

# The Significance of Artificial Intelligence Platforms in Anatomy Education: An Experience With ChatGPT and Google Bard

Review began 09/06/2023

Review ended 09/15/2023

Published 09/15/2023

© Copyright 2023

Ilgaz et al. This is an open access article distributed under the terms of the Creative Commons Attribution License CC-BY 4.0., which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Hasan B. Ilgaz <sup>1</sup>, Zehra Çelik <sup>1</sup>

1. Anatomy, Hacettepe University Faculty of Medicine, Ankara, TUR

**Corresponding author:** Hasan B. Ilgaz, hasanbilgaz@hacettepe.edu.tr

---

## Abstract

This study evaluated the use of two large language models (LLMs), ChatGPT and Google Bard, in anatomy education. The models were asked to answer questions, generate multiple-choice questions, and write articles on anatomy topics. The results showed that the models were able to perform these tasks with varying degrees of accuracy. ChatGPT and Google Bard did not differ significantly in terms of answering questions. Both models were able to generate multiple-choice questions with a high degree of accuracy. However, the performance of the models in article writing was not yet at a sufficient level. The study also found that the use of LLMs in medical education requires caution. This is because LLMs are still under development and they can sometimes generate inaccurate or misleading information. It is important to carefully evaluate the output of LLMs before using them in educational settings. Overall, the study found that LLMs have the potential to be valuable tools for anatomy education. However, more research is needed to improve the accuracy of the models and to better understand how they can be used effectively in educational settings.

---

**Categories:** Medical Education, Anatomy

**Keywords:** chatgpt, large language models, google bard, artificial intelligence, anatomy education

## Introduction

The development of large language models (LLMs) is a major advance in artificial intelligence (AI) and natural language processing (NLP). LLMs have the capacity to process and understand natural language data with extraordinary skill and are already used in a variety of applications such as chatbots, machine translation, and question-answering [1-3]. Some of the most popular LLMs are ChatGPT and Google Bard, both of which have garnered a lot of attention for their capacity to understand textual information and generate contextually relevant responses [4]. These applications are being tried to be used in medical education. Recently, numerous articles have emerged about AI applications in different fields of medicine such as radiology, dermatology, physiology, hematology, ophthalmology, biochemistry, parasitology, neurosurgery, forensic medicine, dental education, etc. [5-14]. These models have been helpful in solving complex medical problems, interpreting radiology reports, being used to diagnose diseases, writing scientific articles or answering and generating different medical exam questions and have shown varying degrees of accuracy in these fields [2-5,7,8,10,11,15]. Although there are not enough articles about its use in the field of anatomy, there are some preliminary evaluations [16-21]. The present study aims to evaluate how various features of open AI platforms such as ChatGPT and Google Bard in their current form can contribute to anatomy education with the perspective of questioning, answering, and writing articles, and aims to respond to some controversial topics.

## Materials And Methods

### Data collection

In order to evaluate the usability of the applications in anatomy education, three stages were first defined. In the first stage, the answers of the applications to the anatomy exam questions would be evaluated. In the second stage, their ability to form questions and in the last stage to write scientific articles would be evaluated. For these processes to be applied sequentially, user accounts were first created for both ChatGPT (<https://chat.openai.com/>) and Google Bard (<https://bard.google.com/>) chat boards. Then for the first stage, ChatGPT and Google Bard platforms were requested to answer the last five years of the Medical Specialty Exam (MSE) questions. MSE is the main specialization exam for medicine in Türkiye and is held twice a year by The Measuring, Selection and Placement Center (OSYM - A Governmental Organization). The anatomy sections of the MSE were about 14 questions (2021-2 MSE questions were given as an example in *Appendix J*). As it could not be reached, the first exam in 2023 and the two exams in 2022 could not be evaluated. In order to evaluate whether the language option affects the success of the chat boards, the questions were also translated into English through them. After the translation process was done by the chat boards, it was checked by two anatomists (one of them is an anatomy specialist and the other one is an anatomy specialist student). The questions that were found to be wrong or insufficient were corrected by the evaluators and

### How to cite this article

Ilgaz H B, Çelik Z (September 15, 2023) The Significance of Artificial Intelligence Platforms in Anatomy Education: An Experience With ChatGPT and Google Bard. Cureus 15(9): e45301. DOI 10.7759/cureus.45301

asked on the chat boards. Two questions in ChatGPT and one question in Google Bard were corrected due to syntax and grammar errors. Since the software cannot evaluate visual content, all questions except those with visual content were asked. A total of 131 questions were asked on the chat boards for each user. The users were ChatGPT Turkish 1, ChatGPT Turkish 2, ChatGPT English, Google Bard Turkish 1, Google Bard Turkish 2, and Google Bard English. The accuracy of the answers was assessed according to the existing literature. The correct answers to Turkish and English questions were compared. A second user was created in both apps to assess whether the responses to the questions were systematic or whether the apps taught random answers. When these answers were compared, it was also observed whether asking the same questions resulted in more accurate answers in the second answer, thanks to machine learning. The accuracy rate of the answers given by the first and second users created in the same chat board was evaluated by comparing them with each other. In the next step, the question-answer capabilities of the software were evaluated. ChatGPT and Google Bard were asked to create and answer MSE-level multiple-choice anatomy questions. Almost all of them were correctly constructed. We asked about 100 questions for 40 different topics and only four of them were wrong. These 40 topics were identified by two anatomists. In order not to expand the boundaries of the article too much, instead of 40 topics, three random topics were sampled in the article. Then, the level was adjusted by giving the commands "make the question harder" or "recreate this question" for those who did not have the appropriate content. The MSE level emerged with the experience of two anatomists studying the exam questions of the last 10 years. The accuracy of the questions and answers at all stages was evaluated according to the current literature. And, finally, the software was asked to describe its ability to write academic papers. The validity of the anatomical information in the paper was checked, and the evaluation of whether it was detailed enough as a scientific article was done by the two different anatomists.

The free version of the LLMs' was used (ChatGPT 3.5 and Google Bard). All the data were collected from ChatGPT and Google Bard.

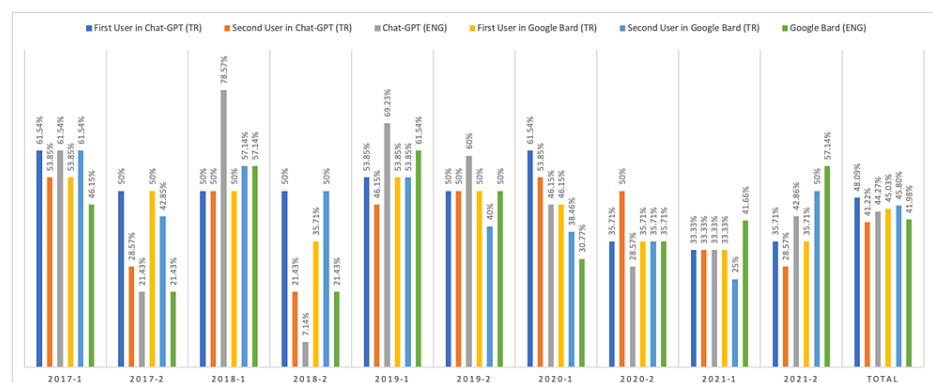
### Statistical analysis

All data were analyzed by using the IBM SPSS (Statistical Package for Social Sciences - Armonk, NY: IBM Corp) version 22.0 package. The data were expressed as numbers and percentages. For the analysis of results, two simple t-tests were used for parametric variables, while the Mann-Whitney U test was used for non-parametric variables. Kruskal Wallis test was used to evaluate the differences between all groups. The significance level was set at  $p < 0.05$ .

## Results

### Answering questions

The scores obtained after asking questions to the applications are shown in Figure 1. The most successful results were seen in the first user of ChatGPT in Turkish (48.09%), the second one was the second user of Google Bard in Turkish (45.80%), and the others were first user in Google Bard in Turkish (45.03%), ChatGPT in English (44.27%), Google Bard in English (41.98%) and first user of ChatGPT in Turkish (41.22%), respectively. None of them scored above 50% on the total correct answers. When the overall results were compared in all groups, no statistically significant difference was found ( $p > 0.05$ ).



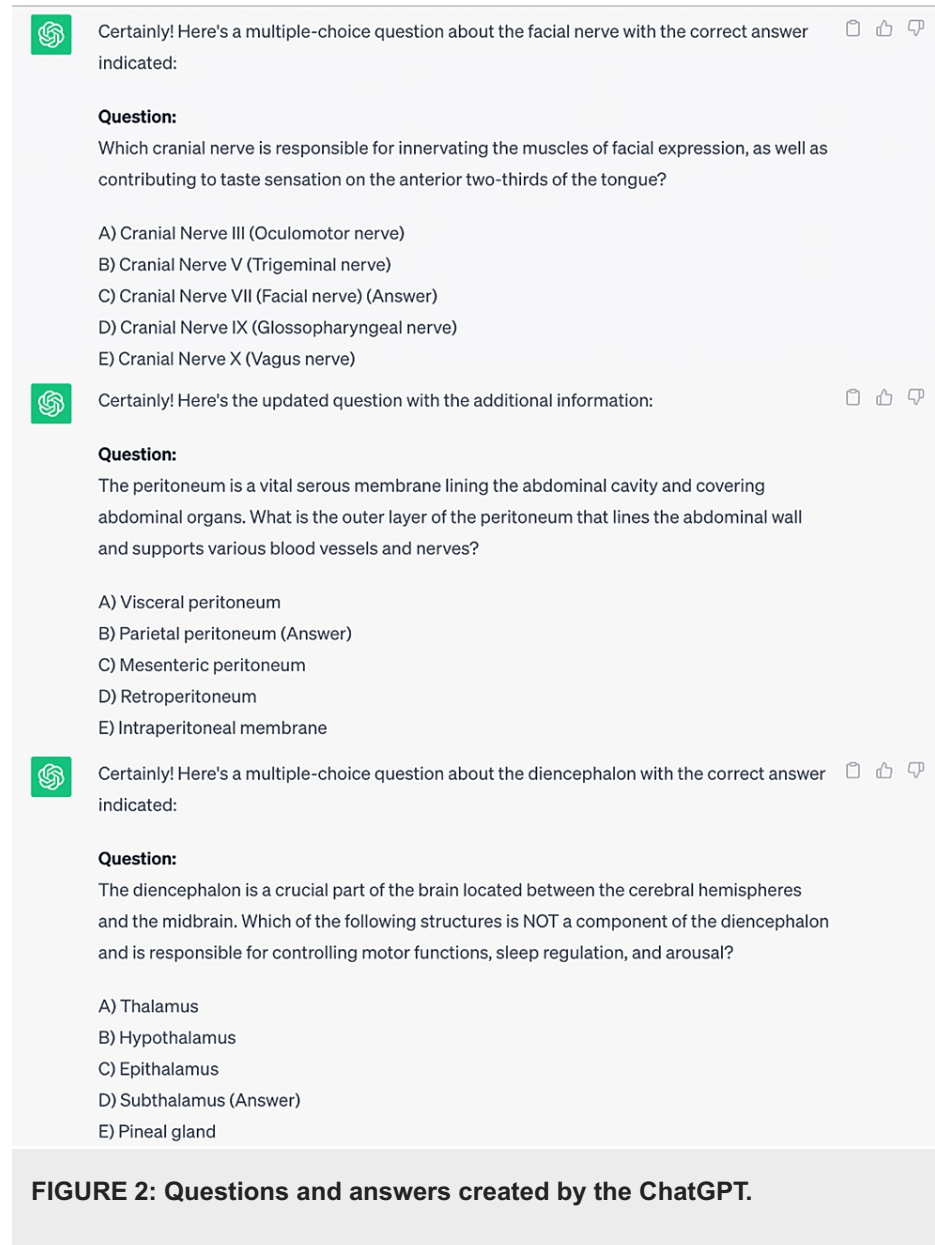
**FIGURE 1: Results of the answering questions.**

(TR: Turkish, ENG: English)

### Question generating

At this stage, the software was tasked with preparing three sample questions about facial nerve, peritoneum, and diencephalon anatomy which were selected by the two anatomists. The questions and answers created

by the ChatGPT are shown in Figure 2.



The screenshot displays three separate ChatGPT responses, each starting with a green circular icon containing a white 'G' and a checkmark. Each response includes a 'Question:' section followed by a list of five multiple-choice options (A-E). The first question is about cranial nerves, the second about the peritoneum, and the third about the diencephalon. In the third question, option D is marked as the correct answer.

Certainly! Here's a multiple-choice question about the facial nerve with the correct answer indicated:

**Question:**  
Which cranial nerve is responsible for innervating the muscles of facial expression, as well as contributing to taste sensation on the anterior two-thirds of the tongue?

A) Cranial Nerve III (Oculomotor nerve)  
B) Cranial Nerve V (Trigeminal nerve)  
C) Cranial Nerve VII (Facial nerve) (Answer)  
D) Cranial Nerve IX (Glossopharyngeal nerve)  
E) Cranial Nerve X (Vagus nerve)

Certainly! Here's the updated question with the additional information:

**Question:**  
The peritoneum is a vital serous membrane lining the abdominal cavity and covering abdominal organs. What is the outer layer of the peritoneum that lines the abdominal wall and supports various blood vessels and nerves?

A) Visceral peritoneum  
B) Parietal peritoneum (Answer)  
C) Mesenteric peritoneum  
D) Retroperitoneum  
E) Intraperitoneal membrane

Certainly! Here's a multiple-choice question about the diencephalon with the correct answer indicated:

**Question:**  
The diencephalon is a crucial part of the brain located between the cerebral hemispheres and the midbrain. Which of the following structures is NOT a component of the diencephalon and is responsible for controlling motor functions, sleep regulation, and arousal?

A) Thalamus  
B) Hypothalamus  
C) Epithalamus  
D) Subthalamus (Answer)  
E) Pineal gland

**FIGURE 2: Questions and answers created by the ChatGPT.**

The answer to the third question generated by ChatGPT was wrong [22].

The questions and answers created by the Google Bard are shown in Figure 3.

◆ Sure, here is a question about the facial nerve with 5 choices:

**Question:** Which of the following is NOT a branch of the facial nerve?

**Answer:** Greater petrosal nerve

**Choices:**

- A. Greater petrosal nerve
- B. Buccal nerve
- C. Zygomatic nerve
- D. Marginal mandibular nerve
- E. Cervical nerve

**Explanation:** The greater petrosal nerve is a branch of the trigeminal nerve, not the facial nerve. The other choices are all branches of the facial nerve.

Is there anything else I can help you with today?

◆ Sure, here is a question about the peritoneum with 5 choices:

**Question:** Which of the following is NOT a function of the peritoneum?

**Answer:** Produces digestive enzymes

**Choices:**

- A. Lines the abdominal cavity
- B. Supports the abdominal organs
- C. Secretes a lubricating fluid
- D. Prevents adhesions between the abdominal organs
- E. Produces digestive enzymes

This question tests the respondent's knowledge of the basic functions of the peritoneum.

Is there anything else I can help you with today?

◆ Sure, here is another question about the diencephalon with 5 choices:

**Question:** Which of the following is NOT a component of the diencephalon?

**Answer:** Corpus callosum

**Choices:**

- Thalamus
- Hypothalamus
- Pineal gland
- Subthalamus
- Corpus callosum

This question tests the respondent's knowledge of the four components of the diencephalon. The corpus callosum is a large bundle of nerve fibers that connects the two hemispheres of the cerebral cortex, and it is not part of the diencephalon.

Is there anything else I can help you with today?

### FIGURE 3: Questions and answers created by the Google Bard.

All the answers to the questions generated by ChatGPT were correct [22].

### Article writing

The models were asked to write an academic paper on the anatomy of the facial nerve. The article written by ChatGPT is shown in Figure 4.



And, the article written by Google Bard is shown in Figure 5.

### ✦ The Facial Nerve: A Review of Anatomy, Function, and Pathology

The facial nerve (cranial nerve VII) is a mixed nerve that controls facial expression, taste sensation, tear production, and salivation. It is the longest of the cranial nerves and has a complex anatomy.

#### Anatomy

The facial nerve originates in the pons and travels through the internal auditory canal. It then exits the skull through the stylomastoid foramen and enters the parotid gland. Within the parotid gland, the facial nerve divides into its terminal branches, which innervate the muscles of facial expression.

The facial nerve is divided into three main segments:

- **Intracranial segment:** This segment of the facial nerve lies within the brainstem and travels through the internal auditory canal.
- **Extracranial segment:** This segment of the facial nerve travels through the temporal bone and exits the skull through the stylomastoid foramen.
- **Cutaneous segment:** This segment of the facial nerve travels through the parotid gland and divides into its terminal branches, which innervate the facial muscles.

#### Function

The facial nerve is responsible for the following functions:

- **Facial expression:** The facial nerve controls the muscles of facial expression, which allow us to make facial expressions such as smiling, frowning, and winking.
- **Taste sensation:** The facial nerve carries taste fibers from the anterior two-thirds of the tongue.
- **Tear production:** The facial nerve stimulates the lacrimal gland to produce tears.
- **Salivary production:** The facial nerve stimulates the parotid gland to produce saliva.

#### Pathology

The facial nerve is susceptible to a number of different pathologies, including:

- **Bell's palsy:** Bell's palsy is the sudden onset of facial paralysis that is thought to be caused by inflammation of the facial nerve.
- **Stroke:** Stroke can damage the facial nerve, resulting in facial paralysis.
- **Tumors:** Tumors of the brainstem or temporal bone can compress the facial nerve, causing facial paralysis.
- **Trauma:** Trauma to the face or head can damage the facial nerve, leading to facial paralysis.
- **Idiopathic facial paralysis:** In some cases, the cause of facial paralysis is unknown.

#### Diagnosis and Treatment

If you experience facial paralysis, it is important to see a doctor right away to determine the underlying cause and get the appropriate treatment. The doctor will perform a physical examination and may order imaging tests such as an MRI or CT scan to assess the facial nerve.

Treatment for facial paralysis will depend on the underlying cause. For example, if Bell's palsy is the cause, the doctor may prescribe steroids or antiviral medication. If stroke is the cause, the doctor may recommend physical therapy or surgery to improve facial movement. If a tumor is the cause, the doctor may recommend surgery to remove the tumor.

#### Prognosis

The prognosis for facial paralysis depends on the underlying cause and the severity of the damage to the facial nerve. For example, Bell's palsy usually resolves on its own within a few weeks. However, if stroke or trauma has caused facial paralysis, the prognosis may be less favorable.

#### Conclusion

The facial nerve is a complex and important nerve that controls a number of important functions. Damage to the facial nerve can result in facial paralysis, which can have a significant impact on a person's quality of life. Early diagnosis and treatment of facial paralysis is important to improve the chances of a full recovery.

#### Additional Information

In addition to the above, there are a number of other important things to know about the facial nerve:

- The facial nerve is the only cranial nerve that has both motor and sensory components.
- The facial nerve is also involved in regulating the production of tears and saliva.
- The facial nerve is very vulnerable to damage, as it travels through a number of narrow passages in the skull.
- Facial paralysis can be a very distressing condition, as it can make it difficult to communicate and express emotions.

## FIGURE 5: The article written by Google Bard.

## Discussion

Although artificial intelligence applications are widely used in the field of medicine, there is not enough evaluation in the field of anatomy and anatomy education yet. Mogali used ChatGPT in the fields of asking questions, engaging like a tutor, and writing assessment items, and as a result, states that although his contribution to students and teachers in type education is not yet at a sufficient level, it has a potential for development [16]. Similar to Mogali's assessment, ChatGPT has significant potential to enhance student's learning experiences and create a more interactive and engaging educational environment, according to Lee, in an article that looked at it by responding to directions and figuring out the advantages and disadvantages of it and designing a curriculum [18]. Totlis et al. on the other hand, by asking 18 questions about anatomy in the fourth version of ChatGPT, stated that this model can be a valuable interactive educational tool that encourages interaction, but that it cannot replace the critical role of educators and should be used as a complementary tool [17]. Abdellatif et al. and Lazarus et al., similar to the evaluations in other studies, also stated that these LLMs have great potential in anatomy education, even if they are not yet at a big efficiency [19,20]. Li et al. customized a database to train an artificial intelligence dialogue system for teaching anatomy to medical students in real-time by using an open-source machine learning architecture,

they conducted a preliminary study. After students' practice sessions, they reported an increase in students' confidence in anatomy knowledge from 2.10 to 3.84 on a five-point Likert scale [21].

Similar to these studies, we observed that ChatGPT and Google Bard, which we evaluated in terms of generating questions, responding to questions, and writing scientific articles in our study, have a significant potential for improvement in anatomy education for medical students and medical teachers in the future.

In addition, when we compared ChatGPT and Google Bard models, we found that there was no difference in problem-solving skills in anatomy. Both chat boards correctly answered about half of the anatomy questions asked. We used ChatGPT-3.5 in our study. Moshirfar et al., on the other hand, used the fourth version of the ChatGPT in their study and found that the fourth version was more successful in answering the questions [9]. The success of the fourth version was both a limitation of our work and proof that chat boards could be developed in a better way. Another limitation of this study was that we could not make visual evaluations, which are crucial in anatomy education. There is a need for future studies on the application of 2D and 3D figures in anatomy education with other artificial intelligence applications.

LLMs are evolving every day. Therefore, more studies are needed to fully explore its capabilities in the future [16-20]. However, these abilities will develop through deep learning. When we tested the performance of the models in different languages, we did not observe any difference between ChatGPT and Google Bard here either. We evaluated that the dataset of these data software did not make any difference in response to the questions in different languages, which showed that the language options were working well. We also tested that the success does not increase if the questions are repeated as a result of machine learning, by creating two different users. Furthermore, we concluded that the algorithm of the software is open to learning from the same user and that no progress was made when the same questions were asked by different users at different times. These results highlight the varying proficiency of language models in reaching accuracy with the same questions at the same time.

When we evaluated the generating questions, we found that both ChatGPT and Google Bard were quite good. Even though ChatGPT generated one incorrect question, these two chat boards can be used to make exam questions of different difficulty and can save teachers time. In fact, Google Bard can be considered a little better in this regard, because while preparing a question, it also suggests different alternative questions.

In terms of creating a scientific article, we saw that both chat boards article writing ability was less detailed, without references, and at a basic level, compared to previously written scientific articles on the subject. ChatGPT has created an article that is more in line with academic norms on this subject, with its general article structure and information divided into subheadings such as embryological development, clinical importance, surgical issues, future directions, and conclusion sections.

The model has the potential to become a more reliable and valuable tool for anatomy education and exam preparation by improving the ability to understand anatomy-related concepts and generate appropriate questions and distractors for different difficulties.

## Conclusions

There were differences in LLM performance, but no significant difference was observed between ChatGPT and Google Bard in terms of answering questions. The question generating of these models was satisfied. It can be useful for exam preparation in educational aspects. The performance of the applications in article writing was not yet at a sufficient level. While LLMs show promise as valuable tools for medical education, their use requires caution, given the complex algorithms and their limitations in dealing with potential inaccuracies. This study highlights the need for continuous improvement and validation of LLMs for reliable healthcare practices. Further research is needed to increase the accuracy of the models and to better understand how they can be used effectively in educational settings.

## Appendices

### Appendix 1

#### 2021-2 MSE Questions

1. The physical examination of the patient, who applied to the emergency clinic with complaints of pain and swelling in his foot after falling from a tree, revealed widespread pain on the sole of the foot and collapse of the medial arch of the foot. The patient's lateral ankle radiography shows that the talar head is dislocated inferiorly.

Which of the following anatomical structures is most likely to be injured in this patient?

A) Plantar aponeurosis

B) Plantar calcaneonavicular ligament

C) Medial talocalcaneal ligament

D) Calcaneocuboid ligament

E) Long plantar ligament

2. Which of the following joints is considered within the plane type articular group, according to the shape and movements of the joint surface?

A) Acromioclavicular joint

B) Proximal radioulnar joint

C) Median atlantoaxial joint

D) Carpometacarpal joint of the thumb

E) Metacarpophalangeal joints

3. In a 68-year-old female patient who underwent a brachial plexus block developed winged scapula after the procedure. On physical examination, the scapula is observed to be elevated and slight medially rotated at rest.

Which of the following nerves is most likely to be injured during the brachial plexus block in this patient?

A) Accessory nerve

B) Dorsal scapular nerve

C) Suprascapular nerve

D) Long thoracic nerve

E) Axillary nerve

4. The inferior angle of the scapula is commonly used as a superficial anatomical landmark in physical examination.

In an adult which stands on the standart anatomical position, at which vertebral level is the angulus inferior located?

A) T3

B) T5

C) T7

D) T9

E) T11

5. Which of the following regions is the most appropriate for palpating the pulse of the posterior tibial artery?

A) Between the lateral and medial malleoli on the anterior side

B) Anterior to the lateral malleolus

C) Posterior to the lateral malleolus

D) Anterior to the medial malleolus



E) Posterior to the medial malleolus

6. Tentorial herniation is detected in a 35-year-old female patient who applied to the hospital with inability to move her right eye up, down and medially.

Which of the following nerves is most likely affected by the compression causing the clinical presentation in this patient?

A) Optic nerve

B) Oculomotor nerve

C) Trochlear nerve

D) Trigeminal nerve

E) Abducens nerve

7. For a minor surgical intervention on the lateral side of the foot, a surgeon injects subcutaneous local anesthetic just behind the lateral malleolus. Which of the following nerves is the most likely target of this block?

A) Posterior tibial nerve

B) Deep fibular nerve

C) Saphenous nerve

D) Superficial fibular nerve

E) Sural nerve

8. Which of the following pathways carries conscious proprioceptive, two-point discrimination, and vibration senses of the body?

A) Anterior spinothalamic tract

B) Gracile and cuneate fasciculi

C) Spinotectal tract

D) Posterior spinocerebellar tract

E) Posterior spinothalamic tract

9. Which of the following sinuses is the dura mater venous sinus, which plays a role in the circulation of the cerebrospinal fluid and has the most “granulationes arachnoideae”?

A) Superior sagittal sinus

B) Inferior sagittal sinus

C) Sigmoid sinus

D) Inferior petrosal sinus

E) Cavernous sinus

10. Which of the following is the free inferior border of the quadrangular membrane?

A) Vocal ligament

B) Vestibular ligament

C) Aryepiglottic fold

D) Thyroepiglottic ligament

E) Elastic cone

11. Where is the most common site of fertilization of the ovum by the spermatozoon?

A) Infundibulum of the uterine tube

B) Fimbriae of the uterine tube

C) Ampulla of the uterine tube

D) Isthmus of the uterine tube

E) Abdominal ostium of the uterine tube

12. Which anatomical structure, entering the inguinal canal from the deep inguinal ring and extending to the subcutaneous tissues of the labia majora pudendi, is a guide region for a minor surgical procedure on the lateral side of the foot?

A) Sacrotuberous ligament

B) Cardinal ligament

C) Rectouterine ligament

D) Pubocervical ligament

E) Round ligament of the uterus

13. During cholecystectomy, when the cystic artery is ligated, which of the following anatomical structures forms the medial border of the Calot's triangle?

A) Common bile duct

B) Cystic duct

C) Common hepatic duct

D) Left hepatic duct

E) Pancreatic duct

14. Which organ and its part are both sites of attachment for the lesser omentum and greater omentum?

A) Liver - Quadrate lobe

B) Stomach - Greater curvature

C) Stomach - Angular notch

D) Duodenum - Superior part

E) Duodenum - Inferior part

## Additional Information

### Author Contributions

All authors have reviewed the final version to be published and agreed to be accountable for all aspects of the work.

**Concept and design:** Hasan B. Ilgaz, Zehra Çelik

**Acquisition, analysis, or interpretation of data:** Hasan B. Ilgaz, Zehra Çelik

**Drafting of the manuscript:** Hasan B. Ilgaz, Zehra Çelik

**Critical review of the manuscript for important intellectual content:** Hasan B. Ilgaz, Zehra Çelik

## Disclosures

**Human subjects:** All authors have confirmed that this study did not involve human participants or tissue.

**Animal subjects:** All authors have confirmed that this study did not involve animal subjects or tissue.

**Conflicts of interest:** In compliance with the ICMJE uniform disclosure form, all authors declare the following: **Payment/services info:** All authors have declared that no financial support was received from any organization for the submitted work. **Financial relationships:** All authors have declared that they have no financial relationships at present or within the previous three years with any organizations that might have an interest in the submitted work. **Other relationships:** All authors have declared that there are no other relationships or activities that could appear to have influenced the submitted work.

## References

- Sallam M: ChatGPT utility in healthcare education, research, and practice: systematic review on the promising perspectives and valid concerns. *Healthcare (Basel)*. 2023, 11:[10.3390/healthcare11060887](https://doi.org/10.3390/healthcare11060887)
- Wang J, Deng H, Liu B, et al.: Systematic evaluation of research progress on natural language processing in medicine over the past 20 years: bibliometric study on pubmed. *J Med Internet Res*. 2020, 22:e16816. [10.2196/16816](https://doi.org/10.2196/16816)
- Kung TH, Cheatham M, Medenilla A, et al.: Performance of ChatGPT on USMLE: potential for AI-assisted medical education using large language models. *PLOS Digit Health*. 2023, 2:e0000198. [10.1371/journal.pdig.0000198](https://doi.org/10.1371/journal.pdig.0000198)
- Rudolph J, Tan S, Tan S: War of the chatbots: Bard, Bing Chat, ChatGPT, Ernie and beyond. The new AI gold rush and its impact on higher education. *Journal of Applied Learning and Teaching*. 2023, 6:[10.37074/jalt.2023.6.1.23](https://doi.org/10.37074/jalt.2023.6.1.23)
- Elkassam AA, Smith AD: Potential use cases for ChatGPT in radiology reporting. *AJR Am J Roentgenol*. 2023, 221:373-6. [10.2214/AJR.23.29198](https://doi.org/10.2214/AJR.23.29198)
- Kluger N: Potential applications of ChatGPT in dermatology. *J Eur Acad Dermatol Venereol*. 2023, 37:e941-2. [10.1111/jdv.19152](https://doi.org/10.1111/jdv.19152)
- Agarwal M, Sharma P, Goswami A: Analysing the applicability of ChatGPT, Bard, and Bing to generate reasoning-based multiple-choice questions in medical physiology. *Cureus*. 2023, 15:e40977. [10.7759/cureus.40977](https://doi.org/10.7759/cureus.40977)
- Kumari A, Kumari A, Singh A, et al.: Large language models in hematology case solving: a comparative study of ChatGPT-3.5, Google Bard, and Microsoft Bing. *Cureus*. 2023, 15:[10.7759/cureus.43861](https://doi.org/10.7759/cureus.43861)
- Moshirfar M, Altaf AW, Stoakes IM, Tuttle JJ, Hoopes PC: Artificial intelligence in ophthalmology: a comparative analysis of GPT-3.5, GPT-4, and human expertise in answering statpearls questions. *Cureus*. 2023, 15:e40822. [10.7759/cureus.40822](https://doi.org/10.7759/cureus.40822)
- Ghosh A, Bir A: Evaluating Chatgpt's ability to solve higher-order questions on the competency-based medical education curriculum in medical biochemistry. *Cureus*. 2023, 15:e37023. [10.7759/cureus.37023](https://doi.org/10.7759/cureus.37023)
- Huh S: Are ChatGPT's knowledge and interpretation ability comparable to those of medical students in Korea for taking a parasitology examination?: a descriptive study. *J Educ Eval Health Prof*. 2023, 20:1. [10.3352/jeehp.2023.20.1](https://doi.org/10.3352/jeehp.2023.20.1)
- Sevgi UT, Erol G, Doğruel Y, Sönmez OF, Tubbs RS, Güngör A: The role of an open artificial intelligence platform in modern neurosurgical education: a preliminary study. *Neurosurg Rev*. 2023, 46:86. [10.1007/s10143-023-01998-2](https://doi.org/10.1007/s10143-023-01998-2)
- Thurzo A, Kosnáčová HS, Kurilová V, et al.: Use of advanced artificial intelligence in forensic medicine, forensic anthropology and clinical anatomy. *Healthcare (Basel)*. 2021, 9:[10.3390/healthcare9111545](https://doi.org/10.3390/healthcare9111545)
- Thurzo A, Strunga M, Urban R, Surovková J, Afrashtehfar KI: Impact of artificial intelligence on dental education: a review and guide for curriculum update. *Education Sciences*. 2023, 13:150. [10.3390/educsci13020150](https://doi.org/10.3390/educsci13020150)
- Gilson A, Safranek CW, Huang T, Socrates V, Chi L, Taylor RA, Chartash D: How does Chatgpt perform on the United States Medical Licensing examination? The implications of large language models for medical education and knowledge assessment. *JMIR Med Educ*. 2023, 9:e45312. [10.2196/45312](https://doi.org/10.2196/45312)
- Mogali SR: Initial impressions of ChatGPT for anatomy education. *Anat Sci Educ*. 2023, [10.1002/ase.2261](https://doi.org/10.1002/ase.2261)
- Totlis T, Natsis K, Filos D, Ediaroglou V, Mantzou N, Duparc F, Piagkou M: The potential role of ChatGPT and artificial intelligence in anatomy education: a conversation with ChatGPT. *Surg Radiol Anat*. 2023, [10.1007/s00276-023-03229-1](https://doi.org/10.1007/s00276-023-03229-1)
- Lee H: The rise of ChatGPT: exploring its potential in medical education. *Anat Sci Educ*. 2023, [10.1002/ase.2270](https://doi.org/10.1002/ase.2270)
- Abdellatif H, Al Mushaiqri M, Albalushi H, Al-Zaabi AA, Roychoudhury S, Das S: Teaching, learning and assessing anatomy with artificial intelligence: the road to a better future. *Int J Environ Res Public Health*. 2022, 19:[10.3390/ijerph192114209](https://doi.org/10.3390/ijerph192114209)
- Lazarus MD, Truong M, Douglas P, Selwyn N: Artificial intelligence and clinical anatomical education: promises and perils. *Anat Sci Educ*. 2022, [10.1002/ase.2221](https://doi.org/10.1002/ase.2221)
- Li YS, Lam CS, See C: Using a machine learning architecture to create an ai-powered chatbot for anatomy education. *Med Sci Educ*. 2021, 31:1729-30. [10.1007/s40670-021-01405-9](https://doi.org/10.1007/s40670-021-01405-9)
- Standring S: Gray's anatomy e-book: the anatomical basis of clinical practice. Standring S (ed): Elsevier Health Sciences, Amsterdam, Netherlands; 2021.