

ASSESSING THE EFFECT OF GENERATIVE ARTIFICIAL INTELLIGENCE
ON ACADEMIC STRESS AND ACADEMIC PERFORMANCE AMONG
MEDICAL STUDENTS OF KHYBER MEDICAL UNIVERSITY

Shafayat Ullah¹, Fatima Manzoor², Ismail Khan³, Altaf Ur Rehman⁴, Shehzad⁵, Sameen Khan⁶,
Muhammad Yaseen Jan⁷, Sahibzada Ahmad Nabi⁸, Shah Faisal Jamal⁹, Ghazal¹⁰

^{1,2,3,4,5,6,7,8,*9}Khyber Medical University, Institute of Paramedical Sciences, Peshawar

¹shafayat9658@gamil.com, ²fatima.ipms@kmu.edu.pk, ³ismailmed612@gmail.com, ⁴altafur235@gmail.com,
⁵shehzad.edu.pk.1122@gmail.com, ⁶samee.afriidi@yahoo.com, ⁷myaseenjan2001@gmail.com,
⁸sahibzadaahmadkhan456@gmail.com, ⁹shahfaisal.ipms@kmu.edu.pk, ¹⁰ghazal.ipms@kmu.edu.pk

DOI: <https://doi.org/10.5281/zenodo.22629018>

Keywords

Generative AI; Academic Stress;
Academic Performance; Multiple
Regression; ANOVA

Article History

Received: 28 June 2026
Accepted: 02 August 2026
Published: 15 August 2026

Copyright @Author

Corresponding Author: *
Shah Faisal Jamal

Abstract

Background: Generative artificial intelligence (AI) is a new reality in the field of medical education that has added new factors to the academic stress and performance of students. This study examines these relationships among students at Khyber Medical University.

Methodology: The study was conducted as a cross-sectional study with 300 Medical and Allied health students and three validated scales. Data analysis consisted of reliability testing (Cronbach's alpha), normality testing, Pearson correlation test, independent-samples t-test, one-way ANOVA test, and simple and multiple linear regression.

Results: The internal consistency of all three scales was very high ($\alpha > 0.97$). Parametric analyses were appropriate, as verified by normality testing. Simple regression found that academic performance was slightly positively correlated with AI use ($R = 0.114$, $p = 0.048$), and academic stress was slightly positively correlated with the use of AI ($R = 0.165$, $p = 0.004$). Multiple regression revealed that academic stress was a stronger predictor of performance ($B = 0.396$; $p < 0.001$) than using AI was when academic stress was included ($B = 0.005$; $p = 0.911$). The mean scores for academic stress and academic performance across academic years were found to be different by one-way ANOVA ($p = 0.028$ and $p = 0.048$), respectively. There were no significant differences between males and females on any of the variables.

Conclusion: Generative AI is an adjunctive learning tool that has little direct impact on learning outcomes. The factors for academic performance are mainly academic stress, and structured policies for AI use and comprehensive mental health support for students are crucial.

INTRODUCTION

AI has seen rapid integration into higher education during the twenty-first century. Generative Artificial Intelligence (GAI), represented by ChatGPT, launched in November

2022 by OpenAI, is one of the most impactful applications, generating original text, code, and research summaries as needed (1). Today, about 86% of all the students at universities around the world are using AI in their studies (2).

The situation with GAI in school is a double-edged sword. On the other hand, it provides individualized learning experiences, immediate feedback, and less cognitive load, allowing students to better understand difficult concepts and optimize their learning process (3). AI-supported teaching strategies have been proven to greatly enhance student motivation, confidence, and learning outcomes over conventional teaching methods (4). Conversely, the use of AI is not without its challenges, as some worry about its reliance, the loss of independent critical thinking, and digital inequality, especially in healthcare professions education, where absolute correctness is essential to the profession (5, 6).

Academic stress is widely recognized as a psychological and physiological reaction to academic demands as perceived by the student, and it has been well established as a determinant of student performance. Moderate stress can help learning and improve attention, but too much stress can lead to difficulty with thinking and concentration and can cause depression, anxiety, and poor school performance (7, 8). Medical students are extremely at risk as they have to learn a lot of material, along with their practical training. The advent of GAI adds a new variable to this equation, which may be a stress reliever as it helps people with tasks and a stress amplifier because of worries related to the content and academic integrity of the tasks (9).

Although the interest in this subject has been growing around the world, empirical evidence is still limited, especially in developing countries like Pakistan. To date, there has been no study that specifically examines the association between the use of generative AI, academic stress, and academic performance at Khyber Medical University (KMU). A comprehensive statistical approach is presented in this study, using normality testing, reliability analysis, Pearson correlation, t-test, ANOVA, and multiple regression to offer a multi-dimensional analysis of these relationships.

AIM AND OBJECTIVES

This study aims to assess the effect of generative artificial intelligence on academic stress and academic performance among medical students at Khyber Medical University. Specific objectives are:

- I. To evaluate academic performance among students using generative artificial intelligence (GAI).
- II. To identify the relationship between AI usage, academic stress, and academic performance.
- III. To compare these variables across gender, academic year, and program of study.

METHODOLOGY

A cross-sectional descriptive study was conducted at the Khyber Medical University, Institute of Allied Health Sciences (KMU-IPMS), Khyber Medical University, Peshawar. It has a number of health sciences divisions, including the Institutes of Allied Health Sciences, Nursing Sciences, Public Health, Mental Health, and Pharmaceutical Sciences. A sample of 300 participants was calculated using G*Power (10) and recruited via non-probability convenience sampling. Inclusion criteria required participants to be currently enrolled undergraduate students at KMU who used generative AI tools and provided informed consent. Students on leave, not enrolled, or who had incomplete questionnaires were not included.

Data were gathered using a structured, self-administered questionnaire containing three validated scales: Generative AI Usage Scale (GAIS: 10 items; 6-point Likert scale from Never to Always), Academic Stress Scale (ASS: 10 items; 5-point Likert scale), and Academic Performance Scale (APS: 9 items; 5-point Likert scale). Before data collection, ethical approval from the Ethics Review Committee at Khyber Medical University (KMU) was obtained under Reference Number "KMU/Dean/AHS-26/2331". Additionally, the study was formally registered with the KMU Office of Research, Innovation, and Commercialization (KMU-ORIC) under Reference Number: "DIR/ORIC/Ref/26/00026". Informed consent was obtained from all participants, and

confidentiality was ensured by anonymization and secure data storage.

Data were analyzed using SPSS 25, and all scales were evaluated with Cronbach's alpha for internal consistency. All continuous variables were analyzed for normality by the Kolmogorov-Smirnov test, validating the use of parametric analyses. To check bivariate relationships between GAIMEAN (AI usage), ASMEAN (academic stress), and APMEAN (academic performance), Pearson correlation was used. An independent-samples t-test with Levene's test for equal variances was used to compare scores between genders. The differences between academic programs and academic years were examined using one-way ANOVA. Simple linear regression was used to evaluate the individual predictive effect of using

AI on stress and performance. Multiple linear regression was subsequently used to test the independent predictive power of the use of AI and academic stress on academic performance. Throughout, the p-value of < 0.05 was defined as a significant level.

RESULTS

The sample was predominantly composed of young students aged 18–25 years (97.0%, $n = 291$), with a slight male majority (53.7%, $n = 161$) and a higher proportion of Allied health sciences students (67.3%, $n = 202$), followed by Nursing (30.3%, $n = 91$) and DPT students (2.3%, $n = 7$). Most participants were in early semesters 1–4 (58.0%, $n = 174$), with 39.7% ($n = 119$) in semesters 5–8 (Table 1).

Table 1. Demographic Distribution of Study Participants (N = 300)

Variable	Category	n	%age
Gender	Male	161	53.7
	Female	139	46.3
Age (years)	18–25	291	97.0
	26–33	9	3.0
Academic Year	Semesters 1–4	174	58.0
	Semesters 5–8	119	39.7
Program	Allied Health	202	67.3
	Nursing	91	30.3
	DPT	7	2.3

All three scales demonstrated excellent internal consistency with Cronbach's alpha values exceeding 0.97 (GAIS: $\alpha = 0.985$; ASS: $\alpha = 0.983$; APS: $\alpha = 0.976$), well above the accepted threshold of 0.70. Kolmogorov-Smirnov normality testing confirmed that all three scale variables were approximately normally distributed (GAIMEAN: $D = 0.040$, $p = 0.705$; ASMEAN: $D = 0.037$, $p = 0.803$; APMEAN: $D = 0.033$, $p = 0.898$), validating the use of parametric statistical procedures.

Descriptive statistics showed moderate mean scores across all variables: GAIMEAN = 3.19 (SD = 0.76), ASMEAN = 3.39 (SD = 0.71), and APMEAN = 3.04 (SD = 0.64).

Pearson correlation analysis revealed significant positive associations among all three variables (Table 2). The strongest relationship was between academic stress and academic performance ($r = 0.409$, $p < 0.001$), representing a moderate association. Generative AI usage showed weaker

but statistically significant positive correlations with both academic performance ($r = 0.114$, $p = 0.048$) and academic stress ($r = 0.165$, $p = 0.004$).

Table 2. Pearson Correlation Matrix (N = 300)

Variable	GAIMEAN	ASMEAN	APMEAN
GAIMEAN	1	.165** (.004)	.114* (.048)
ASMEAN	.165** (.004)	1	.409** (< .001)
APMEAN	.114* (.048)	.409** (< .001)	1

Note. ** $p < 0.01$; * $p < 0.05$ (two-tailed).

Independent samples t-tests revealed no statistically significant gender differences for any of the three variables. Male and female students reported comparable levels of AI usage ($t = 0.920$, $p = 0.359$), academic stress ($t = -1.851$, $p = 0.065$), and academic performance ($t = -0.988$, $p = 0.324$). Levene's test confirmed homogeneity of variances in all comparisons ($p > 0.05$). Similarly, one-way ANOVA found no significant differences across academic programs (nursing, allied health sciences, DPT) for GAIMEAN ($F = 0.907$, $p =$

0.405), ASMEAN ($F = 1.662$, $p = 0.192$), or APMEAN ($F = 0.262$, $p = 0.770$). However, when comparing by academic year, significant differences emerged for academic stress ($F = 4.857$, $p = 0.028$) and academic performance ($F = 3.936$, $p = 0.048$), with early-year students (Semesters 1–4) reporting higher mean stress ($M = 3.46$ vs. $M = 3.27$) and higher performance scores ($M = 3.09$ vs. $M = 2.94$) than later-year students. No significant difference was found for AI usage across academic year ($F = 1.344$, $p = 0.247$) (Table 3).

Table 3. One-Way ANOVA by Academic Year

Variable	Sem 1–4 M (SD)	Sem 5–8 M (SD)	F	p
GAIMEAN	3.15 (0.71)	3.25 (0.83)	1.344	.247
ASMEAN	3.46 (0.71)	3.27 (0.69)	4.857	.028*
APMEAN	3.09 (0.62)	2.94 (0.66)	3.936	.048*

Note. * $p < 0.05$. $df = 1, 291$.

Simple linear regression confirmed that AI usage individually predicted both academic performance ($R = 0.114$, $R^2 = 0.013$, $p = 0.048$) and academic stress ($R = 0.165$, $R^2 = 0.027$, $p = 0.004$), though the variance explained was minimal in both cases. Multiple linear regression, however, revealed a markedly different picture when both GAIMEAN and ASMEAN were entered simultaneously as

predictors of APMEAN (Table 4). The combined model was statistically significant ($F(2, 297) = 35.412$, $p < 0.001$) and explained 19.3% of variance in academic performance ($R^2 = 0.193$, Adj. $R^2 = 0.187$). Critically, AI usage was no longer a significant predictor when controlling for stress ($B = 0.005$, $SE = 0.045$, $p = 0.911$), whereas academic stress remained a strong and

independent predictor of performance ($B = 0.396$, $SE = 0.047$, $\beta = 0.438$, $p < 0.001$).

Table 4. Multiple Linear Regression: Predicting Academic Performance (APMEAN)

Predictor	B	SE	β	t	p
Constant	1.677	0.205	—	8.175	< .001
GAIMEAN	0.005	0.045	0.006	0.112	.911
ASMEAN	0.396	0.047	0.438	8.348	< .001

Note. $R = 0.439$, $R^2 = 0.193$, Adj. $R^2 = 0.187$, $F(2, 297) = 35.412$, $p < .001$. β = standardized coefficients.

DISCUSSION

The age group distribution of the study sample was mostly young (97% aged 18-25 years), which is typical for Allied health programs at the undergraduate level at the majority of institutions in Pakistan, where students enter into professional programs immediately after secondary education (11). Close to gender parity (53.7% male) indicates rising numbers of women enrolling in allied health sciences in South Asian countries. A significant proportion of the students from Allied health sciences (67.3%) is representative of the institutions' enrolment structure in IPMS KMU and should be taken into consideration when generalizing the results. The internal consistency of all three scales was excellent ($\alpha > 0.97$), and all variables were found to be normally distributed using Kolmogorov-Smirnov testing, which provided a sound basis for future statistical analysis.

Simple regression showed that the use of AI had relatively low R^2 values in predicting academic performance ($R^2=0.013$) and academic stress ($R^2=0.027$) which were statistically significant. The results align with some recent literature that indicate that AI tools, like ChatGPT, improve the efficiency of learning but are not inherently a successful study tool (12). The low impact may be due to there being a wide range of ways students use AI, from relying on it to reacting critically to it. These tools have been shown to have more impact on student performance for those with higher AI literacy and self-regulated learning skills (13). This positive correlation between the use of

AI and academic stress is consistent with the notion that AI can alleviate immediate cognitive strain, but also can create anxiety about content accuracy, academic integrity, and reliance (9).

The multiple regression model is the most clinically and educationally relevant finding of this study. The inclusion of academic stress in the model, along with AI usage, rendered the latter non-significant ($p = 0.911$), and academic stress was found to be the only significant predictor of academic performance, accounting for about 19.3% of the variance ($\beta = 0.438$, $p < 0.001$). This suggests that the bivariate relationship found between the use of artificial intelligence and performance in simple regression was heavily moderated by academic stress. This discovery validates the existing literature that reports academic stress to be a major factor in student success in health professions education (7, 14). The moderate stress-performance correlation ($r = 0.409$) is due to the dual nature of academic stress: when it is of moderate intensity, it can boost academic performance and create motivation; when it is intense, it can negatively affect academic performance and decision-making ability (8).

The results showed no difference in the use of AI for students of both genders or in their perception of stress or performance among them, indicating that male and female students at KMU have similar academic experiences and access to technology. This is in contrast with some studies which reported gender gaps in the utilization of technology (11) and might indicate the increasing digital equity among health science institutions in

Pakistan. Likewise, there were no differences found between the programs for nursing, allied health, and DPT, which suggests, despite the curricular differences, relatively similar academic experiences within KMU-IPMS.

There were significantly different scores between academic years for both stress and performance. The early-year students (Semesters 1-4) indicated a greater level of academic stress and confidence in their performance level than the latter-year students. This is in line with studies showing that freshmen and sophomores are at a higher risk of feeling anxious about school, have greater academic uncertainty, and use digital devices more than any other age group, and have higher anxiety during their first professional training experience (11). With time, senior students may have acquired more effective coping mechanisms and normalized academic expectations, which can help to alleviate their stress, but also lead to a decrease in self-evaluation of their performance.

LIMITATION

It is important to recognize certain limitations in this study. This cross-sectional design prevents causal inference. Self-reported data can have response and social desirability biases. There is limited generalizability because the sample is skewed by allied health and junior-year students. Simple regression models with very low R^2 values suggest that other factors, like self-efficacy, motivation, and socioeconomic status, play significant independent roles. Lastly, the change of AI tools could make it possible that the behavioral patterns described in this report are temporary.

Conclusion

The current study aims to perform a detailed multi-test statistical analysis for the relationships between usage of generative AI, academic stress, and academic performance among 300 medical and allied health students of Khyber Medical University. The results as a whole indicate that the use of generative AI has statistically significant links with stress and performance in bivariate analyses, but has very little predictive power for academic performance when academic stress is controlled for in multiple regression. Academic

stress is found to be the main and most dominant predictor of academic performance with a p -value < 0.001 , thus accounting for the highest percentage of variance in the model. Academic year effects are significant, meaning that students in the early semesters are especially susceptible to academic stressors related to stress. The results showed that there were no gender or program-level differences in terms of equitable AI exposure and academic experiences, indicating that student groups at IPMS KMU may have experienced comparable levels of exposure to AI and academic experiences. Combined, the results indicate that generative AI is not a direct path to academic outcomes, and that the issue of students' well-being and specifically stress management should be the primary focus of educational policy and intervention within medical and allied health programs.

RECOMMENDATIONS

The following recommendations are offered to academic institutions, teachers, and **policymakers** based on this study. Universities need to institutionalize the use of generative AI in educational processes and learning assessment with ethical rules and guidelines to support the appropriate and conscious use of AI and avoid dependence on it, as well as provide structured AI literacy education that enables students to analyze AI-generated content. Considering academic stress as the most significant factor in academic performance, institutions should focus on providing increased access to mental health services, the provision and expansion of counselling services, and the provision of stress reduction workshops, especially for students who exhibit high levels of academic stress in the early semester. All faculty should be provided with professional development resources to effectively utilize AI tools in their pedagogy and redesign assessments to focus on critical thinking and clinical reasoning. Last but not least, future studies could adopt a longitudinal design, and other psychological factors like self-efficacy, cognitive load, and AI dependency should be incorporated to gain a more thorough understanding of how

generative AI influences learning outcomes within medical education settings.

REFERENCES

- Khan MJ, Altaf S, Kausar H. Effect of perceived academic stress on students' performance. *FWU Journal of Social Sciences*. 2013;7(2).
- Eltahir ME, Babiker FM. The influence of artificial intelligence tools on student performance in e-learning environments: Case study. *Electronic Journal of e-Learning*. 2024;22(9):91-110.
- Namjoo F, Liaghat E, Shabaziasl S, Modabernejad Z, Morshedi H. Students' experience on self-study through AI. *AI and Tech in Behavioral and Social Sciences*. 2023;1(3):35-42.
- Imaji PJ, Mwaniki CN, Mabeya TM. Exploring Students' Artificial Intelligence Competence and its Impact on their Learning in Public Junior Secondary Schools in Nigeria. *Journal of Popular Education in Africa*. 2025;9(8):221-43.
- Wright K. Challenges and limitations of AI in education: Strategies for addressing racial inequity in ChatGPT. *American Journal of STEM Education*. 2025;4:66-74.
- Chen Z, Zeng J, Liao J, Guo C. The association between academic stress and problematic internet use among adolescents: A three-level meta-analysis. *Journal of Youth and Adolescence*. 2025;54(7):1779-99.
- Wachelka D, Katz RC. Reducing test anxiety and improving academic self-esteem in high school and college students with learning disabilities. *Journal of Behavior Therapy and Experimental Psychiatry*. 1999;30(3):191-8.
- Kovari A. Ethical use of ChatGPT in education—Best practices to combat AI-induced plagiarism. *Frontiers in Education*. 2025;9:1465703.
- Faul F, Erdfelder E, Buchner A, Lang AG. Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*. 2009;41(4):1149-60.

- Paroginog F, Alpuerto MD, Pilongo LW. Social skills and academic performance among grade 10 students of the University of Bohol. *ACADEME University of Bohol*. 2018;12(1):1-31.
- Ng DT, Leung JK, Chu SK, Qiao MS. Conceptualizing AI literacy: An exploratory review. *Computers and Education: Artificial Intelligence*. 2021;2:100041.
- Gao Y, Liu S, Yang L. Artificial intelligence and innovation capability: A dynamic capabilities perspective. *International Review of Economics & Finance*. 2025;98:103923.
- Hussain M. Operationalizing Cognitive Load Theory for Students with Learning Disabilities: A Phase-Based Instructional Design Framework. *ACADEMIA International Journal for Social Sciences*. 2025;4(2).
- Knekta E, Runyon C, Eddy S. One size doesn't fit all: Using factor analysis to gather validity evidence when using surveys in your research. *CBE—Life Sciences Education*. 2019;18(1):rm1.
- Chen YH, Zhang K. Impact of basic artificial intelligence course on understanding concepts, literacy, and empowerment among students. *Computer Applications in Engineering Education*. 2025;33(1):e22806.
- Kampa RK, Padhan DK, Karna N, Gouda J. Identifying the factors influencing plagiarism in higher education: An evidence-based review. *Accountability in Research*. 2025;32(2):83-98.
- Li J, Qin J. Effect of teachers' knowledge sharing behavior on students' entrepreneurial motivation in social media environments. *International Journal of Emerging Technologies in Learning*. 2022;17(2):143-57.
- Mirzeoglu AD. The effects of peer teaching on university students' achievements in volleyball courses. *Educational Research and Reviews*. 2014;9(9):262-71.

Marandi B. A study on artificial intelligence and its applications. *International Journal of Engineering and Computer Science*. 2025;7:48-51.

Liao X, Zhu X, Cheong CM. Direct and indirect effects of independent language skills on integrated writing performance. *Reading and Writing*. 2021;34(10):2529-57.

Di Bitetti MS, Ferreras JA. Publish (in English) or perish: The effect on citation rate of using languages other than English. *Ambio*. 2017;46(1):121-7.

Melo G, Monteza D, Colson G, Zhang YY. How to assess? Student preferences for methods to assess experiential learning. *PLoS One*. 2022;17(10):e0276745.

Caselli G, Spada MM. The desire thinking questionnaire: Development and psychometric properties. *Addictive Behaviors*. 2011;36(11):1061-7.