome of N-PF as low because the pass/fail status is determined based on the participants' exact score. This judgment might be reconsidered if any reference or additional analysis was made to the accuracy of the self-scoring on N-PF. The scores on the mock tests (MTs) were also judged as low risk because the scoring was done objectively by each institution. As a result, about half of the articles were considered high risk for detection bias. Then, attrition bias, that is, a bias that reduces the number of participants in the analysis due to inadequate setting of the outcome, is unlikely to occur in the present systematic review. It is considered low risk because the number of participants is unlikely to be reduced if the outcome is the result of a national examination or an MT. However, if the number of participants recruited is clearly stated and the number of participants analyzed is not stated, it is considered unclear for attrition bias. Regarding reporting bias, we expected two main outcomes of this systematic review; NS and N-PF. If both data were presented and the article stated the statistical difference and theoretical background for which that outcome was chosen, we judged as low/unclear if they were not stated, and high if one outcome was measured with no reason mentioned. The outcome from MTs was judged low as the primary outcome could be the score of the MT. In addition, we judged a study as low risk if the primary outcomes were defined in the method section, and described in the result section, and we could obtain the study protocol from each article. As a result, 65% of the articles were judged as a low risk for reporting bias. Overall, a low-risk bias was obtained in some areas, but for other areas we were undecided or some revealed high-risk bias.

4. Discussion

In the present study, we investigated published research articles on the predictors of performance in the national therapy licensure examination in Japan. Most of the articles included participants studying PT, and there was a relatively small number of articles about OT or ST.

4.1 Predictors of success in the national therapy licensure examination

We clarified significant independent variables; scores for a variety of tests (i.e., periodic tests, mock tests, tests immediately after entrance), admission examinations, admission categories, GPAs, and others.

The most prevalent variable extracted in this review was the scores on tests taken during college, for example, the periodic tests for each level, periodic tests of fundamental information and/or specialized tests. Among them, 13 articles showed significant differences or correlations with the outcome, and seven articles did not reveal a significant effect. It should also be noted that some articles showed both results[9,10,12,15,24,30]. An example of this inconsistency can be found in Motooka et al.[10] where grade course participants in the specialized subjects of PT and ST significantly influenced the N-PF, while there was no difference in the foundational subjects among ST course participants. Additionally, Kitamura et al.[30] found that the scores on periodic tests in the second year were a significant factor determining the scores on MTs, while continuous assessment of first year PT and OT course participants and foundational subjects in the second year for OT course participants, were significant. There is also a difference according to the statistical analysis; in Kitamura et al.'s study[24], where PC or TT made a significant difference to P-S or P-PF by the scores on the periodic tests of first or second years, while LL did not make a significant difference. It is assumed that compared to PC or TT, the multivariable analysis could result in difficulty in extracting meaningful variables for LL. In sum, the inconsistency of these findings could be explained by examining when the test was performed, which subjects were analyzed, that is, foundational or specialized subjects, and statistical tests that were used, that is, univariable or multivariable analysis. Therefore, predictions of national

examination results based on test scores need to be considered, including the possible influence of these background factors and the rationale for these factors.

Then, MTs were commonly extracted as the independent variable. Eight articles showed significant difference or correlation to the outcome, and five articles did not reveal a significant effect. Of these, five articles showed both results[17,18,22,24,30]. There could be some background factors that influence the relationship between the scores on MTs and success in the national examination, which would need to be explored, to resolve this inconsistency. It is suggested that the factor of when the MT is taken in the participants' final year is significant[17,18,24]. The times that predict the success in national examinations are: November and January[17], late January[18], early December, and early February[24]. In contrast, the times of year that were not linked to a prediction of success in the national examinations were: September, October, and December[17], and mid-October, mid-November, early December, and early January[22,24]. The subject is also suggested as a factor. In Inomoto et al.[18] the scores on MTs were predicted according to the subjects of kinesiology, cerebral vascular disease, and psychology by ML. Therefore, in a similar pattern to some tests-other than MTs mentioned before-the inconsistency of these findings could be explained by the timing of the test and which subjects were analyzed. Moreover, the range of timing of the MTs used in each study design could affect the results. It is necessary to improve the accuracy of predicting the success in national examination by the scores on specific subjects and the timing of MTs.

Regarding scores on the tests immediately after the entrance to courses, four articles showed significant differences or correlations to the outcome, and two articles did not show significant effects. Of these, one article showed both results[24]. Concerning admission examinations, three articles showed significant differences or correlations to the outcome (two of them were admission examination scores and the rest of them were admission categories). Six articles did not show a significant effect (one of them concerned admission examination scores and the rest of them were admission categories). Hence, the factors of entrance examinations, or the tests immediately after the entrance, are possibly not as robust as periodic tests or MTs. In fact, the scores on these periodic or mock tests, and GPAs, which will be discussed in the next paragraph, frequently have an impact, so it is suggested that we should focus on these indicators.

Concerning GPAs, three articles showed significant differences or correlations to the outcome[14,17,21], and one article showed both significant (GPAs of second- or third- year students) and non-significant (GPAs of first- or fourth- year students) correlations to the outcome[23]. The summary of these results indicates that GPA had a significant correlation with success in the national examination, when considering the GPAs of the second and third-year students.

Other independent variables reported in each article include: levels of satisfaction with learning or living conditions[19]; the quantity of self-learning indicated by online drills[28]; mood, measured by POMS[29]; and the timing of MTs[31]. As these variables were reported in single articles, further studies are needed to verify the influence of these factors on national examinations.

To summarize the systematic review, we can say that although there are some inconsistencies in the factors that predict success in the national examination-taking into consideration the background factors of when the variables were collected and which subjects were collected-scores on the periodic tests, scores on mock tests, and GPAs were suggested to be the predictive factors. In contrast, admission examination (scores and/or categories) or tests immediately after entrance did not seem to predict success in national examinations.

4.2 Quality assessment

The systematic review revealed that the risk of bias in studies that are designed to predict success in national examinations is dependent on the outcome of each study measured. The success of students in the national examinations (N-PF or N-S) could significantly influence the selection process. Those who fail to advance to the next year or had poor scores would be excluded from taking the national examination, which indicates that selection bias (selection of participants) was high. Additionally, when N-S is used as an

outcome, the detection bias (blinding of outcome assessment) is estimated as high because N-S is measured by the self-reporting of participants. To avoid this bias, it might be recommended that N-PF is used as an outcome. However, it would be necessary to be careful with the statistical analysis, if the pass rate of the national examination is high, because of the biased distribution of the participants between pass and fail groups. The reason for using an outcome (N-PF or N-S) in the statistical analysis should be settled in advance as this could also lead to reporting bias.

In light of the above, it may be useful to use the scores on MTs, that simulated the national examinations, as outcomes rather than using the N-PF or N-S as indicators. This is because the scores on the MTs are continuous values, and unless this is self-reported, the measurements do not affect the outcome and the selection effect could be small. Tsubota et al.[11] addressed this concern about the uncertain reliability of the N-S. In order to use the scores on the MTs to predict success in the national examination, the relationship between these two examinations should be sufficiently clear in advance. For example, by using receiver operating characteristic analysis (ROC) to set cutoff values[24,27,28,30]. The use of the scores on MTs based on the results of such an analysis could be recommended as a basis for the design for the prediction of success in the national examinations.

4.3 Limitations and further implications

In the present study, we used p-value to determine whether factors could predict success in the national examination. However, this leads to the possibility of type 1 or type 2 error because the p-value is influenced by the sample size: the larger the sample size, the smaller the p-values. To prevent these errors, effect size (ES), which is a standardized measure independent of sample size, is a more beneficial index than the p-value. Ideally, it is preferable to estimate the required sample size utilizing a power analysis at the stage of designing the study protocol, but no article addressed this issue. Although we did not include the meta-analysis of this systematic review, the effect size can be calculated by using some statistical values that are usually described in each article, and then, the statistical power also can be calculated by post-hoc power analysis. Our quantitative systematic review has a future implication for these statistical analyses.

Furthermore, this study showed that scores on periodic tests or MTs could be related to the success of the national examination, but the causal relationship is not clear, that is, these scores directly influence the success in the national examination or the existence of pseudo-correlation; the third variable that has yet been accounted for in the analysis affects scores in the tests in school and national examinations. School tests are just the result of the third variable. This causal relationship relates to the strategy that students use to achieve success in the national examination. They need to choose whether to study hard to achieve success in the national examination or aim to achieve good grades in the periodic tests, assuming that this will also assist with the national examination. This raises the question of whether the periodic tests should be designed with the national examinations in mind. The relationships of these variables should be further explored.

Finally, the scope of our systematic review is limited to the national examinations for therapists in Japan. There is diversity in the national examinations and systems of higher educational institutions in different countries. However, as other countries have also reported similar factors, for example, GPAs[33], it is necessary to deepen our discussion by comparing our results with those in other countries. It would be necessary for future studies to accumulate the related literature that was designed to lower the risk of biases and to further develop the result of our study into robust findings.

5. Conclusion

Our systematic review revealed the variables that predict the success of students in the Japanese national therapist licensure examination; the scores from periodic tests, scores from mock tests, and GPAs were found to be the predictive factors, accounting for the background factors of when the variables were collected, and which subjects were collected. Admission examinations (score and/or categories) or tests

immediately after entrance did not seem to predict the success of students in the national examination to the same extent.

Acknowledgments

We wish to thank Editage for English language editing.

References

1. Japanese physical therapy association: *Statistics*. http://www.japanpt.or.jp/about/data/statistics/, 2020. (December 11, 2020) (In Japanese, translated by the author of this article)
2. Japanese Association of Occupational Therapists: *The Journal of Japanese Association of Occupational Therapists (JJAOT)*. https://www.jaot.or.jp/files/page/kankobutsu/pdf/ot-news2020/2020-12.pdf, 2020. (December 11, 2020) (In Japanese, translated by the author of this article)
3. Japanese association of Speech-Language-Hearing Therapists: *What is a Speech-Language Pathologist?* https://www.japanslht.or.jp/what/, 2020. (December 11, 2020) (In Japanese, translated by the author of this article)
4. Ministry of Education, Culture, Sports, Science and Technology: *Selection for University Admission, Current State of University Education*. http://www. kantei. go. jp/jp/singi/kyouikusaisei/dai11/sankou2.pdf, 2013. (December 11, 2020) (In Japanese, translated by the author of this article)
5. Moher D, Liberati A, Tetzlaff J, Altman DG and The PRISMA Group : Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Med*, 6, e1000097, 2009.
6. Kojimahara N, Nakayama T, Morizane T, Yamaguchi N and Yoshida M eds. : *Minds Manual for Guideline Development 2017*. Japan Council for Quality Health Care, Tokyo, 2017. (In Japanese)
7. Kim SY, Park JE, Lee YJ, Seo HJ, Sheen SS, Hahn S, Jang BH and Son HJ : Testing a tool for assessing the risk of bias for nonrandomized studies showed moderate reliability and promising validity. *Journal of Clinical Epidemiology*, 66, 408-414, 2013.
8. Cochrane Collaboration: *Cochrane Handbook for Systematic Reviews of Interventions, Cochrane Collaboration, Version 5.1.0 March 2011*. http://www.cochranehandbook.org/051, 2011. (December 11, 2020)
9. Yanagisawa K, Nitta O, Kasai H, Nekota Y, Iida Y, Kikuchi E, Osada M, Fukushi M, Saitoh H and Fukuda K : Statistical analysis of the results of national examinations relevant to entrance examination results and colleges performance level. *The Journal of Tokyo Academy of Health Sciences*, 2, 16-21, 2000. (In Japanese with English abstract)
10. Motooka N, Iwatani K, Manabu S, Shiromoto O and Domoto T : Relationship between scores of entrance examinations and college performance study at Hiroshima Prefectural College of Health and Welfare. *Journal of Hiroshima Prefectural College of Health Science, Humanity and Science*, 3, 95-104, 2003. (In Japanese with English abstract)
11. Tsubota Y, Kishimoto M, Sakai K and Togashi S : The relationship between scores of examinations of physiology class and practice examination of the national board examination for physical therapist at Osaka Kawasaki Rehabilitation University, and estimation of the scholastic performance of the lower classes. *Journal of Osaka Kawasaki Rehabilitation University*, 5, 63-69, 2011. (In Japanese)
12. Kochi M, Yamasaki H, Hiraga Y, Katayama K, Shigeshima K and Yamamoto S : The relationship simulated national exam scores and the scores of memory and calculation tasks. *Journal of Kochi Rehabilitation Institute*, 13, 39-41, 2012. (In Japanese)
13. Murata K, Hori H and Saito H : Relationship between admission test results and national examination results: Analysis of rehabilitation students in first-third grade students. *Nittazuka Iryou Fukushi Senta Zasshi*, 9, 25-28, 2012. (In Japanese, translated by the author of this article)
14. Akagi M and Hibino I : An investigation of the academic achievement of the physical therapy students for the national examination: The difference of the entrance examination category. *The Nagoya Gakuin Daigaku Ronshu; Journal of Nagoya Gakuin University; Humanities and Natural Sciences*, 49, 7-15, 2013. (In

Japanese)
15. Murao H : Relationship between scores in the national examination for physical therapists and grades of compulsory specialized courses in Kobe Gakuin University: Based on the results of a survey involving graduates of the year 2013. *Kobegakuin Journal of Rehabilitation Research*, 9, 13-20, 2014. (In Japanese with English abstract)
16. Ashiga H and Ohdaira Y : The relationship between national examination and academic proficiency, academic achievement and clinical practice. *Bulletin of Meirin College*, 18, 1-6, 2015. (In Japanese)
17. Goto T, Haruhara N, Tateishi M, Miyamoto S, Imatomi S, Kobayashi T, Saito S, Tsuzuki S, Takasaki J, ... Sadakari J : Analysis of the grade point average and scores on a mock examination to predict results of the national license examination for speech-language-hearing therapists. *Mejiro Journal of Health Care Sciences*, 8, 37-42, 2015. (In Japanese with English abstract)
18. Inomoto A, Ishibashi T, Nakato K, Takahashi S and Hashimoto T : Factors that affect the mark of the physical therapy national license examination: Examination in the trial examination that vendors made. *Bulletin of Kyushu Nutrition Welfare University*, 12, 39-44, 2015. (In Japanese)
19. Kubo A, Kuramoto-Ahuja T, Kobayashi K, Kurosawa K and Maruyama K : Influence of satisfaction with learning and living at the end of the junior and senior college years of PT students on the results of the national examination for physical therapists. *Rigakuryoho Kagaku*, 30, 115-117, 2015. (In Japanese with English abstract)
20. Murao H : Relationships between scores of entrance examinations for the physical therapy course of Kobe Gakuin University and national physical therapy examinations. *Kobegakuin Journal of Rehabilitation Research*, 10, 129-136, 2015. (In Japanese with English abstract)
21. Otaka K and Narita A : Ascertaining the likelihood of failing the national exam and measure to take the early stages of countermeasures. *Memoirs of Osaka Shin-ai College*, 51, 1-5, 2017. (In Japanese, translated by the author of this article)
22. Inomoto A, Ishibashi T, Nakato K, Umeno K and Ohmine S : Influence of anatomy, physiology, and kinematics on national physical therapy license examination. *Bulletin of Kyushu Nutrition Welfare University*, 14, 119-124, 2017. (In Japanese)
23. Murao H and Iwai N : Relationship between scores in the national examination for physical therapists and grade point average (GPA) every semester: Analysis of relationship between scores in the national examination for physical therapist and GPA estimated with required subjects that have positive correlation statistically (GPA-RSPC). *Rehabilitation Kyouiku Kenkyu*, 22, 248-252, 2017. (In Japanese)
24. Kitamura M, Nakamura K and Isobe E : Relationship between undergraduate tests and national examinations for undergraduate students at five rehabilitation colleges. *Igaku Kyoiku*, 49, 127-134, 2018. (In Japanese with English abstract)
25. Kubo A, Horimoto Y, Han H and Sadakiyo K : Association between computer-based testing and national examination results among physical therapy students. *The Journal of Physical Therapy Science*, 30, 790-793, 2018.
26. Matsuura T, Kato M, Yamashita H and Ogawa Y : The report of learning behavior survey for the 2016 national exam. *Bulletin of Aichi Medical College*, 9, 84-90, 2018. (In Japanese)
27. Kitamura M, Isobe E, Murayama M, Kawase N, Kuwahara K, Nakazaki M, Nakata T, Hadano H and Sasaki K : Relationship between a national examinations prediction test and undergraduate tests at five rehabilitation colleges for physical therapists and occupational therapist. *Rigakuryoho Kagaku*, 34, 387-391, 2019. (In Japanese with English abstract)
28. Koshino Y : Relationship between simulated test results and amount of self-learning in national physical therapy examination scores. *Journal of Allied Health Sciences*, 10, 99-106, 2019. (In Japanese with English abstract)
29. Serita T, Hirabayashi S, Toriyama M, Suganuma K, Toyota A, Shimbo M, Kaneko C and Tsukada E : Research on physiotherapy students: relationship between mood stated before national examination and

the result. *Bulletin of Teikyo University of Science and Technology*, 15, 41-47, 2019. (In Japanese with English abstract)

30. Kitamura M, Isobe E, Murayama M, Kawase N, Kuwahara K, Nakazaki M, Nakata T, Hadano H and Sasaki K : Relationship between a national examinations predictions test and undergraduate tests in students of Departments of Physical Therapy and Occupational Therapy with poor grades. *Rigakuryoho Kagaku*, 35, 147-151, 2020. (In Japanese with English abstract)
31. Yakabi A, Onoda K, Ishizaka M, Sadakiyo K and Kubo A : Effect of differences in clinical practice from on grade scores on problems related to practical skills corresponding to the national examination: comparison between the conventional practice and the clinical clerkship. *Rigakuryoho Kagaku*, 335, 367-370, 2020. (In Japanese with English abstract)
32. Heuchert JP and McNair DM : *Profile of Mood States Second Edition*. Multi-Health System, Tronto, 2013.
33. Zadnik M, Lawson S, DeLany JV, Parente F and Archer KR : Analysis of objective factors related to a successful outcome on the National Examination for Occupational Therapists. *Rehabilitation Research, Policy, and Education*, 31, 50-60, 2017.