



ELSEVIER

Educación Médica

www.elsevier.es/edumed



ORIGINAL

Puntajes en pruebas de progreso como predictores del desempeño en el Examen Nacional de Medicina del Perú



Franco Romaní* y César Gutiérrez

Facultad de Medicina Humana, Universidad de Piura, Lima, Perú

Recibido el 11 de abril de 2025; aceptado el 7 de julio de 2025

Disponible en Internet el 5 de agosto de 2025

PALABRAS CLAVE

Aprendizaje basado en problemas;
Educación médica;
Rendimiento académico;
Estudiantes de Medicina;
Habilidades para tomar exámenes;
Perú

Resumen

Introducción: las pruebas de progreso son evaluaciones sumativas que permiten monitorizar la evolución del aprendizaje en estudiantes de Medicina. Su uso como predictor del desempeño en exámenes nacionales de licenciamiento ha sido estudiado principalmente en universidades de los EE. UU. Este estudio evaluó si el rendimiento y la trayectoria de puntajes en múltiples pruebas de progreso predicen el desempeño en el Examen Nacional de Medicina del Perú (ENAM).

Métodos: se realizó un estudio de cohorte retrospectiva con información de estudiantes de una escuela de medicina en Lima, Perú. Se utilizaron datos de la cohorte de ingresantes de 2017 y 2018 que rindieron el ENAM en 2023 y 2024. Las variables predictoras fueron el puntaje en las pruebas de progreso, el promedio ponderado al sexto año, el número de pruebas de progreso y el sexo. Aplicamos modelos lineales generalizados (GLM, por sus siglas en inglés) para predecir el puntaje en el ENAM.

Resultados: se analizaron los datos de 51 estudiantes. El promedio de las 5 primeras pruebas de progreso mostró correlación lineal positiva con el puntaje en el ENAM ($r = 0,544$; IC 95%: 0,316-0,713); además, el promedio ponderado estuvo correlacionado con el puntaje del ENAM ($r = 0,551$; IC 95%: 0,325-0,718). Por cada unidad de incremento en el promedio de las 5 primeras pruebas de progreso, el puntaje en el ENAM aumentó en 0,51 (IC 95%: 0,29-0,73).

Conclusión: el rendimiento en las pruebas de progreso en una escuela de medicina del Perú predice de manera positiva y significativa el puntaje obtenido en el ENAM.

© 2025 El Autor/Los Autores. Publicado por Elsevier España, S.L.U. Este es un artículo Open Access bajo la licencia CC BY-NC (<http://creativecommons.org/licenses/by-nc/4.0/>).

* Autor para correspondencia.

Correo electrónico: franco.romani@udep.edu.pe (F. Romaní).

KEYWORDS
 Problem-Based learning;
 Medical education;
 Academic performance;
 Medical Students;
 Test taking skills;
 Peru

Progress test scores as predictors of performance in the Peruvian National Medical Examination

Abstract

Introduction: Progress tests are summative assessments used to monitor the evolution of learning in medical students. Their use as predictors of performance on national licensing exams has been mainly studied in U.S. universities. This study evaluated whether performance and score trajectories in multiple progress tests predict outcomes on the Peruvian National Medical Exam (ENAM).

Methods: We conducted a retrospective cohort study using data from medical students at a school in Lima, Peru. We analyzed cohorts that entered in 2017 and 2018 and took the ENAM in 2023 and 2024. Predictor variables included progress test scores, sixth-year GPA, number of progress tests taken, and sex. We applied generalized linear models (GLMs) to predict ENAM scores.

Results: We analyzed data from 51 students. The average score of the first five progress tests showed a positive linear correlation with ENAM performance ($r = 0.544$; 95% CI: 0.316 to 0.713). Additionally, the GPA was correlated with ENAM scores ($r = 0.551$; 95% CI: 0.325 to 0.718). For each one-point increase in the average of the first five progress tests, the ENAM score increased by 0.51 points (95% CI: 0.29 to 0.73).

Conclusion: Performance on progress tests in a Peruvian medical school is a significant positive predictor of ENAM scores.

© 2025 The Author(s). Published by Elsevier España, S.L.U. This is an open access article under the CC BY-NC license (<http://creativecommons.org/licenses/by-nc/4.0/>).

Introducción

Las escuelas de medicina forman y evalúan conocimientos y competencias médicas en sus estudiantes; posteriormente, se requiere de una evaluación rigurosa y transparente de las competencias profesionales necesarias para la práctica médica. Estas evaluaciones son denominadas Exámenes Nacionales de Licenciamiento (ENL)¹, ejemplos internacionalmente reconocidos son el Examen de Licenciamiento Médico de los Estados Unidos de América (USMLE, por sus siglas en inglés), el Examen de Calificación del Consejo Médico de Canadá o la Evaluación del Licenciamiento Médico del Reino Unido.

Los puntajes de los ENL están asociados con factores académicos, a su vez, estos factores pueden ser modificables o no modificables. Entre los factores potencialmente modificables, se ha visto que el puntaje en el segundo USMLE (*step 2*) está asociado positivamente con el puntaje del primer USMLE (*step 1*), con el promedio académico ponderado, con las calificaciones en exámenes clínicos objetivos estructurados (ECOE), entre otras evaluaciones sumativas^{2,3}. Un factor no modificable, pero positivamente asociado con el primer y segundo USMLE, es el puntaje en el examen de admisión a las escuelas de medicina⁴.

Las pruebas de progreso (PP) en las escuelas de medicina facilitan que los estudiantes integren nuevos conocimientos con aprendizajes previos, mediante su aplicación en escenarios clínicos^{5,6}. Las PP evalúan progresivamente el conocimiento acumulado en relación con los objetivos del plan de estudio⁷. Esta evaluación supera el enfoque fragmentado y la enseñanza por departamentos en que se organizan las escuelas de medicina^{8,9}; debido a que su formulación se basa en la

progresión de contenidos según el plan de estudio, brindan a los estudiantes una monitorización sobre su ruta de aprendizaje^{10,11} y perfilan sus expectativas al momento de graduarse¹². La exposición repetitiva a PP constituye una preparación integral, además, proporciona retroalimentación sobre las áreas que requieren mejora⁷. Por ello, es plausible plantear que el desempeño en las PP podría ser un predictor independiente del rendimiento en un futuro ENL.

Diversos estudios han evaluado esta hipótesis; uno de los primeros en la Universidad de McMaster, en Canadá, encontró que la implementación de PP incrementó los puntajes del ENL de sus estudiantes¹³. Otro estudio en la Universidad de Michigan, EE. UU., encontró que los estudiantes con un peor desempeño según su trayectoria de puntajes en las PP, tuvieron un menor incremento en el puntaje del primer USMLE (*step 1*)¹². En contextos donde se aplican las Evaluaciones Integrales de Ciencias Básicas (CBSE, por sus siglas en inglés) o las Autoevaluaciones Integrales de Ciencias Clínicas (CCSSA, por sus siglas en inglés) se ha reportado que estas modalidades de PP son un buen predictor del segundo USMLE (*step 2*)³.

En Perú, el Examen Nacional de Medicina (ENAM) se aplica desde 2003 como una evaluación tipo ENL que mide conocimientos en ciencias básicas, clínicas y salud pública en estudiantes por egresar. Respecto a las PP, a diferencia de los EE. UU. y Canadá que disponen del CBSE o CCSSA administradas por organizaciones o asociaciones reconocidas para garantizar que los médicos en formación posean los conocimientos y habilidades necesarios para la práctica clínica, en Perú estas evaluaciones no están implementadas de manera sistemática y estandarizada para todas las escuelas de medicina. Las universidades de los EE. UU.^{12,14}

y otros países^{15,16} utilizan estas evaluaciones integrales de ciencias básicas y clínicas como una PP, mientras que en Perú las PP son iniciativas institucionales y aisladas^{17,18}.

En Perú, se han realizado diversos estudios para identificar los factores asociados al puntaje del ENAM. El promedio ponderado ha sido el factor más estudiado y está positivamente asociado al resultado del ENAM^{17,19,20}. Sin embargo, el rol predictor de los resultados en sucesivas PP en estudiantes de Medicina peruanos no ha sido evaluado. Por ello, el objetivo de nuestro estudio fue analizar si el rendimiento y la trayectoria de puntajes en múltiples pruebas de progreso predicen el desempeño en el ENAM en estudiantes de una escuela de medicina del Perú.

Material y métodos

Diseño y ámbito de estudio

Aplicamos un diseño de cohorte retrospectiva con datos generados por la gestión académica de una escuela de medicina humana ubicada en Lima, Perú. Esta escuela tuvo su primera cohorte de ingresantes en 2017 y, hasta diciembre de 2024, tuvo 2 cohortes (ingresantes de 2017 y 2018) que culminaron el internado médico (7.º año de la carrera) y rindieron el ENAM.

Durante el periodo de estudio, las cohortes señaladas pudieron realizar 6 PP y un ENAM. La PP es usada en esta escuela de medicina como una evaluación sumativa. Esta evaluación contiene 250 preguntas de opción múltiple con una respuesta correcta; desde 2017 hasta 2021 fueron utilizados 4 distractores, y de 2022 en adelante fueron usados 3 distractores. La prueba comprende 3 cuadernillos, cada uno contiene entre 4 y 6 casos clínicos, extraídos de reportes de casos publicados en revistas medidas arbitradas. A partir de los casos clínicos, se formulan preguntas de asignaturas que el estudiante ha completado hasta la fecha del examen. Esta es una evaluación anual, tomada en fecha única en la segunda semana de diciembre, al finalizar el año académico. La participación de los estudiantes es voluntaria¹⁸.

Las cohortes de ingresantes 2017 y 2018 rindieron el ENAM 2023 y 2024, respectivamente. El ENAM está dirigido a los estudiantes de Medicina del último año (internado médico), así como a los médicos titulados en el Perú y el extranjero, es usualmente aplicada en diciembre de cada año²¹. El ENAM es diseñado y conducido por la Asociación Peruana de Facultades de Medicina (ASPEFAM).

Participantes

El criterio de inclusión fue: estudiante de Medicina Humana ingresante en las cohortes 2017 y 2018, y que haya rendido el ENAM de los años 2023 y 2024. Fueron excluidos aquellos que no rindieron PP entre 2017 y 2024. Realizamos el análisis con los datos de toda la población de estudio, no fue requerido un diseño muestral.

Definición de variables

La variable predictora de interés fue el puntaje en las PP. Este puntaje varía de 0 a 250 puntos con base en el número

de respuestas correctas; este puntaje fue convertido a escala vigesimal. Las respuestas incorrectas no restaron puntos. La primera cohorte (2017) rindió las PP de los años 2017 a 2022, mientras que la segunda (2018) rindió las PP de 2018 a 2023. Debido a que la prueba es voluntaria, no todos los estudiantes rindieron las 6 PP posibles. El número de PP rendidas fue considerada como una variable. Además, identificamos el número de orden de la PP rendida, desde la primera hasta la sexta posible.

Otra variable predictora fue el promedio ponderado al completar el sexto año de estudio, es decir, sin considerar la calificación del internado médico. El promedio ponderado fue construido a partir de las calificaciones registradas en el sistema de gestión académica de la escuela. Este promedio se obtuvo ponderando las calificaciones de las asignaturas por el número de créditos. Este puntaje tiene una escala vigesimal.

La variable de respuesta fue el puntaje en el ENAM, y fue reportada también en una escala vigesimal. Usamos la calificación calculada por la ASPEFAM y entregada a la escuela de medicina.

Análisis de datos

Realizamos un análisis descriptivo de los participantes, según sexo y número de PP rendidas mediante frecuencias. Los puntajes de las PP, el promedio ponderado y el ENAM fueron analizadas en su distribución con la prueba de Shapiro-Wilk, en caso de distribución normal utilizamos medias y desviaciones estándar (DE), caso contrario mediana y rango intercuartílico.

Para evaluar la asociación cruda de las variables explicativas con el puntaje del ENAM aplicamos comparación de medias con la prueba t de Student y el análisis de variancia de un factor, según corresponda. Para ambos casos aplicamos la comprobación de los supuestos de homogeneidad de variancias con la prueba de Levene y normalidad de los residuos con la prueba Shapiro-Wilk. Se estimó el coeficiente de correlación de Pearson para evaluar correlaciones en los puntajes.

Debido a que los puntajes de la PP son de carácter longitudinal y corresponden a mediciones repetidas, realizamos un análisis de clasificación de tendencias basado en coeficientes de regresión lineal de cada estudiante. Se clasificó a los estudiantes en 3 grupos de tendencias: creciente, estable y decreciente. Este análisis fue realizado en STATA 16, con los comandos «`statsby beta=_b[tiempo], by(id) clear:regress puntaje tiempo`» y «`cluster kmeans beta, k(3) name (cluster_trends)`». Comparamos la media del ENAM entre los grupos de tendencias usando el análisis de variancia de un factor.

Aplicamos modelos lineales generalizados (GML, por sus siglas en inglés) para predecir el puntaje en el ENAM, considerando como predictores el promedio del puntaje en las 5 primeras PP, y ajustamos el modelo por el promedio ponderado al sexto año. Para este fin, usamos modelos con función de enlace de identidad, y desenlace de distribución gaussiana, además, la constante del modelo fue centrada. Valoramos el ajuste de los modelos con el seudo R². Consideramos el promedio de las 5 PP, pues 28 estudiantes no rindieron la sexta PP.

El análisis estadístico descriptivo y los GML fueron realizados en Jamovi versión 2.4.8.0. Además, replicamos el GML para predecir el puntaje del ENAM, según el predictor

promedio ponderado para cada agrupación de tendencia en los puntajes de la PP. Los gráficos fueron elaborados en GraphPad Prism 10.4.1. El nivel de significación estadística fue establecido en 5%.

Aspectos éticos

El estudio fue aprobado por el Comité Institucional de Ética en Investigación de la universidad. Solicitamos los puntajes de las PP, del ENAM, el sexo del estudiante y el promedio ponderado a la escuela de medicina. La base incluyó el código del estudiante, que fue reemplazado por un identificador para realizar el análisis estadístico en una base de datos anónimizada.

Resultados

Características de la muestra

Analizamos los datos de 51 estudiantes, 28 de ellos fueron varones (54,9%). Los estudiantes rindieron entre 3 y 6 PP: uno (2,0%) rindió 3 pruebas, 4 (7,8%) rindieron 4 pruebas, 27 (52,9%) rindieron 5 pruebas y 19 (37,3%) dieron 6 pruebas. Veintiséis (51,0%) estudiantes realizaron el ENAM en 2023 y 25 (49,0%) en 2024. La media del promedio ponderado fue 15,1 con una desviación estándar (DE) de 0,87. Respecto a los puntajes del ENAM, la media fue 15,3 (DE=1,30), el puntaje mínimo fue 12,7 y el máximo 17,9. La media de las 5 primeras PP fue 11,1 (DE=1,38).

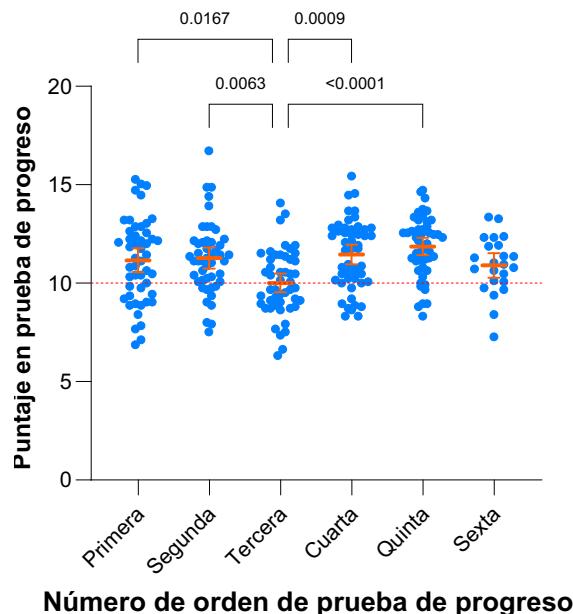
Considerando la progresión en la rendición de las PP, en la quinta y cuarta PP los estudiantes lograron las mayores medias de puntaje, con 11,9 (DE=1,53) y 11,5 (DE=1,76), respectivamente. En la tercera PP, la media fue 10,0 (DE=1,66). Se corroboraron diferencias significativas entre las medias de los puntajes, según el número de orden de la PP (fig. 1).

Variables asociadas al puntaje en el ENAM a nivel crudo

La media del puntaje en el ENAM, según sexo, no tuvo diferencias significativas: 15,4 (DE = 1,21) en varones y 15,1 (DE = 1,41) en mujeres (t de Student = 0,764, p =0,449). Tampoco hubo diferencias en las medias según el número de PP rendidas: 15,3 (DE=0,56) en los que dieron 3 o 4 pruebas; 15,4 (DE=1,24) en los que dieron 5 pruebas y 15,0 (1,52) en los que dieron 6 (F =0,322, p =0,729).

El promedio de las 5 primeras PP tuvo una correlación lineal positiva con los puntajes en el ENAM (r = 0,544; IC 95%: 0,316-0,713), observamos que el promedio ponderado del estudiante estuvo correlacionado con el puntaje del ENAM (r = 0,551; IC 95%: 0,325-0,718). Sin embargo, al evaluar la concordancia, encontramos que los puntajes en el ENAM estuvieron en promedio 4,14 puntos por encima del promedio de las PP (IC 95%: 1,62-6,65), mientras que el promedio ponderado tuvo una mejor concordancia con los puntajes del ENAM (diferencia = 0,135; IC 95%: -2,01-2,28) (fig. 2).

La tercera y cuarta PP tuvieron los coeficientes de correlación más grandes respecto al puntaje del ENAM y el promedio ponderado. En toda la muestra, y considerando el



Número de orden de prueba de progreso

Figura 1 Gráfico de puntos según secuencia de pruebas de progresos rendidas y comparación de medias entre grupos. Líneas centrales y barras de error representan media e intervalo de confianza al 95%. Las barras entre los conjuntos de datos muestran aquellas comparaciones significativas con la prueba de comparaciones múltiples de Tukey (valores de p). Prueba ANOVA (F = 6,126; p < 0,001). La línea roja entrecortada indica la mitad del puntaje posible en una escala vigesimal.

promedio de las 5 primeras PP, encontramos una correlación lineal directa con el promedio ponderado (r = 0,842; p < 0,001). En general, los puntajes de las PP tuvieron una correlación lineal directa más intensa con el promedio ponderado (fig. 3).

Clasificamos a los estudiantes en 3 grupos: 11 (21,6%) como tendencia creciente, 27 (52,9%) como estables y 13 (25,5%) como decrecientes. Entre los estudiantes con tendencia creciente la media (DE) del ENAM fue 14,8 (1,34), en los estables fue 15,3 (1,34) y en los de tendencia decreciente fue 15,6 (1,17). No hubo diferencias significativas entre las medias (F = 1,00; p = 0,382). La media de la primera PP fue mayor entre los estudiantes de patrón decreciente (13,3 [1,62]) comparados con los estables (11,2 [1,51]) y crecientes (8,86 [1,13]) (F = 30,5; p < 0,001).

Modelo lineal generalizado

En el modelo 1, por cada unidad de incremento en el promedio de las 5 primeras PP, un estudiante incrementó en 0,51 el puntaje del ENAM (IC 95%: 0,29-0,73). El promedio ponderado al sexto año también tuvo una asociación positiva y significativa con el ENAM. Sin embargo, cuando ambas variables fueron incluidas simultáneamente, ninguna fue un predictor significativo. En el modelo 4, el promedio ponderado tuvo una asociación positiva y significativa con el puntaje del ENAM en el grupo estable (β = 1,36; IC 95%: 0,77-1,96; p < 0,001) y en el grupo decreciente (β = 0,59; IC 95%: 0,07-1,11; p = 0,026). En el grupo creciente, aunque el coeficiente fue positivo (β = 0,49), no hubo significación estadística

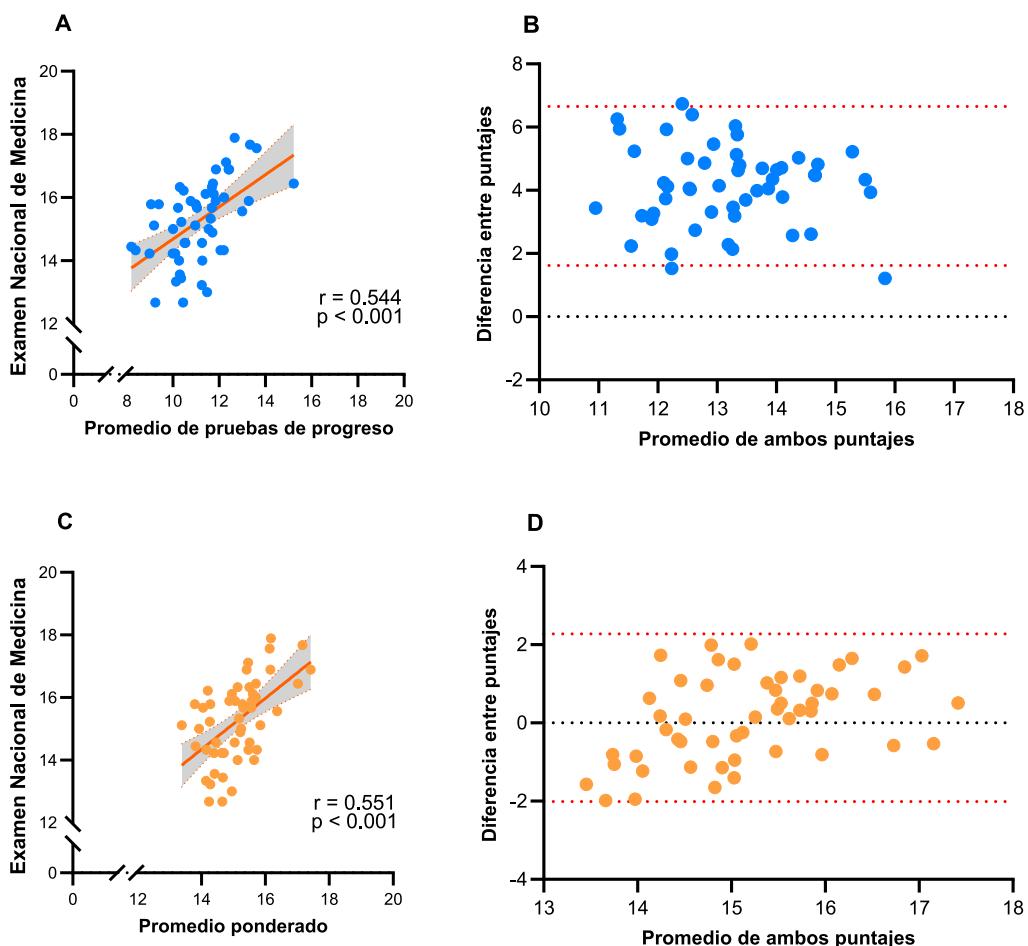


Figura 2 A) Gráfico de dispersión de los puntajes promedio de las pruebas de progreso y el Examen Nacional de Medicina. B) Gráfico de Bland–Altman para evaluar la concordancia de los puntajes de las pruebas de progreso y el Examen Nacional de Medicina (coeficiente de correlación de concordancia = 0,093; IC 95%: 0,041–0,144). C) Gráfico de dispersión de los puntajes del promedio ponderado y el Examen Nacional de Medicina. D) Gráfico de Bland–Altman para evaluar la concordancia de los puntajes del promedio ponderado y el Examen Nacional de Medicina (coeficiente de correlación de concordancia = 0,507; IC 95%: 0,299–0,669). Las líneas rojas entrecortadas en los gráficos de Bland–Altman reflejan los límites inferior y superior de los intervalos de confianza al 95%.

(tabla 1). En un modelo con el patrón de tendencias de las PP y el promedio ponderado como predictores, encontramos que el patrón de tendencia en las PP no estuvo asociado con el puntaje del ENAM (material suplementario: tabla 1).

Discusión

En este estudio, el promedio de las 5 primeras PP fue un predictor independiente del rendimiento en el ENAM. Sin embargo, la trayectoria de los puntajes en las PP no resultó ser un predictor significativo de dicho rendimiento. Encontramos que el tipo de trayectoria modificó el efecto predictor del promedio ponderado. El mayor efecto en el ENAM por unidad de incremento del promedio ponderado se dio entre los estudiantes con un patrón estable en los puntajes de la PP, mientras que en los estudiantes con un patrón creciente el rol predictor desapareció.

Nuestros hallazgos son consistentes con estudios que analizaron la asociación entre rendimiento en las PP y el ENL. La mayoría de los estudios en los EE. UU. ha evaluado la

correlación entre el resultado del CBSE y el primer y segundo USMLE (step 1 y step 2). En la Universidad de Florida Central, el coeficiente de correlación entre el CBSE y el primer USMLE (evaluación de ciencias básicas) varió de 0,85 en la primera PP hasta 0,78 en la quinta PP¹⁴. En otro estudio, los puntajes de estudiantes de 62 escuelas de medicina de los EE.UU., entre 2011 y 2014, encontraron que un examen de medicina administrado por el *National Board of Medical Examiners* (NBME) predijo independientemente el rendimiento en el segundo USMLE (evaluación de ciencias clínicas)²². Este hallazgo también fue observado en ámbitos diferentes, en la Universidad Estatal Paulista, en Brasil, el rendimiento en 6 PP tuvo correlaciones moderadas o fuertes (coeficientes rho de 0,402 a 0,570, dependiendo de la cohorte) con el puntaje de la evaluación de ciencias básicas del ENL de Brasil. Además, hubo una clara tendencia en la proporción de desaprobados en el ENL, según rendimiento en las PP: 57,9% en estudiantes con puntajes de PP por debajo de la desviación estándar, 30,8% en aquellos dentro de la desviación estándar y 8,3% entre los estudiantes por encima de la desviación estándar²³.

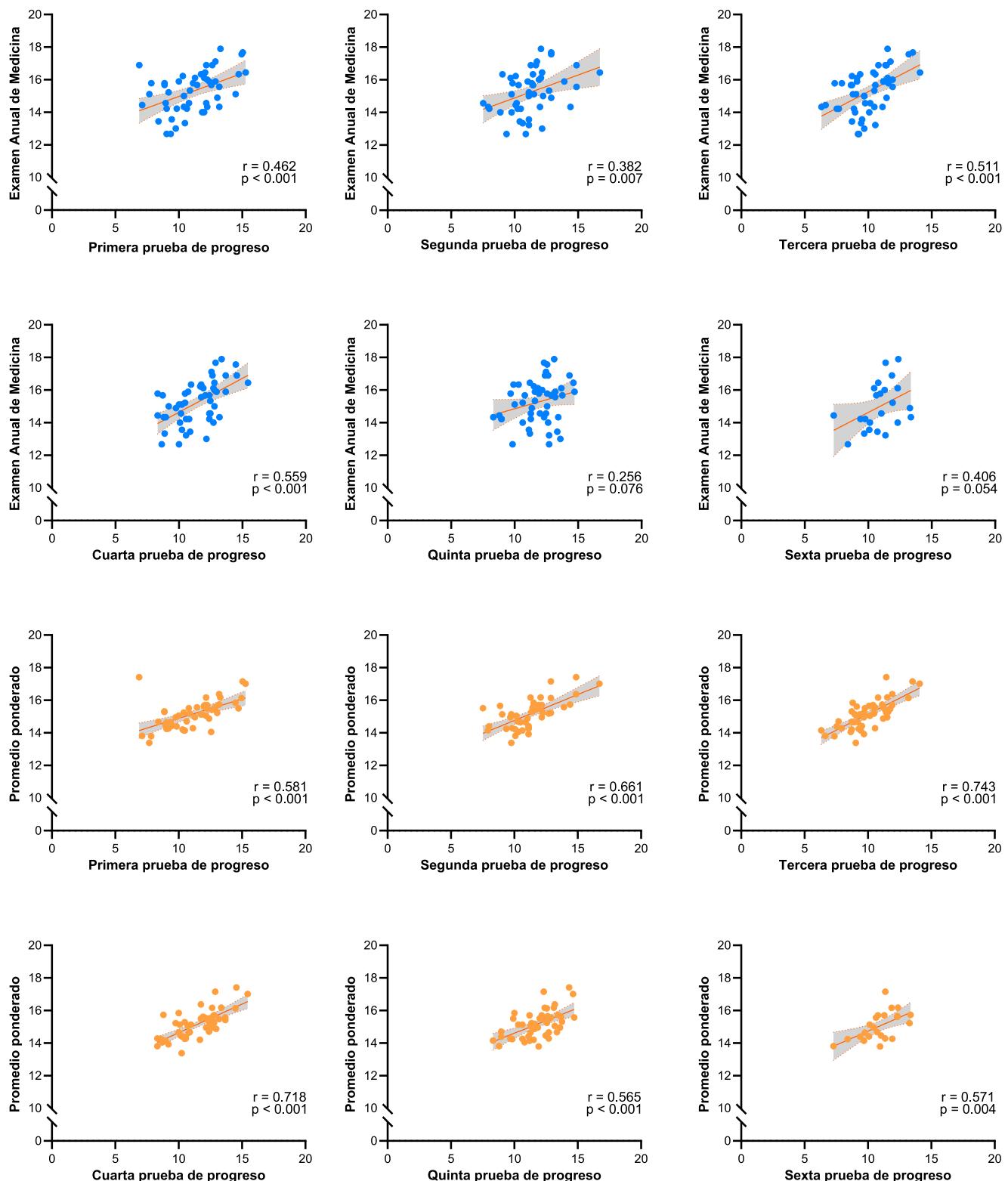


Figura 3 Matriz de correlaciones entre las 6 pruebas de progreso con los puntajes en el Examen Nacional de Medicina y el promedio ponderado.

Otros estudios han brindado evidencia adicional sobre esta asociación. Un estudio de intervención demostró que, luego de la implementación de una PP en la Universidad de McMaster, los puntajes en el ENL de Canadá tuvieron un

incremento significativo, con un cambio en el sentido de la pendiente, de negativo a positivo, y con un mayor incremento a medida que más cohortes fueron expuestas a esta PP¹³. En nuestro estudio, la PP fue implementada desde

Tabla 1 Modelo lineal generalizado para predecir el puntaje del Examen Nacional de Medicina teniendo como predictores los puntajes en las pruebas de progreso y el promedio ponderado

Variables	Modelo 1		Modelo 2		Modelo 3		Modelo 4					
	β (IC 95%)	P	β (IC 95%)	P	β (IC 95%)	P	Creciente	Estable	Decreciente			
							β (IC 95%)	P	β (IC 95%)	P	β (IC 95%)	P
Constante	15,25 (14,95-15,55)	< 0,001	15,25 (14,95-15,55)	< 0,001	15,25 (14,95-15,55)	< 0,001	14,85 (14,08-15,62)	< 0,001	15,25 (14,87-15,63)	< 0,001	15,59 (15,04-16,14)	< 0,001
Promedio de pruebas de progreso	0,51 (0,29-0,73)	< 0,001	—	—	0,26 (-0,15-0,66)	—	—	—	—	—	—	—
Promedio ponderado	—	—	0,82 (0,47-1,16)	—	0,47 (-0,16-1,11)	—	0,49 (-0,29-1,27)	—	1,36 (0,77-1,96)	—	0,59 (0,07-1,11)	—
Observaciones	51	51	51	51	51	51	11	27	27	13	13	13
R ²	0,296	0,303	0,326	0,143	0,143	0,143	0,448	0,448	0,448	0,312	0,312	0,312

Modelo 1: incluyó solo el promedio de las 5 primeras pruebas de progreso. Modelo 2: incluyó el promedio ponderado al 6.º año. Modelo 3: incluyó el promedio de las 5 primeras pruebas de progreso y el promedio ponderado al 6.º año (Diagnóstico de colinealidad con el VIF: promedio de pruebas de progreso [VIF = 3,43] y promedio ponderado [VIF = 3,43]). En el modelo 4: la única predictora fue el promedio ponderado al 6.º año.

los orígenes de la escuela de medicina, por lo tanto, no contamos con un control histórico; sin embargo, en las 2 cohortes analizadas no hubo desaprobados en el ENAM.

La aplicación regular de las PP genera puntajes longitudinales por estudiante, lo que permite analizar la trayectoria del rendimiento en las PP. Un estudio en la Universidad Estatal de Michigan evaluó 10 PP (CBSE), definiendo 4 patrones de tendencia. Este análisis encontró que los estudiantes con patrones ascendentes lograron mayores aumentos de puntaje en el primer USMLE¹². En esta misma universidad, se identificaron trayectorias de los puntajes del NBME, aplicado a estudiantes de primer año, y se encontró que tener una primera puntuación intermedia, pero una pendiente de crecimiento pronunciada, predijo mayores puntajes en el primer USMLE²⁴. Otro estudio, en la Universidad de Colonia (Alemania), observó que el desempeño en el ENL estuvo influenciado en mayor medida por la trayectoria de puntajes de las PP de los años de cursos de ciencias clínicas, y en menor medida por los primeros años de preclínicas²⁵.

En nuestro análisis, no corroboramos que la trayectoria de puntajes en las primeras 5 PP estuviera asociada con el desempeño en el ENAM. Sin embargo, observamos que aquellos con patrón decreciente tuvieron una tendencia no significativa de incrementar el puntaje del ENAM, comparados con los de tendencia creciente. Esta aparente contradicción se debe a que los estudiantes con un patrón decreciente iniciaron con un mayor puntaje en la PP, mientras que los estudiantes con tendencia creciente tuvieron un peor desempeño en la primera PP, lo que les permitió un mayor margen de mejora en sus puntajes. Al estratificar el modelo según el patrón de puntajes, el promedio ponderado predijo positivamente el puntaje en el ENAM solo en estudiantes con patrón estable o decreciente, pero no en los de patrón creciente.

Nuestros hallazgos muestran que, desde la primera PP, esta evaluación tiene una correlación moderada con el promedio ponderado al sexto año (coeficiente rho = 0,581), y que la magnitud de la correlación fue mayor con los

puntajes de la tercera y cuarta PP (coeficientes de 0,743 y 0,718, respectivamente). Esta correlación entre predictores explicaría que cuando ambas variables son ingresadas a la regresión para modelar el puntaje ENAM, ninguna resultara significativa. En el análisis de regresión, considerando solo el promedio ponderado al sexto año, esta variable generó un mayor incremento en el puntaje ENAM comparado con el promedio de las PP. En estudios realizados en Perú, se ha demostrado una correlación lineal positiva moderada o fuerte entre el promedio ponderado y el puntaje del ENAM¹⁹. Una asociación similar ha sido observada entre el GPA (Grade Point Average) y los resultados en los 3 pasos del USMLE, tanto a nivel crudo⁴ como ajustado².

Según nuestros hallazgos, el promedio ponderado al sexto año y el promedio de las PP presentan una correlación de similar magnitud con el puntaje del ENAM. Sin embargo, el promedio ponderado tuvo una concordancia significativa de puntajes con el ENAM, mientras que la PP tuvo en promedio 4,14 puntos menos que el puntaje obtenido por los estudiantes en el ENAM. Este hallazgo refleja la dificultad promedio de las PP respecto al ENAM. Este ENL tiene preguntas con un índice de dificultad promedio de 0,60, lo que podría considerarse como mediano bajo la siguiente clasificación: fácil de 0,66 a 1,00, mediano de 0,34 a 0,65 y difícil de 0 a 0,33.

No corroboramos nuestra hipótesis: que a mayor número de PP, mayor puntaje en el ENAM. Esta falta de asociación también ha sido observada en estudiantes de los EE. UU., pues el número de NBME en ciencias clínicas no tuvo asociación con el puntaje en el segundo USMLE²². Esto podría deberse a que en nuestra escuela de medicina las PP no son obligatorias, y muchos estudiantes optan por no rendirlas o no prepararse para esta evaluación. Un estudio mostró que solo el 45,5% consideraba importante prepararse para estas pruebas²⁶, mientras que la preparación para el ENAM puede ser más exhaustiva y específica, optando por medidas alternativas y extracurriculares –como academias de preparación–.

Nuestro estudio tiene limitaciones. La decisión de rendir las PP es voluntaria, por ello, no todos los estudiantes realizaron las 6 PP posibles antes del internado médico; esto pudo afectar la clasificación de los estudiantes según las tendencias de sus puntajes. Solo incluimos a aquellos estudiantes que culminaron el plan de estudio y rindieron el ENAM; esto puede limitar la posibilidad de generalizar los resultados a estudiantes irregulares que culminaron los estudios en más de 7 años cronológicos, y que suelen tener un menor rendimiento académico. La fuente de información limitó la posibilidad de estudiar otros predictores, como el puntaje del examen de ingreso, el tipo de colegio secundario de procedencia, o la preparación en academias que brindan entrenamiento específico para el ENAM.

Concluimos que el promedio de las 5 primeras PP fue un predictor positivo del rendimiento en el ENAM: por cada punto adicional en dicho promedio, el puntaje del ENAM aumentó en 0,51 puntos. Además, el patrón de desempeño en las PP moduló el rol predictor del promedio ponderado al sexto año; específicamente, entre los estudiantes con un patrón de rendimiento estable, se observó un mayor impacto del promedio ponderado sobre el puntaje del ENAM. Finalmente, los estudiantes en promedio obtuvieron 4,14 puntos adicionales en el ENAM por sobre el promedio en sus PP.

Responsabilidades éticas

El estudio fue aprobado por el Comité Institucional de Ética en Investigación de la Universidad de Piura. El estudio no requirió la aplicación de un consentimiento informado a los sujetos de estudio, pues se usó información generada por actividades académicas rutinarias. Solicitamos a la facultad de medicina una base de datos que contenía el código del alumno y las variables solicitadas. El código del alumno fue intercambiado por un identificador alfanumérico una vez realizada la limpieza de la base de datos.

Financiación

Este estudio fue financiado por la Universidad de Piura (código de financiamiento PI2502).

Conflicto de intereses

Los autores declaran que tienen vínculo laboral con la Universidad de Piura.

Anexo A. Dato suplementario

Se puede consultar material suplementario (tabla 1) en su versión electrónica disponible en <https://doi.org/10.1016/j.edumed.2025.101092>.

Bibliografía

- Price T, Lynn N, Coombes L, Roberts M, Gale T, Regan De Bere S, et al. The international landscape of Medical Licensing Examinations: typology derived from a systematic review. *Int J Health Policy Manag.* 2018;7(9):782–90. doi:10.15171/ijhpm.2018.32.
- Rubright JD, Jodoin M, Barone MA. Examining demographics, prior academic performance, and United States Medical Licensing Examination scores. *Acad Med.* 2019;94(3):364–70. doi:10.1097/ACM.0000000000002366.
- Jacobparayil A, Ali H, Pomeroy B, Baronia R, Chavez M, Ibrahim Y. Predictors of performance on the United States Medical Licensing Examination step 2 clinical knowledge: a systematic literature review. *Cureus.* 2022;14(2):e22280. doi:10.7759/cureus.22280.
- Ghaffari-Rafi A, Lee RE, Fang R, Miles JD. Multivariable analysis of factors associated with USMLE scores across U.S. medical schools. *BMC Med Educ.* 2019;19(1):154. doi:10.1186/s12909-019-1605-z.
- Cecilio-Fernandes D, Kerdijk W, (Debbie) C Jaarsma AD, Tio RA. Development of cognitive processing and judgments of knowledge in medical students: analysis of progress test results. *Med Teach.* 2016;38(11):1125–9. doi:10.3109/0142159X.2016.1170781.
- Romaní-Romaní F, Gutiérrez C, Azurin-Salazar J. Tendencia en la retención de conocimientos de ciencias básicas en una prueba de progreso entre estudiantes de Medicina. *Educ Médica.* 2023;24(4):100830. doi:10.1016/j.edumed.2023.100830.
- Neeley SM, Ulman CA, Sydelko BS, Borges NJ. The value of progress testing in undergraduate medical education: a systematic review of the literature. *Med Sci Educ.* 2016;26(4):617–22. doi:10.1007/s40670-016-0313-0.
- Schuwirth LWT, van der Vleuten CPM. The use of progress testing. *Perspect Med Educ.* 2012;1(1):24–30. doi:10.1007/s40037-012-0007-2.
- Wrigley W, Van Der Vleuten CP, Freeman A, Muijtjens A. A systemic framework for the progress test: strengths, constraints and issues: AMEE guide no. 71. *Med Teach.* 2012;34(9):683–97. doi:10.3109/0142159X.2012.704437.
- Görlich D, Friederichs H. Using longitudinal progress test data to determine the effect size of learning in undergraduate medical education – a retrospective, single-center, mixed model analysis of progress testing results. *Med Educ Online.* 2021;26(1):1972505. doi:10.1080/10872981.2021.1972505.
- Alamro AS, Alghasham AA, Al-Shobaili HA, Alhomaidan HT, Salem TA, Wadi MM, et al. 10 years of experience in adopting, implementing and evaluating progress testing for Saudi medical students. *J Taibah Univ Med Sci.* 2023;18(1):175–785. doi:10.1016/j.jtumed.2022.07.008.
- 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(1):200. doi:10.1186/s12909-021-02633-8.
- Norman G, Neville A, Blake JM, Mueller B. Assessment steers learning down the right road: impact of progress testing on licensing examination performance. *Med Teach.* 2010;32(6):496–9. doi:10.3109/0142159X.2010.486063.
- Johnson TR, Khalil MK, Peppler RD, Davey DD, Kibble JD. Use of the NBME comprehensive basic science examination as a progress test in the preclerkship curriculum of a new medical school. *Adv Physiol Educ.* 2014;38(4):315–20. doi:10.1152/advan.00047.2014.
- Bicudo AM, Hamamoto Filho PT, Abbade JF, Hafner MDLMB, Maffei CML. Teste de progresso em consórcios para todas as escolas médicas do Brasil. *Rev Bras Educ Médica.* 2019;43(4):151–6. doi:10.1590/1981-52712015v43n4rb20190018.
- Majeed GM, Islam J, Nandakumar G, Phoong K. Progress testing in UK medical education: evaluating its impact and potential. *Cureus.* 2024;16(1):e52607. doi:10.7759/cureus.52607.
- Flores J. Asociación entre el promedio ponderado universitario y exámenes de progreso de ciencias básicas y ciencias clínicas frente al puntaje obtenido del ENAM 2020 en internos de medicina de la Universidad Privada de Tacna. Universidad Privada de Tacna; 2021 [consultado 29 Nov 2024], Disponible en: <https://repositorio.upt.edu.pe/handle/20.500.12969/1940>.

18. Romaní-Romaní F, Gutiérrez C. Correlación entre una evaluación sumativa escrita y el promedio ponderado en estudiantes de medicina humana. *Investig En Educ Médica*. 2022;11(43):37–50. doi:[10.22201/fm.20075057e.2022.43.22422](https://doi.org/10.22201/fm.20075057e.2022.43.22422).
19. Flores J. Factors associated with medical students scores on the National Licensing Exam in Peru: a systematic review. *J Educ Eval Health Prof*. 2022;19:38. doi:[10.3352/jeehp.2022.19.38](https://doi.org/10.3352/jeehp.2022.19.38).
20. Obeso R. Asociación entre el promedio ponderado universitario y puntaje obtenido en el Examen Nacional de Medicina en internos de la Universidad Nacional de Cajamarca, 2021–2022. Universidad Nacional de Cajamarca; 2023 [consultado 29 Nov 2024], Disponible en: <https://repositorio.unc.edu.pe/handle/20.500.14074/5666>.
21. Mendoza G, Calla M, Ramos K, Mejía C. Examen Nacional de Medicina (ENAM): análisis de la última década de evaluaciones teóricas en los futuros médicos del Perú. *Acta Medica Peru*. 2021;38(3). doi:[10.35663/amp.2021.383.2164](https://doi.org/10.35663/amp.2021.383.2164).
22. Fitz M, Adams W, Heincelman M, Haist S, Whelan K, Cox L, et al. The impact of internal medicine clerkship characteristics and NBME subject exams on USMLE step 2 clinical knowledge exam performance. *J Gen Intern Med*. 2022;37(9):2208–16. doi:[10.1007/s11606-022-07520-6](https://doi.org/10.1007/s11606-022-07520-6).
23. Hamamoto Filho PT, de Arruda Lourenço PLT, do Valle AP, Abbade JF, Bicudo AM. The correlation between students' progress testing scores and their performance in a residency selection process. *Med Sci Educ*. 2019;29(4):1071–5. doi:[10.1007/s40670-019-00811-4](https://doi.org/10.1007/s40670-019-00811-4).
24. Wang L, Laird-Fick H, Parker C, Liao Z, Solomon D. Growth in basic science knowledge in first-year medical school and USMLE step 1 results: a longitudinal investigation at one school. *Adv Health Sci Educ Theory Pract*. 2022;27(3):605–19. doi:[10.1007/s10459-022-104-y](https://doi.org/10.1007/s10459-022-104-y).
25. Karay Y, Schaubert SK. A validity argument for progress testing: examining the relation between growth trajectories obtained by progress tests and national licensing examinations using a latent growth curve approach. *Med Teach*. 2018;40(11):1123–9. doi:[10.1080/0142159X.2018.1472370](https://doi.org/10.1080/0142159X.2018.1472370).
26. Pachas-Mu A, Bouroncle-Derteano B, Romaní-Romaní F. Percepciones y satisfacción sobre una prueba de progreso en estudiantes de medicina. *Educ Médica*. 2025;26(3):101020. doi:[10.1016/j.edumed.2024.101020](https://doi.org/10.1016/j.edumed.2024.101020).