When can we identify the students at risk of failure in the national medical licensure examination in Japan using the predictive pass rate?

Toshiki Shioiri[1], Michiyo Nakashima[1] and Koji Tsunekawa[1,2*]

Abstract

Context Failure of students to pass the National Medical Licensure Examination (NMLE) is a major problem for universities and the health system in Japan. To assist students at risk for NMLE failure as early as possible after admission, this study investigated the time points (from the time of admission to graduation) at which predictive pass rate (PPR) can be used to identify students at risk of failing the NMLE.

Methods Seven consecutive cohorts of medical students between 2012 and 2018 ($n=637$) at the Gifu University Graduate School of Medicine were investigated. Using 7 variables before admission to medical school and 10 variables after admission, a prediction model to obtain the PPR for the NMLE was developed using logistic regression analysis at five time points, i.e., at admission and the end of the 1st, 2nd, 4th, and 6th grades. All students were divided into high (PPR < 95%) and low (PPR ≥ 95%) risk groups for failing the NMLE at the five time points, respectively, and the movement between the groups during 6 years in school was simulated.

Results Medical students who passed the NMLE had statistically significant factors at each of the 5 time points, and the number of significant variables increased as their grade in school advanced. In addition, two factors extracted at admission were also selected as significant variables at all other time points. Especially, age at entry had a consistent and significant effect during medical school.

Conclusions Risk analysis based on multiple variables, such as PPR, can inform more effective intervention compared to a single variable, such as performance in the mock exam. A longer prospective study is required to confirm the validity of PPR.

Keywords National Medical Licensure Examination, Logistic regression analysis, Time points

*Correspondence:
Koji Tsunekawa
kojit1979@gmail.com
[1]Department of Psychiatry and Psychotherapy, Gifu University Graduate School of Medicine, Gifu 467-8601, Japan
[2]Department of Medical Career and Professional Development, Nagoya City University School of Medical Sciences, 1, Kawasumi, Mizuho-cho, Mizuho-ku, Nagoya 467-8601, Japan

Introduction

Worldwide, one of the many goals of medical schools is to provide newly graduating physicians with a foundation of knowledge, technical skills, reasoning ability, and empathy [1]. To this end, standardized examinations are designed to objectively measure the student's performance throughout medical education and at graduation. Several countries have adopted their respective National Medical Licensure Examination (NMLE), such as the United States Medical Licensing Examination (USMLE) [2], the Medical Council of Canada Qualifying Examination (MCCQE) [3], the National Competence-Based Learning Objectives for Undergraduate Medical Education (Nationaler Kompetenzbasierter Lernzielkatalog Medizin: NKLM) in Germany, [4] the Federal Licensing Examination in human medicine in Switzerland [5], and the Medical National Exam (Lekarski Egzamin Panstwowy: LEP) in Poland [6]. In the United Kingdom, a qualification exam called Medical Licensing Assessment is scheduled to start from 2024 [7]. However, the failure of students to pass the NMLE is a big challenge [8–10]. Despite the current reality, few studies have aimed to create models to predict successful or unsuccessful applicants in NMLE [8, 9].

In our recent study, we identified six predictors for passing the NMLE in Japan at the first attempt and created a model to obtain the predictive pass rate (PPR) using logistic regression analysis. The PPR was successfully applied to identify students likely to fail the NMLE [11].

In contrast, many students who are not successful fail to perform well during their first year in university [8, 11–14]. Kies and Freund highlighted that all medical schools face the issue of poor-performing M-1 students [15]. Therefore, medical schools need to identify as soon as possible after the start students who are likely to fail in the NMLE.

In the current study, we used the PPR at five time points (at admission and the end of 1st, 2nd, 4th, and 6th grades) to identify the time points at which students at risk of failure in the NMLE can be identified.

Current situation of NMLE and the number of doctors in Japan

NMLE is the only national board paper test to obtain doctor's license in Japan. This test evaluates the knowledge of clinical and social medicine. Undergraduate medical education in Japanese medical schools typically lasts 6 years [16, 17], including 4 years of pre-clinical medical sciences and 2 years of clinical training. Graduates from these medical schools can take the NMLE. Content-wise, NMLE includes a part of USMLE Step 1, all of USMLE Step 2 clinical knowledge, and only medical knowledge of USMLE Step 3. [18]

Reflecting the Japanese university culture of difficult enrolment and easy progression, and the future prediction of the number of physicians by the Japanese government, the annual number of graduating medical students is the smallest among all Organization for Economic Cooperation and Development (OECD) countries [19] in 1990s and 2000s. Consequently, Japan is ranked 28th in terms of the number of practicing doctors among the 31 OECD countries [20]. Since 2008, the number of medical students has gradually increased, but the shortage of doctors continues to persist. Therefore, the failure of students to clear NMLE in the first attempt has adverse repercussions for the Japanese healthcare system by aggravating the shortage and regional maldistribution of doctors [11].

Methods

Participants

To develop a reliable PPR for NMLE, seven consecutive cohorts of 637 students (2012–2018, 6th academic year) of the Gifu University School of Medicine (GUSM) were included in this study. GUSM is one of the 51 public schools that are supported by the Japanese government for entrance and school fees. These cohorts comprised 78, 69, 84, 97, 110, 93, and 106 students, respectively. All data were anonymized by the academic affairs office of GUSM before granting access to the research team. The study participants had provided comprehensive consent by opting out. Ethical approval was granted by the GUSM Ethics Committee. Anonymity and confidentiality for students were guaranteed (date: 05/13/2020, reference number: 2020-039).

Variables

The dependent variable was "failing to pass the NMLE." This data was obtained from the Department of Academic Affairs in GUSM. Explanatory variables included pre-admission variables such as (i) gender, (ii) age at admission, (iii) location of high schools (HS) (2 categories; neighborhood and distant; the former includes Gifu and Aichi prefectures which account for 60% of the students in GUSM, and the latter includes other areas in Japan), (iv) type of HS (public/private), (v) Academic level of HS (see Table 1 for more information), (vi) HS Grade Point Average (GPA; 5-grade evaluation), and (vii) achievement (percentage of correct answers) in the common entrance examination for university (National Center Test for University Admissions, NCTUA). Post-admission variables were (viii) Test of English as a Foreign Language (TOEFL) score, (ix) academic performance (percentage) in liberal arts (LA), (x) total score (percentage) in basic sciences in the first year, xi) total score (percentage) in basic biomedical sciences in the second year, xii) total score (percentage) in pre-clinical medical

Shioiri *et al. BMC Medical Education*   (2024) 24:930

Page 3 of 8

Table 1 Characteristics of participants in the 2012–2018 group

Variables		2012–2018 ($n = 637$)		
		Failure (%)	**Passer (%)**	***P*-value**
Gender	Male	38 (7.9)	446 (92.1)	**0.008**
	Female	3 (2.0)	150 (98.0)	
Age at admission (years)		27.83 ± 8.03	19.58 ± 3.57	**< 0.0001**
HS Location		6 (1.6)	376 (98.4)	**< 0.0001**
Neighborhood (Gifu and Aichi Pref.)		35 (13.7)	220 (86.3)	
Distant (Other areas)				
Type of HS		21 (7.3)	267 (92.7)	0.517
Private		20 (5.7)	329 (94.3)	
Public				
Academic level of HS		68.63 ± 6.07	69.13 ± 5.6	0.587
GPA in HS		4.19 ± 0.46	4.47 ± 0.45	**0.0001**
NCTUA score		84.24 ± 4.92	86.46 ± 3.85	**0.0005**
TOEFL score		517.9 ± 28.6	521.6 ± 25.0	0.37
Academic performance in liberal arts		70.10 ± 20.06	71.29 ± 16.47	0.659
Total score in basic sciences in the 1st year.		73.78 ± 4.25	74.93 ± 4.65	0.121
Total score in basic biomedical sciences in the 2nd year.		66.61 ± 4.28	72.56 ± 6.82	**< 0.0001**
Total score in pre-clinical medical sciences in 3rd to 4th year.		73.96 ± 4.6	75.77 ± 5.44	**0.038**
CBT-IRT score in the 4th year.		48.16 ± 8.61	59.7 ± 10.16	**< 0.0001**
Pre-CC OSCE score in the 4th year.		4.23 ± 0.42	4.5 ± 0.39	**< 0.0001**
Performance in clinical clerkship in the 5th to 6th year.		3.77 ± 0.73	4.04 ± 0.44	**0.0003**
Graduation examination in the 6th year.		-1.44 ± 0.74	0.1 ± 0.94	**< 0.0001**
Holdover during the 1st to 6th year.		12 (21.1)	45 (78.9)	**< 0.0001**
+		29 (5.0)	551 (95.0)	
-				

Note HS: High School; NCTUA: The National Center Test for University Admissions; TOEFL: Test of English as a Foreign Language; CBT-IRT; Computer-Based Testing with Item Response Theory; OSCE: Objective Structured Clinical Examination. Academic level of HS shows an average of values quantifying the difficulty of entrance exams in each HS. A higher level means a more difficult entrance exam

sciences during the third to fourth years, xiii) score of nationwide Computer-Based Testing with Item Response Theory (CBT-IRT) which assesses all pre-clinical education in the fourth year, xiv) average score (six-point scale) of Pre-Clinical Clerkship Objective Structured Clinical Examination (Pre-CC OSCE) in the fourth year, xv) achievement (standardized deviation values) in the graduation examination in the sixth year, xvi) performance of clinical clerkship during fifth to sixth years, and xvii) with or without holdover during the first to sixth years (see, Table 1). These data were also obtained from the Office of Academic Affairs in GUSM. For missing values, the average value calculated from the data not including the missing value was substituted.

Data analysis

First, we used Fisher's exact test and independent *t*-test to compare the demographic characteristics and some of the above-mentioned variables before and after attending university between those who failed and passed the NMLE in the 2012–2018 cohort (see, Table 1).

Second, logistic regression was used to predict the likelihood of passing the NMLE at five time points, i.e., at admission and the end of first, second, fourth, and sixth grades, after simultaneously controlling for potential confounders. The results are presented as odds ratios (ORs) and 95% confidence intervals (95% CIs). The explanatory variables varied according to the five time points because of the different variables available for use at each point (see, Table 2). In this logistic regression analysis, the PPR was calculated without covariate selection in multivariable model.

Third, at each time point, we created the PPR, which is a prediction formula, for the pass rate of NMLE using logistic regression analysis with all possible models, respectively. Using the five PPR, we divided all students into two groups; high (PPR < 95%, which indicates the average pass rate for our graduates) and low (PPR ≥ 95%) risk groups for failing the NMLE at five time points, respectively, and run a simulation of movement between the groups during the 6 years in school. The reliability of PPR was assessed using the actual NMLE results and the binary classification.

SPSS ver. 24.0 Japan for Windows (SPSS Inc., Chicago, IL, USA) was used for statistical analyses. Two-tailed p-values < 0.05 were considered indicative of statistical significance.

Table 2 Variables used in the logistic regression analysis at five time points

Variables	At admission	The end of			
		1st grade	2nd grade	4th grade	6th grade
Gender	X	X	X	X	X
Age at admission	X	X	X	X	X
HS Location: Neighborhood	X	X	X	X	X
Type of HS: Public	X	X	X	X	X
Academic level of HS	X	X	X	X	X
GPA in HS	X	X	X	X	X
NCTUA score	X	X	X	X	X
TOEFL score		X	X	X	X
Academic performance in liberal arts		X	X	X	X
Basic sciences			X	X	X
Basic biomedical sciences				X	X
Pre-clinical medical sciences				X	X
CBT-IRT				X	X
Pre-CC OSCE				X	X
Performance in clinical clerkship					X
Graduation exam					X
Holdover		X	X	X	X

Note Holdover means whether the student has at least one experience staying back a year at the end of each grade

Table 3 Predictors of the likelihood of passing the NMLE using logistic regression analysis

Time		Variables	*P*-values	Odd ratio (OR)	95% CI
At admission		Age at admission	< 0.0001	0.834	0.789 ± 0.882
		Neighborhood	0.00426	4.684	1.767 ± 12.41
The end of	1st grade	Age at admission	< 0.0001	0.812	0.762 ± 0.865
		Neighborhood	0.00154	4.960	1.841 ± 13.36
		Basic sciences	0.01361	1.144	1.028 ± 1.272
	2nd grade	Age at admission	< 0.0001	0.810	0.755 ± 0.869
		Basic biomedical sciences	0.00017	1.252	1.114 ± 1.408
		Neighborhood	0.01175	3.925	1.355 ± 11.37
	4th grade	CBT-IRT	< 0.0001	1.214	1.124 ± 1.311
		Age at admission	< 0.0001	0.809	0.742 ± 0.882
		Neighborhood	0.00554	5.27	1.628 ± 17.06
		Pre-clinical medical sciences	0.04473	0.855	0.734 ± 0.996
	6th grade	Graduation exam	< 0.0001	6.337	2.643 ± 15.19
		CBT-IRT	0.00024	1.184	1.082 ± 1.295
		Pre-clinical medical sciences	0.00066	0.696	0.565 ± 0.857
		Age at admission	0.00349	0.863	0.782 ± 0.953
		Performance in clinical clerkship	0.00730	3.291	1.378 ± 7.857
		Neighborhood	0.01915	5.052	1.303 ± 19.60

Results

Characteristics of the students who failed or passed NMLE

Table 1 shows significant differences between students who failed or passed the NMLE with respect to some demographic characteristics and achievements before and after university. Regarding pre-admission variables, the students who failed the NMLE had a higher proportion of males, a higher age at admission, HS located in any prefecture other than Gifu and Aichi prefectures, lower GPA in HS, and lower NCTUA score. Additionally, after university, they showed significant tendencies for achieving lower grades for basic biomedical and clinical medical sciences, CBT-IRT, Pre-CC OSCE, and CC, and they were predisposed to repeat a year in the medical school (see Table 1).

Logistic regression analysis

Table 3 shows only significant variables that predicted the likelihood of passing the NMLE at admission and the end of each grade. All factors are listed in Supplementary file 1 in the Supplementary Information zone. These results show that medical students who passed the NMLE had significant factors at each of the 5 time points, and the number of these factors increased with the advancement of their grade in school (i.e., 2 factors at admission, 3 at 1st and 2nd, 4 at 4th, and 6 at 6th grade) (see Table 3). Interestingly, the 2 factors extracted at admission were also selected as significant variables at all other 4 time points (1st, 2nd, 4th, and 6th grade). Especially, the age at admission showed a consistently strong influence with the likelihood of passing the NMLE.

Prediction formula for the pass rate of the NMLE in Japan

Given that the predictive pass rate of NMLE is p/100, we developed five logistic regression formulae, i.e., at admission, 1st, 2nd, 4th, and 6th grades (Supplementary file 2 in the Supplementary Information zone).

Shioiri *et al. BMC Medical Education*   (2024) 24:930

Page 5 of 8

Simulation of division into high- and low-risk groups by the PPR and confirmation of the reliability of PPR

According to the above-described 6 formulae, subsequent predictions for student failures in the NMLE were made at admission, 1st, 2nd, 4th, and 6th grades in school by dividing them into high- and low-risk groups using the PPR (see Supplementary file 3 in the Supplementary Information zone). This shows a simulation of the movements of all 637 students during 6 years in school. We examined whether the students in the high-risk group, as estimated by the PPR at admission, remained in the same group. Although approximately 30% (193 of 637) students were determined "high-risk group" at admission, the percentage had decreased by 46.1% at the end of 6th grade (from 193 to 104). Among 104 students in the "high-risk group" at the end of 6th grade, the pass rate for NMLE was 64.4% (67/104) compared to 99.2% (529/533) in the "low-risk group." There were also some swings between the two groups. Moreover, we found that 61.2% of all students (390/637) continued to remain in the "low-risk group" during 6 years in school, but not all of them passed the NMLE (the pass rate was 99.5%).

To confirm the reliability of the PPR, Table 4 shows the sensitivity, specificity, false-negative rate (FNR), and negative predictive value (NPV). As shown in the table, the numbers in the "high-risk group" gradually decreased, and the sensitivity, specificity, and likelihood ratio edged upward together during the 6 years in school. By contrast, the FNR and the NPV were maintained at very low and high levels, respectively.

Discussion

There are two important findings of this study. Using the logistic regression analysis, we identified 2 to 6 significant predictors of passing NMLE at the five time points in school, including at admission, 1st, 2nd, 4th, and 6th grade, respectively. Second, using the PPR calculated by the logistic regression formula, we identified the "high-risk group" for failing the NMLE at each of the five time points and showed that using the PPR at an earlier time point (even at admission) can potentially identify the students at risk of failure in the NMLE.

To our knowledge, a few studies have sought to identify or predict student failure in NMLE [9, 21, 22]. However, our results cannot be directly compared to these studies because their models were based on only academic performance and did not include the background characteristics of students.

Significant predictors of passing NMLE at each time point

In the current study, we identified some significant predictors for passing NMLE at each of the five time points. Especially, age at admission and location of the HS in the neighborhood were identified as external predictors. The number of available variables increased with the advance in the grade, increasing the accuracy of the analysis.

Among the variables of academic performance in university such as total score of clinical medicine, CBT-IRT score, Performance in clinical clerkship, and score of graduation examination, only CBT was a nationwide examination. It is administered by the Japanese Ministry of Education, Culture, Sports, Science and Technology (MEXT) for all Japanese medical schools in the fourth year using a computer before clinical clerkship to assess the ability for clinical clerkship. Factors that are significant in multivariate analysis at the sixth year have been discussed in a previous paper [11]. In Brief, CBT as a predictor is very similar to that of USMLE step 1. Elder medical students are more likely to have insufficient study time because of their duty time such as family or working part-time, decline in memory, burnout, and life events such as childbirth and parental care. As for age and gender, these factors should not be used as criteria for selection because they are important issues related to discrimination and bias [23] and should be seen as non-modifiable factors. However, making these at-risk students aware of their higher risk of failure would help mitigate the risk. It is unclear why medical students who

Table 4 Numbers of students with lower PPR (≤ 95%) at five time points during medical education, numbers of passers and failures in the actual NMLE, and the results using the binary classification

		At admission	The end of			
			1st grade	2nd grade	4th grade	6th grade
No. of students with lower the PPR for NMLE (95% or less) = **"high-risk group"**		199	176	154	127	102
Results of the actual NMLE	No. of Fail	34	34	36	38	38
	No. of Pass	165	142	118	89	64
Sensitivity (x)		82.9%	82.9%	87.8%	92.7%	92.7%
Specificity (y)		72.3%	76.2%	80.2%	85.1%	89.3%
Likelihood ratio (x/[1-y])		3.0	3.5	4.4	6.2	8.6
False-negative rate (FNR)		1.6% (7/431)	1.5% (7/454)	1.0% (4/490)	0.6% (3/507)	0.6% (3/532)
Negative predictive value (NPV)		98.4%	98.5%	99.0%	99.4%	99.2%

Note NMLE: National Medical Licensure Examination in Japan

belonged to a neighborhood HS have a better chance of passing the NMLE, but similar results have been reported in a previous study in a Japanese medical school [24].

In the multivariate analysis of the early grades, academic performance at school such as basic sciences and basic biomedical sciences was a significant positive factor, while in the multivariate analysis of the fourth and sixth years, clinical medicine score was a dominant negative factor. Although this seems to be a contradictory result, in the univariate analysis, the number of failures was significantly lower for all factors of academic performance. This may also be attributable to the fact that the clinical medicine score is similar to the graduation examinations, which have a stronger positive influence as an explanatory force. As for the other predictors of medical school performance, previous studies have shown that the Medical College Admission Test (MCAT) score is a useful indicator of the performance of students in medical schools [25–27]. In particular, Wiley and Koenig found that MCAT scores had a slightly stronger correlation with medical school grades than did undergraduate GPA [26]. Koenig et al. indicated that MCAT scores, alone and in combination with undergraduate GPA, are good predictors of medical school performance, but not perfect [29]. Julian also investigated the validity of MCAT scores for predicting medical school performance (medical school grades, USMLE step scores, and academic distinction or difficulty) and found that MCAT scores were better predictors of USMLE step scores than were undergraduate GPAs, and the combination only slightly outperformed the MCAT scores alone [29]. Recently, Zhao et al. reported that the repeaters for MCAT are expected to achieve lower Step 1 scores than non-repeaters [30]. In Denmark, O'Neill et al. evaluated the predictive validity of non-grade-based admission testing versus grade-based admission relative to subsequent dropout and found that admission test students had a lower relative risk for dropping out of medical schools within 2 years of admission (OR 0.56, 95% CI 0.39–0.80) [31]. In our study, on the other hand, there were differences between passers and failures for NML-J in terms of the scores of NCTUA and GPA before admission in medical schools in the univariate analysis, but not in the multivariate analysis. The differences between our results and others may be attributable to the differences in examination period between NMLE and Steps 1 and 2. In addition, the lack of significant differences in our study may be because the averages and ranges of TCTUA and GPA are higher and narrower in medical students passing the admission exam. This may also be due to the distinct context of the entrance examination system for medical schools in Japan. For instance, the medical school admission quota in Japan is very small compared to the US and European countries, and medical school candidates are forced to work hard

to compete with each other (i.e., maximal magnification ratio for the last decade in GUSM was about one hundred which was highest in Japan). Donnon et al. conducted a meta-analysis of published studies to determine the predictive ability of the MCAT score for medical school performance and medical board licensing examinations. They found that the predictive ability of the MCAT score ranged from small to medium [32]. Lastly, Ramsbottom-Lucier et al. reported a modest gender difference for the NBME I, with the men performing better than the women [33]. Recently, McDougle et al. indicated that the relative risk of first attempt STEP 1 failure for medical school graduates was 3.2 for women (95%CI: 1.8–5.8, $p=0.0001$) [34]. Contrarily, in the study by Koenig et al., sex and race were not included in the subsequent prediction equations [28]. In the present study, gender showed a significant influence on passing NMLE in the univariate analysis, with the women outperforming men, but not in the multivariate analysis. Further research is required because of the inconsistent results in terms of the influence of gender.

Predicting NMLE with data in lower grades

Our previous study suggested the possibility of predicting NMLE in the early grades, and as shown in Table 3, we found similar results in multivariate analysis in the early grades. This reveals that some degree of risk analysis is possible using a similar method not only at graduation but also at lower grade levels and even at admission.

Baars et al. also developed a model for the early prediction of students who fail to pass the first year of the undergraduate medical curriculum within two years after the start [8]. In their study, independent variables included 5 pre-admission and 4 or 5 post-admission variables, and the predictions for failure in the first-year curriculum were made at 0, 4, 6, 8, 10, and 12 months by logistic regression analyses [8] Their results showed that students who had passed all exams at 4, 6, or 8 months (so-called "optimals") had a 99% chance of passing the first-year curriculum. The earliest time point with the highest specificity to predict student failure in the first-year curriculum was 6 months; however, additional factors are needed to improve this prediction or to bring forward the predictive moment [8]. It is well-known that the majority of students who are not successful fail to perform well during their first year in university [8, 13, 35]. All MSs are faced with the issue of poor-performing M-1 students [15]. The challenge is to encourage these students to take remedial programs that address their academic problems and assist them in becoming high-performing physicians [15]. Kies and Freund indicated that medical students who decompress their M-1 year prior to M-1 year failure outperform those who fail their first year and then repeat it. They suggested the need for

Shioiri *et al. BMC Medical Education*   (2024) 24:930

Page 7 of 8

careful monitoring of the performance of M-1 students and implementing early intervention and counseling of struggling students [6].

Improvement of actual pass rate for NMLE after intervention in 2018

In GUSM, from 2012 to 2017, specific intervention was implemented for students who performed poorly in the mock exam (ME) conducted approximately four months before the actual NMLE exam. However, this intervention was not effective because not all poorly performing students took the ME as participation in the ME was voluntary. Additionally, some young students with poor performance in the ME were able to pass the actual NMLE, while some older students with good performance in the ME did not pass the actual NMLE.

Therefore, as discussed previously, using PPR and a new sample in 2018, we picked up the 15 candidates who had lower PPR for NMLE (≤95%), indicating a strong likelihood to fail NMLE at the first attempt, to confirm the validity of the formula (see Table 3). Additionally, the students were provided adequate guidance by the competency committee to overcome their shortcomings before taking the actual NMLE. This dramatically improved the actual pass rate for NMLE at the first attempt in 2018. Moreover, PPR predicted all 5 failures who were included in 15 candidates. This suggested that performing risk analysis based on several variables, such as PPR, can lead to more effective intervention compared to a single variable, i.e., performance in the ME. Further prospective studies are needed in other cultural settings to confirm the validity of PPR.

Strengths and limitations

Some limitations of this study should be acknowledged. First, our results cannot be directly compared with those of previous studies because of the different independent variables used. Second, our results may be influenced by the differences with respect to the selection of medical students and the medical education system in Japan compared to other countries. Third, the applicability of our results to other Japanese MSs is not clear because no similar studies have been conducted in other schools and the duration of the prospective study was only one year. Finally, there are inherent pitfalls of using prediction models such as overfitting [36].

Implications for future research

Improving the reliability of PPR developed in the current study may help reduce the number of failures in NMLE, USMLE, or the undergraduate medical curriculum. As for the next steps, we are planning a new prospective study lasting at least several years to obtain more robust evidence of the applicability of PPR. A consistent program of support needs to be developed for students at high risk of failure when they enter the program. In addition, these data would allow for more targeted studies using area under the curves and conditional inference trees.

Conclusions

To our knowledge, this is the first study to identify six significant predictors of passing NMLE at the first attempt and the possibility of decreasing the number of NMLE failures prospectively using PPR which was developed by a logistic regression formula. Adopting a similar approach in other MSs may help address a major challenge regarding medical schools' performance, and help reduce the failure rate in national exams. To confirm these results, however, further studies are needed because of no similar trial until this point. To take measures against the failures, the next step, we immediately expect to move into action (i.e., repeat a course).

Data availability
Our data are not on a data repository. The datasets used and/or analyzed during the current study are available from the corresponding author upon reasonable request. Only coded data may be shared.

Declarations

Ethics approval and consent to participate
Ethical approval was granted by the GUSM Ethics Committee. Anonymity and confidentiality for students were guaranteed (date: 05/13/2020, reference number: 2020-039). The participants of this study had provided comprehensive consent by opting out, and Informed Consent obtained from participant.

Consent for publication
(Not applicable)

References

1. Casey PM, Palmer BA, Thompson GB, et al. Predictors of medical school clerkship performance: a multispecialty longitudinal analysis of standardized examination scores and clinical assessments. BMC Med Educ. 2016;16:128.
2. USMLE. United States Medical Licensing Examination. www.usmle.org. Accessed Nov 24, 2023.
3. Medical Council of Canada. Medical Council of Canada Qualifying Examination Part I. https://mcc.ca/examinations/mccqe-part-i. Accessed Nov 24, 2023.
4. Zavlin D, Jubbal KT, Noé JG, Gansbacher B. A comparison of medical education in Germany and the United States: from applying to medical school to the beginnings of residency. GMS German Med Sci. 2017;15:Doc15.
5. Bonvin R, Nendaz M, Frey P, et al. Looking back: twenty years of reforming undergraduate medical training and curriculum frameworks in Switzerland. GMS J Med Educ. 2019;36:Doc64.
6. Janczukowicz J. Medical education in Poland. Med Teach. 2013;35:537–43.
7. General Medical Council. Medical licensing assessment. https://www.gmc-uk.org/education/medical-licensing-assessment. Accessed Nov 24, 2023.
8. Baars GJA, Stijnen T, Splinter TAW. A model to predict student failure in the first year of the undergraduate medical curriculum. Health Professions Educ. 2017;3:5–14.
9. Coumarbatch J, Robinson L, Thomas R, Bridge PD. Strategies for identifying students at risk for USMLE step 1 failure. Fam Med. 2010;42:105–10.
10. Winnie WU, Garcia K, Chandrahas S, et al. Predictors of performance on USMLE step 1. Southwest Respiratory Crit Care Chronicles. 2021;9:63–72.
11. Tsunekawa K, Suzuki Y, Shioiri T. Identifying and supporting students at risk of failing the national medical licensure examination in Japan using a predictive pass rate. BMC Med Educ. 2020;20:419.
12. Murtaugh PA, Burns LD, Schuster J. Predicting the retention of medical students. Res High Educt. 1999;40:355–71.
13. Arulampalam W, Naylor R, Smith J. Factors affecting the probability of first year medical student dropout in the UK: a logistic analysis for the intake cohorts of 1980-92. Med Educ. 2004;38:492–503.
14. Peng P, Yang WF, Liu Y, et al. High prevalence and risk factors of dropout intention among Chinese medical postgraduates. Med Educ Online. 2022;27:2058866.
15. Kies SM, Freund GG. Medical students who decompress during the M-1 year outperform those who fail and repeat it: a study of M-1 students at the university of Illinois college of medicine at Urbana-Champaign 1988–2000. BMC Med Educ. 2005;5:18.
16. Suzuki Y, Gibbs T, Fujisaki K. Medical education in Japan: a challenge to the healthcare system. Med Teach. 2008;30:846–50.
17. Imafuku R, Saiki T, Suzuki Y. Developing undergraduate research in Japanese medical education. Council Undergrad Res Q. 2016;37:34–40.
18. United States Medical Licensing Examination. https://www.usmle.org/step-exams Accessed Jul 1, 2024.
19. Organization for Economic Co-operation and Development. OECD Data: medical graduates. https://data.oecd.org/healthres/medical-graduates.htm. Accessed Nov 24, 2023.
20. Organization for Economic Co-operation and Development. OECD Data: doctors. https://data.oecd.org/healthres/doctors.htm. Accessed Nov 24, 2023.
21. Monteiro KA, George P, Dollase R, Dumenco L. Predicting United States medical licensure examination step 2 clinical knowledge scores from previous academic indicators. Adv Med Educ Pract. 2017;8:385–91.
22. Wang L, Laird-Fick HS, Parker CJ, Solomon D. Using Markov chain model to evaluate medical students' trajectory on progress tests and predict USMLE step 1 scores—a retrospective cohort study in one medical school. BMC Med Educ. 2021;21:200.
23. Matsui T, Sato M, Kato Y, Nishigori H. Professional identity formation of female doctors in Japan – gap between the married and unmarried. BMC Med Educ. 2019;19:55.
24. Tokuda Y, Goto E, Otaki J, et al. Undergraduate educational environment, perceived preparedness for postgraduate clinical training, and pass rate on the national medical licensure examination in Japan. BMC Med Educ. 2010;10:35.
25. Mitchell K, Haynes R, Koenig J. Assessing the validity of the updated medical college admission test. Acad Med. 1994;69:394–401.
26. Swanson DB, Case SM, Koenig J, Killian CD. Preliminary study of the accuracies of the old and new medical college admission tests for predicting performance on USMLE Step 1. Acad Med. 1996;71:S25–7.
27. Wiley A, Koenig JA. The validity of the medical college admission test for predicting performance in the first two years of medical school. Acad Med. 1996;71:S83–5.
28. Koenig JA, Sireci SG, Wiley A. Evaluating the predictive validity of MCAT scores across diverse applicant groups. Acad Med. 1998;73:1095–106.
29. Julian ER. Validity of the medical college admission test for predicting medical school performance. Acad Med. 2005;80:910–7.
30. Zhao X, Oppler S, Dunleavy D, Kroopnick M. Validity of four approaches of using repeaters' MCAT scores in medical school admissions to predict USMLE Step 1 total scores. Acad Med. 2010;85:S64–7.
31. O'Neill L, Hartvigsen J, Wallstedt B, Korsholm L, Eika B. Medical school dropout–testing at admission versus selection by highest grades as predictors. Mededical Educ. 2011;45:1111–20.
32. Donnon T, Paolucci EO, Violato C. The predictive validity of the MCAT for medical school performance and medical board licensing examinations: a meta-analysis of the published research. Acad Med. 2007;82:100–6.
33. Ramsbottom-Lucier M, Johnson MM, Elam CL. Age and gender differences in students' preadmission qualifications and medical school performances. Acad Med. 1995;70:236–9.
34. McDougle L, Mavis BE, Jeffe DB, et al. Academic and professional career outcomes of medical school graduates who failed USMLE Step 1 on the first attempt. Adv Health Sci Education: Theory Pract. 2013;18:279–89.
35. Nieuwoudt JE, Pedler ML. Student retention in higher education: why students choose to remain at university. J Coll Student Retention: Res Theory Pract. 2023;25:326–49.
36. Steyerberg EW, Vergouwe Y. Towards better clinical prediction models: seven steps for development and an ABCD for validation. Eur Heart J. 2014;35:1925–31.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.