BG Pulse Journal USA
ISSN:1947-2536 http://www.bgpulseusa.com/index.php/bg

DESIGNING AN AI-SUPPORTED BILINGUAL MEDICAL TERMINOLOGY LEARNING ECOSYSTEM: INTEGRATING ENGLISH, LATIN, AND CLINICAL CONTEXTS

Djalilova Zarnigor Obidovna
Professor, Department of Fundamental Medical Sciences
Asia International University, Bukhara, Uzbekistan
Email: djalilovazarnigorobidovna@oxu.uz

Abstract: The article proposes and experimentally substantiates an artificial intelligence-supported bilingual ecosystem for teaching medical terminology through the coordinated integration of English, Latin and clinical contexts. The ecosystem treats English as the language of international professional communication, Latin as the structural and nomenclatural basis of medical terms, and clinical scenarios as the environment in which terminology acquires functional meaning. The empirical study involved 596 first-year medical students. The experimental model combined an English-Latin terminology repository, adaptive AI-generated tasks, morphological and semantic analysis of terms, virtual clinical cases, immediate feedback, and teacher-controlled learning analytics. After the intervention, the proportion of students at a high competence level increased from 18.3% to 49.0% in the experimental group, whereas the control group increased from 18.6% to 22.3%. Low-level performance in the experimental group decreased from 36.2% to 12.4%. The findings indicate that an AI-supported ecosystem can strengthen terminological accuracy, cross-language transfer, professional reading and clinically situated communication when technological automation is combined with methodological supervision and medical validation.

Keywords: artificial intelligence, Medical English, medical Latin, bilingual terminology, clinical context, adaptive learning, virtual patients, professional-linguistic competence.

INTRODUCTION

Medical terminology is a multilingual and interdisciplinary system rather than a simple list of equivalents. English provides access to international research, clinical guidelines and professional communication, while Latin supplies the historical, morphological and nomenclatural framework of anatomy, histology, pharmacology and clinical term formation. First-year medical students therefore need to recognize not only what a term means, but also how it is formed, how its components change across languages, and how it functions in a clinical statement. When English and Latin are taught as disconnected subjects, students may memorize isolated vocabulary yet remain unable to interpret complex terms, construct accurate phrases or transfer terminology into professional communication.

Artificial intelligence offers a mechanism for connecting these layers. Current medical-education research identifies intelligent tutoring, adaptive assessment, simulation, clinical reasoning support and feedback as major areas of AI application (Gordon et al., 2024). Studies of English medical terminology also show that structured generative-AI prompts can increase performance, confidence and learner autonomy when they supplement rather than replace classroom instruction (Ajbar, 2026). At the same time, virtual-patient research demonstrates that AI-supported scenarios can provide repeated opportunities for communication and clinical reasoning, although their educational value depends on validated content, standardized evaluation and human supervision (Plackett et al., 2022; Zeng et al., 2025).

The concept of a learning ecosystem is useful because it shifts attention from a single chatbot or application to the interaction of content, pedagogy, assessment, feedback, analytics and teacher governance. In the proposed ecosystem, AI does not merely translate an English term into Latin. It identifies term components, relates them to anatomical or clinical concepts, generates graded tasks, embeds terms in realistic cases, analyzes student errors, and recommends the next learning action. The teacher controls the terminology database, verifies clinical accuracy and determines when automated feedback is sufficient and when direct explanation is necessary.

The aim of this study was to design and evaluate an AI-supported bilingual medical terminology learning ecosystem that integrates English, Latin and clinical contexts. The study addressed three questions: whether such an ecosystem improves students' professional-linguistic and terminological competence; which competence levels change most strongly; and which design principles are necessary for responsible implementation in medical higher education.

CONCEPTUAL DESIGN OF THE BILINGUAL ECOSYSTEM

The ecosystem was organized around six interconnected modules. The first module was a bilingual terminology repository containing English terms, standardized Latin equivalents, word-building elements, grammatical characteristics, definitions and links to relevant medical disciplines. The second module was an AI semantic-morphological analyzer that decomposed terms into roots, prefixes and suffixes, compared English and Latin forms, and highlighted medically significant distinctions. For example, students could connect the English expression inflammation of the liver with the Latin anatomical form hepar and the clinical term hepatitis, thereby understanding both lexical meaning and derivational structure.

The third module was an adaptive practice engine. Diagnostic tasks identified weaknesses in vocabulary recognition, Latin case agreement, plural formation, translation, term construction and contextual use. The system then adjusted task difficulty and repetition intervals. The fourth module consisted of virtual clinical micro-scenarios: taking a short history, matching symptoms with organ systems, completing a case note, interpreting an anatomical description, selecting a pharmaceutical form or reformulating a patient-friendly explanation. These scenarios moved terminology from declarative knowledge to professional action.

The fifth module provided immediate multilevel feedback. A lexical error received a corrected equivalent; a morphological error was linked to the relevant Latin rule; a semantic error triggered a contrastive explanation; and a clinical-context error prompted the learner to reconsider the appropriateness of the term. The sixth module was a teacher dashboard displaying recurrent errors, competence dynamics and task-completion patterns. This governance layer was essential because AI-generated medical content can contain inaccuracies, context loss or plausible but incorrect explanations. The ecosystem was therefore designed according to a human-in-the-loop principle: automation supported personalization, while teachers retained responsibility for curricular alignment, medical validity and ethical use.

OPERATIONAL LEARNING CYCLE

The operational cycle began with a short diagnostic activity linked to the current medical module. Students first identified terms, translated selected units, analyzed word-building elements and completed a brief context task. The system used the response pattern to assign an individual pathway rather than merely calculating a total score. A student who recognized terminology but misused Latin endings entered a morphology pathway; a student who translated accurately but failed to interpret the term in a clinical sentence entered a contextualization pathway; and a student

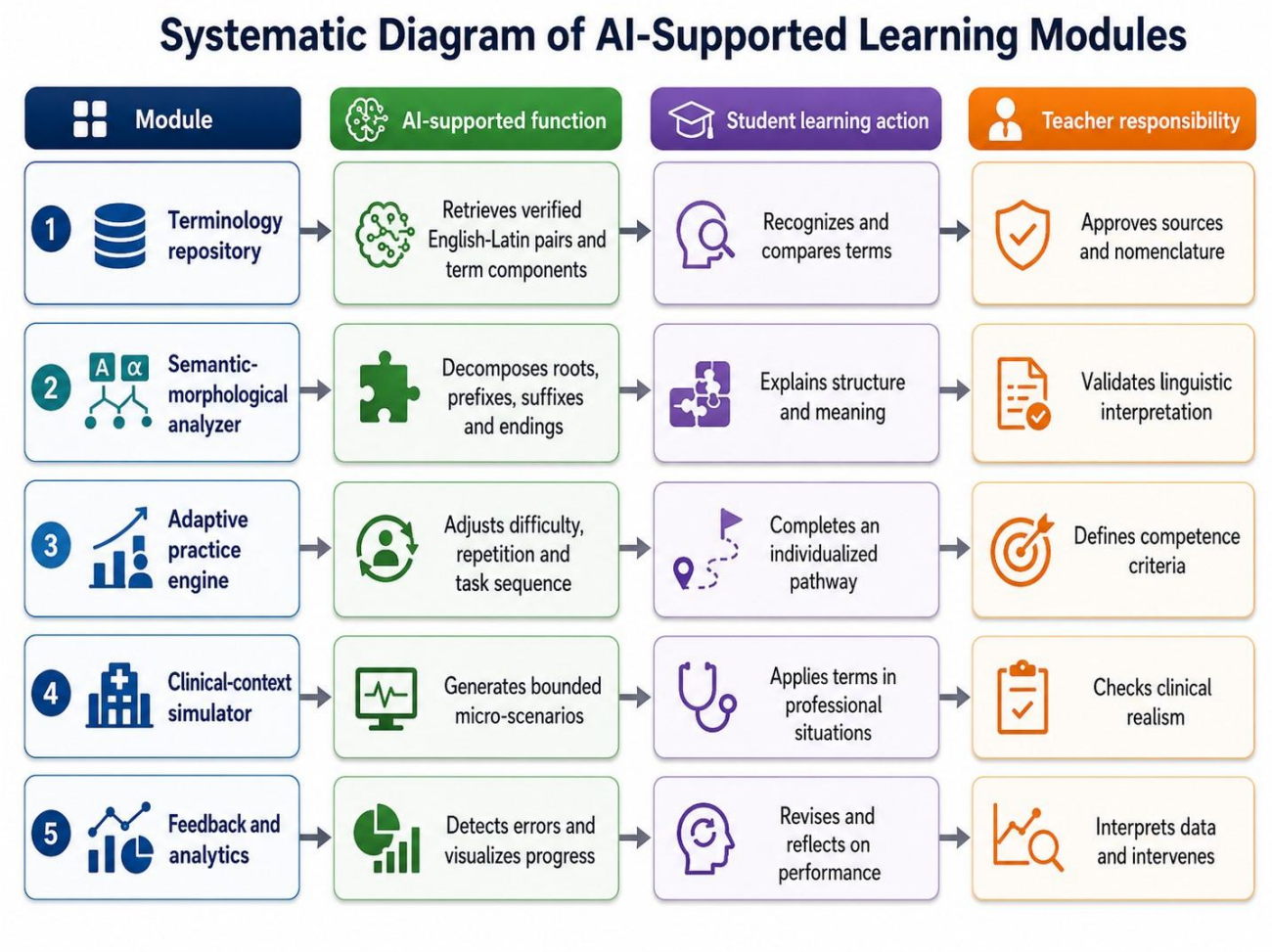
with weak recognition received a vocabulary-and-retrieval pathway. In this way, assessment directly informed instruction.

The second phase combined controlled bilingual mapping with generative practice. Learners moved between English definitions, Latin nomenclature, roots of Greek-Latin origin and short clinical descriptions. AI-generated examples were constrained by the verified repository, and tasks were organized from recognition to production: selecting an equivalent, completing a term, explaining its components, applying it in a phrase and finally using it in a clinical micro-scenario. This progression prevented the common methodological error of asking beginners to generate complex clinical language before they had acquired structural control of the terminology.

The third phase was reflective consolidation. After immediate feedback, students compared their initial and corrected responses, identified the source of the error and completed a parallel transfer task. Teachers then used dashboard data to select errors for group discussion. Consequently, individual automation and collaborative classroom work were not treated as competing methods. AI managed repeated low-level practice and rapid feedback, while the teacher focused on conceptual distinctions, clinical appropriateness and the formation of academic language culture.

Table 1. Functional architecture of the AI-supported bilingual ecosystem

BG Pulse Journal USA
ISSN:1947-2536 http://www.bgpulseusa.com/index.php/bg
BG PULSE JOURNAL



MATERIALS AND METHODS

A mixed pedagogical-experimental design was used. The empirical base included 596 first-year medical students from four medical higher education institutions, distributed across 50 academic groups. The control group and experimental group each consisted of 298 students. The preparatory stage was conducted in July-August 2025, the formative intervention in September-October 2025, and the final assessment in November 2025. Both groups followed the approved English and Latin curricula; however, the experimental group additionally used the integrated AI-supported ecosystem, whereas the control group received predominantly traditional classroom instruction and standard digital materials.

The diagnostic framework combined five domains: knowledge of English and Latin medical vocabulary; morphological and semantic analysis of terminology; reading and interpretation of professional texts; production of short written medical statements; and use of terminology in doctor-patient or clinical-educational situations. Performance was classified at low, medium and high levels. The experimental tasks included adaptive quizzes, English-Latin matching, term-component analysis, error correction, scenario-based dialogue, short clinical documentation, and AI-assisted revision. Teachers reviewed task banks and sampled AI feedback for medical and linguistic accuracy.

Quantitative analysis included percentage comparison, Pearson chi-square, Fisher's criterion, Spearman correlation and Kendall concordance procedures. Because the reported percentages were rounded to one decimal place, the post-test chi-square comparison reported below was calculated from the nearest whole-number frequencies for the two equal-sized groups and should be interpreted as an approximation. Qualitative analysis focused on recurrent error patterns, the transfer of terms between English and Latin, and students' ability to use terminology appropriately in clinical contexts.

RESULTS

Initial diagnostics confirmed that the groups were comparable. Before the intervention, high-level competence was observed in 18.6% of the control group and 18.3% of the experimental group. The largest proportion of students in both groups occupied the medium level, while more than one third remained at the low level. After instruction, both groups showed positive movement, but the magnitude and structure of change differed substantially.

Table 2. Distribution of professional-linguistic and terminological competence levels (%)

Group and stage	Low	Medium	High	Pedagogical interpretation
Control group - before	34.6	46.8	18.6	Comparable baseline
Control group - after	28.5	49.2	22.3	Moderate improvement
Experimental group - before	36.2	45.5	18.3	Need for guided support
Experimental group - after	12.4	38.6	49.0	Strong shift toward independent performance

Note. Post-test comparison based on rounded frequencies: $\chi^2(2) = 52.98$, $p < .001$; Cramer's V = .30, indicating a moderate between-group association.

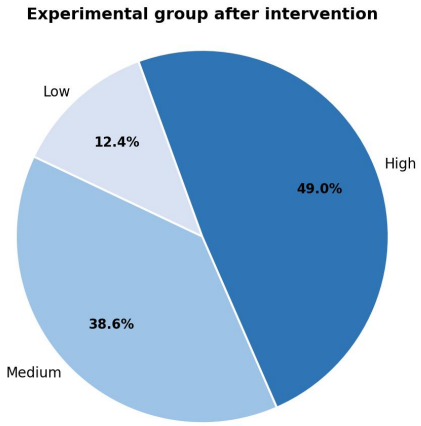


Figure 1. Final competence distribution in the experimental group.

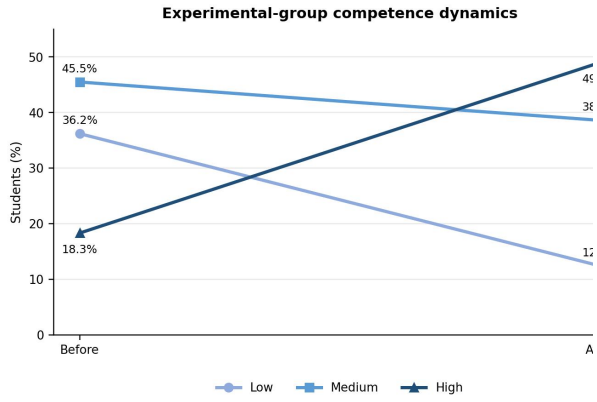


Figure 2. Changes in experimental-group competence levels.

The experimental group demonstrated a 30.7-percentage-point increase in high-level competence, compared with a 3.7-point increase in the control group. Simultaneously, the low-level category in the experimental group decreased by 23.8 points, whereas the control group decreased by only 6.1 points. The medium category declined slightly in the experimental group because many students moved upward into the high category; in the control group, it remained the dominant level. These shifts show that the ecosystem did more than produce small improvements within the same performance band: it changed the overall competence structure.

Component analysis showed the strongest outcomes in integrated English-Latin knowledge, where mastery reached 95-100%, and in terminological classification, which reached 66.8%. Professional-text reading reached 63.7%, written medical statements 58.5%, academic presentation 42.6%, scientific-text interpretation 39.1%, and doctor-patient communication 28.9%. The lower productive-communication indicators were expected because they require not only terminology but also interactional fluency, clinical reasoning and pragmatic sensitivity. Nevertheless, classroom observations indicated that students increasingly used terms as semantic units within clinical statements rather than as isolated translations.

The AI analytics also revealed a qualitative change in errors. Early mistakes were dominated by direct word-for-word translation, confusion of Latin endings and inability to identify term-forming elements. Later errors were more often related to contextual precision, register and the selection of the most clinically appropriate expression. This shift suggests movement from elementary recognition toward higher-order terminological decision making.

DISCUSSION

The findings support the view that bilingual medical terminology should be taught through coordinated cognitive operations: recognition, structural analysis, cross-language mapping, contextual interpretation and professional production. English and Latin perform different but complementary functions. English connects students with current scientific discourse and clinical communication; Latin stabilizes nomenclature and supports the analysis of complex medical words. Clinical context links both systems to the future professional activity of the learner. The ecosystem's effectiveness therefore derives not from AI alone, but from the orchestration of these three knowledge environments.

Adaptive sequencing was particularly important. Traditional terminology exercises usually present the same task to the whole group, even when students have different difficulties. In the experimental ecosystem, a learner who confused Greek-Latin roots received component-analysis tasks, a learner who made agreement errors received targeted Latin morphology, and a learner with adequate vocabulary but weak contextual use received clinical micro-scenarios. Such personalization corresponds to the broader potential of AI for differentiated instruction and formative assessment identified in medical-education research (Gordon et al., 2024; Preiksaitis & Rose, 2023).

The virtual clinical layer also contributed to transfer. Systematic reviews show that virtual patients can support clinical reasoning and feedback, although evidence remains heterogeneous and implementation quality is decisive (Plackett et al., 2022; Jay et al., 2025). More recent LLM-based systems offer flexible dialogue and repeated practice, but current studies emphasize that they should supplement, not replace, standardized patients and teacher feedback (Sheth et al., 2025; Zeng et al., 2025). In the present model, virtual cases were intentionally short and terminology-centered. Their purpose was not to simulate full diagnosis, but to create meaningful conditions for selecting, interpreting and using medical terms.

Immediate feedback was another central mechanism. AI-assisted feedback can process large quantities of student language and identify recurring patterns, but automated explanations require contextual control and human oversight (Khan et al., 2025). The ecosystem therefore used a layered feedback model: the system provided instant correction and explanation, while teachers reviewed high-risk content, ambiguous cases and persistent misconceptions. This division of labor reduced repetitive correction without transferring academic responsibility to the algorithm.

The results are also consistent with emerging research on AI-supported terminology learning. Ajbar (2026) reported that structured ChatGPT prompts improved English medical terminology performance and offered a low-anxiety environment for autonomous practice. The present study extends this logic by integrating a second terminological system, Latin, and by embedding terminology in clinical contexts. Bilingual medical language models such as BiMediX further demonstrate the technical feasibility of aligning domain knowledge across languages, although educational ecosystems require explicit pedagogy, assessment and teacher governance beyond model performance alone (Pieri et al., 2024).

VALIDITY AND ETHICAL GOVERNANCE

The validity of an AI-supported terminology ecosystem depends on the quality of its reference base and on the separation of pedagogical assistance from clinical decision making. The platform used educational cases and de-identified or synthetic examples rather than real patient records. Prompts restricted the system to the curricular topic, required the use of approved terminology and instructed it to signal uncertainty instead of inventing an answer. Teachers periodically compared generated tasks with reference textbooks and accepted nomenclature. These safeguards are especially important because fluent language can conceal factual error and may lead novice students to overestimate the authority of an automated response.

Ethical implementation also requires transparency. Students were informed when feedback was automated, were taught not to enter identifiable patient information, and were encouraged to challenge questionable outputs. Learning analytics were used to support instruction rather than to make high-stakes decisions without human review. This governance model follows the principle that educational AI should increase access to practice and feedback while preserving teacher accountability, learner autonomy and data protection.

PEDAGOGICAL IMPLICATIONS AND LIMITATIONS

For curriculum design, the ecosystem can be implemented as a sequence of weekly learning cycles: diagnostic assessment, adaptive terminology practice, clinical-context application, immediate feedback, teacher review and reflective consolidation. Each cycle should use a limited set of verified terms connected with the current anatomy, histology, physiology or clinical module. AI outputs should be generated from an institutionally approved terminology base, and students should be taught to verify definitions, distinguish nomenclature from clinical usage, and recognize the possibility of algorithmic error.

The study has several limitations. The competence results were reported at group level, so detailed individual transition patterns could not be reconstructed. The intervention focused on first-year students and may not fully represent advanced clinical communication. Productive doctor-patient performance remained lower than receptive and terminological indicators, showing the need for longer interventions, peer interaction and standardized-patient practice. Finally, generative AI systems change rapidly; therefore, the pedagogical model should be evaluated independently of any single commercial platform and should prioritize transparent criteria, data protection and teacher control.

CONCLUSION

An AI-supported bilingual medical terminology learning ecosystem can effectively integrate English, Latin and clinical contexts within a unified methodological system. The experimental results demonstrate a marked shift from low and medium competence toward high-level professional-linguistic and terminological performance: the high-level share in the experimental group rose from 18.3% to 49.0%, while the low-level share fell from 36.2% to 12.4%. The model is most effective when AI is used for adaptive practice, term analysis, contextualization and immediate feedback, while teachers ensure medical accuracy, curricular coherence and ethical implementation. The proposed ecosystem offers a scalable basis for digital textbooks, intelligent terminology modules and integrated language instruction in medical higher education.

REFERENCES

1. Ajbar, O. (2026). Enhancing English medical terminology learning through ChatGPT-driven prompts. In M. Khaldi (Ed.), Proceedings of the E-Learning and Smart Engineering Systems (ELSES 2025) (pp. 217-231). Atlantis Press. https://doi.org/10.2991/978-94-6239-634-0_18
2. Gordon, M., Daniel, M., Ajiboye, A., Uraiby, H., Xu, N. Y., Bartlett, R., et al. (2024). A scoping review of artificial intelligence in medical education: BEME Guide No. 84. Medical Teacher, 46(4), 446-470. https://doi.org/10.1080/0142159X.2024.2314198
3. Jay, R., Sandars, J., Patel, R., Leonardi-Bee, J., Ackbarally, Y., Bandyopadhyay, S., et al. (2025). The use of virtual patients to provide feedback on clinical reasoning: A systematic review. Academic Medicine, 100(2), 229-238. https://doi.org/10.1097/ACM.0000000000005908
4. Khan, S. A., Taiyara, J., Zary, N., & Otaki, F. (2025). Artificial intelligence in narrative feedback analysis for competency-based medical education: A review. Studies in Health Technology and Informatics, 327, 1423-1427. https://doi.org/10.3233/SHTI250637
5. Pieri, S., Mullappilly, S. S., Khan, F. S., Anwer, R. M., Khan, S., Baldwin, T., & Cholakkal, H. (2024). BiMediX: Bilingual medical mixture of experts LLM. arXiv. https://doi.org/10.48550/arXiv.2402.13253
6. Djalilova, Z. O. (2023). Advancing pedagogical approaches: Leveraging artificial intelligence technologies to enhance the integration of English and Latin language instructional methods. International Bulletin of Medical Sciences and Clinical Research, 3(12), 54–60. DOI: 10.5281/zenodo.10426821.
7. Djalilova, Z. (2024). Ingliz va lotin tillarini integrativ o‘qitishda sun’iy intellekt texnologiyalarining samaradorligini matematik-statistik tahlil qilish. Central Asian Journal of Education and Innovation, 3(12), 56–64. DOI: 10.5281/zenodo.14509570.
8. Djalilova, Z. O. (2024). The pedagogical-psychological aspects of artificial intelligence technologies in integrative education. International Journal of Literature and Languages, 4(3), 13–19. DOI: 10.37547/ijll/Volume04Issue03-03.
9. Djalilova, Z. O. (2026). Artificial intelligence–driven integrative teaching of English and Latin for medical students. Journal of Multidisciplinary Sciences and Innovations, 5(3), 1826–1833. DOI: 10.5281/zenodo.19233890.