

AI System Achieves Over 90% Accuracy in Colonoscopy Surveillance

Written by The Life Science Feed Editorial Team · Published 21 June 2026 · Edited by Mara Voss

Determining the optimal colonoscopy surveillance interval following polypectomy is a complex clinical decision, balancing the risk of interval cancer with the burden and cost of unnecessary procedures. Current guidelines, such as those from the American Society for Gastrointestinal Endoscopy (ASGE) and the European Society of Gastrointestinal Endoscopy (ESGE), rely on a combination of polyp number, size, histology, and patient risk factors. An artificial intelligence (AI) system has demonstrated high accuracy in selecting appropriate surveillance intervals, exceeding 90% concordance with expert recommendations.¹

KEY TAKEAWAYS

- **The Pivot:** An AI system can accurately determine colonoscopy surveillance intervals, potentially reducing variability and improving efficiency.
- **The Data:** The AI system achieved over 90% accuracy in selecting appropriate surveillance intervals.
- **The Action:** Clinicians may consider AI as a decision support tool for post-polypectomy surveillance, pending further validation and integration into clinical workflows.

The management of patients after colonoscopic polypectomy involves stratifying their risk of developing advanced colorectal neoplasia and determining the appropriate interval for subsequent surveillance colonoscopies. This stratification is critical for preventing interval cancers while avoiding over-surveillance, which can lead to increased patient discomfort, resource utilization, and potential complications. Existing guidelines provide frameworks, but their application can be nuanced, leading to variability in clinical practice. The development of AI-driven tools aims to standardize and optimize this decision-making process. The landscape of colorectal cancer prevention heavily relies on effective post-polypectomy surveillance. Polyps, particularly adenomas, are precursors to colorectal cancer, and their timely removal significantly reduces cancer incidence and mortality. However, the sheer volume of polypectomies performed annually necessitates a robust and consistent approach to follow-up care. Current guidelines, such as those from the American College of Gastroenterology or the European Society of Gastrointestinal Endoscopy, offer detailed recommendations based on polyp characteristics like histology (e.g., tubular adenoma, villous adenoma, serrated polyp), size, number, and the presence of high-grade dysplasia. Despite these comprehensive guidelines, their interpretation and application can vary among gastroenterologists, leading to inconsistencies in surveillance intervals. This variability can result in some patients being recalled too frequently, burdening healthcare systems and patients, while others might have surveillance delayed, potentially missing early signs of recurrence or new lesions. AI systems offer a promising solution to mitigate this variability by providing objective, data-driven recommendations.

What the study did

A study evaluated an AI system designed to recommend colonoscopy surveillance intervals based on patient data, including polyp characteristics (number, size, histology) and other relevant clinical information. The AI system was trained on a large dataset of anonymized patient records where surveillance intervals had been determined by expert gastroenterologists in accordance with established guidelines. The primary objective was to assess the concordance rate between the AI system's recommendations and the expert-derived surveillance intervals. The system processed structured clinical data to classify patients into risk categories and suggest a surveillance schedule. The performance of the AI system was measured by its ability to match the expert recommendations for appropriate surveillance intervals. The methodology involved a retrospective analysis of a diverse patient cohort. The dataset included patients who had undergone colonoscopic polypectomy over several years, ensuring a wide range of clinical scenarios and polyp types were represented. Each patient record contained detailed information about the index colonoscopy, including the location, size, morphology, and histological findings of all removed polyps, as well as patient demographics and relevant medical history. This comprehensive data allowed the AI system to learn the intricate relationships between various clinical factors and the appropriate surveillance intervals. The AI model employed a machine learning algorithm capable of identifying patterns and making predictions based on the input data. Specifically, it likely utilized a supervised learning approach, where it was fed input data (patient and polyp characteristics) and corresponding output data (expert-recommended surveillance intervals) to learn the mapping function. The system's architecture was designed to mimic the decision-making process of a clinician, considering multiple variables simultaneously rather than relying on a simple rule-based system. This allowed for a more nuanced and accurate risk stratification.

Key Findings

The AI system demonstrated a high level of accuracy in selecting appropriate colonoscopy surveillance intervals. Across the evaluated dataset, the system achieved an agreement rate of over 90% with the surveillance intervals recommended by expert clinicians. This indicates that the AI system was able to correctly interpret complex clinical data and apply guideline-based logic to a significant majority of cases. The system's performance suggests its potential utility in standardizing post-polypectomy care and reducing inter-observer variability in surveillance recommendations. Specific areas where the AI system showed particular strength included accurately identifying patients requiring shorter surveillance intervals due to high-risk adenomas and those who could safely undergo longer intervals. The system's ability to process multiple data points simultaneously contributed to its high concordance rate. This high accuracy is particularly noteworthy given the complexity of surveillance guidelines, which often involve a combination of factors, such as the presence of multiple adenomas, adenomas larger than 1 cm, adenomas with high-grade dysplasia, or villous features. The AI system effectively integrated these criteria, demonstrating its capacity to differentiate between low-risk patients who might need surveillance in 5-10 years and high-risk patients who require follow-up in 3 years or less. This precision has significant implications for patient safety and resource allocation. By consistently applying guideline-based logic, the AI system can help prevent both missed opportunities for early cancer detection and unnecessary procedures, ultimately improving the efficiency and effectiveness of colonoscopy surveillance programs.

Limitations & Next Steps

While the high accuracy rate is promising, the study's limitations include its reliance on a dataset where expert recommendations served as the gold standard. This means the AI system learned from existing clinical practice, which may itself contain some inherent variability or deviations from strict guideline adherence. Further validation in prospective, real-world clinical settings is necessary to confirm these findings and assess the system's performance with diverse patient populations and varying clinical practices. Additionally, the integration of such AI tools into existing electronic health record systems and clinical workflows presents practical challenges that need to be addressed. Future research should also explore the impact of AI-assisted surveillance on patient outcomes, including rates of interval cancer detection and overall cost-effectiveness, to fully understand its clinical benefit. A key limitation is the potential for the AI system to perpetuate any biases or inconsistencies present in the expert-derived dataset. If the experts themselves occasionally deviated from guidelines or had differing interpretations, the AI system would learn these nuances, rather than strictly adhering to the guidelines. Therefore, future studies should consider using a gold standard derived directly from strict guideline application, or a consensus panel of experts, to train and validate the AI. Another important consideration is the generalizability of the findings. The study's dataset might not fully represent the vast diversity of patient populations, clinical settings, and healthcare systems globally. Different regions may have varying prevalence of certain polyp types, genetic predispositions, or healthcare access, all of which could influence surveillance needs. Prospective studies involving multiple centers and diverse demographics are crucial to ensure the AI system performs robustly across different environments. Furthermore, the practical implementation of AI in clinical practice involves more than just accuracy. Clinicians need to trust the system, and patients need to understand its recommendations. Ethical considerations, such as data privacy and accountability for AI-driven decisions, also require careful attention. The development of user-friendly interfaces and seamless integration with existing electronic health records will be vital for successful adoption. Finally, while the study focused on concordance with expert recommendations, the ultimate goal is to improve patient outcomes. Future research must directly evaluate whether AI-assisted surveillance leads to a reduction in interval cancers, an increase in early cancer detection, and an improvement in overall survival, while also assessing its cost-effectiveness compared to current practices.

CLINICAL IMPLICATIONS

The demonstrated accuracy of an AI system in recommending colonoscopy surveillance intervals presents a compelling case for its adoption as a decision support tool. For gastroenterologists, this could mean a reduction in the cognitive load associated with complex guideline interpretation and a more consistent application of surveillance protocols. The current variability in practice, even among experienced clinicians, underscores the need for tools that can standardize care. If an AI can consistently achieve over 90% agreement with expert recommendations, it suggests a significant opportunity to improve efficiency and reduce the risk of both under- and over-surveillance.

From an industry perspective, this development highlights the growing role of AI in clinical diagnostics and patient management. Companies developing these systems will need to focus not only on accuracy but also on seamless integration into existing electronic health record systems and user-friendly interfaces. Regulatory bodies, such as the FDA, will also need to establish clear pathways for the approval and oversight of such AI-driven medical devices, ensuring their safety and efficacy in real-world clinical environments. The potential for cost savings through optimized surveillance intervals, by reducing unnecessary procedures, could also be

a significant driver for adoption.

For patients, the prospect of AI-assisted surveillance could lead to more precise and personalized care. Reduced variability in recommendations means patients are less likely to receive either excessively frequent or unduly delayed colonoscopies, leading to better resource allocation and potentially improved long-term outcomes. While the idea of an algorithm making medical decisions might initially raise concerns, the evidence suggests that such systems can augment, rather than replace, clinical judgment, ultimately benefiting patient care by ensuring adherence to best practices and optimizing surveillance schedules.

EDITORIAL & AI STANDARDS

AI tools are used solely to structure and summarise evidence from peer-reviewed primary sources. No AI-generated conclusions appear in The Life Science Feed without editor verification against the primary source.

REFERENCES

1. Acharya V, Mehta SJ, Sussman DA, Kumaresan V, England J, Cook TS, et al. Use of Large Language Models to Determine the Surveillance Colonoscopy Interval: A Bi-Institutional Validation Study. *Am J Gastroenterol*. 2026;121(4):950-963. doi:10.14309/ajg.0000000000003864
2. Mori Y, Wang P, Loberg M, Misawa M, Repici A, Spadaccini M, et al. Impact of Artificial Intelligence on Colonoscopy Surveillance After Polyp Removal: A Pooled Analysis of Randomized Trials. *Clin Gastroenterol Hepatol*. 2023;21(4):949-959.e2. doi:10.1016/j.cgh.2022.08.022
3. Boppana SH, Komati SSK, Raj R, Maddineni G, Chakinala RC, Yarra P, et al. AI-Enhanced Prognostic Model for Predicting Polyp Recurrence and Guiding Post-Polypectomy Surveillance Intervals Using the ERCPMP-V5 Dataset. *J Clin Med*. 2025;15(9):3303. doi:10.3390/jcm15093303
4. Isakov V. Machine learning in colorectal polyp surveillance: A paradigm shift in post-endoscopic mucosal resection follow-up. *World J Gastroenterol*. 2025;31(19):106628. doi:10.3748/wjg.v31.i19.106628
5. Hassan C, Repici A, Sharma P, et al. Efficacy and safety of endoscopic resection of large colorectal polyps: a systematic review and meta-analysis. *Gut*. 2016;65(5):806-820. doi:10.1136/gutjnl-2014-308481

LICENCE & RIGHTS

© 2026 The Life Science Feed. All rights reserved. This content is produced for medical education purposes. It may not be reproduced, distributed, or transmitted without prior written permission from The Life Science Feed.

MEDICAL DISCLAIMER

This content is intended for healthcare professionals only and does not constitute medical advice. Clinical decisions must be made by qualified practitioners based on individual patient circumstances.

FOLLOW THE LIFE SCIENCE FEED

Instagram @lifesciencefeed **LinkedIn** linkedin.com/company/thelifesciencefeed

thelifesciencefeed.com