MODESTUM EURASIA Journal of Mathematics, Science and Technology Education, 2026, 22(10), em2937
ISSN:1305-8223 (online)
OPEN ACCESS Research Paper https://doi.org/10.29333/ejmste/19427

A medical physics elective course: A quasi-experimental study of artificial intelligence-supported learning in Kazakhstan

Aizada Kambarbekova [1], Bakytgali Rakhashev [1], Dauletbay Berdaliyev [1], Torebay Turmambekov [1], Bayan Ualikhanova [1*]

[1] South Kazakhstan Pedagogical University named after Ozbekali Zhanibekov, Shymkent, KAZAKHSTAN

Received 23 June 2026 ▪ Accepted 12 September 2026

Abstract

Modern school physics education requires stronger interdisciplinary and career-oriented approaches, especially for students considering medical fields. Although physical laws underlie many diagnostic and therapeutic medical technologies, the role of physics in school-based medical career guidance often remains insufficiently emphasized. This study assessed the outcomes of a 34-hour elective course in medical physics for grade 10-11 students in Kazakhstan and examined the additional difference associated with artificial intelligence (AI)-supported implementation. A preliminary survey involved 152 physics teachers and career guidance specialists. The main stage was organized as a quasi-experiment with 103 students from four existing classes: 51 completed the course without systematic AI support, whereas 52 studied the same content using teacher-mediated materials prepared with ChatGPT, Canva AI, Gamma, and digital simulations. Knowledge was measured before and after the course; final questionnaires assessed perceived usefulness, interest, and career-related relevance. Test scores increased significantly in both groups. After controlling baseline knowledge and grade level, the AI-supported group achieved a higher post-test score. The findings indicate the educational potential of the elective course and show higher adjusted post-test performance in the AI-supported condition, although the small number of intact classes limits causal interpretation of this between-group difference.

Keywords: medical physics, elective course, artificial intelligence, quasi-experimental study, career guidance, secondary education

INTRODUCTION

Modern medicine relies heavily on physical laws, measurement methods, and high-technology equipment. Medical physics encompasses the application of physical principles to diagnostic imaging, treatment, nuclear medicine, dosimetry, and radiation protection (International Atomic Energy Agency, n.d.). Therefore, its elements can be used to demonstrate the practical significance of school physics.

Despite the fundamental role of physics in modern medicine, school students' career orientation toward medical fields is often associated primarily with biology and chemistry. As a result, students may have an insufficient understanding of the technological basis of diagnosis and treatment, as well as a limited awareness of professions located at the intersection of medicine, physics, engineering, and digital technologies. Modern career guidance involves not only informing students about professions but also introducing them to real-life activities, educational pathways, and the knowledge necessary for their future professional fields. Research on adolescent career readiness shows that uncertainty about future pathways remains a common problem, and that participation in meaningful career guidance activities can help students develop more informed ideas about their future (Organization for Economic Co-operation and Development, n.d.).

Medical physics can serve as a meaningful bridge between school physics and the medical professional

✉ kambarbekova.aizada@okmpu.kz ✉ rakhashev@mail.ru ✉ berdaliyev@mail.ru ✉ tore_bai@mail.ru
✉ ualikhanova.bayan@okmpu.kz **(*Correspondence)**

Contribution to the literature

- This study presents an empirical evaluation of a systematic 34-hour elective course in medical physics for grade 10-11 students, whereas previous studies have mainly examined individual modules and short-term outreach or educational interventions.
- The findings distinguish between the substantial increase in knowledge after completing the shared course content and the higher adjusted post-test performance observed in the condition implemented with pedagogically organized artificial intelligence (AI) support.
- The study extends the literature on context-based physics education by providing evidence from Kazakhstan and by combining test results, students' perceived usefulness of the course, and its career-related relevance.

field. Context-based learning connects scientific concepts with real-life and technological situations and can support more positive attitudes toward science among students (Bennett et al., 2007). A detailed theoretical justification for this approach is presented in the next section. Career-oriented learning scenarios can also support interest in science and enhance students' understanding of the diversity of STEM careers (Archer et al., 2015). One organizational form of such learning is an elective course. Unlike a stand-alone educational event or a single topic-based lesson, an elective course allows for the systematic exploration of the connections between several areas of school physics and their medical applications. Previous school modules and outreach programs have confirmed the accessibility of medical physics for adolescents and its potential to broaden understanding of the applications of physics and related professions (Sudirman et al., 2023; Yorke et al., 2025).

However, most studies have focused on individual modules, short-term outreach programs, or teacher training. Full-length school elective courses with pre-test/post-test assessments and comparisons of different teaching conditions have been studied considerably less frequently (Buckley, 2016; Cetnar et al., 2025).

AI, digital visualizations, and interactive simulations offer additional opportunities for such a course. Generative systems can be used by teachers to develop additional explanations, learning questions, projects, non-standard tasks, and visually organized materials. Digital simulations, in turn, allow for the presentation of processes that cannot be directly observed in the classroom, such as radiation propagation, medical image formation, changes in wave parameters, or the interaction of a magnetic field with matter.

A systematic review by Heeg and Avraamidou (2023) found that AI is used in school science education for adaptive learning, automated feedback, assessment, and analysis of learning data. However, studies vary in the technologies used, duration, and pedagogical models. A quasi-experimental study by Alarbi et al. (2024) also demonstrated the potential benefit of targeted use of ChatGPT in teaching physics to high school students. However, the positive results of a single tool cannot be automatically generalized to any use of AI or any subject content.

Generative AI (GenAI) can produce convincingly formulated but factually incorrect explanations, incorrect solutions, and inaccurate visual representations. Its use with minors also raises issues of privacy, age appropriateness, and personal data protection. Therefore, UNESCO recommends a human-centered approach in which the teacher retains responsibility for learning objectives, content verification, interpretation of results, and student safety (Miao & Holmes, 2023). In this study, AI was not considered as an autonomous teacher or independent course content. It was used primarily by the teacher as a supplementary tool for preparing explanations, visual materials, projects, and assignments.

It should also be noted that AI-supported learning is not the result of a single application, but rather a comprehensive pedagogical scenario. The use of ChatGPT, Canva AI, Gamma, and digital simulations was accompanied by content selection, teacher review, task assignment, discussion, and feedback. Therefore, the potential benefit of such learning should be attributed to a pedagogically organized support package rather than the isolated effect of a specific language model or digital service.

Kazakhstan's educational context provides organizational opportunities for developing such courses. At the level of general secondary education, the science and mathematics track includes in-depth study of physics and other natural science subjects. This creates a basis for introducing elective content that links school physics with medicine and career pathways. However, the presence of an elective component alone does not resolve issues related to course content, pedagogical model, or assessment. In Kazakhstan, the senior stage of general secondary education includes specialized tracks and variable components of curriculum content, which creates a regulatory framework for the development of interdisciplinary elective courses (Ministry of Education of the Republic of Kazakhstan, 2022). However, empirical studies comparing the same school medical physics course in formats with and without AI support remain limited. This study aims to address this gap.

The aim of the study was to evaluate the course outcomes and the additional differences associated with the AI-supported form of its implementation.

The study addressed the following research questions (RQs):

RQ1. How did the knowledge test scores of grade 10-11 students change after completing a 34-hour elective course in medical physics?

RQ2. Did the post-test scores of students who completed the course with AI support differ from those of students who completed the same course without systematic AI support, after accounting for baseline knowledge and grade level?

RQ3. How did students rate the usefulness of the course, their understanding of the application of physics in medicine, their interest in medical physics, and its potential role in career self-determination?

For the quantitative component, the following hypotheses (Hs) were tested:

H1. After completing the elective course, knowledge test scores will be higher than baseline scores in both groups.

H2. After accounting for baseline knowledge and grade level, students in the AI-supported group will have higher post-test scores than students in the comparison group.

LITERATURE REVIEW

Medical Physics in Secondary Education

The applied nature of medical physics allows school physics content to be integrated with clinical technologies. Hands-on programs in X-ray imaging and radiotherapy have shown that safe models, low-cost equipment, and problem-based tasks can be used to introduce students to the physical principles of diagnosis and treatment (Fagerstrom et al., 2019; Schleicher et al., 2005). Outreach programs also increase students' awareness of medical physics and other healthcare-related STEM fields (Santoso et al., 2021; Yorke et al., 2025).

However, the evidence base remains heterogeneous. Many initiatives are one-time events, summer schools, or short modules. Even longer programs often focus on teacher professional development rather than on controlled comparisons of student outcomes (Cetnar et al., 2025). Therefore, studies are needed in which the content, duration, and instructor remain the same, while the additional technological component is compared more systematically.

The closest study to the present work is a comparative study by Sudirman et al. (2023), which focused on the teaching of a medical physics module to 14- to 16-year-old students in schools in Indonesia and Ireland. The module was structured around real-world applications of physical laws in medicine, and the study examined the combination of direct and inquiry-based instruction, students' attitudes toward the module, and the challenges of its implementation. The results confirmed the educational appeal of the medical context, but the study focused primarily on pedagogical approaches and students' perceptions of the module rather than on comparing the outcomes of one extended course delivered in formats with and without AI support. Thus, a quasi-experimental evaluation of full-length school medical physics courses with control for baseline knowledge remains necessary.

Context-Based Learning and the Professional Relevance of Physics

Context-based learning connects scientific concepts to real-world situations, socially relevant problems, and professional practices. Research synthesis shows that such approaches can improve students' attitudes toward science without degrading subject achievement (Bennett et al., 2007). Medical technologies provide a particularly rich context: physical laws become tools for explaining the operation of familiar diagnostic instruments rather than only objects of abstract calculation.

A professionally and personally relevant context can influence not only students' understanding of content but also their attitudes toward science. A systematic review by Potvin and Hasni (2014) found that interest, motivation, and attitudes toward science during the school years depend on the content of instruction, the perceived usefulness of the subject, and the nature of the learning experience. Experimental studies of utility value also suggest that tasks encouraging students to connect scientific content with their own lives and future goals can support interest and academic achievement (Hulleman & Harackiewicz, 2009). Physics identity, the expected usefulness of physics, self-confidence, and ideas about future professions are particularly important in physics education. These factors are associated with intentions to continue studying physics and to consider related educational pathways (Hazari et al., 2010; Mujtaba & Reiss, 2014; Sheldrake et al., 2019). Research on science capital also shows a link between students' access to science-related resources, their social environment, their perceptions of science, and their aspirations to continue education in physics and engineering fields (Moote et al., 2020). In this regard, medical physics can enhance the perceived usefulness of the subject by demonstrating specific connections between school physics laws, medical technologies, and professional activities.

From a science capital perspective, not only knowledge but also an understanding of where and by whom science is used is important. Awareness of careers and a sense of connection between science and one's own future trajectory can support students' STEM aspirations (Archer et al., 2015). However, in the present study, career orientation was assessed only through post-course self-reports, so the results should be interpreted as a measure of perceived career-related significance rather than as a proven change in career choice.

Artificial Intelligence in School Science and Physics Education

AI is used in school education for adaptive learning, automated feedback, analysis of learning data, question generation, and support for personalized explanations.

Research on GenAI in science education describes its potential for explaining concepts, generating questions, developing instructional materials, and engaging in dialogue with students (Cooper, 2023; Lo, 2023). Bibliometric evidence also indicates a substantial expansion of research on AI in science education across computer science, social science, science education, and technology and engineering education (Akhmadieva et al., 2023).

The emergence of GenAI has also expanded the range of available educational functions. Large language models can generate explanations, questions, examples, lesson plans, and assignment options. A systematic review by Marzano (2025), which included 197 publications on GenAI in K-12 education, identified opportunities for personalization, motivational support, instructional design, and improved assessment. The review also highlighted the need for teacher training, regulatory frameworks, privacy protection, and pedagogically sound use of the technology. In physics education, ChatGPT has been used to explain concepts, generate questions, analyze solutions, and organize critical work with system responses (Bitzenbauer, 2023; Liang et al., 2023).

Alarbi et al. (2024) used a quasi-experimental design to study the use of ChatGPT in teaching physics to high school students. The study demonstrated the potential benefit of the targeted use of the system when learning complex physics content. However, such results cannot be automatically generalized to any subject or any method of using AI. The effect depends on the duration of instruction, the structure of the tasks, students' initial level, the role of the teacher, and whether the system is used directly by students or indirectly through teacher-prepared materials.

One of the main challenges of large language models is the ability to generate plausible but scientifically inaccurate responses. Kasneci et al. (2023) summarized the opportunities and risks of using large language models in education. Benefits include adapting explanations, generating feedback, and supporting teacher preparation. Risks include factual errors, bias, academic dishonesty, overreliance on the system, and privacy concerns. Research by Polverini and Gregorcic (2024) further demonstrates that the high linguistic persuasiveness of a model does not necessarily translate into a robust ability to solve physics problems. When working with kinematics graphs, ChatGPT was able to offer productive reasoning strategies but struggled with interpreting visual information. A subsequent study of tasks on electricity and magnetism also revealed problems with spatial coordination, application of physical rules, and analysis of visual representations. These findings are particularly relevant for medical physics, where instruction often relies on device schematics, images, waveform graphs, and spatial models.

In physics and STEM education, ChatGPT has been studied as a tool for developing instructional tasks and as an object for organizing critical reasoning. These studies identified both the system's potential and limitations related to the quality of generated tasks, the reliability of information, and the need for expert review (Küchemann et al., 2023; Vasconcelos & Santos, 2023). More general research also points to risks of factual errors, dependence on prompt wording, academic dishonesty, breaches of confidentiality, and reduced student autonomy (Farrokhnia et al., 2024; Tlili et al., 2023).

Generative visual tools also require critical scrutiny. Cooper and Tang (2024) showed that depictions of science classrooms and educators generated by GenAI can reproduce simplified and stereotypical notions of science. In the context of medical physics, risks include not only social stereotypes but also incorrect depictions of equipment design, radiation direction, anatomy, or physical processes. Therefore, the visual appeal of materials should not be taken as evidence of their scientific accuracy.

UNESCO guidance on GenAI in education recommends a human-centered approach in which the educator retains responsibility for learning objectives, content validation, data protection, and the age appropriateness of materials (Miao & Holmes, 2023). This approach is consistent with the model of the present study: AI was used primarily by the teacher to prepare and adapt materials, rather than as an autonomous teacher or an unsupervised source of student responses. Therefore, the relevant question is not whether AI is effective in general, but whether pedagogically organized AI support provides an additional advantage when course content, teacher, and duration are comparable. This distinction is necessary for a more accurate interpretation of intergroup results.

Taken together, the emerging literature indicates that AI can serve multiple functions in science education, including personalized support, assessment, instructional assistance, and the development of learning materials. However, its educational value depends not only on the capabilities of a particular AI tool but also on how the technology is pedagogically

Table 1. Distribution of participants in the quasi-experimental stage

Condition of training	Class	Direction of study	n
Group A	10 B	Natural sciences and mathematics	25
Group A	11 B	Natural sciences and mathematics	26
Group B: AI-assisted course	10 C	Natural sciences and mathematics	26
Group B: AI-assisted course	11 C	Natural sciences and mathematics	26
Total			103

Note. Assignment was made at the level of existing classes; Group A was the comparison group, not the group without educational intervention

integrated, how its outputs are validated, and what role the teacher retains in the learning process (Fayzullina et al., 2025). This is particularly important in physics education, where GenAI may produce plausible reasoning while still making conceptual errors or struggling with visual and spatial information (Aldazharova et al., 2024).

Recent mapping of AI-related research in science education also shows growing attention to pedagogical applications, learner engagement, and classroom practice, while important questions concerning implementation and teacher-AI collaboration remain open (Karampelas, 2026). What remains insufficiently examined is whether a systematically organized package of AI-supported instruction—combining teacher-prepared generative explanations, visual materials, digital simulations, learning tasks, and teacher validation—provides an additional benefit across an extended secondary-school course when the core content, course duration, and teacher are held comparable. The present study addresses this gap in the context of a full-length medical physics elective course.

Digital Simulations and Multimodal Representation of Physical Processes

Digital simulations allow students to manipulate parameters, observe the dynamics of physical processes, and connect mathematical models with visually represented phenomena. Research reviews indicate that simulations can support conceptual understanding and complement traditional explanations, but their effectiveness depends on the design of the learning task, teacher facilitation, and integration with discussion, prediction, and subsequent reflection (Rutten et al., 2012; Smetana & Bell, 2012). Similar principles underline PhET simulations, which are designed to visualize physical models and support inquiry-based learning (Wieman et al., 2008).

In the present study, digital simulations were one component of an integrated support package, along with generative explanations, visual materials, and additional feedback. Therefore, the observed between-group difference should be interpreted as reflecting the entire pedagogically organized package rather than the isolated effect of the language model or a single digital tool.

METHOD

Study Design and Participants

The study was conducted during the 2025-2026 academic year and included two sequential stages. During the preliminary stage, a survey was conducted among 152 physics teachers and career guidance specialists from secondary education organizations in Kazakhstan. Its purpose was to elicit opinions on the need to integrate medical physics into school education and the potential role of an elective course in students' cross-curricular learning and students' career orientation.

The main stage was designed as a quasi-experimental study with nonequivalent groups and pre- and post-test measurements. Individual randomization was not conducted, as the study involved four existing classes following the science and mathematics track at a secondary school in Shymkent, Kazakhstan.

A total of 103 students from four existing classes participated in the quasi-experiment: 51 in group A and 52 in group B. The class distribution is presented in **Table 1**. Both groups completed the same 34-hour elective course in medical physics. Group A studied the course without systematically integrated AI support, whereas in group B, AI support was integrated across all course topics. Both groups were taught by the same teacher.

The sample size was determined by the availability of the four intact classes participating in the elective course rather than by an a priori power calculation. With the achieved sample of 51 students in group A and 52 students in group B, a two-sided comparison at $\alpha = .05$ and 80% power corresponded to a nominal minimum detectable standardized between-group effect of approximately Cohen's $d = 0.56$. This sensitivity estimate assumes independent individual observations and should thus be interpreted cautiously given the nesting of students within only four intact classes; the effective power may be lower when classroom clustering is considered.

Elective Course and Conditions for Its Implementation

The central pedagogical intervention was a 34-hour elective course, "medical physics," designed for students in grade 10 and grade 11. Its goal was to demonstrate the

Table 2. Structure of the 34-hour elective course "medical physics"

No	Thematic section	LH	Basic physical concepts	Medical context and technology
1	Introduction to medical physics	2	Subject, history and applications of medical physics	Development of medical diagnostics and treatment technologies
2	Mechanics, fluids and gases in medicine	5	Pressure, Bernoulli's law, viscosity, gas laws	Blood pressure measurement, blood flow, breathing techniques, ventilators
3	Acoustics and ultrasound technologies	5	Sound waves, frequency, ultrasound propagation, Doppler effect	Ultrasound diagnostics and Dopplerography
4	Electrical diagnostic methods	5	Electrical signals, potential, recording of bioelectrical activity	ECG, EEG, EMG, electrical stimulation
5	Optics and laser technologies	5	Light, optical properties of tissues, laser radiation	Optical diagnostics, laser medicine, endoscopy and optical tomography
6	Thermal phenomena in medicine	3	Temperature, heat transfer and thermoregulation	Thermography and therapeutic thermal procedures
7	Electromagnetic field and imaging technologies	4	Electromagnetic field and radiation, magnetic properties of matter, obtaining layer-by-layer images	MRI, radiography, computed tomography and tomographic methods
8	Nuclear physics in medicine	5	Radioactivity, radioisotopes, ionizing radiation and dosimetry	PET, radionuclide diagnostics, radiation safety and nuclear medicine technologies
	Total	34		

Note. LH: Lesson hours

application of school-level physical laws in modern diagnostics and treatment, strengthen interdisciplinary connections, and support students' career orientation toward medical and related scientific and technical fields.

The course included eight thematic sections with a total duration of 34 academic hours (**Table 2**).

The course was implemented from September 2025 to May 2026. In group A, instruction was based on teacher explanations, teaching materials, illustrations, discussions, projects, and calculation-based and non-standard problems in medical physics.

In group B, AI support was integrated throughout the 34-hour course. The primary user of the digital tools was the teacher, who used the web versions of ChatGPT, Canva AI, and Gamma available during the study period, as well as digital simulations. These tools were used to prepare explanations, visualize the operating principles of medical equipment, and develop presentations, projects, and non-standard tasks. Students primarily worked with materials prepared by the teacher, completed projects, analyzed medical applications of physical laws, and discussed solutions.

Exact software version numbers were not recorded because the web services were updated throughout the academic year. Therefore, the study reports the period of use, pedagogical function, and primary user of the tools but does not assign unconfirmed model numbers.

Materials prepared with AI support were checked by the teacher for alignment with the lesson topic, scientific accuracy, comprehensibility for high school students, and correct description of medical technologies. The course design and educational materials were also discussed with the academic supervisor and specialists from the university research unit. The intervention description was structured with reference to the TIDieR recommendations (Hoffmann et al., 2014) (**Table 3**).

An Example of an AI-Supported Lesson

To illustrate the pedagogical implementation of the course, the lesson "imaging methods based on electromagnetic radiation: radiography and computed tomography" was selected from the section "electromagnetic field and imaging technologies." The lesson lasted one academic hour and was conducted in both groups based on the same learning objectives: explaining the physical nature of X-rays, the main components of an X-ray tube, the interaction of radiation with tissues, and the principle of obtaining a cross-sectional image in computed tomography.

In group A, the content was taught through teacher explanations, static illustrations, instructional text, discussion, and problem solving. In group B, the same content was accompanied by interactive material prepared by the teacher using Canva AI and other digital tools. The material included a sequential presentation of the structure of an X-ray tube, the process of radiation generation, the interaction of X-rays with body tissues, and the reconstruction of a computed tomography image.

Students compared the physical properties of X-rays with their diagnostic applications, analyzed differences in radiation attenuation by tissues, and explained how a series of projections is converted into a cross-sectional image. The AI-supported material did not replace the teacher's explanation but was used as structured visual support for discussion and problem solving (**Figure 1**).

Table 3. Comparison of learning conditions

Intervention component	Group A: Basic version of the course	Group B: AI-assisted course
Course content	All eight sections of medical physics	All eight sections of medical physics
Duration	34 academic hours	34 academic hours
Implementation period	September 2025-May 2026	September 2025-May 2026
Teacher	The same physics teacher	The same physics teacher
The main way of explanation	Teacher's explanations and teaching materials	Teacher explanations supplemented with AI-enabled materials
Educational materials	Text materials, illustrations, discussions and tasks	Interactive visual materials, structured presentations, simulations, and AI-assisted tasks
ChatGPT	Not used systematically	Preparation of explanations, questions, educational situations and non-standard tasks
Canva AI	Not used systematically	Preparation of visual and interactive materials
Gamma	Not used systematically	Structuring content and creating presentation materials
AI primary user	Not applicable	Teacher
The role of students	Studying the content, solving problems and discussing	Working with prepared materials, projects, analysis and solving non-standard problems
Content check	Teacher	Teacher

Note. AI was an additional pedagogical tool for group B and did not replace explanation, feedback and assessment by the teacher

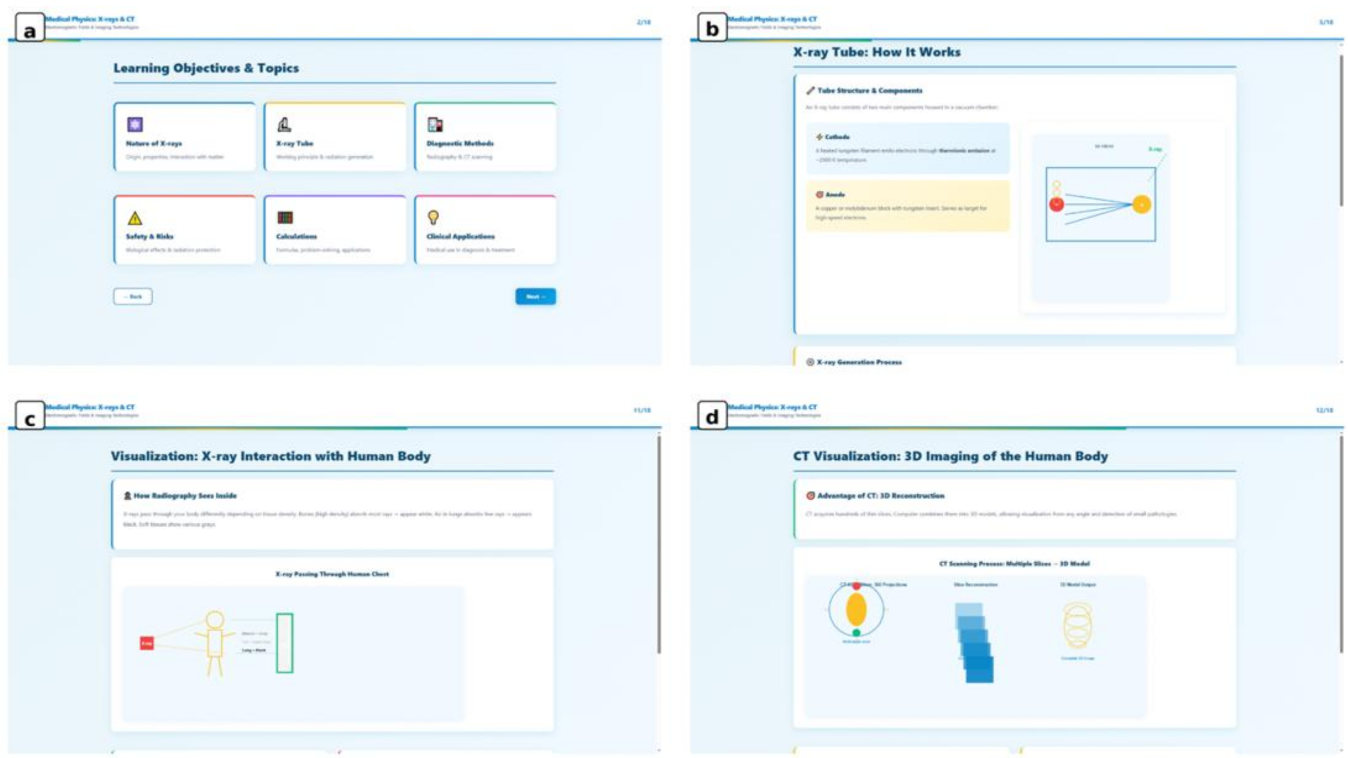

Figure 1. An example of AI-supported visual material on radiography and computed tomography: (a) learning objectives and topic structure; (b) X-ray tube structure; (c) interaction of X-ray radiation with tissues; (d) principle of reconstruction of a computed tomographic image (this material was prepared by the teacher using Canva AI and digital tools, verified for scientific accuracy, and used as visual support for explanation, discussion, and problem-solving)

Data Collection Tools and Procedure

The preliminary stage used a questionnaire consisting of five statements rated on a five-point Likert scale, from 1 = strongly disagree to 5 = strongly agree. The statements addressed the interdisciplinary role of medical physics, career guidance, the importance of physics for medical technologies, and the need for a specialized elective course. The results from this stage were used only to contextualize the development of the course.

The primary instrument of the quasi-experiment was a medical physics knowledge test, consisting of 25 multiple-choice items with four answer options. A correct answer was awarded 1 point, whereas an incorrect or missing answer was awarded 0 points; the total score ranged from 0 to 25 points. The same test was used before and after the course. The items covered ultrasound, bioelectrical diagnostic methods, optics, medical imaging, nuclear medicine, and therapeutic applications of physical phenomena. The full version of the test and its content specification are provided in the supplementary materials. Pre- and post-test results were linked at the individual level using a pre-established fixed order of students in the class lists. In group A, the first 25 entries corresponded to grade 10B and the next 26 to grade 11B; in group B, the first 26 entries corresponded to grade 10C and the next 26 to grade 11C. The order was the same for both measurements and was checked by the teacher. After matching, participants were assigned anonymous research codes.

After completing the course, both groups completed eight-item questionnaires on a five-point scale. Six items were identical and addressed interest in the connection between physics and medicine, understanding of the application of physics in medicine, perceived usefulness of the course, reflection on a future career, interest in medical professions, and desire to continue studying medical physics.

Two items differed between the questionnaires. In group A, these items assessed general understanding of medical devices and the connection between physics and medicine; in group B, they assessed the perceived contribution of AI support to these outcomes. Therefore, direct between-group comparison was performed only for the six identical items, whereas the two AI-specific statements were analyzed separately. Because the student questionnaire was administered only after the course, the results were interpreted as post-course perceptions rather than as evidence of a proven change in interest or career choice.

Quality of Instruments and Data Analysis

The test and questionnaires were developed by the course teacher under the scientific and methodological supervision of the second author. Internal content validation included assessment of alignment with the course program, scientific accuracy, clarity of wording, and accuracy of the answer keys. Because the developer and the supervisor were members of the research team, this procedure was not considered independent external validation.

The content distribution of the test was uneven: electrical diagnostic methods and nuclear physics were represented more extensively, whereas the section on mechanics, fluids, and gases did not have a separate item. Therefore, the total score was interpreted as an indicator of performance on 25 predefined items rather than as a comprehensive assessment of all course sections.

The internal consistency of the dichotomous test was assessed using KR-20, whereas the questionnaires were assessed using Cronbach's alpha. The unstable internal consistency of the test was considered when interpreting the results; items were not removed after the experiment solely to increase the reliability coefficient. For the preliminary stage, frequencies and percentage distributions of expert responses were calculated. For **RQ1**, changes in scores within each group were analyzed using a paired-samples t-test; mean gain, 95% confidence interval (CI), and Cohen's d_z were also calculated (Taber, 2018).

For **RQ2**, baseline comparability between the groups was assessed using the pre-test. The primary analysis was conducted using ANCOVA (Lakens, 2013; Van Breukelen, 2006), with post-test score as the dependent variable, instructional condition as the main factor, pre-test score as the covariate, and grade level as a categorical control factor. Because of heterogeneity in the residuals, HC3 heteroscedasticity-robust standard errors were used. In addition, a secondary model including the interaction between instructional condition and grade level was fitted to examine whether the between-group difference varied between grade 10 and grade 11 students.

Because the instructional condition was implemented at the level of intact classes, students were nested within four classrooms rather than being fully independent observations. Only two classes were available in each condition, with one intact class representing each grade-by-condition combination. Given this limited cluster structure, class-level variation could not be reliably separated from variation associated with the instructional condition. Therefore, the ANCOVA was retained as an individual-level analysis, and classroom nesting was explicitly considered as a statistical limitation when interpreting the between-group results.

As a sensitivity analysis, the ANCOVA was repeated after excluding the single participant in group B whose post-test score was one point lower than the pre-test score. This analysis was conducted only to assess the robustness of the between-group result; the participant was retained in all primary analyses because there was no evidence that the observation was erroneous.

Student responses to **RQ3** were analyzed at the statement level. Frequencies, medians, and interquartile ranges were calculated for the six identical items; between-group differences were tested using the Mann-Whitney U test. Items about the role of AI in group B were analyzed descriptively. The level of statistical significance was set at $p < .05$.

Ethical Considerations

The study received approval from the local ethics committee of the university, protocol no. 1 dated September 15, 2025. Participation was voluntary; for students who were minors, informed consent from parents or legal representatives and student assent were required. The results were anonymized and did not affect official school grades. Students' personal or medical data were not entered into GenAI systems. Responsibility for the content of AI-supported materials and their use remained with the teacher.

Table 4. Results of a preliminary survey of physics teachers and career guidance specialists (n = 152)

Statement	M	SD	Agree or strongly agree: n (%)
Integration of medical physics promotes interdisciplinary connections	3.95	.81	117 (77.0)
Medical physics promotes early career guidance	4.11	.85	126 (82.9)
The role of physics in medical career guidance can be strengthened by an elective course	3.93	.83	108 (71.1)
The operation of modern medical technologies is based on physical laws.	4.18	.79	134 (88.2)
Medical physics elective programs are underrepresented in schools	4.07	.86	120 (78.9)

Note. Response scale: 1 – Strongly disagree & 5 – Strongly agree

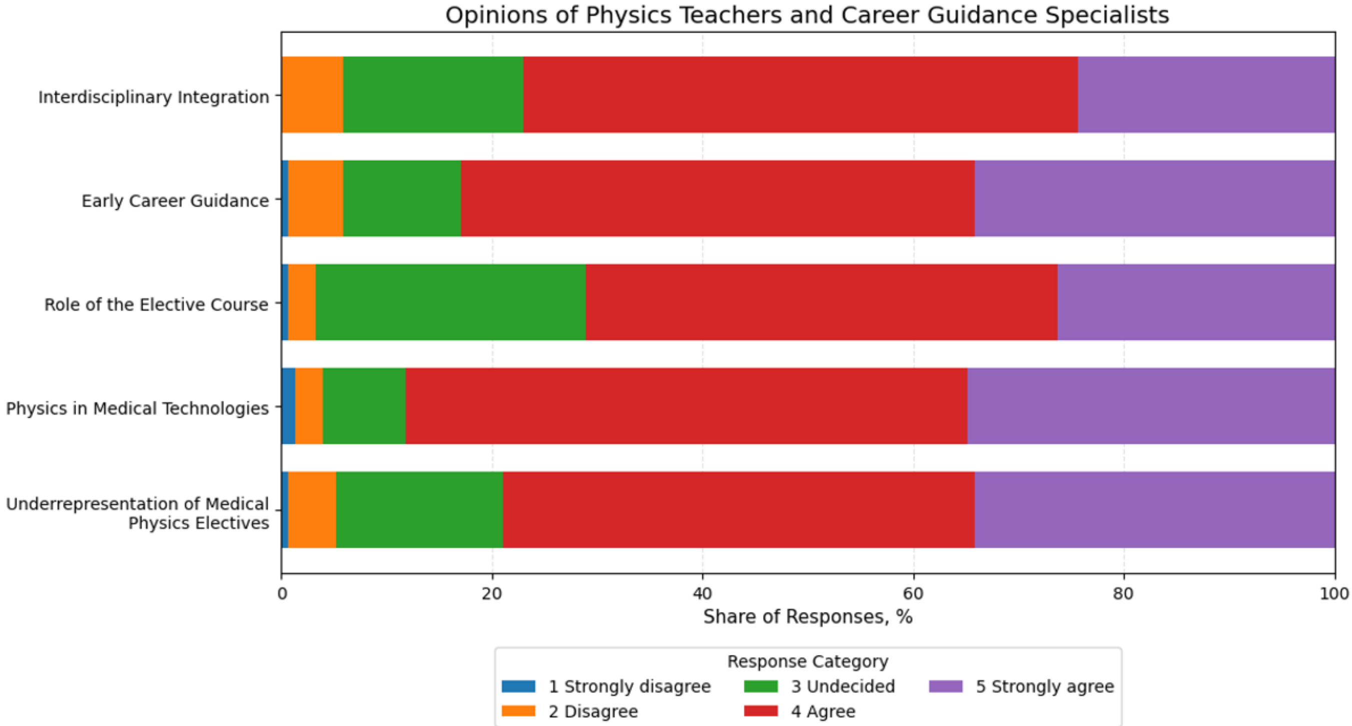

Figure 2. Distribution of teachers' and career guidance specialists' responses to the five preliminary survey statements (the authors' own elaboration)

RESULTS

Preliminary Justification for the Need for an Elective Course

A preliminary survey of physics teachers and career guidance specialists was used to contextualize the development of the course rather than as a standalone RQ in the quasi-experimental phase. For all five statements, the predominant responses were "agree" and "strongly agree" (**Table 4**).

The highest percentage of agreement was obtained for the statement about the connection between modern medical technologies and the laws of physics, with 88.2% of respondents agreeing or strongly agreeing. More than 80% of participants also agreed or strongly agreed that studying medical physics can support early career orientation. The combined mean (M) across the five statements was M = 4.05, and the internal consistency of the questionnaire was high, α = .871. These results were used as a preliminary justification for the educational need, not as evidence of the effectiveness of the course (**Figure 2**).

Change in Knowledge After Completing an Elective Course

RQ1 examined changes in students' knowledge after completing the 34-hour elective course. Before the course, the M scores of the two groups were similar: M = 11.24 in group A and M = 11.35 in group B.

After the course, the M score for group A increased to 19.71 points, whereas the M score for group B increased to 21.04 points. Descriptive statistics and paired comparison results are presented in **Table 5**.

Scores increased significantly in both groups. In group A, all 51 students showed positive gains. In group B, 51 of 52 students improved their scores, whereas one participant's post-test score was one point lower than the pre-test score. This result was retained in the analysis because there was no reason to consider it erroneous (**Figure 3**).

Table 5. Changes in knowledge test results within groups

Group	n	M ± SD		Mean gain: 95% CI	t(df)	p	Cohen's d_z
		Pre-test	Post-test				
A: The course does not have systematic AI support	51	11.24 ± 1.88	19.71 ± 1.59	8.47 [7.89; 9.05]	29.47 (50)	< .001	4.13
B: AI-enabled course	52	11.35 ± 1.69	21.04 ± 2.28	9.69 [9.06; 10.32]	31.01 (51)	< .001	4.30

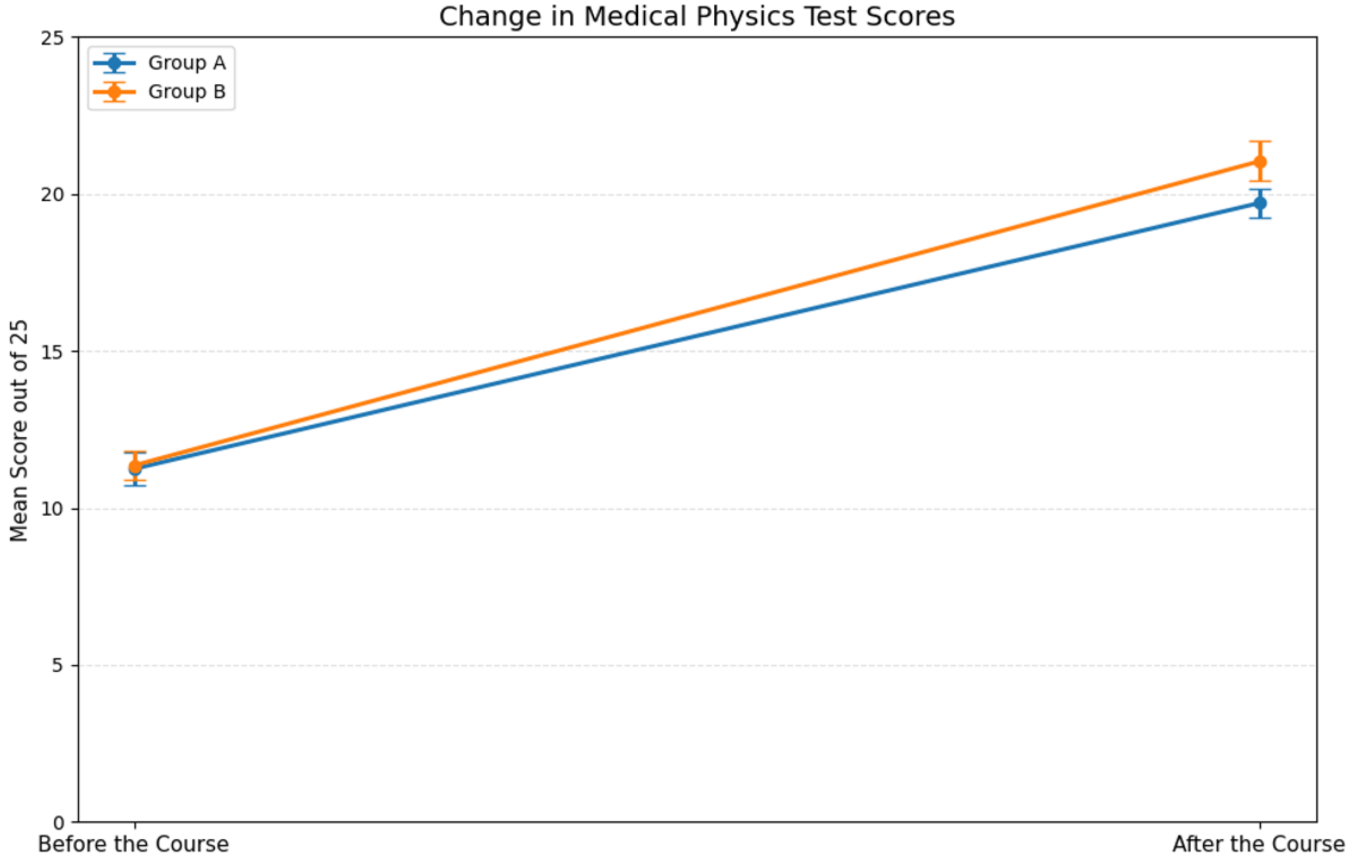

Figure 3. Mean pre- and post-test scores in group A and group B with 95% CIs (the authors' own elaboration)

Table 6. Intergroup comparisons of test results

Comparison	Difference B-A	95% CI	p	Cohen's d
Pre-test	.11	[-.59; .81]	.754	.06
Post-test	1.33	[.56; 2.10]	< .001	.68
Individual growth	1.22	[.38; 2.06]	.005	.57

Note. Independent t-tests were used for between-group comparisons; Cohen's d was calculated as the standardized difference between groups

An Additional Difference Related to Artificial Intelligence Support

RQ2 examined whether the two groups' performance differed after controlling baseline knowledge.

On the pre-test, no statistically significant difference was found between the groups: the mean difference was 0.11 points ($p = .754$, $d = .06$). After completing the course, group B scored, on average, 1.33 points higher than group A. The average individual gain was also higher in group B by 1.22 points (**Table 6**).

The primary ANCOVA, controlling pre-test score and grade level, showed a statistically significant difference between instructional conditions, $F(1, 99) = 15.30$, $p = .000168$. The adjusted advantage of group B was 1.31 points, 95% CI [0.65, 1.98]. The adjusted post-test means were approximately 19.72 points for group A and 21.03 points for group B.

Grade level was also significantly associated with post-test performance after controlling for baseline score and instructional condition, $F(1, 99) = 22.13$, $p = .000008$. Eleventh-grade students scored, on average, 1.57 points higher than tenth-grade students, 95% CI [0.91, 2.23]. In the secondary model including the instructional condition × grade level interaction, the interaction was not statistically significant, $F(1, 98) = 0.23$, $p = .629$. Thus, there was no evidence that the between-group difference varied between grade 10 and grade 11 students.

The sensitivity analysis excluding the single group B participant whose post-test score decreased produced

Table 7. Students' responses to comparable statements after the course

Statement	Group A: median [IQR]; agreement, %	Group B: median [IQR]; agreement, %	pHolm	r_{rb}
Interest in the relationship between physics and medicine	4 [3-4]; 64.7	4 [4-4]; 76.9	.336	.14
Understanding the application of physics in medicine	4 [3-4]; 64.7	4 [4-5]; 100	< .001	.56
Usefulness of classes	4 [4-4]; 80.4	4 [4-5]; 100	< .001	.39
The course helped me think about my future profession	4 [4-4]; 80.4	4 [4-5]; 100	< .001	.47
Interest in medical professions	4 [3-4]; 72.5	4 [4-4]; 76.9	.488	.07
Desire to continue studying medical physics	4 [4-4]; 78.4	4 [4-5]; 100	.002	.35

Note: Between-group differences were tested using the Mann–Whitney U test. P values were adjusted using the Holm method. r_{rb} is the rank-biserial correlation. Agreement includes responses 4 and 5

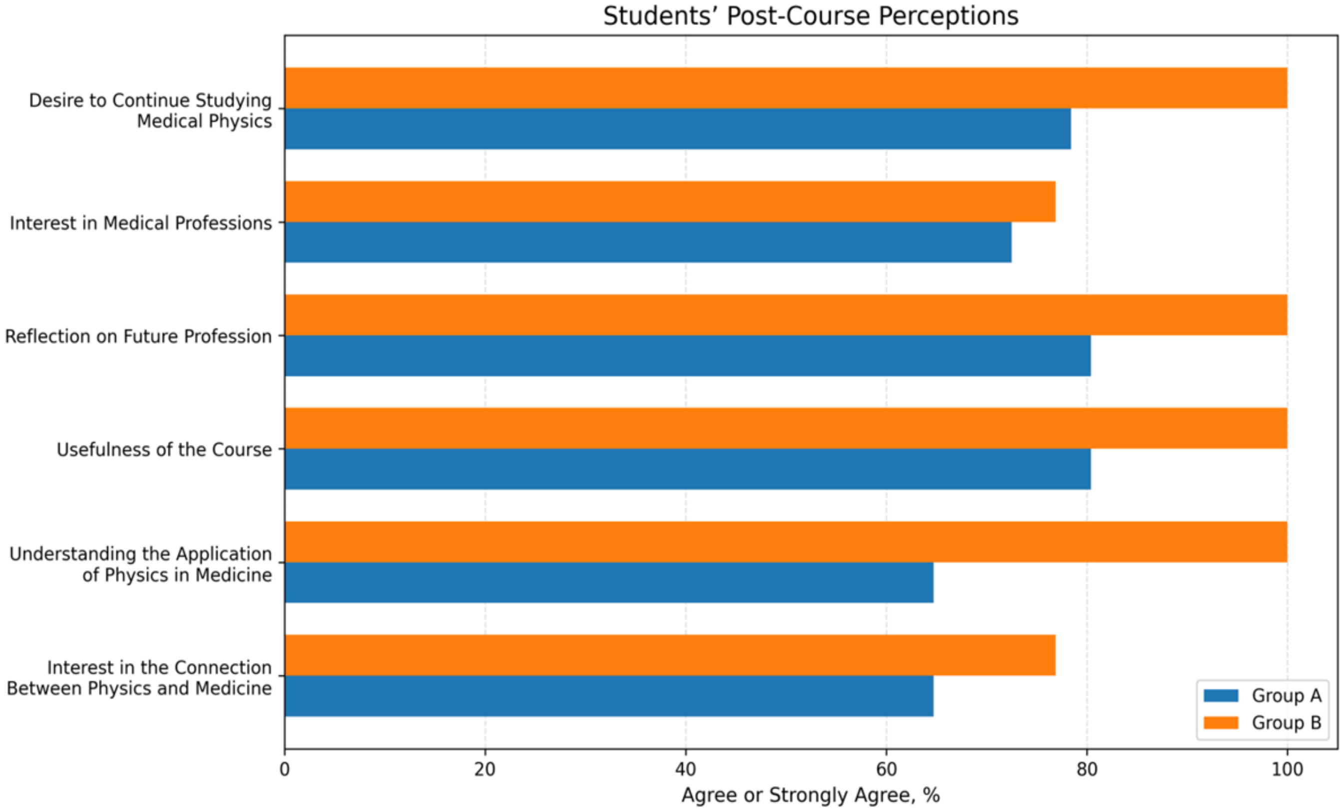

Figure 4. Percentage of students who agreed or strongly agreed with the six comparable post-course statements (the authors' own elaboration)

the same substantive conclusion. With this participant excluded, the adjusted advantage of group B was 1.51 points, 95% CI [0.97, 2.05], $F(1, 98) = 31.13$, $p < .001$. Thus, the observed between-group difference was not dependent on the inclusion of this single observation.

Perception of the Course, Interest and Career Orientation

RQ3 concerned students' post-course perceptions of the course. Between-group analysis was conducted only for the six identical items in the two questionnaires. The results are presented in **Table 7**.

In both groups, the majority of students positively rated the usefulness of the course, its career-oriented function, and the opportunity for further study of medical physics. Group B demonstrated higher ratings for understanding the application of physics in medicine, the usefulness of the course, the impact of the course on students' reflections about their future profession, and the desire to continue studying medical physics.

No statistically significant differences were found between the groups for general interest in the connection between physics and medicine or for interest in medical professions. This indicates that the higher results in group B did not apply equally to all aspects of students' attitudes.

The two AI-specific items in group B were analyzed separately. For each of the two items, 76.9% of students agreed or strongly agreed that AI support helped them better understand the operation of medical devices and the connection between physics and medicine. For both items, the median was 4 and the M was 3.88 (standard deviation [SD] = .58) (**Figure 4**).

DISCUSSION

This study aimed to evaluate a 34-hour elective course in medical physics for grade 10-11 students and to examine the additional difference associated with the AI-supported form of its implementation. The results provide a consistent picture. Physics teachers and career guidance specialists generally supported the need for the course; knowledge test scores increased significantly in both groups after course completion; higher adjusted post-test performance was observed in the AI-supported condition; and most students positively rated the usefulness of the course and its connection to future career pathways. However, these findings require cautious interpretation given the quasi-experimental design, the absence of a group that did not take the course, the nesting of students within only four intact classes, and the specific characteristics of the researcher-developed test.

Perceived Need for an Elective Course in Medical Physics

The results of the preliminary survey revealed a high level of agreement among specialists that modern medicine is closely linked to physics and that an elective course can enhance interdisciplinary learning and students' career orientation. The highest level of agreement was obtained from the statement about the physical foundations of modern medical technologies. This result is understandable because diagnostic imaging, bioelectric signal recording, ultrasound methods, laser technologies, nuclear medicine, and radiation protection are directly based on physical laws.

Educators' opinions also indicate that medical physics is perceived as an underrepresented but potentially significant area of school education. A similar issue has been noted in international outreach studies, where many students had limited understanding of medical physics and related professional opportunities before participating in specialized programs. In particular, Yorke et al. (2025) showed that practical exposure to medical physics among high school students expanded their awareness of the applications of physics in healthcare. Santoso et al. (2021) also emphasized that medical physics outreach programs should not only demonstrate technologies but also link content to the needs of the target audience and to career pathways.

At the same time, the results of our preliminary survey cannot be considered as evidence of the objective absence of medical physics programs in all schools in Kazakhstan. The survey reflects the perceptions of 152 participants, not the results of a national curriculum audit. Its significance lies elsewhere: it confirms that the idea of the course corresponds to the professional needs of teachers and career guidance specialists and was not developed solely as a research construct disconnected from school practice.

Change in Knowledge After Completing an Elective Course

The main result of the study is a significant increase in mean test scores in both groups. This pattern is consistent with the educational potential of the 34-hour elective course, although the absence of a no-course comparison group does not allow the observed gains to be attributed exclusively to the course.

This interpretation is consistent with the main logic of the study: the central pedagogical intervention was the 34-hour elective course itself rather than AI. Group A was not left without instruction and did not serve as a passive control group. Students in group A completed the same course, which included eight thematic sections, projects, discussions, and non-routine problems. Therefore, the gains observed in both groups are consistent with learning during the systematic study of medical physics across the academic year, while other factors, including repeated testing and concurrent learning, cannot be completely excluded.

One possible mechanism for the observed growth is context-based design of the course. Physical concepts were not studied in isolation but through real medical processes and technologies. Pressure was linked to blood flow and medical measurements, mechanical waves to ultrasound diagnostics, bioelectrical signals to ECG and EEG, electromagnetic phenomena to MRI and computed tomography, and radioactivity to PET and radiation safety. This way of organizing the material could have helped students see functional connections between a physical law and its practical application.

This finding is consistent with the synthesis by Bennett et al. (2007), which found that context-based teaching can support more positive attitudes toward science without reducing mastery of subject content. It is also consistent with the results of Fagerstrom et al. (2019), who showed that complex ideas in medical physics and radiation therapy can be adapted for students through models and hands-on activities.

The international study by Sudirman et al. (2023) complements this interpretation. In that study, a medical physics module was used as a way to introduce physics through real-world applications and stimulate students' interest. The authors showed that perceptions of the module depended on the pedagogical approach, type of school, equipment, and language difficulty. This is particularly important for the present study: the medical context should not be considered automatically effective. Its effectiveness is determined by the structure of the course, the sequence of explanation, the accessibility of terminology, and how actively students work with the content.

These findings also support a broader point that learning outcomes in physics depend on the structure and sequencing of instruction as well as on students' prior knowledge. In upper-secondary physics, Lazos et al. (2025) demonstrated the value of a deliberately structured teaching-learning sequence for identifying students' difficulties and refining instruction around measurement and data processing. Although conducted with first-year university physics students rather than secondary-school students, Gleichmann et al. (2025) similarly found that learning gains from a mathematics pre-course were related to students' prior knowledge, with particularly clear improvement among students with intermediate baseline knowledge. These studies do not concern AI, but they reinforce the importance of interpreting the present knowledge gains in relation to structured course design and students' starting levels rather than attributing improvement to a technological component alone.

However, the absence of a no-course group, the repeated use of the same test, and the limited internal consistency of the test make it difficult to attribute all gains solely to the course.

Between-Group Difference Associated with AI-Supported Learning

After controlling baseline and grade level, group B demonstrated a higher post-test score, but the between-group difference was considerably smaller than the overall gain in both groups.

This result does not change the main conclusion about the value of the course, but it indicates that higher adjusted post-test performance was observed in the condition implemented with pedagogically organized AI support. The difference between the groups was considerably smaller than the overall gain within each group. Therefore, it would be incorrect to claim that AI alone determined the primary educational outcome. A more reasonable interpretation is that both groups benefited from the course, while group B received additional support in mastering the same content.

A possible mechanism for this advantage is related to the multimodal representation of complex processes. Many phenomena in medical physics cannot be directly observed in a classroom. Students cannot see X-ray transmission through tissues, the formation of a tomographic slice, the response of hydrogen nuclei in a magnetic field, or the transformation of bioelectrical signals into graphical representations. In group B, such processes were presented through teacher-prepared explanations, visual materials, simulations, and structured presentations.

Reviews by Rutten et al. (2012) and Smetana and Bell (2012) demonstrate that digital simulations can support conceptual understanding, especially when accompanied by tasks, prediction, discussion, and teacher guidance. This is consistent with the model of our course: digital materials were not used as stand-alone demonstrations but were integrated into projects, analysis of medical technologies, and problem solving.

An additional mechanism may have been the variability of explanations. ChatGPT was used by the teacher to prepare alternative formulations, questions, and tasks. Canva AI and Gamma helped structure and visualize the content. This use may have reduced the gap between an abstract physics concept and its medical application, especially for students for whom a single mode of explanation was insufficient.

These results are partially consistent with Alarbi et al. (2024), who found an advantage of using ChatGPT in high school physics learning in a quasi-experimental study. However, there is a significant difference between the studies. The present study did not test the effect of ChatGPT in isolation. Group B received an integrated support package that included a language model, visual materials, presentation tools, simulations, projects, non-standard problems, and teacher review of content.

Therefore, the between-group difference cannot be attributed solely to ChatGPT, Canva AI, Gamma, or the simulations. It is related to the entire form of course implementation. In this regard, the results are consistent with the review by Heeg and Avraamidou (2023), which showed that the impact of AI in school science education depends not only on the type of technology but also on the pedagogical function, the organization of learning, and the role of the teacher.

Teacher agency was likely an important factor. In the study, students primarily worked not with uncontrolled system responses but with materials pre-selected and reviewed by the teacher. This is consistent with UNESCO recommendations for the human-centered use of GenAI, in which the teacher retains responsibility for the goals, content, accuracy, and safety of learning.

Recent research in science education supports this cautious, pedagogically oriented interpretation of AI use. A systematic review by Fayzullina et al. (2025) found that AI applications in science education can support student performance, motivation, engagement, and personalized learning, while also identifying challenges related to system reliability, technological infrastructure, ethical issues, and teacher preparation. Bibliometric evidence further indicates that AI-supported science teaching, including inquiry-oriented applications, is a rapidly expanding area of research (Akhmadieva et al., 2023; Karampelas, 2026). In physics specifically, Aldazharova et al. (2024) showed that GPT-4 performed successfully on many force concept inventory items but experienced difficulties with questions requiring interpretation of figures and spatial reasoning. These findings strengthen the rationale for the teacher-mediated model used in the present study: AI-generated materials should not be treated as inherently reliable instructional resources but should be selected, checked, and integrated within a pedagogically structured learning process. Accordingly, the present findings are better interpreted as evidence associated with a teacher-mediated AI-support package than as evidence of the independent effectiveness of any single AI tool.

Consequently, both versions of the course were associated with substantial performance gains, while higher post-test performance was observed in the AI-supported condition after adjustment for baseline knowledge and grade level. However, because the instructional condition was implemented within only four intact classes, this between-group difference cannot be fully separated from unmeasured class-level characteristics.

Perception of the Course and Its Career Guidance Significance

The majority of students in both groups positively rated the usefulness of the course, the connection between physics and medicine, and the opportunity for further study of medical physics. This suggests that the medical context was perceived not only as an additional set of physics problems but also as content related to real-world technologies and future educational pathways.

This positive perception is consistent with studies by Sudirman et al. (2023) and Yorke et al. (2025), in which medical physics also stimulated students' interest and broadened their understanding of physics applications. Medical technologies have high perceived practical value: students are familiar with ultrasound, X-ray imaging, MRI, and ECG as elements of real healthcare systems, but they do not always understand the physical mechanisms by which these technologies operate.

This result can be interpreted through the concept of perceived usefulness. Hulleman and Harackiewicz (2009) demonstrated that establishing connections between scientific content, personal life, and future goals can support interest and academic achievement. In our course, this connection was formed through the question not only "What law applies?" but also "Where is this law used in medicine, and what professional activities does it enable?"

Career-related relevance is also linked to students' perceptions of physics as part of their future identity. Hazari et al. (2010) showed that interest, recognition, confidence, and expectations of the usefulness of physics are associated with physics identity and future career-related choices. Medical physics can potentially expand this choice because it demonstrates that physics is applicable not only to traditional engineering professions but also to healthcare, diagnostics, medical imaging, and radiation therapy.

In group B, students gave higher ratings for their understanding of the application of physics in medicine, the usefulness of the course, the impact of the course on their reflections about their future profession, and their desire to continue studying medical physics. These differences are consistent with the higher post-test scores and may indicate that visual and interactive support made medical applications more understandable and accessible.

However, there were no differences for two important indicators: general interest in the connection between physics and medicine and interest in medical professions. This indicates that AI support was not associated with higher ratings on all motivational outcomes. Students may have understood the material better and found the course useful while still maintaining previously formed career preferences.

Career preferences are influenced by multiple factors and should not be viewed as a direct consequence of a single course.

Because measurement was conducted only after the course, the results reflect students' post-course perceptions rather than a confirmed change in career intentions.

Limitations of the Study

The study results must be interpreted with several limitations in mind.

First, the quasi-experiment was conducted in a single school and included only four intact classes. Students were nested within classrooms, and the instructional condition was assigned at the class level rather than at the individual-student level: group A comprised classes 10B and 11B, whereas group B comprised classes 10C and 11C. Consequently, observations from students within the same classroom cannot be assumed to be fully independent. Moreover, because only one intact class represented each grade-by-condition combination, class-level characteristics could not be fully separated from the instructional condition. The individual-level ANCOVA adjusted for baseline score and grade level but did not model between-class clustering. Thus, the uncertainty surrounding the estimated between-group difference may be greater than that indicated by the individual-level analysis, and the observed advantage of group B should not be interpreted as a pure causal effect of AI-supported instruction. Future studies should include a substantially larger number of classes within each condition to permit reliable multilevel or cluster-adjusted analyses.

Second, the study did not include a group that did not take the elective course. Therefore, the significant improvement in scores in group A and group B cannot be interpreted as a strictly established causal effect of the course. Repeated testing, naturalistic learning throughout the academic year, and other unaccounted factors could have influenced the results. However, the identical implementation period, course content, and teacher strengthen the comparability of the two study conditions.

Third, all classes were taught by a single teacher. This reduced variability due to differences between teachers but also limited the transferability of the results to other schools and teaching practices. The effectiveness of the course may depend on the teacher's subject preparation, experience working with medical content, and ability to use AI tools pedagogically.

Fourth, the same 25-item test was administered before and after the course. Repeated use of identical items may have resulted in familiarization with the format and content of the questions. Furthermore, the test unevenly covered the eight sections of the course: electrical diagnostic methods and nuclear physics were represented most extensively, whereas the section on mechanics, fluids, and gases did not have a separate test item.

The internal consistency of the test was limited and varied across groups and measurement stages. This may be due to the thematic heterogeneity of the items, the narrow range of post-test scores, and differences in the difficulty of individual items. Therefore, the total score was interpreted as performance on a specific set of 25 predefined items rather than as a highly reliable measure of a single latent construct of medical physics.

Fifth, course interest, usefulness, and career orientation were assessed only after course completion. The lack of a baseline measurement precludes any assertion that the course increased interest or changed career intentions. The data obtained reflect students' post-course perceptions and the perceived career-related relevance of the course. A delayed measurement was also not conducted, so the persistence of knowledge and the impact of the course on actual choice of subsequent educational pathway remain unknown.

The two versions of the student questionnaire were not completely equivalent: two items in group B assessed the perceived contribution of AI, whereas the corresponding items in group A were worded without mentioning the technology. For this reason, direct between-group comparisons were conducted only for the six identical statements.

Sixth, the AI-supported condition included a combination of tools and learning activities: ChatGPT, Canva AI, Gamma, digital simulations, visual materials, projects, and non-standard tasks. Therefore, it is impossible to determine which specific component contributed most to the between-group difference. The results apply to the pedagogically organized support package as a whole.

The exact versions of the web services used and the full prompt history for each lesson were not recorded. Because the tools were updated during the 2025-2026 academic year, the reproducibility of individual technical details is limited. However, the course content, the pedagogical functions of the tools, and the teacher's role were recorded.

Future studies should cover multiple schools and regions, include a larger number of classes in each condition, add a separate group without the elective course, conduct independent external validation of the instruments, and include delayed outcome measurement. It would also be useful to compare individual components of AI support and examine whether differences persist after course completion.

CONCLUSION

The study developed and evaluated a 34-hour elective course in medical physics for grade 10-11 students.

A preliminary survey showed that physics teachers and career guidance specialists generally positively assessed the integration of medical physics into school education and the potential role of such a course in revealing the connection between physics, medical technologies, and future professions.

After course completion, knowledge test scores increased significantly in both groups. This is the primary finding of the study and indicates the educational potential of systematically studying medical physics throughout the academic year. However, the absence of a group that did not take the course precludes strict causal inference regarding the effect of the elective course.

After controlling baseline knowledge and grade level, higher post-test performance was observed in the AI-supported condition. This between-group difference was smaller than the overall gains observed in both groups and should be interpreted cautiously because the instructional conditions were implemented within only four intact classes. Therefore, the findings are consistent with a possible additional benefit of the pedagogically organized AI-supported implementation, but they do not establish a causal effect of AI support independently of class-level characteristics. Students positively rated the usefulness and career-related relevance of the course, but the study design did not allow conclusions about changes in career choice.

These results support the further development of medical physics as a school elective and the conduct of multicenter studies with a more rigorous design and improved assessment instruments.

Author contributions: AK: conceptualization, data curation, visualization, writing – original draft; **BR:** formal analysis, validation; **DB:** conceptualization, methodology, writing – review & editing; **TT & BU:** conceptualization, formal analysis, methodology, supervision, validation, writing – review & editing. All authors agreed with the results and conclusions.

Funding: No funding source is reported for this study.

Acknowledgments: The authors would like to thank the administration, teachers, and students at Bolashak School in Shymkent for their assistance in conducting the study. The authors would also like to thank the specialists from the research department at South Kazakhstan Pedagogical University named after Ozbekali Zhanibekov for their scientific and methodological discussions of the course curriculum and teaching materials, as well as the physics teachers and career guidance specialists who participated in the preliminary survey.

Ethical statement: This study was approved by the Institutional Ethics Committee at South Kazakhstan Pedagogical University named after Ozbekali Zhanibekov on 15 September 2025 with approval number 1. Participation was voluntary. Prior to data collection, written informed consent was obtained from parents or legal guardians of minor participants, and student assent was obtained. Participants were informed of the purpose of the study, procedures, data use, and the right to withdraw from participation without academic or other adverse consequences. The collected data were anonymized and used solely for research purposes.

AI statement: Generative AI tools—ChatGPT, Canva AI, and Gamma—were used as part of the pedagogical intervention in group B, according to the procedure described in the methods section. They were used primarily by the teacher to prepare explanations, visual materials, presentations, projects, and non-standard assignments. All learning materials created with these tools were reviewed and edited by the authors for scientific accuracy, pedagogical appropriateness, and consistency with the course content. The authors state that after writing the initial draft of the article, it was checked for grammar and style using the latest version of the AI language model.

Declaration of interest: No conflict of interest is declared by the authors.

Data sharing statement: Data supporting the findings and conclusions are available upon request from the corresponding author.

REFERENCES

Akhmadieva, R. S., Udina, N. N., Kosheleva, Y. P., Zhdanov, S. P., Timofeeva, M. O., & Budkevich, R. L. (2023). Artificial intelligence in science education: A bibliometric review. *Contemporary Educational Technology, 15*(4), Article ep460. https://doi.org/10.30935/cedtech/13587

Alarbi, K., Halaweh, M., Tairab, H., Alsalhi, N. R., Annamalai, N., & Aldarmaki, F. (2024). Making a revolution in physics learning in high schools with ChatGPT: A case study in UAE. *Eurasia Journal of Mathematics, Science and Technology Education, 20*(9), Article em2499. https://doi.org/10.29333/ejmste/14983

Aldazharova, S., Issayeva, G., Maxutov, S., & Balta, N. (2024). Assessing AI's problem solving in physics: Analyzing reasoning, false positives and negatives through the force concept inventory. *Contemporary Educational Technology, 16*(4), Article ep538. https://doi.org/10.30935/cedtech/15592

Archer, L., Dawson, E., DeWitt, J., Seakins, A., & Wong, B. (2015). "Science capital": A conceptual, methodological, and empirical argument for extending Bourdieusian notions of capital beyond the arts. *Journal of Research in Science Teaching, 52*(7), 922-948. https://doi.org/10.1002/tea.21227

Bennett, J., Lubben, F., & Hogarth, S. (2007). Bringing science to life: A synthesis of the research evidence on the effects of context-based and STS approaches to science teaching. *Science Education, 91*(3), 347-370. https://doi.org/10.1002/sce.20186

Bitzenbauer, P. (2023). ChatGPT in physics education: A pilot study on easy-to-implement activities. *Contemporary Educational Technology, 15*(3), Article ep430. https://doi.org/10.30935/cedtech/13176

Buckley, L. (2016). SU-F-E-08: Medical physics as a teaching tool for high school science curriculum. *Medical Physics, 43*(6Part5), Article 3354. https://doi.org/10.1118/1.4955694

Cetnar, A. J., Woollard, J., & Ding, L. (2025). Assessment of a medical physics educational program for science teachers. *Journal of Applied Clinical Medical Physics, 26*(5), Article e70087. https://doi.org/10.1002/acm2.70087

Cooper, G. (2023). Examining science education in ChatGPT: An exploratory study of generative artificial intelligence. *Journal of Science Education and Technology, 32*, 444-452. https://doi.org/10.1007/s10956-023-10039-y

Cooper, G., & Tang, K. S. (2024). Pixels and pedagogy: Examining science education imagery by generative artificial intelligence. *Journal of Science Education and Technology, 33*, 556-568. https://doi.org/10.1007/s10956-024-10104-0

Fagerstrom, J. M., Gao, W., & Robertson, G. E. (2019). A hands-on introduction to medical physics and radiation therapy for middle school students. *Journal of Applied Clinical Medical Physics, 20*(4), 148-153. https://doi.org/10.1002/acm2.12569

Farrokhnia, M., Banihashem, S. K., Noroozi, O., & Wals, A. (2024). A SWOT analysis of ChatGPT: Implications for educational practice and research. *Innovations in Education and Teaching International, 61*(3), 460-474. https://doi.org/10.1080/14703297.2023.2195846

Fayzullina, A. R., Filippova, A. A., Garnova, N. Y., Astakhov, D. V., Kalmazova, N., & Zaripova, Z. F. (2025). Artificial intelligence in science education: A systematic review of applications, impacts, and challenges. *Contemporary Educational Technology, 17*(4), Article ep613. https://doi.org/10.30935/cedtech/17519

Gleichmann, J., Kubitschke, H., & Schnauß, J. (2025). Impact of a mathematical pre-course on first-year physics students. *European Journal of Science and Mathematics Education, 13*(3), 172-190. https://doi.org/10.30935/scimath/16363

Hazari, Z., Sonnert, G., Sadler, P. M., & Shanahan, M. C. (2010). Connecting high school physics experiences, outcome expectations, physics identity, and physics career choice: A gender study. *Journal of Research in Science Teaching, 47*(8), 978-1003. https://doi.org/10.1002/tea.20363

Heeg, D. M., & Avraamidou, L. (2023). The use of artificial intelligence in school science: A systematic literature review. *Educational Media International, 60*(2), 125-150. https://doi.org/10.1080/09523987.2023.2264990

Hoffmann, T. C., Glasziou, P. P., Boutron, I., Milne, R., Perera, R., Moher, D., Altman, D. G., Barbour, V., Macdonald, H., Johnston, M., Lamb, S. E., Dixon-Woods, M., McCulloch, P., Wyatt, J. C., Chan, A.-W., & Michie, S. (2014). Better reporting of interventions: Template for intervention description and replication (TIDieR) checklist and guide. *BMJ, 348*, Article g1687. https://doi.org/10.1136/bmj.g1687

Hulleman, C. S., & Harackiewicz, J. M. (2009). Promoting interest and performance in high school science classes. *Science, 326*(5958), 1410-1412. https://doi.org/10.1126/science.1177067

International Atomic Energy Agency. (n.d.). Medical physics. *IAEA*. https://www.iaea.org/resources/hhc/medical-physics

Karampelas, K. (2026). Artificial intelligence in inquiry-based science teaching: A bibliometric study. *European Journal of Science and Mathematics Education, 14*(1), 104-116. https://doi.org/10.30935/scimath/17626

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., ... Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences, 103*, Article 102274. https://doi.org/10.1016/j.lindif.2023.102274

Küchemann, S., Steinert, S., Revenga, N., Schweinberger, M., Dinc, Y., Avila, K. E., & Kuhn, J. (2023). Physics task development of prospective physics teachers using ChatGPT. *Physical Review Physics Education Research, 19*(2), Article 020128. https://doi.org/10.1103/PhysRevPhysEducRes.19.020128

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for *t*-tests and ANOVAs. *Frontiers in Psychology, 4*, Article 863. https://doi.org/10.3389/fpsyg.2013.00863

Lazos, P., Kateris, A., Tsoukos, S., Tzamalis, P. G., & Velentzas, A. (2025). Teaching learning sequence for introductory instruction in the physics laboratory in upper secondary school. *Contemporary Mathematics and Science Education, 6*(2), Article ep25016. https://doi.org/10.30935/conmaths/17308

Liang, Y., Zou, D., Xie, H., & Wang, F. L. (2023). Exploring the potential of using ChatGPT in physics education. *Smart Learning Environments, 10*, Article 52. https://doi.org/10.1186/s40561-023-00273-7

Lo, C. K. (2023). What is the impact of ChatGPT on education? A rapid review of the literature. *Education Sciences, 13*(4), Article 410. https://doi.org/10.3390/educsci13040410

Marzano, D. (2025). Generative artificial intelligence (GAI) in teaching and learning processes at the K-12 level: A systematic review. *Technology, Knowledge and Learning, 31*, 789-829. https://doi.org/10.1007/s10758-025-09853-7

Miao, F., & Holmes, W. (2023). *Guidance for generative AI in education and research*. UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000386693

Ministry of Education of the Republic of Kazakhstan. (2022). On approval of the state compulsory standards of preschool education and training, primary, basic secondary and general secondary, technical and vocational, post-secondary education (order no. 348, August 3, 2022). *Ministry of Education of the Republic of Kazakhstan*. https://www.refworld.org/legal/decreees/natlegbod/2022/ru/150741

Moote, J., Archer, L., DeWitt, J., & MacLeod, E. (2020). Science capital or STEM capital? Exploring relationships between science capital and technology, engineering, and maths aspirations and attitudes among young people aged 17/18. *Journal of Research in Science Teaching, 57*(8), 1228-1249. https://doi.org/10.1002/tea.21628

Mujtaba, T., & Reiss, M. J. (2014). A survey of psychological, motivational, family and perceptions of physics education factors that explain 15-year-old students' aspirations to study physics in post-compulsory English schools. *International Journal of Science and Mathematics Education, 12*(2), 371-393. https://doi.org/10.1007/s10763-013-9404-1

Organization for Economic Co-operation and Development. (n.d.). Career guidance. *OECD*. https://www.oecd.org/en/topics/sub-issues/career-guidance.html

Polverini, G., & Gregorcic, B. (2024). How understanding large language models can inform the use of ChatGPT in physics education. *European Journal of Physics, 45*(2), Article 025701. https://doi.org/10.1088/1361-6404/ad1420

Potvin, P., & Hasni, A. (2014). Interest, motivation and attitude towards science and technology at K-12 levels: A systematic review of 12 years of educational research. *Studies in Science Education, 50*(1), 85-129. https://doi.org/10.1080/03057267.2014.881626

Rutten, N., van Joolingen, W. R., & van der Veen, J. T. (2012). The learning effects of computer simulations in science education. *Computers & Education, 58*(1), 136-153. https://doi.org/10.1016/j.compedu.2011.07.017

Santoso, A. P., Jupitz, S., & Lin, C. (2021). A framework for developing community-focused medical physics outreach programs. *Journal of Applied Clinical Medical Physics, 22*(10), 305-314. https://doi.org/10.1002/acm2.13413

Schleicher, S., Hunt, A., Brophy, S., Garay, C., Paschal, C., & Klein-Gardner, S. (2005). Teaching X-ray imaging in the high school physics classroom: Safe, inexpensive and hands-on instruction. In *Proceedings of the 2005 Annual Conference & Exposition*. https://doi.org/10.18260/1-2--14159

Sheldrake, R., Mujtaba, T., & Reiss, M. J. (2019). Students' changing attitudes and aspirations towards physics during secondary school. *Research in Science Education, 49*, 1809-1834. https://doi.org/10.1007/s11165-017-9676-5

Smetana, L. K., & Bell, R. L. (2012). Computer simulations to support science instruction and learning: A critical review of the literature. *International Journal of Science Education, 34*(9), 1337-1370. https://doi.org/10.1080/09500693.2011.605182

Sudirman, S., Kennedy, D., & Soeharto, S. (2023). The teaching of physics at upper secondary school level: A comparative study between Indonesia and Ireland. *Frontiers in Education, 8*, Article 1118873. https://doi.org/10.3389/feduc.2023.1118873

Taber, K. S. (2018). The use of Cronbach's alpha when developing and reporting research instruments in science education. *Research in Science Education, 48*, 1273-1296. https://doi.org/10.1007/s11165-016-9602-2

Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments, 10*, Article 15. https://doi.org/10.1186/s40561-023-00237-x

Van Breukelen, G. J. P. (2006). ANCOVA versus change from baseline had more power in randomized studies and more bias in nonrandomized studies. *Journal of Clinical Epidemiology, 59*(9), 920-925. https://doi.org/10.1016/j.jclinepi.2006.02.007

Vasconcelos, M. A. R., & Santos, R. P. D. (2023). Enhancing STEM learning with ChatGPT and Bing Chat as objects to think with: A case study. *Eurasia Journal of Mathematics, Science and Technology Education, 19*(7), Article em2296. https://doi.org/10.29333/ejmste/13313

Wieman, C. E., Adams, W. K., & Perkins, K. K. (2008). PhET: Simulations that enhance learning. *Science, 322*(5902), 682-683. https://doi.org/10.1126/science.1161948

Yorke, A. A., Schandorf, M. T., Quaye, A. N. M., Twum, P. T., Sengupta, B., Nkansah-Poku, K., Kplorfia, J. A., & Fagerstrom, J. (2025). Empowering young minds through STEM education: Engaging high schoolers in Ghana through medical physics. *Journal of Applied Clinical Medical Physics, 26*(6), Article e70126. https://doi.org/10.1002/acm2.70126