




ChatGPT Validation on Medical Education for **Medical Students**

Thank you for taking part in this validation process on ChatGPT targeted **medical students** as the end user. Your expertise is invaluable in helping us to determine the model's effectiveness and identify its major limitations in assisting users within Medical Education. As such is your available time, this form was prepared to be simple and swift (hopefully) towards facilitating your medical work:

 Here, you will face 5 tasks conducted by ChatGPT to review its results (present inside a PDF file that you may have received from the same email that links you to here), which were from prompting the model in diverse manners (prompt engineering omitted).

 For demonstrative purposes, the experiments lean on the *Infective Endocarditis* (IE) disease.

 Please click the button below to start and read the task descriptions carefully. Put your medical hat on, and focus on the **medical student** profile!

Task 1 - IE Concept Explication in Text

As future medical professionals, a comprehensive recognition of evidence, guidelines, and disease implications to patient care is crucial for education. In this task, ChatGPT tries to explain our target disease to such students (via text) in less than 100 words.

Validation Methodology

Please adhere to the following convention when validating ChatGPT performance throughout this task:

 From **1 as "Unusable"** to **5 as "Perfect"**

How well does ChatGPT explain the Infective Endocarditis disease to **medical students** (in text)? *

	1	2	3	4	5
T1-1	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
T1-2	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
T1-3	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
T1-4	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
T1-5	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
T1-6	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Did you find any non-factual information in the description given by ChatGPT you would like to report? If so, which? *

T1-4- designação dos critérios de diagnóstico já não é a referida desde 2015. Mortalidade 10% é errado.
T1-5 - endocardite não é prevalente em Portugal. doença rara.
Na generalidade linguagem não médica

Which of the below do you think is the student's academic level best aligns with the ChatGPT result **T1-6**? *

- ☐ Regular Medical Student (3rd year of graduation)
- ☐ Advanced Medical Student (5th year of graduation)
- ☐ Inconclusive based on available information
- ☒ Outra: alunos não médicos

Do you have any extra observations you want to provide? (Optional)

Linguagem não médica

Task 2 - IE Theoretical Slide Builder

Preparing theoretical lectures for medical students involves explaining foundational biomedical concepts and exploring existing frameworks while studying a disease. In this task, as medical teachers, we confront ChatGPT in building 5 theoretical slides about our target disease for our medical students.

Validation Methodology

As for the previous task, please adhere to the following convention when validating ChatGPT performance throughout this task:

🚩 From 1 as "Unusable" to 5 as "Perfect"

How can ChatGPT assist in making theoretical slides about Infective Endocarditis for **medical * students**?

	1	2	3	4	5
T2-1	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
T2-2	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
T2-3	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
T2-4	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
T2-5	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>

Did you find any non-factual information present among ChatGPT-built slides you would like to report? If so, which?

T2-1 - slide 2 - nao ha dados para portugal por regioes; slide 5 - a designação dos criterios de diagnostico não esta atualizada.

T2-3 e T2-4 - não apresenta os dados

Which of the below do you think is the student's academic level best aligns with the ChatGPT * result **T2-5**?

- ☒ Regular Medical Student (3rd year of graduation)
- ☐ Advanced Medical Student (5th year of graduation)
- ☐ Inconclusive based on available information
- ☐ Outra:

Do you have any extra observations you want to provide? (Optional)

NA

Task 3 - IE Evaluation with Multiple-Choice Questions

One evaluation format commonly used by educators in quickly assessing medical students is the multiple-choice question. Besides responding directly, these questions have to cover the evaluated medical subject by containing independent options, one of which being the correct (or most suitable) answer to its question. In this task, ChatGPT was ordered to create 5 multiple-choice questions alongside an explanation of the correct option.

Validation Methodology

As for the previous task, please adhere to the following convention when validating ChatGPT performance throughout this task:

🚩 From 1 as "Unusable" to 5 as "Perfect"

How well does ChatGPT assist in generating multiple-choice questions for assessing Infective Endocarditis knowledge of **medical students**? *

	1	2	3	4	5
T3-1	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
T3-2	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
T3-3	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
T3-4	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>

Did you encounter any inappropriate academic questions, ambiguous options, or non-factual explanations from ChatGPT you would like to report? If so, which?

T3-1 - questao 1 e 3 são perolas pedagogicas - o CHAT GPT errou!

T3-2 - Questao 2 tem duas respostas certas. Questao 3 e 5 opções consideradas corretas estão erradas

Which of the below do you think is the student's academic level best aligns with the ChatGPT * result **T3-4**?

- ☒ Regular Medical Student (3rd year of graduation)
- ☐ Advanced Medical Student (5th year of graduation)
- ☐ Inconclusive based on available information
- ☐ Outra:

Do you have any extra observations you want to provide? (Optional)

na

Task 4 - IE Evaluation with Open-Ended Questions

Sometimes, multiple-choice questions are inadequate to analyze or discuss a medical topic. As often approached by educators, open-ended questions are used instead, offering a more nuanced methodology and allowing for in-depth exploration and demonstration of knowledge. In this task, ChatGPT attempts to elaborate 3 such thought-provoking questions, each one with a standard answer.

Validation Methodology

As for the previous task, please adhere to the following convention when validating ChatGPT performance throughout this task:

🚩 From 1 as "Unusable" to 5 as "Perfect"

How well does ChatGPT assist in generating open-ended questions for assessing Infective Endocarditis knowledge of **medical students**? *

	1	2	3	4	5
T4-1	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
T4-2	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
T4-3	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
T4-4	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>

Did you encounter any inappropriate academic questions, or non-factual explanations from ChatGPT you would like to report? If so, which? *

A explicação é na generalidade muito simples para alunos de medicina.

Which of the below do you think is the student's academic level best aligns with the ChatGPT * result **T4-4**?

- ☒ Regular Medical Student (3rd year of graduation)
- ☐ Advanced Medical Student (5th year of graduation)
- ☐ Inconclusive based on available information
- ☐ Outra:

Do you have any extra observations you want to provide? (Optional)

na

Task 5 - IE Clinical Case Study Evaluation

One enriched exercise created by medical teachers is to place students in clinically analyzing a prepared case study of a demonstrative patient condition and then answering a series of case-related questions. In this task, we prompted ChatGPT to generate an evaluation sheet based on a generated clinical case study.

Validation Methodology

As for the previous task, please adhere to the following convention when validating ChatGPT performance throughout this task:

🚩 From 1 as "Unusable" to 5 as "Perfect"

Is ChatGPT capable of creating a clinical case-based evaluation of Infective Endocarditis disease for **medical students**? *

	1	2	3	4	5
T5-1	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
T5-2	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
T5-3	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
T5-4	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>

Did you see the hypothetical case study as dubious, inappropriate academic questions, or any inconsistencies between both you would like to report? If so, which? *

Na generalidade a observação do doente descrita é demasiada simples para um aluno de medicina. Ha erros - em T5-1 nao se pode dizer que o doente ouve um murmur - isso é obtido na avaliação pelo medico. Em T5-3 pela historia clinica o sopro seria mais provavelemnte no foco aortico e nao tricuspide pela historia.

Which of the below do you think is the student's academic level best aligns with the ChatGPT result **T5-4**? *

- ☒ Regular Medical Student (3rd year of graduation)
- ☐ Advanced Medical Student (5th year of graduation)
- ☐ Inconclusive based on available information
- ☐ Outra: _____

Do you have any extra observations you want to provide? (Optional)

NA

Este conteúdo não foi criado nem aprovado pela Google.

Google Formulários