

CONTENT ACCURACY OF OPTIONS OF MULTIPLE-CHOICE QUESTIONS (MCQS) DEVELOPED BY PULMONOLOGIST AND MEDICAL EDUCATIONIST VERIFIED WITH CHATGPT

Rano Mal Piryani,¹ Rajesh Piryani,² Shomeeta Piryani³

ABSTRACT

INTRODUCTION

The vetting, a process of review of multiple-choice questions (MCQs- Items) including stem, lead in questions and options (correct answer and distractors) is done by panel of experts. The objective of this study was to assess the content accuracy of the options (correct answers and distractors) of MCQs developed by pulmonologist and medical educator with generated by Chatbot Generative Pretrained Transformer (ChatGPT) using the same stems with lead in questions and assess the rationale of the options.

MATERIAL AND METHODS

In first step, one best answer type (Type A) MCQs at undergraduate level were developed by pulmonologist & medical educator following item writing guidelines and utilizing checklist of item quality. During second step, the options (correct answers and distractors) of the developed MCQs generated by ChatGPT with rationale using concise, contextual and relevant prompt. In third step, correct answers of those MCQs developed by medical educator whose correct answer different than created by ChatGPT were verified with ChatGPT twice. In fourth step, review of content accuracy of options & correct answer of all MCQs with ChatGPT was done. Finally, percentage of accuracy of options and correct answer generated by ChatGPT was calculated.

RESULTS

ChatGPT 4.1 free version confirmed content of 91% options (correct answers and distractors) accurate correct and acceptable and 9% possibly incorrect & less specific and less plausible. The rationales generated by ChatGPT were acceptable.

CONCLUSION

The ChatGPT 4.1 version may be considered as an expert for confirming the accuracy of content of options including correct answer and distractors of MCQs with acceptable rationales.

KEYWORDS

AI, Content accuracy, ChatGPT, Items, MCQs, Options.

1. Bilawal Medical College for Boys, Liaquat University of Medical & Health Sciences, Jamshoro, Sindh, Pakistan.
2. Akkodis Research, Akkodis, Blagnac, France.
3. Consultant Radiologist, Memon Medical Institute, Karachi, Sindh, Pakistan

<https://doi.org/10.3126/jucms.v13i03.88834>

For Correspondence

Dr. Rano Mal Piryani
Professor of Pulmonology & Medical Education,
Director Medical Education
Bilawal Medical College for Boys,
Liaquat University of Medical and Health Sciences,
Jamshoro, Sindh, Pakistan.
Email: r_piryani@hotmail.com,
rmpiryani@gmail.com

INTRODUCTION

For assessing the students' knowledge and understanding, medical colleges use multiple choice questions (MCQs), a selection method of assessment of cognition.¹⁻² Because of power of assessing higher order of thinking skills, wide range of content and greater number of learners at a time; and having appropriate validity, reliability, objectivity and scalability, MCQs stays as an essential component of health professions education assessments.³⁻⁶

Each MCQ is termed as an item. The parts of item are stem, lead in question, and answer options. One of the options is a key option i.e. correct answer and others are distractors. The process of developing MCQs is known as an item writing.¹⁻² For developing quality MCQs, especially A and modified A type, extensive resources including faculty development training is required.^{1,4-5} The process of review and evaluation of items (MCQs) including stem, lead in questions and options (correct answer and distractors) is called as vetting. Items are rectified by ensuring the content of items are correct, identifying item writing flaws and eliminating these and detecting errors of grammar and language and correcting these errors. This process is done by the experts (item writing author/s, content expert/s and medical educationist) for improving the quality of items.⁶

For over last few years large language models (LLMs) specifically Chatbot Generative Pretrained Transformer (ChatGPT), open artificial intelligence (AI) tool is an alternative experimented for creating MCQs.^{1,5,7-8} ChatGPT has been utilized for developing MCQs using prompt.¹ The quality of prompt affects the quality of item (MCQ) and its parts (stem, lead in question, and options) generated by AI.^{5,7-9} Prompt must be concise with relevant information and context, as to get relevant and accurate response.^{2,10}

The different studies documented the different level of content errors in the MCQs generated by ChatGPT and in some substantial inaccuracies were noted.⁹⁻¹¹ Validating content parts of MCQs is vital as quality of assessment influences competence including clinical competence that is critical for future healthcare professionals.¹¹

This study verified the accuracy of content of the options including correct answer and distractors of MCQs developed by pulmonologist and medical educator with generated by ChatGPT using prompt and assess the rationale for the options given by ChatGPT. This may be the first study of its kind.

MATERIAL AND METHODS

Objectives of the study: The overall objective of this study was to assess the content accuracy of the options and correct answer to the multiple-choice questions (MCQs) items with generated by ChatGPT. The specific objectives were to: 1) create by ChatGPT the options and correct answer for the stems of MCQs developed by pulmonologist and medical educator; and 2) assess the content accuracy of the options and correct answer of the stems of MCQs developed by pulmonologist and medical educator with generated by ChatGPT.

Process of this study: The first author (Pulmonologist & Medical Educator) developed 11 MCQs of pulmonology for medical undergraduate level (assessment of final year

MBBS) following guidelines for the development of MCQs. Figure 1 illustrates the step wise process of the study.

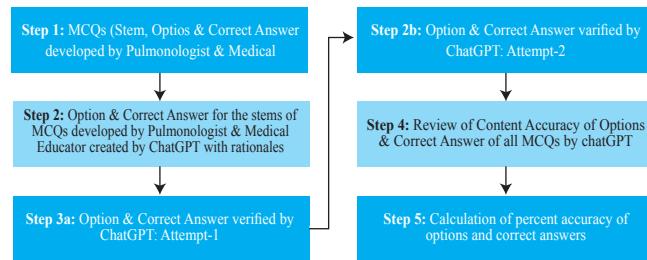


Figure 1. Stepwise Process

The stepwise process of this study is given below:

Step 1: Development of pulmonology MCQs

One best answer type (Type A) eleven MCQs at undergraduate level were developed by Experienced Pulmonologist & Medical Educator. The developed items were evaluated by author following item writing guidelines used by Netere AK et al.¹² and utilizing the checklist for MCQs item quality documented in the study conducted by Rezigalla A. A.⁴

Step 2: Creation of Options & Correct Answer of MCQs with rationales by ChatGPT

Options & Correct Answer for the stems of MCQs developed by Pulmonologist & Medical Educator were created by ChatGPT with rationales.

Step 2: Prompt

“Given the following 10 vignettes for MCQs provide 5 options for each, mention the correct answer option and explain why answer option is correct and why other options are incorrect.”

Step 3: Verification of Options & Correct Answer

Options & Correct Answer of those MCQs developed by medical educator whose correct answer other (different) than created by ChatGPT were verified by ChatGPT twice.

Step 3: Prompt

“Given the following vignettes for MCQs with 5 options for each, and the correct answer option. Please verify the correct answer? if not correct then provide reason.”

Step 4: Verification of Options & Correct Answer (review of Content Accuracy)

Options & Correct Answer of all MCQs developed by medical educator entered to ChatGPT with prompt given below:

Step 4: Prompt

“Verify, the options are acceptable & answer option is correct. If options are not acceptable & answer option is not correct, then mention the reason.”

Step 5: Calculation of percent accuracy of options and correct answers

Calculation of percentage of accuracy of options and correct answer generated by ChatGPT.

Ethical Approval: This study was not done on human beings and did not use any patient data, or personal

identifiers, hence, this study does not require ethical approval. This study is based on the demonstration of a methodological approach using AI ChatGPT 4.1 free version for use (generation of options for item including distractor and correct answer).

RESULTS

The results are documented in tables:

Comparison of Options of Multiple-Choice Questions (MCQs), one best answer (Type-A) developed by pulmonologist and medical educator and generated by ChatGPT utilizing same stems are documented in Table 1. For reference ChatGPT link is given below:

https://drive.google.com/file/d/1DTjyrTD5IK_Sc6LXHgucHOp3X8xRDRLK/view?usp=share_link

Verification of options and correct answer of those MCQs whose correct answer and/or options created by ChatGPT were different then created by pulmonologist and medical educator is shown in table 2. For reference ChatGPT links (first attempt & second attempt are given below:

https://drive.google.com/file/d/1DTjyrTD5IK_Sc6LXHgucHOp3X8xRDRLK/view?usp=share_link

https://drive.google.com/file/d/1rXBOxmjBwp8CvohsLv7pcD_OXdw8O5Dm/view?usp=sharing

Verification (review & validation) by ChatGPT of options and correct answer of all MCQs developed by medical educator & pulmonologist is given in table 3. For reference ChatGPT link is given below:

https://drive.google.com/file/d/1Q6O4QQfKJzaidaRRn9WbqdbRnqzLQN9/view?usp=share_link

Content Validity (Accuracy) of Options, Correct Answer and Distractor of Eleven MCQs verified by ChatGPT is depicted in Table 4.

Table 1. Comparison of Options of Multiple-Choice Questions (MCQs), one best answer (Type-A) developed by pulmonologist and medical educator and generated by ChatGPT utilizing same stems

Stem developed by Medical Educator & Pulmonologist	Options and correct answer developed by Medical Educator & Pulmonologist	Options and correct answer with reason generated /created by ChatGPT Free Version 4.1
MCQ-1 A 25-year-old male presented in OPD having fever, cough with sputum and chest pain for five days and no other significant history revealed. Temperature was 101°F and pulse 110 beats/minute. On auscultation bronchial breathing and increased vocal resonance on anterior right chest detected. Rest of examination unremarkable. The most likely diagnosis is	Options: 1. Acute Bronchitis 2. Community Acquired Pneumonia 3. Acute exacerbation of Bronchial Asthma 4. Acute exacerbation of Bronchiectasis 5. Acute exacerbation of COPD Correct Answer: 2. Community Acquired Pneumonia Explanation: • Correct (B): Lobar pneumonia presents with fever, cough, bronchial breath sounds, and increased vocal resonance. • A: Asthma typically causes wheezing, not bronchial breathing or localized signs. • C: TB has a more chronic course, typically without sudden onset fever and productive cough. • D: Lung abscess causes foul-smelling sputum and is more subacute. • E: Pleural effusion causes reduced breath sounds, not bronchial breathing.	
MCQ-2 A 20-year-old, thin tall presented in emergency with right sided chest pain of 3 hours duration. His general physical examination was normal except tachycardia (pulse 106 beats/minute). Chest examination revealed hyper-resonant percussion on right side with decreased breathing sound. The first test you choose for diagnosis	Options: 1. Complete blood count 2. Electrocardiogram 3. X-Ray Chest PA View 4. Arterial Blood Gas 5. Echocardiograph Correct Answer: 3. X-Ray Chest PA View Explanation: • Correct (C): Chest X-ray is the initial test to confirm pneumothorax. • A: CT is more sensitive but not first-line. • B: ECG is irrelevant. • D: ABG may show hypoxia but doesn't confirm diagnosis. • E: Ultrasound can detect pneumothorax but is not standard first-line.	

MCQ-3 A 30-year-female presented with history of cough for 3 weeks and chest pain on left side for one week. Examination consistent with pleural effusion and confirmed with ultrasound. You have aspirated the fluid which one of the following tests you send to confirm the fluid is likely exudative	Options: 1. Pleural fluid protein 2. Pleural fluid glucose 3. Pleural fluid RBCs 4. Pleural fluid Eosinophils 5. Pleural fluid Monocytes Correct Answer: 1. Pleural Fluid Protein	Options: A. Total WBC count B. Light's criteria C. Serum albumin D. Fluid color E. Chest X-ray Correct Answer: B. Light's criteria Explanation: • Correct (B): Light's criteria is the standard for differentiating transudate vs exudate. • A: WBC count helps with infection, not exudate/transudate. • C: Albumin levels alone aren't enough. • D: Fluid color is subjective. • E: X-ray shows effusion, not its type.
MCQ-4 A 50-year female developed relapse of Pulmonary Tuberculosis. Consultant prescribed her five anti-TB drugs (Streptomycin, Isoniazid, Rifampicin, Ethambutol and pyrazinamide). After one month she developed tingling sensation in legs. She went to consultant who told her she developed peripheral neuropathy. Peripheral neuropathy is common side effect of which drug	Options: 1. Streptomycin 2. Isoniazid 3. Rifampicin 4. Ethambutol 5. Pyrazinamide Correct Answer: 2. Isoniazid	Options: A. Rifampicin B. Isoniazid C. Pyrazinamide D. Ethambutol E. Streptomycin Correct Answer: B. Isoniazid Explanation: • Correct (B): INH causes pyridoxine deficiency leading to neuropathy. • A: Rifampicin causes hepatotoxicity, not neuropathy. • C: Pyrazinamide causes hyperuricemia. • D: Ethambutol causes optic neuritis. • E: Streptomycin causes ototoxicity/nephrotoxicity.
MCQ-5 As per WHO definition patients is having Multidrug Resistant (MDR) MTB if patient develop resistant to which of the following drugs	Options: 1. Ethambutol & Isoniazid 2. Ethambutol & pyrazinamide 3. Isoniazid & Rifampicin 4. Rifampicin & pyrazinamide 5. Isoniazid & Pyrazinamide Correct Answer: 3. Isoniazid and Rifampicin	Options: A. Isoniazid and Ethambutol B. Rifampicin and Streptomycin C. Isoniazid and Rifampicin D. Pyrazinamide and Isoniazid E. Ethambutol and Rifampicin Correct Answer: C. Isoniazid and Rifampicin Explanation: • Correct (C): MDR-TB is defined as resistance to both isoniazid and rifampicin. • Other options involve incorrect combinations.
MCQ-6 While examining chest of 60-year-old patient in outpatient department (OPD), you detected hyper resonant percussion note on right side of the chest. Due think which is likely gross pathology in right side of thorax	Options: 1. Pleural Effusion 2. Consolidation 3. Large Cyst 4. Pneumothorax 5. Chylothorax Correct Answer: 4. Pneumothorax	Options: A. Consolidation B. Pleural effusion C. Lung collapse D. Pneumothorax E. Pulmonary edema Correct Answer: D. Pneumothorax Explanation: • Correct (D): Pneumothorax leads to hyper resonant percussion. • A/C: These cause dullness, not hyperresonance. • B: Effusion = stony dullness. • E: Usually dullness, not hyperresonance
MCQ-7 A 25-year-old male having no comorbidity developed pneumonia which is typical community acquired pneumonia; the most common microorganism causing typical community acquired pneumonia (CAP) in this young adult is	Options: 1. Moraxella catarrhalis 2. Streptococcus pneumoniae 3. Haemophilus influenzae 4. Klebsiella pneumoniae 5. Mycoplasma pneumoniae Correct Answer: 2. Streptococcus pneumoniae	Options: A. Klebsiella pneumoniae B. Haemophilus influenzae C. Streptococcus pneumoniae D. Mycoplasma pneumoniae E. Legionella pneumonia Correct Answer: C. Streptococcus pneumoniae Explanation: • Correct (C): Most common typical CAP pathogen. • A: More common in alcoholics. • B: Seen in COPD patients. • D: Atypical pneumonia. • E: Legionella is rare and atypical
MCQ-8 A 50-year-old female who has bronchial asthma for last 10 years developed hypertension. Doctor prescribed her one antihypertensive drug; she could not remember the name. After taking drugs for few weeks her asthma symptoms have started worsening. The drug triggering asthma is likely	Options: 1. Amlodipine 2. Hydrochlorothiazide 3. Prazosin 4. Losartan 5. Spironolactone Correct Answer: 4. Losartan	Options: A. Amlodipine B. Enalapril C. Losartan D. Atenolol E. Hydrochlorothiazide Correct Answer: D. Atenolol Explanation: • Correct (D): Beta-blockers (especially non selective like atenolol) worsen asthma. • A: CCBs are safe in asthma. • B: ACE inhibitors may cause cough, not bronchospasm. • C: ARBs are generally safe. • E: Diuretics don't worsen asthma.
MCQ-9 A 55-year-old male is suffering from reversible obstructive airway disease. Do you think which one among the following diseases he is suffering from	Options: 1. COPD 2. Bronchiectasis 3. Bronchial Asthma 4. Primary Ciliary Dyskinesia 5. Emphysema Correct Answer: 3. Bronchial Asthma	Options: A. COPD B. Bronchiectasis C. Asthma D. Lung cancer E. Sarcoidosis Correct Answer: C. Asthma Explanation: • Correct (C): Asthma is the only one with reversible airway obstruction. • A: COPD is irreversible. • B: Not fully reversible, with purulent sputum. • D/E: Not typically obstructive or reversible.
MCQ-10 During examination of the patient of 45-year-old male in ward you detected trachea is grossly shifted on the same side of pathology i.e. on right side. In your opinion what is likely cause	Options: 1. Consolidation of right lung 2. Pneumothorax on right side 3. Collapse of right lung 4. Hydrothorax on right side 5. Cyst in right lung Correct Answer: 3. Collapse of right lung	Options: A. Right-sided pleural effusion B. Right-sided pneumothorax C. Right-sided lung fibrosis D. Right-sided lung tumor E. Left-sided pleural effusion Correct Answer: C. Right-sided lung fibrosis Explanation: • Correct (C): Fibrosis pulls trachea toward the same side. • A/B: Effusion and pneumothorax push it away. • D: Tumor may push unless there's collapse. • E: Unrelated side.

MCQ-II	<p>Options:</p> <ol style="list-style-type: none"> 1. Sputum for AFB Smear 2. Chest X-Ray PA view 3. GenXpert test of sputum 4. Mycobacterial Culture of sputum 5. Tuberculin Test <p>Correct Answer: 4 Mycobacterial Culture of sputum</p>	<p>Options:</p> <ol style="list-style-type: none"> A. Chest X-ray B. Sputum AFB smear C. GeneXpert D. Sputum culture (Lowenstein-Jensen medium) E. ESR <p>Correct Answer: D. Sputum culture (LJ medium)</p> <p>Explanation:</p> <ul style="list-style-type: none"> • Correct (D): Culture is the gold standard. • A/B/C: Faster, but less definitive than culture. • E: Nonspecific.
--------	--	---

Table 2. Verification of options and correct answer of those MCQs whose correct answer and/or options created by ChatGPT different then created by educator.

Educator created stem and lead in question with options and correct Answer	ChatGPT generated options and correct Answer on first attempt	ChatGPT generated options and correct Answer on second attempt
MCQ-I A 25-year-old male presented in OPD having fever, cough with sputum and chest pain for five days and no other significant history revealed. Temperature was 101°F and pulse 110 beats/minute. On auscultation bronchial breathing and increased vocal resonance on anterior right chest detected. Rest of examination unremarkable. The most likely diagnosis is Options: 1. Acute Bronchitis 2. Community Acquired Pneumonia 3. Acute exacerbation of Bronchial Asthma 4. Acute exacerbation of Bronchectasis 5. Acute exacerbation of COPD Correct Answer: 2. Community Acquired Pneumonia	MCQ-I Vignette Summary: • 25-year-old male • Fever, productive cough, chest pain • Febrile (101°F), tachycardic • Bronchial breath sounds and increased vocal resonance on the right anterior chest Answer Given: 2. Community Acquired Pneumonia <input checked="" type="checkbox"/> Correct Reasoning: • Bronchial breath sounds and increased vocal resonance are classic signs of lung consolidation , most commonly due to pneumonia . • Other options like bronchitis, asthma, bronchiectasis, and COPD do not typically show focal bronchial breathing or increased vocal resonance.	MCQ-I Vignette Summary: Young male with fever, productive cough, pleuritic chest pain, tachycardia, bronchial breath sounds, and increased vocal resonance on right anterior chest. Correct Answer Given: 2 This is CORRECT. Explanation: • Fever + cough + sputum + pleuritic chest pain + bronchial breath sounds and increased vocal resonance → all are classic signs of lobar pneumonia (a type of CAP). • Other options: Acute bronchitis – usually viral, dry cough, auscultation normal or with wheeze. Asthma exacerbation – no fever or vocal resonance changes. Bronchiectasis exacerbation – chronic, usually with past history, copious purulent sputum. COPD exacerbation – older patients, history of smoking, wheeze likely.
MCQ-3 A 30-year-old female presented with history of cough for 3 weeks and chest pain on left side for one week. Examination consistent with pleural effusion and confirmed with ultrasound. You have aspirated the fluid which one of the following tests you send to confirm the fluid is likely exudative Options: 1. Pleural fluid protein 2. Pleural fluid glucose 3. Pleural fluid RBCs 4. Pleural fluid Eosinophils 5. Pleural fluid Monocytes Correct Answer: 1. Pleural Fluid Protein	MCQ-3 Vignette Summary: • 30-year-old female • Cough (3 weeks), left chest pain (1 week) • Pleural effusion confirmed on ultrasound • Pleural fluid aspirated Question: Test to confirm fluid is exudative? Answer Given: 1. Pleural Fluid Protein <input checked="" type="checkbox"/> Correct Correct Answer: <input checked="" type="checkbox"/> Incorrect <input checked="" type="checkbox"/> Correct answer should be: **Use of Light's criteria (which uses pleural fluid protein as one component, not alone) Since "Light's Criteria" isn't an option here, let's assess option 1 Since "Light's Criteria" isn't an option here, let's assess option 1: Pleural fluid protein >3 g/dL suggests exudate but is not sufficient alone to definitively classify it as exudative. Light's Criteria (which compares pleural fluid and serum protein and LDH levels) is the standard . Conclusion: Option 1 is the best available answer in the absence of "Light's Criteria" — but incomplete. <input checked="" type="checkbox"/> Conditionally accepted due to lack of better option	MCQ-3 Vignette Summary: Young woman with pleuritic chest pain, pleural effusion confirmed on ultrasound, pleural fluid aspirated. Correct Answer Given: 1. Pleural Fluid Protein <input checked="" type="checkbox"/> This is CORRECT. Explanation: • Light's Criteria is used to differentiate exudate vs. transudate. Exudate if any of the following is true: 1. Pleural fluid protein / serum protein > 0.5 2. Pleural fluid LDH / serum LDH > 0.6 3. Pleural fluid LDH > 2/3 upper limit of normal serum LDH • So, pleural fluid protein is a key test for determining exudative nature. • Other options are less specific: Glucose can be low in TB, empyema, rheumatoid, but doesn't define exudate. RBCs, eosinophils, monocytes help with etiology, not exudate classification.
MCQ-8 A 50-year-old female who has bronchial asthma for last 10 years developed hypertension. Doctor prescribed her one antihypertensive drug; she could not remember the name. After taking drugs for few weeks her asthma symptoms have started worsening. The drug triggering asthma is likely Options: 1. Amlodipine 2. Hydrochlorothiazide 3. Prazosin 4. Losartan 5. Spironolactone Correct Answer: 4 Losartan	MCQ-8 Vignette Summary: • 50-year-old female with bronchial asthma develops hypertension • Starts new antihypertensive • Asthma worsens after starting the drug Answer Given: 4. Losartan <input checked="" type="checkbox"/> Incorrect Correct Answer: <input checked="" type="checkbox"/> Incorrect <input checked="" type="checkbox"/> Correct answer should be: None of these directly—but commonly it's non-selective beta blockers (e.g., atenolol, propranolol) or ACE inhibitors (like enalapril). Analysis: • Losartan is an ARB (angiotensin receptor blocker) and is considered safe in asthma patients. • If Enalapril (an ACE inhibitor) were listed, it would be correct due to risk of cough and bronchospasm. • Among these, none are typical asthma triggers. Conclusion: Given options are suboptimal — if forced to choose, losartan is not correct.	MCQ-8 Vignette Summary: Known asthmatic started antihypertensive → asthma worsened. Answer Given: 4. Losartan <input checked="" type="checkbox"/> INCORRECT. Correct Answer: **None of the options are classic triggers — but the most likely culprit should be a non-selective beta blocker like propranolol or sometimes ACE inhibitors (due to cough). But among these options, the closest suspect for asthma worsening would be: • ACE inhibitors (e.g., enalapril, lisinopril) — cause cough in 10-20% of patients and can exacerbate asthma. • Losartan is an ARB, less likely to cause cough and does not worsen asthma. So technically, none of these listed options clearly cause asthma worsening. <input checked="" type="checkbox"/> Correct answer should be revised or clarified.
MCQ-10 During examination of the patient of 45-year-old in ward you detected trachea is grossly shifted on the same side of pathology i.e. on right side. In your opinion what is likely cause Options: 1. Consolidation of right lung 2. Pneumothorax on right side	MCQ-10 Vignette Summary: • Trachea shifted toward right side (same side of pathology) Answer Given: 3. Collapse of right lung <input checked="" type="checkbox"/> Verification: <input checked="" type="checkbox"/> Correct Reasoning: • Tracheal deviation toward the side of the	MCQ-10 Vignette Summary: Tracheal deviation toward right side. Correct Answer Given: 3. Collapse of right lung <input checked="" type="checkbox"/> This is CORRECT. Explanation: • Tracheal deviation toward side of pathology occurs with volume loss , such as: o Lung collapse

3. Collapse of right lung 4. Hydrothorax on right side 5. Cyst in right lung Correct Answer: 3 Collapse of right lung	lesion = volume loss , as seen in lung collapse or fibrosis . • Pneumothorax and pleural effusion typically push trachea away from the lesion. • Consolidation doesn't usually shift the trachea. <input checked="" type="checkbox"/>	<ul style="list-style-type: none"> o Fibrosis • Tracheal deviation away from pathology occurs with: o Tension pneumothorax o Large pleural effusion o Massive tumor <p>So:</p> <ul style="list-style-type: none"> • Collapse of right lung → trachea pulled to right <input checked="" type="checkbox"/> • Pneumothorax or effusion → trachea pushed to opposite side <input checked="" type="checkbox"/>
<input checked="" type="checkbox"/> Final Summary of Verifications:	<input checked="" type="checkbox"/> Summary of Accuracy	

Table-3. Verification (review & validation) by ChatGPT of options and correct answer of all MCQs developed by medical educator & pulmonologist

Educator created stem and lead in question and options and key or correct answer	ChatGPT verified options and key or correct answer	Comments by educator
MCQ-1 Diagnosis: Community Acquired Pneumonia (CAP) Options: 1. Acute Bronchitis 2. Community Acquired Pneumonia 3. Acute exacerbation of Bronchial Asthma 4. Acute exacerbation of Bronchectasis 5. Acute exacerbation of COPD Correct Answer: 2. Community Acquired Pneumonia	MCQ-1 Diagnosis: Community Acquired Pneumonia (CAP) • Findings: Fever, productive cough, pleuritic chest pain, bronchial breath sounds, increased vocal resonance. • Correct Answer: 2. Community Acquired Pneumonia • Options Appropriateness: Acceptable and distinct. <input checked="" type="checkbox"/> Answer and options are correct.	Valid Options and Correct Answer are correct.
MCQ-2 Presentation: Tall, young male with sudden chest pain and hyporespiration → suggests spontaneous pneumothorax. Options: 1. Complete blood count 2. Electrocardiogram 3. X-Ray Chest PA View 4. Arterial Blood Gas 5. Echocardiograph Correct Answer: 3. X-Ray Chest PA View	MCQ-2 Presentation: Tall, young male with sudden chest pain and hyporespiration → suggests spontaneous pneumothorax. • Initial Investigation: Chest X-ray (PA view) to confirm. • Correct Answer: 3. X-Ray Chest PA View <input checked="" type="checkbox"/> • Options Appropriateness: All other options are either irrelevant (e.g., ECG, Echo) or secondary. <input checked="" type="checkbox"/> Answer and options are correct.	Valid Options and Correct Answer are correct.
MCQ-3 Diagnosis: Pleural effusion Options: 1. Test for Exudate vs. Transudate: Light's criteria → Pleural fluid protein is one criterion. 2. Pleural fluid protein 2. Pleural fluid glucose 3. Pleural fluid RBCs 4. Pleural fluid Eosinophils 5. Pleural fluid Monocytes Correct Answer: 1. Pleural Fluid Protein	MCQ-3 Diagnosis: Pleural effusion • Test for Exudate vs. Transudate: Light's criteria → Pleural fluid protein is one criterion. • Correct Answer: 1. Pleural fluid protein <input checked="" type="checkbox"/> • Options Appropriateness: Only 1 is relevant to Light's criteria. Others are misleading. <input checked="" type="checkbox"/> Options could be improved by including: • Pleural fluid LDH • Serum protein • Serum LDH <input checked="" type="checkbox"/> Correct answer is valid.	Valid Correct Answer is correct Options are also correct but comparably less specific
MCQ-4 Side Effect: Peripheral neuropathy due to Isoniazid → Pyridoxine deficiency. Options: 1. Streptomycin 2. Isoniazid 3. Rifampicin 4. Ethambutol 5. Pyrazinamide Correct Answer: 2 Isoniazid	MCQ 4 A 50-year-old female developed tingling sensation in legs. She went to consultant who told her she developed peripheral neuropathy. Peripheral neuropathy is common side effect of which drug Options: 1. Streptomycin 2. Isoniazid 3. Rifampicin 4. Ethambutol 5. Pyrazinamide Correct Answer: 2 Isoniazid	Valid Options and Correct Answer are correct.
MCQ-5 Definition of MDR-TB: Resistant to Isoniazid and Rifampicin. Options: 1. Ethambutol & Isoniazid 2. Ethambutol & pyrazinamide 3. Isoniazid & Rifampicin 4. Rifampicin & pyrazinamide 5. Isoniazid & Pyrazinamide Correct Answer: 3 Isoniazid and Rifampicin	MCQ-5 As per WHO definition patients is having Multidrug Resistant (MDR) MTB if patient develop resistant to which of the following drugs Options: 1. Ethambutol & Isoniazid 2. Ethambutol & pyrazinamide 3. Isoniazid & Rifampicin 4. Rifampicin & pyrazinamide 5. Isoniazid & Pyrazinamide Correct Answer: 3 Isoniazid and Rifampicin	Valid Options and Correct Answer are correct.

MCQ-6	MCQ-6 Finding: Hyperresonance → Pneumothorax most likely. • Correct Answer: 4. Pneumothorax <input checked="" type="checkbox"/> • Options • Appropriateness: Acceptable. <input checked="" type="checkbox"/> Answer and options are correct.	Valid Options and Correct Answer are correct.																																																				
MCQ-7	MCQ-7 Typical CAP organism in young adult: Streptococcus pneumoniae • Correct Answer: 2. Streptococcus pneumoniae <input checked="" type="checkbox"/> • Options Appropriateness: Good mix of typical and atypical organisms. <input checked="" type="checkbox"/> Answer and options are correct.	Valid Options and Correct Answer are correct.																																																				
MCQ-8	MCQ-8 Asthma worsening after antihypertensive use: • Likely culprit: Beta blockers (but not in options). • Losartan is an ARB, not a common asthma trigger. • Amlodipine, Prazosin, HCTZ are safer for asthma. • Possibly wrong answer. <input checked="" type="checkbox"/> Correct Answer Should Be: None listed clearly triggers asthma — if it were a non-selective beta-blocker, that would be a better option. <input checked="" type="checkbox"/> Correction needed. Replace option 4 or include a beta-blocker (e.g., Propranolol).	Propranolol, a beta blocker is not used as a antihypertensive drug but atenolol, metoprolol, bisoprolol are used for treatment of hypertension. So, the correct answer could be replaced with either atenolol or metoprolol, or bisoprolol																																																				
MCQ-9	MCQ-9 Reversible airway disease: Classical definition of Bronchial Asthma. • Correct Answer: 3. Bronchial Asthma <input checked="" type="checkbox"/> • Options Appropriateness: Valid. <input checked="" type="checkbox"/> Answer and options are correct.	Valid Options and Correct Answer are correct.																																																				
MCQ-10	MCQ-10 Tracheal shift to same side: Classic for lung collapse or fibrosis. • Correct Answer: 3. Collapse of right lung <input checked="" type="checkbox"/> • Options • Appropriateness: Acceptable. <input checked="" type="checkbox"/> Answer and options are correct.	Valid Options and Correct Answer are correct.																																																				
MCQ-11	MCQ-11 Gold Standard for TB diagnosis: Mycobacterial culture • Correct Answer: 4. Mycobacterial Culture of sputum • Options Appropriateness: Valid; GenXpert is rapid but not the gold standard. <input checked="" type="checkbox"/> Answer and options are correct.	Valid Options and Correct Answer are correct.																																																				
<table border="1"> <thead> <tr> <th colspan="4">Summary Table</th> </tr> <tr> <th>MCQ</th> <th>Correct Answer</th> <th>Verdict</th> <th>Comments</th> </tr> </thead> <tbody> <tr> <td>1</td> <td>2</td> <td><input checked="" type="checkbox"/> Correct</td> <td></td> </tr> <tr> <td>2</td> <td>3</td> <td><input checked="" type="checkbox"/> Correct</td> <td></td> </tr> <tr> <td>3</td> <td>1</td> <td><input checked="" type="checkbox"/> Partially adequate</td> <td>Add LDH/serum markers for full Light's criteria</td> </tr> <tr> <td>4</td> <td>2</td> <td><input checked="" type="checkbox"/> Correct</td> <td></td> </tr> <tr> <td>5</td> <td>3</td> <td><input checked="" type="checkbox"/> Correct</td> <td></td> </tr> <tr> <td>6</td> <td>4</td> <td><input checked="" type="checkbox"/> Correct</td> <td></td> </tr> <tr> <td>7</td> <td>2</td> <td><input checked="" type="checkbox"/> Correct</td> <td></td> </tr> <tr> <td>8</td> <td>4</td> <td><input checked="" type="checkbox"/> Possibly incorrect</td> <td>None of the listed drugs rarely cause asthma include beta-blocker</td> </tr> <tr> <td>9</td> <td>3</td> <td><input checked="" type="checkbox"/> Correct</td> <td></td> </tr> <tr> <td>10</td> <td>3</td> <td><input checked="" type="checkbox"/> Correct</td> <td></td> </tr> <tr> <td>11</td> <td>4</td> <td><input checked="" type="checkbox"/> Correct</td> <td></td> </tr> </tbody> </table>			Summary Table				MCQ	Correct Answer	Verdict	Comments	1	2	<input checked="" type="checkbox"/> Correct		2	3	<input checked="" type="checkbox"/> Correct		3	1	<input checked="" type="checkbox"/> Partially adequate	Add LDH/serum markers for full Light's criteria	4	2	<input checked="" type="checkbox"/> Correct		5	3	<input checked="" type="checkbox"/> Correct		6	4	<input checked="" type="checkbox"/> Correct		7	2	<input checked="" type="checkbox"/> Correct		8	4	<input checked="" type="checkbox"/> Possibly incorrect	None of the listed drugs rarely cause asthma include beta-blocker	9	3	<input checked="" type="checkbox"/> Correct		10	3	<input checked="" type="checkbox"/> Correct		11	4	<input checked="" type="checkbox"/> Correct	
Summary Table																																																						
MCQ	Correct Answer	Verdict	Comments																																																			
1	2	<input checked="" type="checkbox"/> Correct																																																				
2	3	<input checked="" type="checkbox"/> Correct																																																				
3	1	<input checked="" type="checkbox"/> Partially adequate	Add LDH/serum markers for full Light's criteria																																																			
4	2	<input checked="" type="checkbox"/> Correct																																																				
5	3	<input checked="" type="checkbox"/> Correct																																																				
6	4	<input checked="" type="checkbox"/> Correct																																																				
7	2	<input checked="" type="checkbox"/> Correct																																																				
8	4	<input checked="" type="checkbox"/> Possibly incorrect	None of the listed drugs rarely cause asthma include beta-blocker																																																			
9	3	<input checked="" type="checkbox"/> Correct																																																				
10	3	<input checked="" type="checkbox"/> Correct																																																				
11	4	<input checked="" type="checkbox"/> Correct																																																				

Table 4. Content Accuracy of Options, Correct Answer and Distractor of Eleven MCQs verified by ChatGPT

Content	Developed by medical educator & pulmonologists	Content confirmed by ChatGPT Correct & Acceptable	Content confirmed by ChatGPT Possibly incorrect & less specific and less plausible	Correct Rationale ChatGPT
Options for the *stem	55	50 (90.9%)	5 (9.1%)	52 (94.5%)
Distractors	44	40 (90.9%)	4 (9.1%)	42 (94.4%)
Correct Answer	11	10 (90.9%)	1 (9.1%)	10 (90.9%)

*Each stem has 5 options (Distractors= 4; correct answer=1), so, 11 MCQs have 55 options (Distractors= 44; correct answer=11) 4

DISCUSSION

The overall objective of this study was to assess the content accuracy of the options (correct answers and distractors) to the multiple-choice questions (MCQs) items developed by pulmonologist and medical educator with generated by ChatGPT using the same stems with lead in questions and assess the rationale generated by ChatGPT. This process before the use of AI and still being done by item author/s, subject expert/s and medical educationist.^{6,13} Study published in first trimester of 2025 by Akram Z et al reported 32% flaws (72% had one flaw and 28% had more than one flaws) in 750 MCQs received from the basic sciences department of private medical colleges of one of the provinces of Pakistan.¹³ Faculty members (content experts) of one of the medical colleges in public sector who participated in the process perceived the vetting activities of MCQs enhances item writing skills, improves content validity and boost quality of assessment. But vetting activities were demanding and exhausting. They further explained demanding in terms of time and commitment of faculty members.⁶ In this study it is done with ChatGPT 4.1 free version which accomplished the task in a considerably very less time and neither it was demanding and nor exhausting.

In this study, ChatGPT 4.1 free version confirmed content accuracy of 91% options (correct answers and distractors) correct and acceptable and around 9% possibly incorrect & less specific and less plausible. The rationales generated by ChatGPT were acceptable.

Study done by Ch'en et al documented that GPT-4 generated accurate, clear and appropriate answer rationales for the pre-existing MCQs in pre-clerkship courses. This indirectly supports the findings of this study.⁷ Zachary H et al in their study concluded that ChatGPT generated pulmonology MCQs had comparable content validity to the pulmonologists-generated MCQs but ChatGPT-generated MCQs were comparably too easy and had inferior psychometric indices.¹⁴ The content accuracy of the options and correct answers to MCQs may maximize the accuracy of assessment. The process done in this study may help to refine the options including correct answer and distractors of MCQs based on evidence for already existing MCQs in MCQs banks. ChatGPT 4.1 version demonstrated substantial potential for verifying content accuracy of options including correct answer and distractors. It may be considered as an expert for confirming the accuracy of content of options including correct answer and distractors of MCQs.

There are limitations of this study. The significant limitation is small number of MCQs used for verification of only content accuracy for options (correct answers and distractors) only, not for detection of other flaws. Another limitation is only pulmonology MCQs items were used developed by one pulmonologist. The findings of this study are fundamentals laying the foundations of further studies.

CONFLICT OF INTEREST

None

FUNDING

No funding required.

REFERENCES

1. Camarata T, McCoy L, Rosenberg R, Temprine Grellinger K, R, Brettschnieder K, & Berman J. LLM-Generated multiple choice practice quizzes for preclinical medical students. *Advances in physiology education* 2025; 49(3): 758–763.
2. Kuusemets L, Parve K, Ain K, Kraav T. Assessing AI-generated (GPT-4) versus human created MCQs in mathematics education: A comparative inquiry into vector topics. *International Journal of Education in Mathematics, Science, and Technology (IJEMST)* 2024; 12 (6): 1538-1558.
3. Agarwal M, Sharma P, Wani P. Evaluating the Accuracy and Reliability of Large Language Models (ChatGPT, Claude, DeepSeek, Gemini, Grok, and Le Chat) in Answering Item-Analyzed Multiple-Choice Questions on Blood Physiology. *Cureus* 2025; 17(4): e81871.
4. Rezigalla A. A. AI in medical education: uses of AI in construction type A MCQs. *BMC medical education* 2024; 24(1): 247.
5. Stadler, M. Horrer A, Fischer M R. Crafting medical MCQs with generative AI: A how-to guide on leveraging ChatGPT. *GMS journal for medical education* 2024; 41(2): Doc20.
6. Aziz A, Sabqat M, Kiran F, Mirza TI. Exploring the experiences of content experts with item vetting during item bank development. *Pak J Med Sci.* 2024;40(6):1241-1246.
7. Ch'en P Y, Day W, Pekson R C, et al. GPT-4 generated answer rationales to multiple choice assessment questions in undergraduate medical education. *BMC Med Educ* 2025; 25 (1): 333.
8. Law et al. AI versus human-generated multiple-choice questions for medical education: a cohort study in a high-stakes examination. *BMC Medical Education* 2025; 25: 208.
9. Kiyak Y S, Emekli E. ChatGPT prompts for generating multiple-choice questions in medical education and evidence on their validity: a literature review. *Postgraduate medical journal* 2024; 100(1189): 858–865.
10. Jiang Z, Feng S. UsmleGPT: An AI application for developing MCQs via multi-agent system Software Impacts 2025; 23:100742.
11. Chauhan A, Khaliq F, Nayak KR. Assessing Quality of Scenario-Based Multiple-Choice Questions in Physiology: Faculty-Generated vs. ChatGPT-Generated Questions among Phase I Medical Students. *Int J Artif Intell Educ* 2025.
12. Netere AK et al. Evaluating the quality of multiple-choice question pilot database: a global educator-created tool for concept-based pharmacology learning. *Pharmacol. Res. Perspect* 2024; 12 (5): e70004.
13. Akram Z, Aman K, Wheed N, Rahim SK, Hanif F, Abdullah Z. Quality Assessment of multiple-choice questions. *Rawal Medical College Journal* 2025; 50 (1): 191-194.
14. Zachary H, Brian E, Andrew G, Kimberly F. ChatGPT VS Subject Matter Expert Generated Pulmonary Multiple-Choice Questions, *CHEST* 2024; 166 (4): Supplement: A3858.