**Original Research**

# Accuracy, Usefulness, and Impact Variability of ChatGPT-4 for COPD Medication Management: A Modified Delphi Study

Paul M. Boylan, PharmD, BCPS[1] Devin L. Lavender, PharmD, BCACP[2] Rebecca H. Stone, PharmD, FCCP, BCPS, BCACP[2] Russ Palmer, PhD[2] Richard L. Lamb, PhD[2,3] Kiley Merritt[2] Joshua Caballero, PharmD, FCCP, BCPP[2]

[1] College of Pharmacy, University of Oklahoma, Oklahoma City, Oklahoma, United States

[2] College of Pharmacy, University of Georgia, Athens, Georgia, United States

[3] College of Veterinary Medicine, University of Georgia, Athens, Georgia, United States

***Address correspondence to:***
Paul Boylan, PharmD, BCPS
1110 N Stonewall Ave CPB 239
Oklahoma City, OK, 73117
Phone: (405) 271-6878 x47255
Email: paul-boylan@ou.edu

***Running Head:*** **Chat-GPT for COPD Medication Management**

***Keywords***: COPD; ChatGPT; artificial intelligence; large language models; clinical decision-making;

***Abbreviations***: AI = artificial intelligence; ChatGPT = Chat Generative Pre-Trained Transformer; COPD = chronic obstructive pulmonary disease; ECOPD = exacerbation of COPD; FDA = US Food and Drug Administration; LLMs = large language models

***Funding Support***: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sections.

***Date of Acceptance***: July 29, 2026 | ***Published Online Date***: August 4, 2026

***Citation***: Boylan PM, Lavender DL, Stone RH, et al. Accuracy, usefulness, and impact variability of ChatGPT-4 for COPD medication management: a modified Delphi study. *Chronic Obstr Pulm Dis*. 2026; Published online August 4, 2026.
https://doi.org/10.15326/jcopdf.2026.0794

***This article has an online supplement.***

**Abstract**

**Background.** Chronic obstructive pulmonary disease (COPD) management is complex and rapidly evolving. ChatGPT is a large language model (LLM) shown to generate treatment plans for chronic conditions, yet its accuracy, usefulness, and consistency for COPD remain poorly characterized. This study evaluated the accuracy, usefulness, and impact variability of simultaneous ChatGPT-4.0 responses to COPD medication management questions.

**Methods.** Five COPD treatment questions were simultaneously entered into three separate computers using ChatGPT-4.0 during a single session, generating 15 total responses. Three residency-trained, board-certified clinical pharmacists rated each response across three domains - accuracy, usefulness, and impact variability - using a 3-point ordinal scale (0 to 2) via a three-round modified Delphi process. Consensus was defined *a priori* as unanimous agreement among all three panelists.

**Results.** Of 45 response-domain ratings, consensus was achieved in 40 (88.9%). Accuracy ranged from poor (0) to good (2), usefulness from somewhat (1) to very useful (2), and impact variability from low (0) to high (2). For one question on stable COPD pharmacotherapy, all three simultaneous responses cited a retired clinical practice guideline, resulting in poor accuracy ratings. For two questions - treating exacerbations and managing a complex case - one response per question was rated higher in usefulness than the others.

**Conclusions**. Simultaneous ChatGPT-4.0 responses to identical COPD prompts differed materially in accuracy and potential clinical impact. LLM outputs should augment rather than replace clinician judgment and must be deployed with current guideline grounding and appropriate local oversight.




Published online August 4, 2026 https://doi.org/10.15326/jcopdf.2026.0794

## Introduction

Chronic obstructive pulmonary disease (COPD) affects an estimated 392 million people, or 6% of all global deaths, with prevalence highest among adults greater than 40 years of age, particularly in low and middle income countries where exposure to tobacco smoke, biomass fuel, and air pollution is common.[1-5] Large population-based analyses estimate global prevalence among individuals aged 30-79 years at 10.3%, with substantial regional variation driven by differences in smoking rates, occupational exposures, and socioeconomic status.[6, 7] COPD prevalence and mortality are rising among women, particularly where smoking and household air pollution are prevalent.[8] Projections suggest that, despite advances in prevention and treatment, COPD prevalence and mortality will continue to rise through 2050, particularly in aging populations and regions with persistent exposure to modifiable risk factors.[8]

In addition to rising prevalence, patients lack access to adequately trained health care professionals, particularly in rural and resource-limited settings.[9] As a result, patients and providers lacking training and experience in COPD may resort to alternative resources for disease management support, including artificial intelligence (AI)-based tools such as large language models (LLMs) to obtain information and guidance outside of traditional clinical care pathways. Ultimately, AI may aid clinicians to deliver patient care by collecting vast and disparate data to support accurate disease diagnosis, clinical reasoning, and treatment decisions.[10] AI could streamline clinicians' workload and enable health care providers to dedicate more time to patient care responsibilities.[11] Over the last decade, the FDA's public list of AI and machine learning-enabled medical devices has grown rapidly, with over 1,300 authorizations across specialties as of the latest update.[10, 12] Asian and European regulatory agencies have similarly




Published online August 4, 2026 https://doi.org/10.15326/jcopdf.2026.0794

convened focus groups, developed toolkits, and enhanced pharmacovigilance programs to facilitate patient safety, increase health care professional awareness, and minimize AI risks.[12]

In controlled comparisons, AI software has outperformed pulmonologists in pulmonary function test interpretation accuracy and diagnostic assessment, underscoring both the promise and need for careful guardrails in respiratory decision support.[13] Chat Generative Pre-Trained Transformer (ChatGPT, OpenAI, San Francisco, CA) is an LLM AI that has demonstrated the potential to answer medical questions, including questions regarding COPD.[14-16] In a study by Imtiaz and colleagues, ChatGPT responses demonstrated high degrees of accuracy, completeness, clarity, and safety when prompted to 21 COPD questions, including six items addressing COPD medications.[14] However, other studies have identified varying degrees of ChatGPT's accuracy responding to complex medication-related questions and noted low inter-response agreement when the same prompts are administered over time.[17, 18] These aforementioned studies discuss single responses at one point in time or several responses over a period of time; however, limited data exist exploring differences between responses when the same question is simultaneously inputted across different devices using the same LLM AI (e.g., ChatGPT).[14-16] Given the prevalence and clinical burdens of COPD, further research on the accuracy, utility, and reliability of ChatGPT responding to COPD questions has been suggested.[10, 19, 20] These recent studies show mixed performance of LLMs on COPD tasks: in a clinician assessment of 21 COPD questions, ChatGPT provided detailed – but at times outdated – advice compared with Bing, while pharmacy-focused evaluations demonstrated limited accuracy and reproducibility on drug information queries, directly motivating ongoing needs to assess simultaneous LLM outputs.




Published online August 4, 2026 https://doi.org/10.15326/jcopdf.2026.0794

The purpose of this study was to evaluate the accuracy, usefulness, and impact variability of ChatGPT-4.0 across three simultaneous responses across different computer devices and to describe reasons for differences in accuracy, usefulness, and impact variability performance.

## Methods

This project was a three round modified Delphi approach with the purpose of forming consensus on ChatGPT's accuracy, usefulness, and impact variability responding to questions on COPD treatment. Given the generative intent of this project, the Delphi method was used to derive consensus rather than use a nominal group technique.[21] This is because the Delphi method explicitly forces convergence and ensures final ratings represent collective clinical judgement. Preliminary results were presented as a poster at the 2025 American College of Clinical Pharmacy Annual Meeting.[22] Reporting adhered to the ACcurate COnsensus Reporting Document (ACCORD) guideline.[23] This study was exempt by the Institutional Review Board for Human Subjects at The University of Georgia. We used ChatGPT-4.0 (OpenAI, San Francisco, CA) on three separate computers with distinct user accounts. Sessions occurred on May 6, 2025 with browsing enabled. Prompts were entered as the questions (verbatim) and are provided in the Supplementary Appendix.

Five educational clinical questions were created by three residency-trained board-certified pharmacists whose primary employment was a faculty position at a college of pharmacy. The pharmacists graduated and completed two years of post-doctoral training (post graduate year-1 and -2 residency training) at different institutions, practiced in settings that differed in delivery of care (i.e., ambulatory care, inpatient), types of facilities (i.e., government, non-profit), patient




Published online August 4, 2026 https://doi.org/10.15326/jcopdf.2026.0794

population (i.e., geriatrics, adults), and had a range of practice experience between 9 to 15 years; therefore, although the pharmacist panelists had similar expertise, they did not all practice at the same practice site (i.e., three distinct locations in Georgia, USA and Oklahoma, USA). All three pharmacists teach COPD pharmacotherapeutics within the didactic and experiential curriculum at their colleges.

The pharmacists used COPD as the primary focus of the questions, addressing domains such as evidence-based management, patient counseling considerations, and contraindications. The questions differed in format, ranging from well-defined structured questions to more open-ended ill-structured scenarios (Supplementary Appendix). Well-structured questions typically provide key details explicitly and narrow the range of acceptable solutions, often mapping neatly onto established guidelines or a single best answer.[24] By comparison, ill-structured questions usually demand more nuanced clinical reasoning, requiring clinicians to synthesize patient-specific variables and weigh trade-offs among multiple defensible options to arrive at a prioritized, context-dependent decision. Each question was entered concurrently into three separate computers using ChatGPT-4.0 to generate responses (output). Notably, ChatGPT-4.0 is a subscription-based program that had internet access and knowledge cutoffs ending in 2023.[18, 25] All generated responses were compiled and evaluated by (the same) three pharmacists using a modified Delphi process, as depicted in Figure 1.

<Figure 1>




Published online August 4, 2026 https://doi.org/10.15326/jcopdf.2026.0794

Pharmacists independently assessed each response across three domains: accuracy, usefulness, and impact variability. Based on previous research on LLM, domains were separated on a 3-point Likert scale.[26, 27] Accuracy reflected the degree to which the response aligned with correct or accepted clinical information (e.g., published guidelines, randomized controlled trials) and was ordinal scale scored as 0 (poor), 1 (borderline), or 2 (good). Usefulness represented an overall judgment of whether the response met the intended goal of the prompt either from a clinician or patient perspective depending on the prompt and was ordinal scale scored as 0 (not useful), 1 (somewhat useful), or 2 (very useful). Impact variability (sometimes referred to as either consistency or reliability) captured the extent to which differences in responses could meaningfully influence patient outcomes (e.g., clinical targets) and was ordinal scale scored as 0 (low), 1 (moderate), or 2 (high).[26, 27] For impact variability, panelists judged whether differences among the three simultaneous responses could change medication management (e.g., inhaled therapy class, corticosteroid exposure, antibiotic use) or key clinical targets (e.g., exacerbation risk reduction). Operational definitions and examples for each level were pre-distributed to panelists (Supplementary Appendix). Individual ratings were subsequently aggregated, anonymized, and redistributed to the pharmacists to allow comparison with group responses and the opportunity to make revisions or share additional commentary with the group. The study moderator summarized feedback and convened one virtual consensus meeting with the three pharmacists to address items with discordant ratings. Consensus was defined as unanimous agreement among all three pharmacists for each response and domain (i.e., accuracy, usefulness, impact variability) and achieved through discussion and agreement. Discordant ratings were reviewed item-by-item and rating changes were not decided by a majority vote. In some instances, discussion led panelists who were initially in agreement to revise toward the




Published online August 4, 2026 https://doi.org/10.15326/jcopdf.2026.0794

discordant rating. Prior to the consensus meeting, Fleiss' kappa was calculated for each domain across Delphi rounds 1 and 2 with 95% confidence intervals via bootstrapping to understand the reliability of each assessment[28, 29]; we also summarized domain-specific kappa by question. Descriptive statistics were used to summarize the findings. Data analysis was performed in SAS version 9.4 (SAS Institute, Cary, NC).

**Results**

Fleiss' kappa for Delphi rounds 1 and 2 are presented in Table 1.[30] Interrater reliability among the three panelists evaluating ChatGPT-4.0's accuracy increased from slight agreement to fair agreement proceeding from Delphi round 1 to Delphi round 2. Poor agreement was observed regarding ChatGPT-4.0's usefulness and impact variability in both Delphi rounds 1 and 2.

<Table 1>

Among five COPD-related questions simultaneously administered across three instances of ChatGPT-4.0, consensus among three board certified pharmacists was achieved on 40 out of 45 responses (88.9% consensus agreement). Consensus was completely achieved regarding ChatGPT's accuracy though not achieved regarding the usefulness of one question and the impact variability of another question. The accuracy, usefulness, and impact variability across questions and responses are presented in Table 2.

<Table 2>




Published online August 4, 2026 https://doi.org/10.15326/jcopdf.2026.0794

ChatGPT-4.0 accuracy ranged from poor to good across the five questions. The consensus panelists agreed that ChatGPT-4.0 was borderline accurate on staging COPD, recommending initial inhaled medications, and assessing and planning a COPD medication treatment plan for a complex patient case. The consensus panel also felt that ChatGPT was borderline accurate responding to a question regarding treatment of an exacerbation of COPD (ECOPD) though one response was scored as good (versus others scored as borderline). The consensus panel agreed that ChatGPT-4.0 demonstrated poor accuracy responding to a question on selecting and recommending medications to treat stable COPD because each ChatGPT-4.0 response referenced an outdated clinical practice guideline.

ChatGPT-4.0 usefulness ranged between somewhat useful to very useful; the consensus panel did not score ChatGPT-4.0 as not useful across any responses (i.e., ChatGPT-4.0 was at least somewhat useful). The consensus panel felt that ChatGPT-4.0 was somewhat useful staging COPD and initiating medications and selecting medications to treat stable COPD. Regarding treating ECOPD and assessing and planning a complex case, the consensus panel scored ChatGPT-4.0 as somewhat useful though one response (in each question) was scored as very useful. Consensus was not achieved among the panelists regarding ChatGPT's usefulness to a question asking to identify medication-related adverse events; herein, the consensus panel scored ChatGPT-4.0 somewhat to very useful.

ChatGPT-4.0 impact variability ranged between low to high. The consensus panel felt that ChatGPT-4.0 displayed high impact variability staging COPD and initiating medication and displayed moderate impact variability treating ECOPD. Regarding medication-related adverse




Published online August 4, 2026 https://doi.org/10.15326/jcopdf.2026.0794

events and complex case assessment and planning, the consensus panel scored ChatGPT-4.0 as moderate to high and low to moderate impact variability, respectively. Consensus was not achieved among the panelists regarding ChatGPT's impact variability to a question regarding selection of medications to treat stable COPD; notably, this was the same and only question the panelists agreed ChatGPT-4.0 displayed poor accuracy because it referenced an outdated clinical practice guideline.

## Discussion

This appears to be one of the first studies evaluating simultaneous responses to clinical questions across three devices using the same LLM AI (i.e., ChatGPT-4.0). Overall, some differences existed across accuracy, usefulness, and impact variability. Using a three-round modified Delphi approach among residency-trained board-certified pharmacist faculty, consensus was achieved on 89% of ChatGPT outputs following COPD prompts. The consensus panel of pharmacists agreed ChatGPT-4.0 accuracy ranged from poor to good, usefulness was at least somewhat useful and ranged between somewhat useful to very useful, and impact variability ranged between low and high. Though Fleiss' kappa measures of interrater reliability ranged from poor-to-fair during Delphi rounds 1 and 2, the pattern of improving agreement and the formation of consensus over time was expected given the use of a modified Delphi approach. Despite having an answer key and instructions, initial variability may reflect differences in the pharmacist panelists' clinical interpretation and reasoning. However, subsequent rounds allowed the panelists to review their colleagues' justification, promoting reflection and possible recalibration of their initial scores. The final consensus meeting provided an opportunity for structured discussion through the moderator to facilitate alignment. This is inherent to Delphi methods




Published online August 4, 2026 https://doi.org/10.15326/jcopdf.2026.0794

which typically results in improved inter-rater agreement over time.[31] Based on the literature, presenting these data are needed, since some studies propose that Delphi studies lack defining interrater reliability and should be included when possible.[32, 33]

Though AI offers the promise of optimizing clinicians' workflow and addressing patients' clinical questions, concerns about AI's static knowledge base have been raised, particularly concerning medications.[11, 17, 18] A study by Imtiaz and colleagues compared two LLM AIs, ChatGPT-3.5 and Bing, and scored ChatGPT-3.5 higher than Bing across accuracy, completeness, clarity, and safety domains among 21 COPD prompts.[14] Though question-level data were not provided, the authors noted that ChatGPT-3.5 referenced outdated clinical practice guidelines and provided incomplete responses with respect to inhaled medications including corticosteroids, long-acting beta agonists, and long-acting muscarinic antagonists. Results from our study identified similar accuracy concerns, with one response displaying poor accuracy concerning medication treatment for stable COPD because ChatGPT-4.0 referenced a retired copy of the Global Initiative for Obstructive Lung Disease (GOLD) Report. These findings are consequential to health care providers and patients alike because international evidence-based guidelines for asthma and COPD, the Global Initiative for Asthma (GINA) and GOLD Reports, are updated annually.[34, 35] ChatGPT's and other chatbots including Bing and Claude inability to capture, assess, and respond to contemporary and dynamic updates in pulmonary medicine are consequential and may lead to patients being either undertreated or overtreated with inhaled medications. Patients who are undertreated are at increased risk for life-threatening ECOPD whereas patients who are overtreated are at increased risk for preventable adverse drug events. A plausible remedy to this problem includes routinely incorporating (uploading) clinical practice




Published online August 4, 2026 https://doi.org/10.15326/jcopdf.2026.0794

guidelines and landmark clinical trials into LLMs in real-time.[14] Future studies should evaluate if prompting, such as asking the model to use only guidelines published within a specified time-bound window, reduces outdated recommendations.[36]

Previous reports have criticized ChatGPT's ability to assess complex clinical cases.[17, 25] In our project, ChatGPT's responses to a complex COPD case demonstrated borderline accuracy, somewhat usefulness, and moderate impact variability. The 2026 GOLD Report strengthened its emphasis on avoiding unnecessary inhaled corticosteroid exposure (i.e., favoring dual inhaled long-acting bronchodilators unless there is clear eosinophilia or exacerbation-prone phenotype present) and clarified pharmacotherapy escalation after one ECOPD.[35] Despite moving away from the recommendation of using a combination of a long-acting beta agonist with an inhaled corticosteroid, without a long-acting muscarinic antagonist several years ago, clinicians continue to prescribe this combination inappropriately. A study conducted in Australia found that more than 50% of patients included were inappropriately prescribed an inhaled corticosteroid.[37] A cross-sectional study conducted within the United States Department of Veteran's Affairs, found that 23.9% of veterans were prescribed an inhaled corticosteroid without an identifiable indication.[38] Even with guideline guidance on appropriate inhaled corticosteroid use and algorithms for transitioning patients from long-acting beta-agonist/inhaled corticosteroid combination therapy to other more effective regimens, these errors persist.

Inappropriate COPD treatment extends beyond inhaled corticosteroids, too. A study conducted in the United States found that 45.3% of patients did not receive general guideline recommended inhaler treatment, further highlighting inappropriate use of medications in COPD.[39] Another




Published online August 4, 2026 https://doi.org/10.15326/jcopdf.2026.0794

facet to COPD management extends beyond stable patients to those presenting with exacerbations. A study of intensivists in France found that adherence to guidelines regarding the use of laboratory biomarkers and short-acting beta agonist medications were moderate and antibiotic prescribing was poor.[40]

Variability in acute COPD management is well-documented, with international guidelines from the European Respiratory Society and American Thoracic Society highlighting strong recommendations and conditional areas precisely where LLM outputs must be cross-checked against current evidence-based guidance.[41] Our data, when combined with existing literature and clinical practice guidelines, shows that COPD management, specifically with inhaled corticosteroids, is complex and requires individualized patient assessment and clinical judgement. Because of this, LLM-generated recommendations with variable accuracy and usefulness can have a large impact on patient care. Prior to implementing these models for the management of COPD, the model should be trained by field experts, validated with updated clinical recommendations, and integrated with adequate safeguards to mitigate potentiation of medication errors. While this may be ideal, many clinicians using LLMs have not received formal training in prompt design, which may limit their ability to elicit accurate, clinically relevant, and appropriately tailored responses.[42, 43] Therefore, our results may reflect the types of responses likely to occur during real-world use, where prompts are often developed without specialized prompt-engineering expertise. Collectively, results from our work and existing projects illustrate how either outdated or ambiguous advice, whether human or LLM-generated, can propagate COPD over- and undertreatment risks.




Published online August 4, 2026 https://doi.org/10.15326/jcopdf.2026.0794

This project possesses limitations warranting discussion. Our data, when combined with existing literature, shows COPD management, specifically with inhaled corticosteroids, is complex and requires individualized patient assessment and clinical judgement. Clear guidance on Delphi panel composition is lacking; though only three board certified pharmacists holding faculty appointments served as experts, increasing the panel size to include between five and fifteen participants may have added robustness.[44] The small panel size also limited how much blinding could be preserved across multiple rounds. To mitigate this risk, discordant ratings were reviewed against operational definitions, and consensus was not treated as an issue of majority vote. Careful, moderated, item-level discussion allowed panelists to reconsider discordant ratings as systematically as possible though panel convergence towards consensus may have been susceptible to some degree of the bandwagon effect.[45, 46] The use of a 3-point ordinal scale (0, 1, 2) to rate accuracy, usefulness, and impact variability may have centered consensus panelists towards the median response (1); rather, a 4-point ordinal scale (1, 2, 3, 4) could have been considered.[47] At the time of this project, ChatGPT-4.0, only available via a paid subscription, was utilized; patients and providers using open-access (free) ChatGPT-3.5 may obtain different outputs, which have been shown to be less accurate than subscription-based versions.[18] Finally, we did not experimentally control computer temperature, a factor that plausibly influences inter-response variability and merits formal study.

## Conclusions

In this modified Delphi evaluation of simultaneous ChatGPT-4.0 responses to COPD medication management questions, variability was observed in accuracy and usefulness across responses generated at the same time using identical questions, and the differences identified may possess




Published online August 4, 2026 https://doi.org/10.15326/jcopdf.2026.0794

clinical implications. Although most responses were at least somewhat useful, the presence of outdated guideline references and differences in impact variability highlight important limitations for clinical applications. These findings suggest, at this time, LLMs should be considered supportive informational tools rather than stand-alone clinical decision-making resources. Given the evolving nature of evidence-based resources, careful clinician oversight remains essential to ensure safe, contemporary, and patient-centered care. Continued evaluation of AI performance, integration of updated clinical standards, and development of responsible implementation frameworks will be critical as AI is incorporated into healthcare delivery. Safe clinical use of LLMs in COPD care requires transparent grounding in current GOLD Report guidance, local oversight, and deployment patterns minimizing risks of outdated or inconsistent recommendations.




Published online August 4, 2026 https://doi.org/10.15326/jcopdf.2026.0794

Ethical approval and informed consent statement: This study was reviewed and determined to be exempt from full review by the Institutional Review Board for Human Subjects at The University of Georgia (Athens, GA, USA) in accordance with 45 CFR §46.104.

Data availability statement: The data and analytic code underlying this study are available from the corresponding author upon reasonable request.

Declaration of interest: The authors declare no conflicts of interest.




Published online August 4, 2026 https://doi.org/10.15326/jcopdf.2026.0794

**References**

1. World Health Organization. Chronic obstructive pulmonary disease (COPD), (2023, accessed Feb 10 2026).

2. Adeloye D, Song P, Zhu Y, et al. Global, regional, and national prevalence of, and risk factors for, chronic obstructive pulmonary disease (COPD) in 2019: a systematic review and modelling analysis. *Lancet Respir Med* 2022; 10: 447–458. 20220310. DOI: 10.1016/S2213-2600(21)00511-7.

3. Buist AS, McBurnie MA, Vollmer WM, et al. International variation in the prevalence of COPD (the BOLD Study): a population-based prevalence study. *Lancet* 2007; 370: 741–750. DOI: 10.1016/S0140-6736(07)61377-4.

4. Ruvuna L and Sood A. Epidemiology of Chronic Obstructive Pulmonary Disease. *Clin Chest Med* 2020; 41: 315–327. DOI: 10.1016/j.ccm.2020.05.002.

5. Labaki WW and Rosenberg SR. Chronic Obstructive Pulmonary Disease. *Ann Intern Med* 2020; 173: ITC17–ITC32. DOI: 10.7326/AITC202008040.

6. Raherison C and Girodet PO. Epidemiology of COPD. *Eur Respir Rev* 2009; 18: 213–221. DOI: 10.1183/09059180.00003609.

7. Safiri S, Carson-Chahhoud K, Noori M, et al. Burden of chronic obstructive pulmonary disease and its attributable risk factors in 204 countries and territories, 1990-2019: results from the Global Burden of Disease Study 2019. *BMJ* 2022; 378: e069679. 20220727. DOI: 10.1136/bmj-2021-069679.

8. Boers E, Barrett M, Su JG, et al. Global Burden of Chronic Obstructive Pulmonary Disease Through 2050. *JAMA Netw Open* 2023; 6: e2346598. 20231201. DOI: 10.1001/jamanetworkopen.2023.46598.




Published online August 4, 2026 https://doi.org/10.15326/jcopdf.2026.0794

9. Shatto JA, Stickland MK and Soril LJJ. Variations in COPD Health Care Access and Outcomes: A Rapid Review. *Chronic Obstr Pulm Dis* 2024; 11: 229–246. DOI: 10.15326/jcopdf.2023.0441.

10. Kaplan A, Cao H, FitzGerald JM, et al. Artificial Intelligence/Machine Learning in Respiratory Medicine and Potential Role in Asthma and COPD Diagnosis. *J Allergy Clin Immunol Pract* 2021; 9: 2255–2261. 20210219. DOI: 10.1016/j.jaip.2021.02.014.

11. Wong A, Flanagan T, Covington EW, et al. Forecasting the impact of artificial intelligence on clinical pharmacy practice. *J Am Coll Clin Pharm* 2025; 8: 302–310. DOI: 10.1002/jac5.70004.

12. Abdalla M, Saad M, Abazia D, et al. Responsible Use of Artificial Intelligence in Health Care: Evidence, Challenges, and Best Practices: An Opinion of the Drug Information Practice and Research Network of the American College of Clinical Pharmacy. *J Am Coll Clin Pharm* 2025; 8: 1333–1361. DOI: 10.1002/jac5.70131.

13. Topalovic M, Das N, Burgel PR, et al. Artificial intelligence outperforms pulmonologists in the interpretation of pulmonary function tests. *Eur Respir J* 2019; 53 20190411. DOI: 10.1183/13993003.01660-2018.

14. Imtiaz A, King J, Holmes S, et al. ChatGPT versus Bing: a clinician assessment of the accuracy of AI platforms when responding to COPD questions. *Eur Respir J* 2024; 63 20240620. DOI: 10.1183/13993003.00163-2024.

15. Merc P, Pirincci CS and Cihan E. Evaluation of AI chatbots for patient education and information on chronic obstructive pulmonary disease. *Heart Lung* 2026; 75: 21–25. 20250911. DOI: 10.1016/j.hrtlng.2025.09.002.




Published online August 4, 2026 https://doi.org/10.15326/jcopdf.2026.0794

16. Yin Y, Riaz Z, Amoro Sanchez R, et al. Evaluating ChatGPT as a Standalone Tool for Patient Education: A Review of Frequently Asked Questions by Patients With Chronic Obstructive Pulmonary Disease. *Cureus* 2025; 17: e92519. 20250917. DOI: 10.7759/cureus.92519.

17. Tran T, Le UM and Phan V. Assessing ChatGPT's Response Accuracy in Pharmacy Education: Insights into Cognitive Complexity. *Am J Pharm Educ* 2024; 88: 51. DOI: 10.1016/j.ajpe.2024.100943.

18. Khatri S, Sengul A, Moon J, et al. Accuracy and reproducibility of ChatGPT responses to real-world drug information questions. *J Am Coll Clin Pharm* 2025; 8: 432–438. DOI: 10.1002/jac5.70038.

19. Antao J, de Mast J, Marques A, et al. Demystification of artificial intelligence for respiratory clinicians managing patients with obstructive lung diseases. *Expert Rev Respir Med* 2023; 17: 1207–1219. 20240125. DOI: 10.1080/17476348.2024.2302940.

20. Mekov E, Miravitlles M and Petkov R. Artificial intelligence and machine learning in respiratory medicine. *Expert Rev Respir Med* 2020; 14: 559–564. 20200317. DOI: 10.1080/17476348.2020.1743181.

21. Caudle KE, Burlison JD, Bourque MS, et al. Research and scholarly methods: Consensus techniques. *J Am Coll Clin Pharm* 2025; 8: 610–618. DOI: 10.1002/jac5.70052.

22. Caballero J, Lavender DL, Stone RH, et al. Assessing ChatGPT's Accuracy and Usefulness in Medication Management. *J Am Coll Clin Pharm* 2025; 8: 1447. DOI: 10.1002/jac5.70148.




Published online August 4, 2026 https://doi.org/10.15326/jcopdf.2026.0794

23. Gattrell WT, Logullo P, van Zuuren EJ, et al. ACCORD (ACcurate COnsensus Reporting Document): A reporting guideline for consensus methods in biomedicine developed via a modified Delphi. *PLoS Med* 2024; 21: e1004326. DOI: 10.1371/journal.pmed.1004326.

24. Jonassen DH. Learning to Solve Problems: A Handbook for Designing Problem-Solving Learning Environments. New York: Routledge, 2011.

25. Tran BAC, Wantuch GA, Mangasar M, et al. Battle of the (Chat)Bots: Comparative Evaluation of April 2025 AI Model Accuracy on Pharmacotherapeutic Cases. *Am J Pharm Educ* 2026; 90: 101922. 20251212. DOI: 10.1016/j.ajpe.2025.101922.

26. Ramasubramanian S, Balaji S, Kannan T, et al. Comparative evaluation of artificial intelligence systems' accuracy in providing medical drug dosages: A methodological study. *World J Methodol* 2024; 14: 92802. DOI: 10.5662/wjm.v14.i4.92802.

27. Varady NH, Lu AZ, Mazzucco M, et al. Understanding How ChatGPT May Become a Clinical Administrative Tool Through an Investigation on the Ability to Answer Common Patient Questions Concerning Ulnar Collateral Ligament Injuries. *Orthop J Sports Med* 2024; 12: 23259671241257516. DOI: 10.1177/23259671241257516.

28. Fleiss JL and Cohen J. The Equivalence of Weighted Kappa and the Intraclass Correlation Coefficient as Measures of Reliability. *Educ Psychol Meas* 1973; 33: 613–619. DOI: 10.1177/001316447303300309.

29. Haukoos JS and Lewis RJ. Advanced statistics: bootstrapping confidence intervals for statistics with "difficult" distributions. *Acad Emerg Med* 2005; 12: 360–365. DOI: 10.1197/j.aem.2004.11.018.

30. Landis JR and Koch GG. The Measurement of Observer Agreement for Categorical Data. *Biometrics* 1977; 33. DOI: 10.2307/2529310.




Published online August 4, 2026 https://doi.org/10.15326/jcopdf.2026.0794

31. Hsu C-C and Sandford BA. The Delphi Technique: Making Sense of Consensus. *Pract Assess Res Eval* 2007; 12. DOI: 10.7275/pdz9-th90

32. Niederberger M and Spranger J. Delphi Technique in Health Sciences: A Map. *Front Public Health* 2020; 8: 457. DOI: 10.3389/fpubh.2020.00457.

33. Spranger J, Homberg A, Sonnberger M, et al. Reporting guidelines for Delphi techniques in health sciences: A methodological review. *Z Evid Fortbild Qual Gesundhwes* 2022; 172: 1–11. DOI: 10.1016/j.zefq.2022.04.025.

34. Global Initiative for Asthma. *Global Strategy for Asthma Management and Prevention, 2025 Report*. May 2025.

35. Global Initiative for Obstructive Lung Disease. *Global Strategy for the Diagnosis, Management, and Prevention of Chronic Obstructive Pulmonary Disease (2026 Report)*. Nov 2025.

36. Meskó B. Prompt Engineering as an Important Emerging Skill for Medical Professionals: Tutorial. *J Med Internet Res* 2023; 25: e50638. DOI: 10.2196/5063.

37. Harrison A, Borg B, Thompson B, et al. Inappropriate inhaled corticosteroid prescribing in chronic obstructive pulmonary disease patients. *Intern Med J* 2017; 47: 1310–1313. DOI: 10.1111/imj.13611.

38. Griffith MF, Feemster LC, Zeliadt SB, et al. Overuse and Misuse of Inhaled Corticosteroids Among Veterans with COPD: a Cross-sectional Study Evaluating Targets for De-implementation. *J Gen Intern Med* 2020; 35: 679–686. DOI: 10.1007/s11606-019-05461-1.

39. Sharif R, Cuevas CR, Wang Y, et al. Guideline adherence in management of stable chronic obstructive pulmonary disease. *Respir Med* 2013; 107: 1046–1052. DOI: 10.1016/j.rmed.2013.04.001.




Published online August 4, 2026 https://doi.org/10.15326/jcopdf.2026.0794

40. Haudebourg L, Faure M, Dres M, et al. Management of severe exacerbations of COPD by French intensivists and adherence to guidelines. *Respir Med Res* 2025; 87: 101159. DOI: 10.1016/j.resmer.2025.101159.

41. Wedzicha JAEC-C, Miravitlles M, Hurst JR, et al. Management of COPD exacerbations: a European Respiratory Society/American Thoracic Society guideline. *Eur Respir J* 2017; 49 DOI: 10.1183/13993003.00791-2016.

42. Liu J, Liu F, Wang C, et al. Prompt Engineering in Clinical Practice: Tutorial for Clinicians. *J Med Internet Res* 2025; 27: e72644. DOI: 10.2196/72644.

43. Shah K, Xu AY, Sharma Y, et al. Large Language Model Prompting Techniques for Advancement in Clinical Medicine. *J Clin Med* 2024; 13. DOI: 10.3390/jcm13175101.

44. Keeney S, Hasson F and McKenna H. The Delphi Technique in Nursing and Health Research. Ames, Iowa, USA: Wiley-Blackwell, 2011.

45. O'Connor N and Clark S. Beware bandwagons! The bandwagon phenomenon in medicine, psychiatry and management. *Australas Psychiatry* 2019; 27: 603–606. DOI: 10.1177/1039856219848829.

46. Rikkers LF. The bandwagon effect. *J Gastrointest Surg* 2002; 6: 787–794. DOI: 10.1016/s1091-255x(02)00054-9.

47. Weinfurt KP and Reeve BB. Patient-Reported Outcome Measures in Clinical Research. *JAMA* 2022; 328: 472–473. DOI: 10.1001/jama.2022.11238.




Published online August 4, 2026 https://doi.org/10.15326/jcopdf.2026.0794

**Table 1.** Fleiss' Kappa by Delphi Round, Question, and Domain, with Bootstrapped 95% Confidence Intervals for Delphi Rounds 1 and 2.

| Delphi Round | Question | Accuracy κ (95% CI) | Usefulness κ (95% CI) | Impact Variability κ (95% CI) |
|---|---|---|---|---|
| 1 | Overall | 0.22 (-0.08 to 0.50) | -0.02 (-0.26 to 0.17) | -0.02 (-0.27 to 0.16) |
| | 1 | 0.10 | -0.50 | -0.13 |
| | 2 | 1.00 | -0.29 | -0.29 |
| | 3 | -0.50 | -0.35 | -0.38 |
| | 4 | 0.00 | 0.36 | -0.35 |
| | 5 | -0.35 | 0.00 | -0.04 |
| 2 | Overall | 0.43 (0.10 to 0.67) | -0.02 (-0.26 to 0.20) | -0.02 (-0.26 to 0.16) |
| | 1 | 0.10 | -0.50 | -0.13 |
| | 2 | 1.00 | -0.29 | -0.29 |
| | 3 | -0.50 | -0.35 | -0.38 |
| | 4 | 0.36 | 0.36 | -0.35 |
| | 5 | -0.13 | 0.00 | -0.04 |

*Note.* Fleiss' kappa was calculated across all responses within each Delphi round (N=15 responses per round, 3 panelists, 3-point ordinal scale). Each question-level kappa was calculated from 3 ChatGPT responses rated by 3 panelists on a 3-point ordinal scale. Due to the small number of subjects per question (N = 3), these estimates should be interpreted with caution.




Published online August 4, 2026 https://doi.org/10.15326/jcopdf.2026.0794

**Table 2.** Accuracy, usefulness, and impact variability of ChatGPT simultaneous responses to COPD questions.

| **Question, Structure, and Topic** | **Accuracy**[*] | **Usefulness**[†] | **Impact variability**[‡] |
|---|---|---|---|
| 1; ill-structured;<br>GOLD staging and initial medication treatment | Response 1: Borderline<br>Response 2: Borderline<br>Response 3: Borderline | Response 1: Somewhat<br>Response 2: Somewhat<br>Response 3: Somewhat | Response 1: High<br>Response 2: High<br>Response 3: High |
| 2; structured;<br>Medication-related adverse events | Response 1: Good<br>Response 2: Good<br>Response 3: Good | Response 1: 2<br>Response 2: Somewhat-to-Very[§]<br>Response 3: Somewhat-to-Very[§] | Response 1: High<br>Response 2: Moderate<br>Response 3: Moderate |
| 3; ill-structured;<br>Stable COPD medication treatment | Response 1: Poor<br>Response 2: Poor<br>Response 3: Poor | Response 1: Somewhat<br>Response 2: Somewhat<br>Response 3: Somewhat | Response 1: Moderate[§]<br>Response 2: Low-to-Moderate[§]<br>Response 3: Moderate[§] |
| 4; ill-structured;<br>ECOPD medication treatment | Response 1: Borderline<br>Response 2: Good<br>Response 3: Borderline | Response 1: Somewhat<br>Response 2: Very<br>Response 3: Somewhat | Response 1: Moderate<br>Response 2: Moderate<br>Response 3: Moderate |
| 5; structured;<br>Complex COPD case requiring assessment and plan | Response 1: Borderline<br>Response 2: Borderline<br>Response 3: Borderline | Response 1: Somewhat<br>Response 2: Somewhat<br>Response 3: Very | Response 1: Low<br>Response 2: Moderate<br>Response 3: Low |

Abbreviations: COPD = chronic obstructive pulmonary disease; ECOPD = chronic obstructive pulmonary disease exacerbation; GOLD = Global Initiative for Obstructive Lung Disease

[*]Accuracy score scale: 0 (poor), 1 (borderline), 2 (good)

[†] Usefulness score scale: 0 (not useful), 1 (somewhat useful), 2 (very useful)

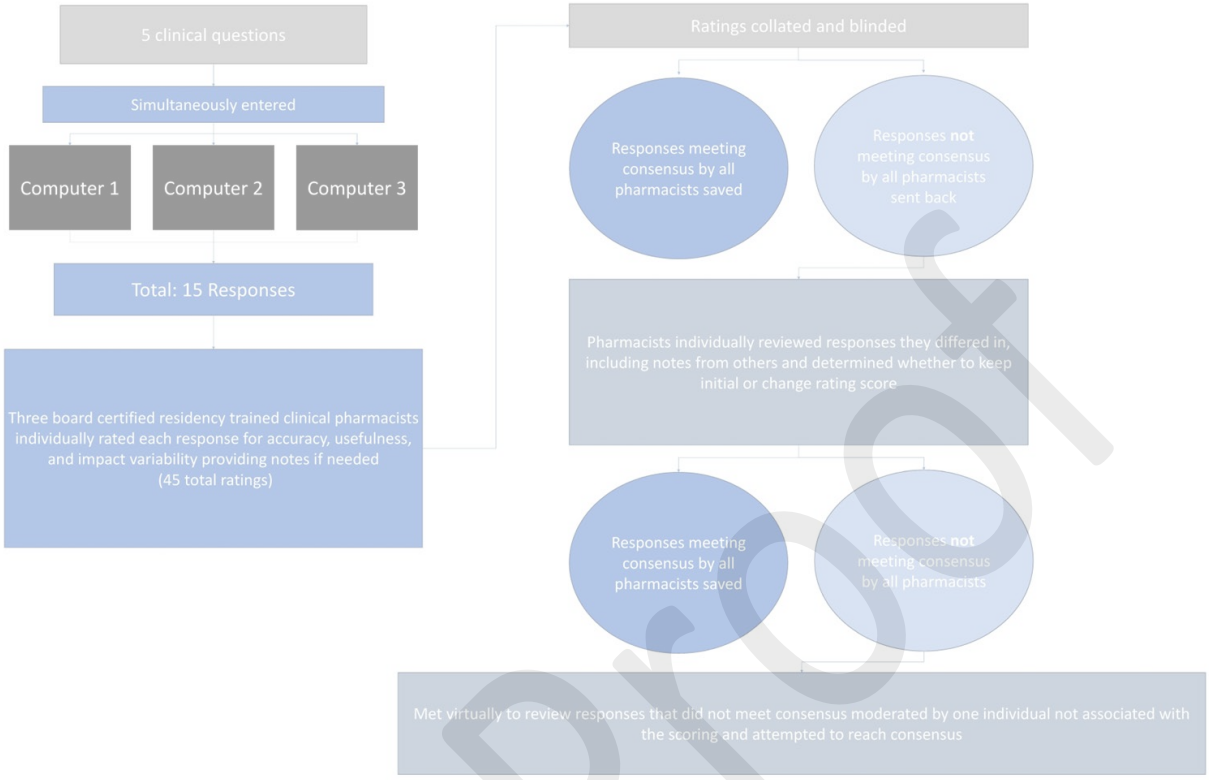
[‡] Impact variability score scale: 0 (low), 1 (moderate), 2 (high)

[§]Did not reach consensus




Published online August 4, 2026 https://doi.org/10.15326/jcopdf.2026.0794

Figure 1







Published online August 4, 2026 https://doi.org/10.15326/jcopdf.2026.0794

Online Supplement

# Appendix

**Appendix A: Questions and rationale**

1. JH is a 62-year-old Hispanic female with a past medical history significant for Hypertension, Tobacco use, Hyperlipidemia, Diabetes, and Gout. Their current medication list includes Lisinopril 10mg daily, HCTZ 25 mg daily, Atorvastatin 80mg daily, Metformin ER 1000mg BID, Empagliflozin 25mg daily, and Allopurinol 300mg daily. They recently have been diagnosed with COPD and completed their initial spirometry (see below). Their primary care provider has consulted you to assist with managing their COPD.

Spirometry:

| | Prebronchodilator | | | Postbronchodilator | |
|---|---|---|---|---|---|
| | Predicted | Measured | % Predicted | Measured | % Predicted |
| FVC | 4.58 | 3.12 | 68 | 3.23 | 70 |
| $FEV_1$ | 3.02 | 1.42 | 47 | 1.45 | 48 |
| $FEV_1$/FVC | | 0.46 | | 0.45 | |

   a. If JH's mMRC score is 3, CAT score is 22, and has a history of 1 exacerbation in the last year (did ***NOT*** required hospitalization) what COPD classification and initial group is this patient according to the 2024 GOLD Guidelines?

Rationale: According to the 2024 GOLD COPD Guidelines, this patient would be classified as GOLD 3/Severe based on their post-bronchodilator $FEV_1$ of 48% ($FEV_1$ >30% but < 50% predicted). He would be in Group B based on their symptoms (mMRC ≥ 1, CAT ≥ 10) and 1 moderate exacerbations.

   b. Based on JH's initial GOLD Assessment group, what medication regimen would be the most appropriate to treat their COPD?

Rationale: According to the 2024 GOLD COPD guidelines, all patients should be initiated on a short acting rescue inhaler for immediate symptoms relief. Based on the patient's ABE group of B they are indicated for LAMA + LABA, such as Albuterol + Olodaterol/Tiotropium (Stiolto). LAMA and SAMA which is not appropriate due to increase risk of anticholinergic side effects.




Published online August 4, 2026 https://doi.org/10.15326/jcopdf.2026.0794

2. MN contacts the community pharmacy and asks about their inhalers and a potential adverse effect. MN states that they started a new inhaler about 1 month ago and have developed dry mouth and dry eyes. They are currently maintained on Albuterol + Umeclidenium (Incruse Ellipta) + Olodaterol (Striverdi). Are these side effects caused by their inhalers? If so, which one is the most likely cause?

Rationale: Umeclidenium is a long-acting muscarinic antagonist with potential to have anticholinergic side effects (Anti-SLUD). These side effects of dry mouth and dry eyes are classic examples of anticholinergic side effects.

3. Which medication is the best recommendation to treat a patient living with GOLD Category 3B COPD?

   Rationale: Umeclidinium/vilanterol

4. A patient living with COPD presents to urgent care with worse-than-usual shortness of breath. The patient denies increased sputum production. However, they are able to produce a sputum sample in urgent care and it is translucent. Which pharmacotherapy plan would be most appropriate for this patient and which key counseling points would you provide?

   Rationale: Albuterol and ipratropium

5.

**History of Present Illness:**

CH is a 59 year old white male who was consulted to the pharmacotherapy clinic by his primary care physician (PCP) for the assessment of his current COPD therapy. Today, he complains of "trouble breathing and being tired all of the time." He reports getting short of breath with activity and that he has difficulty keeping up with his wife when they are shopping. He is always concerned with finding places to sit down and rest while away from home. He was first diagnosed with COPD six months ago and was discharged after management of an exacerbation 2 weeks ago. He denies history of any other exacerbations. He denies muscle aches, chest pain, fever, or night sweats. Upon walking from the waiting room to the exam room, the patient was so short of breath that he had to rest. When taking the CAT Assessment, his score is 22.




Published online August 4, 2026 https://doi.org/10.15326/jcopdf.2026.0794

| |
|---|
| He currently takes a Combivent Respimat inhaler for his COPD. The prescription is for 1 puff four times daily as needed. However, he admits that he doesn't "always use my inhaler" as intended.<br><br>When asked to demonstrate inhaler technique, the patient placed the inhaler in his mouth actuated the inhalation while exhaling, inhaled quickly and failed to hold breath. The patient then quickly inhaled a second time and held breath for 10 seconds before exhaling. |
| **Past Medical History**: Diabetes mellitus type 2, Diabetic neuropathy, Dyslipidemia, Hypertension, Chronic Obstructive Pulmonary Disease |
| **Social History**: Tobacco: 2 ppd x 35 years; Quit 9 years ago<br>Alcohol: 1 drink ~ 3x/week<br>Denies any illicit drug use |

**Allergies: NKDA**

**Current Medications:**

| Rx # | Rx Name and Strength | Quantity | Refill History | Directions | Prescriber |
|---|---|---|---|---|---|
| 245876 | Albuterol 100mcg/ Ipratropium 20mcg | 30 day Supply | 2 mo ago<br>6 mo ago | Inhale one puff by mouth four times daily prn | Smith |
| 245870 | Gabapentin 300mg | 30 day supply | 1 mo ago<br>2 mo ago<br>3 mo ago | Take one (1) capsule by mouth three times daily for nerve pain | Smith |
| 245872 | Glipizide 5mg | 30 day supply | 1 mo ago<br>2 mo ago<br>3 mo ago | Take one (1) tablet by mouth twice daily with food for diabetes | Smith |
| 245871 | Insulin NPH 100 U/mL | 30 day supply | 1 mo ago<br>2 mo ago<br>3 mo ago | Administer eight (8) units subcutaneously twice daily before meals for diabetes | Smith |
| 245873 | Metformin 500mg | 30 day supply | 1 mo ago<br>2 mo ago<br>3 mo ago | Take two (2) tablets by mouth twice daily for diabetes | Smith |




Published online August 4, 2026 https://doi.org/10.15326/jcopdf.2026.0794

| | | | | | |
|---|---|---|---|---|---|
| 245874 | Simvastatin 80mg | 30 day supply | 1 mo ago<br>2 mo ago<br>3 mo ago | Take one-half (1/2) tablet by mouth every evening for cholesterol | Smith |
| 245941 | Lisinopril 5 mg | 30 day supply | 1 mo ago<br>2 mo ago<br>3 mo ago | Take one-half (1/2) tablet by mouth daily for blood pressure | Smith |
| OTC | Aspirin 81mg | 100 | | Take one tablet daily | |

**Physical Exam**

Vitals: Today BP 137/68 mmHg left arm sitting; HR 87 bpm; RR 22; O2 Saturation 95%

Ht: 69” TBW: 173lbs BMI: 25.6

**(from physician visit last month)**

General: Appears in respiratory and physical distress upon minimal exertion

HEENT (Head, eyes, ears, nose and throat): PERRLA (pupils equal round and reactive to light and accommodation), EOMI (extra ocular movements intacts), TM (tempanic membranes) intact

Neck: (-) JVD (jugular venous distention)

Cardiovascular: RRR (regular rate and rhythm); normal S1 and S2 sounds, no S3 or S4

Lungs: Diffuse bilateral inspiratory and expiratory wheezing, (-) rales and rhonchi

Gastrointestinal: Normal bowel sounds and movement

Abdomen: Soft, NT/ND (non-tender/non-distended)

Genitourinary: Unremarkable

Skin: Warm, dry

Neurological: Alert & Oriented x 3; CN (cranial nerves) II-XII intact

| **Spirometry** | | | | | **Lung Volumes** | | | |
|---|---|---|---|---|---|---|---|---|
| **6 months ago** | Pred | Actual | % Pred | | | Pred | Actual | % Pred |
| FVC (L) | 4.66 | 3.17 | 68 | | SVC (L) | 4.66 | 3.08 | 66 |
| FEV1 (L) | 3.55 | 1.53 | 43 | | TLC (L) | 6.79 | 7.66 | 113 |
| FEV1/FVC (%) | 76 | 48 | 64 | | RV (L) | 2.15 | 4.57 | 213 |




Published online August 4, 2026 https://doi.org/10.15326/jcopdf.2026.0794

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| FEV3/FVC (%) | | 72 | | | RV/TLC (%) | 32 | 60 | 187 |
| FEF 25-75% (L/sec) | 2.98 | 0.52 | 18 | | FRC (L) | 3.5 | 6.9 | 197 |

Labs from last month

| Lab | Value | Units | Reference | | Lab | Value | Units | Reference |
|---|---|---|---|---|---|---|---|---|
| Na | 138 | meq/L | 133-145 | | Glucose | 242 | mg/dL | 70-110 |
| K | 4.2 | meq/L | 3.3-5 | | Urea | 19 | mg/dL | 6-26 |
| Cl | 100 | meq/L | 96-108 | | Creatinine | 1.0 | mg/dL | 0.6-1.5 |
| CO2 | 31 | meq/L | 24-32 | | Est. GFR | 76 | mL/min | > 60 |
| Ca | 9.4 | mg/dL | 8.4-10.6 | | | | | |

**Review of Systems:**

- (+) shortness of breath with productive cough of clear sputum; (+) weakness and fatigue; (-) myalgia
- Denies any chest pain, fever, or night sweat

**Based on this information, develop an assessment and plan to address this patient's COPD and related preventative care issues.**

**A:**

1. **GOLD 3 Severe:** FEV1 between 30 – 49% predicted
2. **Patient Group E:** mMRC grade 2 OR 3 and 1 exacerbation requiring hospitalization; LAMA + LABA (preferably single inhaler therapy for convenience) OR LAMA + LABA + ICS if Eosinophils ≥ 300 indicated as initial therapy
3. **Non-adherence** to ipratropium/albuterol therapy which is not appropriate therapy based on exacerbation risk and symptoms.
4. **Inappropriate inhaler technique**

**P:** Remember to include recommendations, monitoring and education (pertinent drug and non-drug therapy) in your plan.

1. **Recommend LAMA + LABA combination**
2. **Change albuterol/ipratropium to albuterol MDI 2 puffs Q4-6 hrs prn.**




Published online August 4, 2026 https://doi.org/10.15326/jcopdf.2026.0794

3. **Educate** patient on **appropriate inhaler technique, adherence** and potential adverse effects of therapy.
4. **Recommend annual influenza vaccine and pneumococcal (and shingles) if not previously vaccinated**
    a. Adults 19-64 years of age with certain chronic conditions (ie COPD, alcoholism, DM, HIV, etc):
        i. Give one dose of PCV15 or PCV20
            1. If PCV is used, then follow with a dose of PPSV23 after one year
            2. If PCV20 used, then PPSV23 NOT indicated
        ii. For those who only received PPSV23
            1. Give one dose of PCV15 or PCV20 (should be administered at least 1 year after most recent PPSV23 vaccine)
            2. Additional PPSV23 dose not recommended regardless if they received PCV15 or PCV20 because they already received it previously.
    b. ≥ 65 years: if never received pneumococcal vaccines/unknown history, then PCV15 or PCV20
        i. If PCV15 used, then follow with PPSV23 one year later
        ii. If PCV20 used, PPSV23 NOT indicated
5. **Pulmonary rehab**
6. **Pt to monitor symptoms and rescue inhaler use**
7. Recommend **follow-up** (time may vary here but must be reasonable). Annual spirometry.




Published online August 4, 2026 https://doi.org/10.15326/jcopdf.2026.0794

**Appendix B: Definitions and explanations of accuracy, usefulness, and impact variability**

| **Accuracy** | | |
|---|---|---|
| **0** | **1** | **2** |
| Response is **poor** (not accurate with significant inaccuracies) | Response is **borderline** (partly accurate and inaccurate) | Response is **good** (completely accurate and error free) |
| Notes if needed: | | |

| **Usefulness** | | |
|---|---|---|
| **0** | **1** | **2** |
| Response is **not** useful | Response is **somewha**t useful | Response is **very** useful |
| Notes if needed: | | |

| **Impactful Variability between outputs** | | |
|---|---|---|
| **0 High impact** | **1 Moderate impact** | **2 Low impact** |
| Responses widely vary between outputs | Responses vary to some degree between outputs | Answers do not have impactful variability |
| Notes if needed: | | |




Published online August 4, 2026 https://doi.org/10.15326/jcopdf.2026.0794