Infection Control & Hospital Epidemiology (2026), **47**, 863–865
doi:10.1017/ice.2026.10511

SHEA

Commentary

Does AI know how to choose wisely? Do not trust and definitely verify when using generative AI in stewardship operations

Amy V. Dora MD[1], Lawrence N. Benjamin MD, PhD[2,3], Christopher J. Graber MD, MPH[1,3] and Westyn Branch-Elliman MD, MMSC[1,3]

[1]Infectious Diseases Section, VA Greater Los Angeles Healthcare System, Los Angeles, USA; [2]Pulmonary & Critical Care Section, VA Greater Los Angeles Healthcare System, Los Angeles, USA and [3]UCLA David Geffen School of Medicine, Los Angeles, USA

Abstract

Antimicrobial stewardship programs (ASP) serve as a valuable resource to healthcare institutions. This commentary explores a real world demonstration of the potential advantages and pitfalls of using artificial intelligence (AI) in ASP, both in supporting the development of institutional guidelines and in providing a cost-effectiveness analysis of these protocols, as well as the caution that should be exercised when involving AI in stewardship.

(Received 23 June 2026 UTC; accepted 24 June 2026 UTC; electronically published 17 July 2026)

Antimicrobial stewardship programs (ASP) achieve impact by improving antimicrobial use to limit the development of resistance and antimicrobial-associated harms. Economic benefit of the programs stems from avoidance of harm and reducing the use of expensive drugs—but most programs lack health economics support to make a robust business case for the programs. Enter ChatGPT? In this issue of ICHE, Enriquez et al. provides a real-world demonstration of how artificial intelligence (AI) tools can be used to support development of institutional guidelines—and the limitations of the AI approach—as well as demonstrating a proof-of-concept of how AI tools can be used to expand upon existing services to support a business use case for stewardship.[1] Ironically, the AI economic analysis gave its own protocol a failing grade.

In this proof-of-concept study, Microsoft Copilot™ was first instructed to generate a postlung transplantation cytomegalovirus (CMV) prophylaxis protocol based on all available Pub-Med indexed articles or guidelines, an inherently retrospective summarization task. Recommendations included in the AI-generated protocol were then compared to the 4th International Consensus Guidelines on the Management of Cytomegalovirus in Solid Organ Transplantation and two physician-developed protocols from Vanderbilt University Medical Center (VUMC). When the four different documents were compared, the AI-generated guideline demonstrated a consistent bias toward errors of commission rather than omission: the AI-generated recommendations favored longer durations of treatment and tended to place less weight on the potential harms of the approach versus the current consensus and clinician-derived guidelines.

Why did AI generate the most conservative protocol? Despite specific guidance to create an evidence-based prophylaxis regimen with citations from a guideline or trial indexed in Pub-Med, Copilot™ heavily relied on non-indexed CMV-prophylaxis institutional guidelines published in 2013 and 2023 from tertiary academic centers[1]—in other words, the system did not anchor on the right information and did not place more weight on newer, stronger evidence. Perhaps most concerningly, it also did not follow the letter of the instructions it was provided with. The American Society of Transplantation Infectious Diseases Community of Practice guidelines published in 2019 recommend 6–12 months of prophylaxis for intermediate (CMV R+) groups[2] (in agreement with the 4th International Consensus Guidelines written in 2025),[3] and at least one month prophylaxis against herpes simplex virus (HSV) in low CMV risk groups (CMV D-/R-).[4] These were not referenced in the AI-generated bibliography. Weighting on the "wrong" documents resulted in the AI's aggressive suggestions, including (1) 12-month prophylaxis for intermediate-risk groups, (2) repeating valganciclovir induction during the maintenance phase of CMV prophylaxis, and 3) continuing HSV prophylaxis for 12 months in low-risk groups (CMV D-/R-).[5] While Copilot™ only substantiated its 12-month CMV prophylaxis recommendation for intermediate-risk groups based on background information present in a valganciclovir pharmacokinetics/pharmacodynamics study,[6] higher-quality evidence (including a multicenter randomized controlled trial demonstrating the benefit of 12 months over three months prophylaxis in intermediate and high-risk groups) was not included in the AI-generated bibliography.[7,8] Copilot™ also did not validate repeat valganciclovir induction during maintenance phase prophylaxis or the extended duration of valacyclovir prophylaxis for low-risk groups. Data supporting shorter prophylaxis durations guided by CMV-specific immunity[9] was notably absent from the AI protocol, perhaps reflecting a general bias toward more aggressive care.

What risks do we take when using AI-backed ASP protocols? This study demonstrates bias of one AI model toward creating a protocol framework based on older, open access institutional

Corresponding author: Amy V. Dora; Email: Amy.Dora@va.gov

Cite this article: Dora AV, Benjamin LN, Graber CJ, Branch-Elliman W. Does AI know how to choose wisely? Do not trust and definitely verify when using generative AI in stewardship operations. *Infect Control Hosp Epidemiol* 2026. 47: 863–865, doi: 10.1017/ice.2026.10511

© The Author(s), 2026. Published by Cambridge University Press on behalf of The Society for Healthcare Epidemiology of America.

https://doi.org/10.1017/ice.2026.10511 Published online by Cambridge University Press

guidelines over current international consensus guidelines, and data cited did not consistently reflect the most current, high-quality, high-grade evidence to support its recommendations. References were incomplete and it was not always clear how they were selected, including in some cases where operator prompts were explicitly disregarded. In the specific question of CMV prophylaxis in postlung transplantation protocols, balancing infection prevention with the harms of antimicrobial exposure and potential for mutagenesis with long-term valganciclovir receipt requires a complex risk-benefit analysis based on institution-specific populations, CMV infection burden amongst donors and recipients, center-specific immunosuppressive regimens, and posttransplant care that cannot be entirely entrusted to an AI model. With longer durations of CMV prophylaxis in intermediate-risk groups (i.e. 12 months rather than six or nine months), we may risk incurring CMV resistance to valganciclovir or drug toxicity such as leukopenia requiring a switch to the more costly letermovir, without necessarily preventing more cases of CMV disease.[10] Unless there is significant oversight and external human validators to verify the quality of studies used, the risk of publication bias can predispose AI to assign more value to older, positive studies resulting in longer durations of antimicrobial therapy, which is antithetical to stewardship goals. More work is needed on how to optimize prompts to limit these risks and on which AI tools and models are best suited for these purposes. Clinicians should beware that AI models are only as good as the size of the data set used to train them and will inherently perform less well for tasks for which the training data is small. The nuances of transplant stewardship certainly fall into this category and offer a cautionary tale for the use of these tools for other rare and unusual events in infectious diseases care—the place where clinicians need help and support the most.

The pervasiveness of the AI bias to recommend more medication and more medical care in general is an avenue that warrants ongoing review and evaluation as, if persistent, would have substantial impacts on how care is provided, particularly in stewardship. A related risk is the need to train physicians in how to both identify and respond to AI recommendations that are inherently prone to specific types of risk aversion—a major reason stewardship programs exist in the first place is to combat this general tendency in medicine, and so it would be ironic if a move toward integration of AI tools for important stewardship tasks set the field backward. The tendency of this tool to synthesize evidence with the bias toward doing more, not less, also raises questions about whether this is a systematic problem with these AI summarization tools or a challenge specific to the application tested. As programs such as Open Evidence are increasingly used in place of textbooks and other resources, understanding more about this bias—and how to combat it—will be a critical question not just for antimicrobial stewardship programs but for the practice of medicine more broadly.

After creating the guideline, Copilot™ was then instructed to perform cost-effectiveness analyses of the four distinct CMV prophylaxis protocols based on simulated patient data, with statistics informed by a cohort of 165 patients who underwent lung transplantation at VUMC. Given the recommendations for longer and more aggressive treatment courses, the AI tool ironically found its own protocol to be the least cost-effective of all four proposed CMV protocols with the Copilot™ cost-analysis.

Although the AI guideline may not have passed muster, Enriquez et al demonstrated an important proof-of-concept for how stewardship programs can use AI tools to *expand* services and build important business cases for institutional support[11]—a perennial challenge in infection prevention, stewardship, and infectious diseases generally. In the specific example of CMV prophylaxis for lung transplant recipients, analyses pinpointed rationales for more extensive costs (based on prophylaxis and treatment choices and durations) and stratified how mean and median patient costs would change from baseline (specifically identifying that longer prophylaxis duration and increased letermovir use in the AI-generated protocol made this the least cost-effective option). Additional analyses could further increase the value of this tool, especially if it can factor in costs associated with drug toxicity such as additional outpatient visits, laboratory monitoring, or hospitalization costs including length of stay, level of care, and coordinated consultative care. Utilizing AI to assess ASP protocols based on cost-effectiveness can offset the significant manpower it may take to perform these complex analyses before ASP protocol rollouts occur.

Antimicrobial stewardship programs add tremendous value in healthcare, including developing protocols that increase uptake of evidence-based practices and preventing harm associated with prolonged antimicrobial use such as adverse patient effects and development of resistance. Although further studies are needed to evaluate how AI can assist with protocol development, from a cost-savings standpoint, AI may serve as a useful tool to validate the existence and dollar value of specific ASP protocols and support the expansion of these practices especially in settings where these data analyses are not readily available.

Great care must be taken in utilizing AI for antimicrobial stewardship in several aspects: selecting the most appropriate model that is best suited to answer the question at hand, crafting prompts that can generate the most specific answer to that question, validating that the answer is accurate, and ensuring that the answer applies to the patient population at hand. Stewardship programs using these tools should be aware of the potential risk of undermining themselves by deploying AI tools that recommend *more*, not less, and anchor on older, less relevant evidence. The Ronald Reagan quote of a Russian proverb translated as "Trust, but verify" comes to mind.[12] However, John Kerry's modification (following a suspected breach of a weapons agreement) is perhaps more apt: "verify and verify."[13] At least for now, humans are still the ones best tasked with "Choosing Wisely."[14,15]

Acknowledgements. None.

Competing interests. The authors report no conflicts of interests.

References

1. Enriquez KT, Dulanto Chiang A, Smiley C, Staub M. Updating CMV protocols in lung transplant patients: a single-center case study modeling use of generative AI for antimicrobial stewardship protocol development and economic impact analysis. *Infect Control Hosp Epidemiol* 2026;1–8.
2. Razonable RR, Humar A. Cytomegalovirus in solid organ transplant recipients-Guidelines of the American Society of Transplantation Infectious Diseases Community of Practice. *Clin Transplant* 2019;33:e13512. doi: 10.1111/ctr.13512.
3. Kotton CN, Kumar D, Manuel O, *et al.* Transplantation Society International CMV Consensus Group. The fourth international consensus guidelines on the management of cytomegalovirus in solid organ transplantation. *Transplantation* 2025;109:1066–1110. doi: 10.1097/TP.0000000000005374.
4. Lee DH, Zuckerman RA, AST Infectious Diseases Community of Practice. Herpes simplex virus infections in solid organ transplantation: guidelines from the American Society of Transplantation Infectious Diseases Community of Practice. *Clin Transplant* 2019;33:e13526. doi: 10.1111/ctr.13526.

5. Ochsner. Guideline for CMV prophylaxis and treatment in lung transplant recipients [PDF]. https://ochsner-craft.s3.amazonaws.com/www/static/Cytomegalovirus_Prevention_and_Treatment_in_Lung_Transplant_Recipients.pdf. Accessed June 2, 2026.
6. Kiser TH, Fish DN, Zamora MR. Evaluation of valganciclovir pharmacokinetics in lung transplant recipients. *J Heart Lung Transplant* 2012;31:159–66. doi: 10.1016/j.healun.2011.11.016.
7. Palmer SM, Limaye AP, Banks M, *et al.* Extended valganciclovir prophylaxis to prevent cytomegalovirus after lung transplantation: a randomized, controlled trial. *Ann Intern Med* 2010;152:761–9. doi: 10.7326/0003-4819-152-12-201006150-00003.
8. Finlen Copeland CA, Davis WA, Snyder LD, *et al.* Long-term efficacy and safety of 12 months of valganciclovir prophylaxis compared with 3 months after lung transplantation: a single-center, long-term follow-up analysis from a randomized, controlled cytomegalovirus prevention trial. *J Heart Lung Transplant*. 2011;30:990–6. doi: 10.1016/j.healun.2011.02.017.
9. Páez-Vega A, Vaquero-Barrios JM, Ruiz-Arabi E, *et al.* Safety and efficacy of immunoguided prophylaxis for cytomegalovirus disease in low-risk lung transplant recipients in Spain: a multicentre, open-label, randomised, phase 3, noninferiority trial. *Lancet Reg Health Eur* 2025;52:101268. doi: 10.1016/j.lanepe.2025.101268.
10. Patel N, Snyder LD, Finlen-Copeland A, Palmer SM. Is prevention the best treatment? CMV after lung transplantation. *Am J Transplant* 2012;12:539–44. doi: 10.1111/j.1600-6143.2011.03837.x.
11. Perencevich EN, Stone PW, Wright SB, Carmeli Y, Fisman DN, Cosgrove SE. Society for Healthcare Epidemiology of America. Raising standards while watching the bottom line: making a business case for infection control. *Infect Control Hosp Epidemiol* 2007;28:1121–33. doi: 10.1086/521852.
12. Murphy B. Suzanne Massie, Reagan's Russian whisperer in Cold War, dies at 94. The Washington Post. https://www.washingtonpost.com/obituaries/2025/01/30/suzanne-massie-russia-reagan-dies/. Accessed June 15, 2026.
13. Chu, H. U.S., Russia agree on a disposal plan for Syria's chemical weapons. Los Angeles Times. https://www.latimes.com/world/la-xpm-2013-sep-14-la-fg-wn-kerry-lavrov-syria-20130914-story.html.
14. Choosing Wisely. Choosing Wisely. Choosingwisely.org, 2017. https://www.choosingwisely.org/. Accessed June 15, 2026
15. Morden NE, Colla CH, Sequist TD, Rosenthal MB. Choosing wisely–the politics and economics of labeling low-value services. *N Engl J Med* 2014:370;589–92. doi: 10.1056/NEJMp1314965.

https://doi.org/10.1017/ice.2026.10511 Published online by Cambridge University Press