

---

# Large Language Models Leverage External Knowledge to Extend Clinical Insight Beyond Language Boundaries

## Supplementary Information

### 1. Supplementary Figures

1.1 Figure 1. The integrated prompt of different enhancements in KFE .....	1
--	---

### 2. Supplementary Tables

2.1 Table 1: Model version and access .....	2
---	---

2.2 Table 2: Model parameters .....	2
-------------------------------------	---

### 3. Supplementary Materials

3.1 Materials 1: A case study of the segmented question and options .....	3
---	---

3.2 Materials 2: A case study of the knowledge enhancement .....	4
--	---

3.3 Materials 3: A case study of the few-shot enhancement .....	5
---	---

3.4 Materials 4: A case study of different strategies for few-shot enhancement .....	6
--	---

3.5 Materials 5: A case study of the predicted answer by GPT3.5 .....	7
---	---

3.6 Materials 6: A case study of KFE via Instruction with inference .....	8
---	---

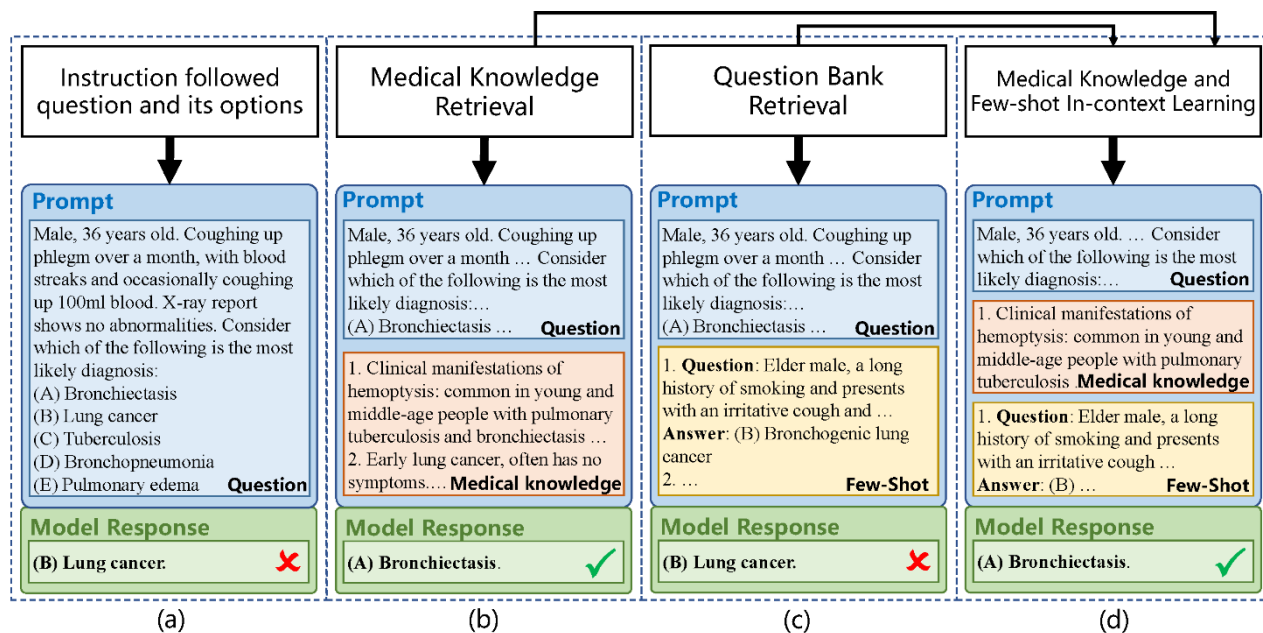
3.7 Materials 7: A case study of inference generated by GPT-3.5 under the basic setting	9
---	---

3.8 Materials 8: A case study of inference generated by GPT-3.5 under the KFE framework	10
---	----

### Note:

*In the actual experiments, all instructions, questions, knowledge, and examples were written in Chinese, without any translation. The English versions in the following images are for easier readability and have been translated using Google Translation.*

# 1. Supplementary Figures



**Supplementary Figure 1. The integrated prompt of different enhancements in KFE.**

(a) list a basic form of prompt that includes the question and options; (b) further includes retrieved related medical knowledge which is in the form of text pieces; (c) includes retrieved pairs of questions and answers as few-shot examples; (d) includes both retrieved knowledge and few-shot examples in prompts.

---

## 2. Supplementary Tables

**Supplementary Table 1. Model version and access**

Model	Version	Access
GPT 3.5 / ChatGPT	gpt-3.5-turbo-0301	<a href="https://platform.openai.com/docs/models/gpt-3-5">https://platform.openai.com/docs/models/gpt-3-5</a>
GPT4	gpt-4-0613	<a href="https://platform.openai.com/docs/models/gpt-4-and-gpt-4-turbo">https://platform.openai.com/docs/models/gpt-4-and-gpt-4-turbo</a>
Baichuan2-7B	Version 2.0	<a href="https://huggingface.co/baichuan-inc/Baichuan2-7B-Chat">https://huggingface.co/baichuan-inc/Baichuan2-7B-Chat</a>
Baichuan2-13B	Version 2.0	<a href="https://huggingface.co/baichuan-inc/Baichuan2-13B-Chat">https://huggingface.co/baichuan-inc/Baichuan2-13B-Chat</a>

**Supplementary Table 2. Model parameters**

Model	Parameter	Version
GPT-series	max_tokens	4096
	temperature	0
	top_p	1
	frequency_penalty	0
	presence_penalty	0
Baichuan-series	model_max_length	4096
	max_new_tokenss	512
	do_sample	False

### 3. Supplementary Materials

#### Question

问题: 男性, 36岁, 咳痰1月余, 痰中有血丝, 偶尔咯血100ml, X线检查显示无异常, 考虑诊断为:

**Question:** Male, 36 years old. Coughing up phlegm over a month, blood-tinged sputum, and occasionally 100ml hemoptysis. X-ray report shows no abnormalities. Consider the following diagnoses:

#### Question (segmented)

男性 36 岁 咳痰 1 月余 痰 中 有 血丝 偶尔 咯血  
Male 36 years-old Coughing up phlegm over one month sputum in blood-tinged occasionally hemoptysis  
100ml X射线 检查 显示 无 异常 考虑 诊断 为 :  
100ml X-ray examination shows no abnormalities. Consider diagnose is :

#### Options

选项: (A)支气管扩张, (B)肺癌, (C)结核, (D)支气管肺炎, (E)肺水肿

**Options:** (A) Bronchiectasis (B) Lung cancer (C) Tuberculosis (D) Bronchopneumonia (E) Pulmonary edema

#### Options (segmented)

支气管扩张 肺癌 结核 支气管肺炎 肺水肿  
Bronchiectasis Lung cancer Tuberculosis Bronchopneumonia Pulmonary edema

#### Supplementary materials 1. A case study of the segmented question and options.

The question and options are segmented with a medical lexicon, which divides sentences according to common Chinese vocabulary and professional clinical terminologies.

### Input of Knowledge Enhancement

以下是关于医学知识的单项选择题，请参考给定的相关背景知识，根据问题输出唯一的答案选项

#### 背景知识:

1. 咳嗽咳痰咯血、呼吸困难与水肿 咳嗽咳痰咯血 5.咯血的临床表现 (1) 年龄青壮年咯血多见于肺结核、支气管扩张、二尖瓣狭窄。40岁以上的吸烟者多见于支气管肺癌。
2. 胸心血管外科 肺部疾病 肺癌 临床表现 1.早期肺癌 特别是周围型肺癌往往无任何症状,大多在行胸部X线检查或胸部CT检查时发现。随着肿瘤的进展,出现不同的症状。 血痰常见于中心型肺癌,通常为痰中带血点、血丝或断续地少量咯血,大量咯血少见。
3. 肺结核 1. 可疑症状患者的筛选大约86活动性肺结核患者和95痰涂片阳性肺结核患者有可疑症状主要可疑症状为咳嗽、咳痰持续2周以上和咯血 其次是午后低热、乏力、盗汗、月经不调或闭经有肺结核接触史或肺外结核上述情况应考虑到肺结核病的可能性要进行痰抗酸杆菌和胸部X线检查
4. 细菌性肺炎早期肺脓肿无论在症状和胸部X线片表现上与细菌性肺炎很相似, 一般细菌性肺炎无大量脓臭痰, 胸部X线片示肺叶或段性实变或呈片状淡薄炎症病变边缘模糊不清, 没有空腔形成, 但当用抗生素治疗高热不退咳嗽、咳痰加剧并咳出大量脓痰时应考虑为继发肺脓肿
5. 临床医学综合 呼吸系统 支气管扩张 诊断与鉴别诊断 4. 先天性肺囊肿继发感染时可发生咳嗽咳痰咯血, X线可见多个边界纤细的圆形或椭圆形阴影, 周围无浸润, 高分辨率CT可明确诊断",

**问题:** 男性, 36岁。咳痰1月余, 痰中有血丝, 偶尔咯血100ml, X线检查无异常, 考虑诊断为: (A)支气管扩张, (B)肺癌, (C)结核, (D)支气管肺炎, (E)肺水肿

**答案:**

### Model Response of ChatGPT

A

### Input of Knowledge Enhancement

Here is a multi-choice question about medical knowledge, please refer to the given relevant background knowledge and examples, and output the unique answer option according to the question.

#### Background knowledge:

1. Coughing up phlegm and blood, difficulty breathing, and edema. Coughing up phlegm and blood 5. Clinical manifestations of hemoptysis: (1) Hemoptysis is common in young and middle-aged people with pulmonary tuberculosis, bronchiectasis, and mitral valve stenosis. In smokers over 40 years old, it is more common in bronchogenic lung cancer.
2. Thoracic and cardiovascular surgery, lung diseases, lung cancer clinical manifestations: 1. Early lung cancer, especially peripheral lung cancer, often has no symptoms and is mostly found during chest X-ray or chest CT examination. As the tumor progresses, different symptoms appear. Bloody sputum is common in central lung cancer, usually with blood spots, blood streaks in the sputum, or intermittent small amounts of hemoptysis, and large amounts of hemoptysis are rare.
3. Pulmonary tuberculosis: 1. Screening of patients with suspicious symptoms, about 86 active pulmonary tuberculosis patients and 95 sputum smear-positive pulmonary tuberculosis patients have suspicious symptoms. The main suspicious symptoms are cough, phlegm lasting for more than 2 weeks, and hemoptysis, followed by afternoon low fever, fatigue, night sweats, menstrual irregularities or amenorrhea. If there is a history of contact with tuberculosis or extrapulmonary tuberculosis, the possibility of pulmonary tuberculosis should be considered, and sputum acid-fast bacillus and chest X-ray examination should be performed.
4. Bacterial pneumonia: Early lung abscess, both in terms of symptoms and chest X-ray manifestations, is very similar to bacterial pneumonia. General bacterial pneumonia does not have a large amount of foul-smelling pus, and chest X-ray shows lobar or segmental consolidation or patchy thin inflammatory lesions with unclear edges, without cavity formation. However, when high fever persists after antibiotic treatment and cough and sputum worsen and a large amount of purulent sputum is coughed up, secondary lung abscess should be considered.
5. Clinical Medicine Comprehensive, Respiratory System, Bronchiectasis Diagnosis and Differential Diagnosis: 4. Congenital pulmonary cysts can cause cough, phlegm, and hemoptysis when secondary infection occurs. X-rays can show multiple thin-bordered round or oval shadows with no infiltration around them, and high-resolution CT can confirm the diagnosis.

**Question:** Male, 36 years old. Coughing up phlegm over a month, blood-tinged sputum, and occasionally 100ml hemoptysis. X-ray report shows no abnormalities. Consider the following diagnoses: (A) Bronchiectasis (B) Lung cancer (C) Tuberculosis (D) Bronchopneumonia (E) Pulmonary edema

**Answer:**

### Model Response of ChatGPT

A

## Supplementary materials 2. A case study of the knowledge enhancement.

Input includes a detailed instruction, related knowledge retrieved from knowledge base, and a testing question.

### Input of Few-shot Enhancement

以下是关于医学知识的单项选择题，请参考给定的相关例题，根据问题输出唯一的答案选项

#### 参考例题：

(1)问题：老年男性。有长期吸烟史，出现刺激性咳嗽，痰中带血，最可能为：(A)支气管扩张，(B)支气管肺癌，(C)肺炎，(D)肺结核，(E)肺水肿

答案：B

(2)问题：患者，女，58岁，无明显诱因出现反复低热、刺激性咳嗽、咳痰2个月余，痰中可见血丝，体检咳嗽呈高调金属音，抗生素治疗效果欠佳，有多年重度吸烟史。该患者最可能的诊断是：(A)慢性支气管炎，(B)支气管肺癌，(C)肺结核，(D)支气管扩张，(E)肺脓肿

答案：B

(3)问题：男性，69岁。持续痰中带血3个月余，既往有慢性支气管炎病史，无支气管扩张、肺结核病史，x线胸片提示右肺中叶占位。该患者咯血的病因是：(A)支气管扩张，(B)肺结核，(C)支气管肺癌，(D)慢性支气管炎，(E)肺脓肿

答案：C

(4)问题：男性，18岁。反复午后发热1个月，体温在37.3~37.8℃，疲乏无力，消瘦。近1周咳嗽，偶尔咯血性痰，夜间盗汗，无胸痛、气短。外院X线检查见右锁骨上斑片状阴影，痰结核菌检查阴性。该患者最可能的诊断是：(A)浸润型肺结核，(B)支气管肺癌，(C)支气管扩张合并感染，(D)军团菌肺炎，(E)真菌性肺炎

答案：A

(5)问题：女，22岁，2年来反复痰中带血，间有大口咯血。体格检查无异常体征，X线胸片示左下肺纹理增粗，紊乱，最可能的诊断是：(A)风心病二尖瓣狭窄，(B)慢性支气管炎，(C)支气管扩张症，(D)支气管肺癌，(E)肺结核

答案：C

问题：男性，36岁。咳嗽1月余，痰中有血丝，偶尔咯血100ml，X线检查无异常，考虑诊断为：(A)支气管扩张，(B)肺癌，(C)结核，(D)支气管肺炎，(E)肺水肿

答案：

Model Response of ChatGPT

B

### Input of Few-shot Enhancement

Here is a multi-choices question about medical knowledge, please refer to the given example question and output the unique answer option according to the question.

#### Examples:

(1)Question: Elderly male with a long history of smoking presents with irritating cough and blood-tinged sputum. The most likely diagnosis is:(A) Bronchiectasis, (B) Bronchogenic carcinoma, (C) Pneumonia, (D) Tuberculosis, (E) Pulmonary edema

Answer: B

(2)Question: Female, 58 years old, presents with recurrent low-grade fever, irritating cough, and expectoration with streaks of blood for over 2 months, unresponsive to antibiotics, and has a history of heavy smoking for many years. The most likely diagnosis is:(A) Chronic bronchopneumonia, (B) Bronchogenic carcinoma, (C) Tuberculosis, (D) Bronchiectasis, (E) Lung abscess

Answer: B

(3)Question: Male, 69 years old, with a history of chronic bronchitis but no history of bronchiectasis or tuberculosis, has had blood in his sputum for over 3 months. X-ray shows an occupying lesion in the right middle lobe of the lung. The cause of his hemoptysis is:(A) Bronchiectasis, (B) Tuberculosis, (C) Bronchogenic carcinoma, (D) Chronic bronchitis, (E) Lung abscess

Answer: C

(4)Question: Male, 18 years old, has had recurrent fever in the afternoon for a month, with temperatures between 37.3-37.8℃, fatigue, weight loss, and a recent cough with occasional bloody sputum. He experiences night sweats, but no chest pain or shortness of breath. An X-ray from another hospital shows patchy shadows above the right clavicle, and sputum tests for tuberculosis are negative. The most likely diagnosis is:(A) Infiltrative tuberculosis, (B) Bronchogenic carcinoma, (C) Bronchiectasis with infection, (D) Legionnaires' disease, (E) Fungal pneumonia

Answer: A

(5)Question: Female, 22 years old, has had blood in her sputum for 2 years, occasionally coughing up large amounts of blood. Physical examination shows no abnormal signs. X-ray shows increased and disordered lung patterns in the lower left lung. The most likely diagnosis is:(A) Rheumatic heart disease with mitral valve stenosis, (B) Chronic bronchitis, (C) Bronchiectasis, (D) Bronchogenic carcinoma, (E) Tuberculosis

Answer: C

Question: Male, 36 years old. Coughing up phlegm over a month, blood-tinged sputum, and occasionally 100ml hemoptysis. X-ray report shows no abnormalities. Consider the following diagnoses: (A) Bronchiectasis (B) Lung cancer (C) Tuberculosis (D) Bronchopneumonia (E) Pulmonary edema

Answer:

Model Response of ChatGPT

B

## Supplementary materials 3. A case study of the few-shot enhancement.

Input includes a detailed instruction, similar questions retrieved from question bank, and testing question.



### Example

问题: 老年男性。有长期吸烟史, 出现刺激性咳嗽, 痰中带血, 最可能为: (A)支气管扩张, (B)支气管肺癌, (C)肺炎, (D)肺结核, (E)肺水肿

#### Question + Options + Correct Answer

答案: **B** ✓

#### Question + Options + Generated Answer

答案: 根据题目和选项, 让我们逐步分析: 患者是老年男性, 有长期吸烟史, 出现刺激性咳嗽, 痰中带血。这些症状和病史提示可能是肺结核。肺结核是一种由结核分枝杆菌引起的传染病, 常表现为刺激性咳嗽、咳痰, 痰中可能带血丝。而支气管扩张通常不会导致痰中带血的状况, 支气管肺癌虽然也与吸烟有关, 但一般表现为反复咳嗽、咳痰、咳血等症状, 但与结核不同, 结核通常不会出现高调金属音的体征。肺炎通常伴随发热等症状, 与患者的病史不符。肺水肿通常与心脏问题有关, 与吸烟史和痰中带血的状况不太相关。因此, 根据题目和选项, **最可能的诊断是(D)肺结核**。✗

#### Question + Options + Generated Correct Answer

答案: 根据题目和选项, 让我们逐步分析: 老年男性有长期吸烟史, 出现刺激性咳嗽, 痰中带血, 这些症状提示可能是肺部疾病。支气管扩张和肺炎一般不会出现痰中带血的状况, 肺水肿也不太可能。肺结核和支气管肺癌都有可能出现刺激性咳嗽和痰中带血的状况, 但是肺结核一般不会出现长期吸烟史的情况。因此, **最终答案是(B)支气管肺癌**。✓

#### Question + Options + Correct Answer + Generated Inference Detail

答案: 根据题目中的长期吸烟史和痰中带血的状况, 可以初步判断为肺部疾病。支气管扩张和肺炎一般不会出现痰中带血的状况, 肺结核和肺水肿也不太可能出现刺激性咳嗽的状况。而支气管肺癌是与吸烟有关的常见疾病, 且常伴随着刺激性咳嗽和痰中带血的状况。因此, **最终答案是(B)支气管肺癌**。✓

### Example

Question: Elderly male with a long history of smoking presents with irritating cough and blood-tinged sputum. The most likely diagnosis is:(A) Bronchiectasis, (B) Bronchogenic carcinoma, (C) Pneumonia, (D) Tuberculosis, (E) Pulmonary edema

#### Question + Options + Correct Answer

Answer: **B** ✓

#### Question + Options + Generated Answer

Answer: Based on the question and options, let's analyze step by step: The patient is an elderly male with a long history of smoking, presenting with irritative cough and blood-tinged sputum. These symptoms and medical history suggest the possibility of tuberculosis. Tuberculosis is a contagious disease caused by the Mycobacterium tuberculosis bacterium, commonly manifesting as an irritative cough, expectoration, and possibly streaks of blood in the sputum. While bronchiectasis typically does not lead to blood-tinged sputum, bronchogenic carcinoma, although also related to smoking, generally presents with repeated coughing, expectoration, and hemoptysis, but unlike tuberculosis, tuberculosis usually does not present with a high-pitched metallic percussion sound. Pneumonia usually accompanies symptoms like fever, which does not match the patient's medical history. Pulmonary edema is generally related to heart problems and is not much related to a history of smoking and blood-tinged sputum. Therefore, **the most likely diagnosis is (D) Tuberculosis**. ✗

#### Question + Options + Generated Correct Answer

Answer: Based on the question and options, let's analyze step by step: An elderly male with a long history of smoking presents with irritative cough and blood-tinged sputum. These symptoms suggest a potential pulmonary disease. Bronchiectasis and pneumonia generally do not present with blood-tinged sputum, and pulmonary edema is also unlikely. Both tuberculosis and bronchogenic carcinoma may present with irritative cough and blood-tinged sputum, but tuberculosis generally does not occur with a long history of smoking. Therefore, **the final answer is (B) Bronchogenic carcinoma**. ✓

#### Question + Options + Correct Answer + Generated Inference Detail

Answer: Based on the long history of smoking and blood-tinged sputum symptoms in the question, it can be preliminarily judged as a pulmonary disease. Bronchiectasis and pneumonia generally do not present with blood-tinged sputum, and both tuberculosis and pulmonary edema are unlikely to present with irritative cough symptoms. Bronchogenic carcinoma, a common disease related to smoking, often accompanies symptoms like irritative cough and blood-tinged sputum. Therefore, **the final answer is (B) Bronchogenic carcinoma**. ✓

## Supplementary materials 4. A case study of different strategies for few-shot enhancement.

<b>Question</b>
<p><b>问题:</b> 男性, 36岁。咳痰1月余, 痰中有血丝,偶尔咯血100ml, X线检查无异常, 考虑诊断为:(A)支气管扩张, (B)肺癌, (C)结核, (D)支气管肺炎, (E)肺水肿</p> <p><b>答案:</b></p>
<b>Zero-shot</b>
B ✖
<b>Knowledge Enhancement</b>
A ✔
<b>Few-shot Enhancement</b>
B ✖
<b>Knowledge and Few-shot enhancement</b>
A ✔

<b>Question</b>
<p><b>Question:</b> Male, 36 years old. Coughing up phlegm over a month, blood-tinged sputum, and occasionally 100ml hemoptysis. X-ray report shows no abnormalities. Consider the following diagnoses: (A) Bronchiectasis (B) Lung cancer (C) Tuberculosis (D) Bronchopneumonia (E) Pulmonary edema</p> <p><b>Answer:</b></p>
<b>Zero-shot</b>
B ✖
<b>Knowledge Enhancement</b>
A ✔
<b>Few-shot Enhancement</b>
B ✖
<b>Knowledge and Few-shot enhancement</b>
A ✔

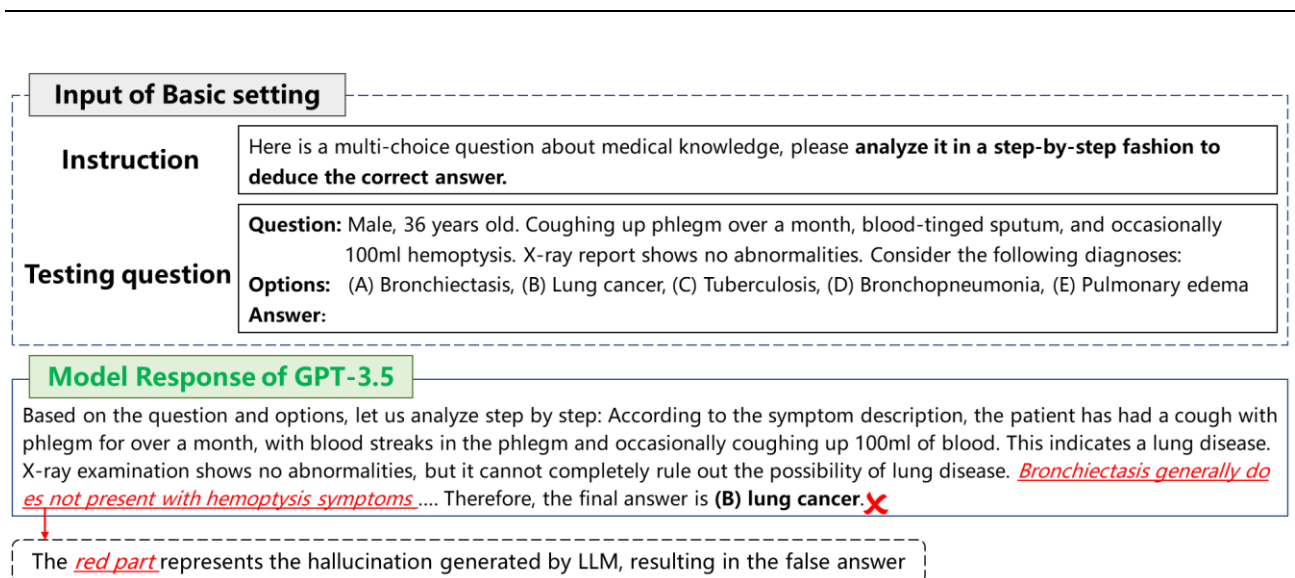
## Supplementary materials 5. Model response of the predicted answer by GPT-3.5



Input of KFE via Instruction with inference	
Instruction	Here is a multi-choice question about medical knowledge, please refer to the given relevant background knowledge and examples, and <b>analyze it in a step-by-step fashion to deduce the correct answer.</b>
Knowledge Enhancement	<b>Knowledge 1:</b> Clinical manifestations of hemoptysis: (1) Hemoptysis is common in young ... <b>Knowledge 2:</b> Early lung cancer, especially peripheral lung cancer, often has no symptoms ... ... ..
Few-shot Enhancement	<b>Question:</b> An elderly male with a long history of smoking presents with an irritating cough and blood-tinged sputum. The most likely diagnosis is: <b>Options:</b> (A) Bronchiectasis, (B) Bronchogenic carcinoma, (C) Pneumonia, (D) Tuberculosis, (E) ... <b>Answer:</b> (B) Bronchogenic carcinoma <b>Generated Inference:</b> Based on the long history of smoking and blood-tinged sputum symptoms in question, it can be preliminarily judged as a pulmonary disease. Bronchiectasis and ... ... ..
Testing question	<b>Question:</b> Male, 36 years old. Coughing up phlegm over a month, blood-tinged sputum, and occasionally 100ml hemoptysis. X-ray report shows no abnormalities. Consider the following diagnoses: <b>Options:</b> (A) Bronchiectasis, (B) Lung cancer, (C) Tuberculosis, (D) Bronchopneumonia, (E) Pulmonary edema <b>Answer:</b>
<b>Model Response of GPT3.5</b> Based on background knowledge and examples, let us analyze step by step: The patient presents with a cough with phlegm for over a month, blood streaks in the phlegm, and occasionally coughing up 100ml of blood. ... Bronchiectasis may cause symptoms of cough with phlegm and bloody phlegm and is more common in young people. Therefore, the most likely diagnosis is <b>(A) bronchiectasis</b> .	

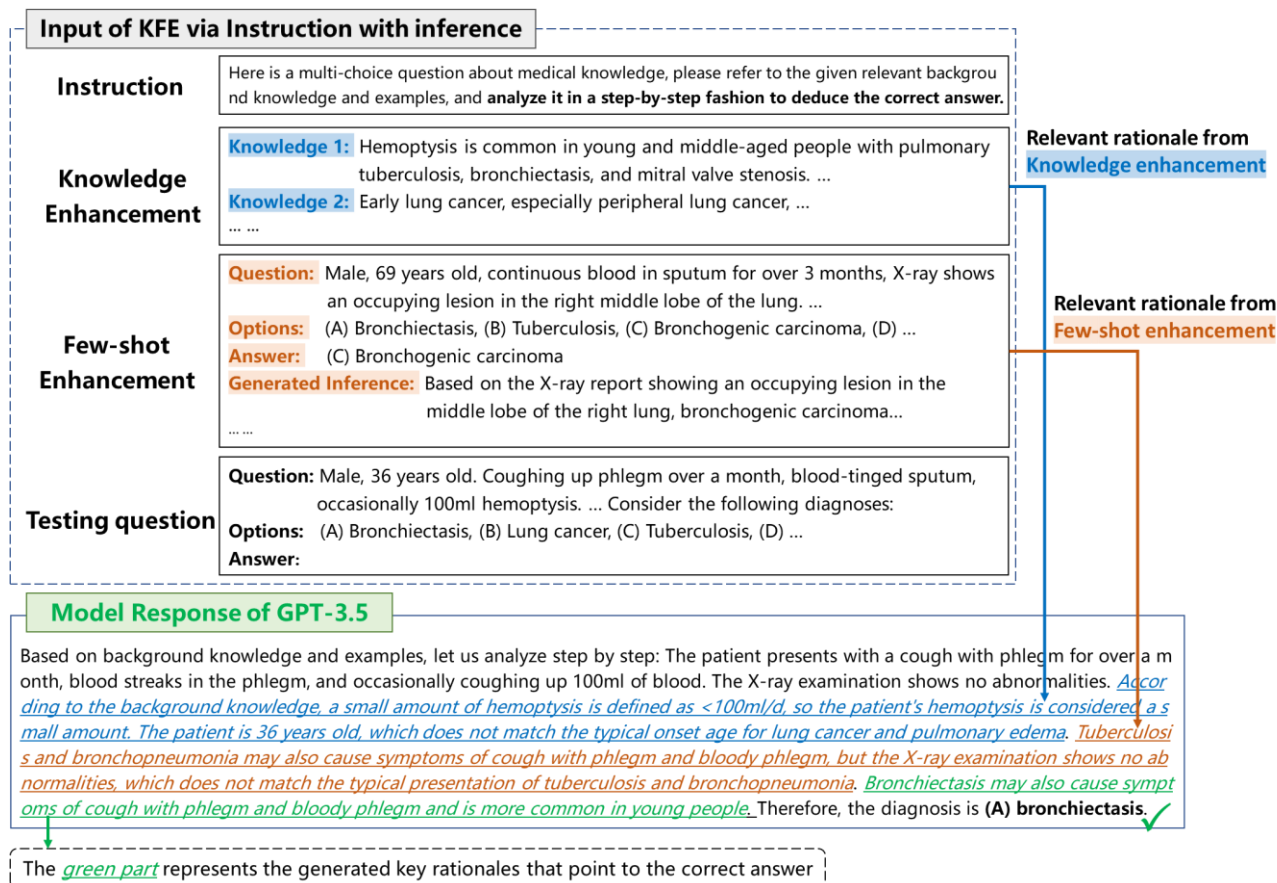
### Supplementary materials 6. A case study of KFE via Instruction with inference

The whole prompt of KFE via Instruction with inference consists of four parts: Instruction with inference, Knowledge enhancement, Few-shot enhancement, and a testing question. After retrieving the relevant knowledge and examples, we instructed the LLM to analyze the testing question step-by-step and deduce the correct answer. As the model response of GPT-3.5 shows, guided by the instruction and reasoning process exemplified in few-shot, the LLM generated the inference details prior to determining the final answer.



### Supplementary materials 7. A case study of inference generated by GPT-3.5 under the basic setting.

In the zero-shot setting, without any support, GPT-3.5 generated the false cognition that "Bronchiectasis generally does not present with hemoptysis," leading to an incorrect diagnosis



## Supplementary materials 8. A case study of inference generated by GPT-3.5 under the KFE framework.

Amidst the support of KFE, knowledge enhancement provided relevant information about the common causes of hemoptysis, while few shot enhancement provided examples showcasing typical X-ray findings associated with lung cancer. The comprehensive KEF empowers GPT-3.5 to integrate essential rationales derived from both knowledge and few-shot enhancements, thereby generating critical cognition needed for problem-solving and ultimately determining the correct diagnosis.