

Literature Review: **Artificial Intelligence Pada Bidang Gastroenterologi**

Hendra Asputra^{1*}, Reza Okta Lestari²

ABSTRACT

Artificial intelligence (AI) is rapidly evolving in medicine, particularly in gastroenterology, a field heavily reliant on image analysis and clinical data. Technologies such as computer-aided detection (CADe) and computer-aided diagnosis (CADx) have been developed to enhance the detection and characterization of gastrointestinal (GI) lesions. This literature review aims to evaluate the roles, applications, benefits, and limitations of AI in gastroenterology. Utilizing a narrative synthesis approach, relevant primary studies from the past 5 to 10 years were retrieved from PubMed, Google Scholar, ScienceDirect, and Wiley Online Library. The findings indicate that AI is predominantly applied in endoscopic procedures, significantly improving the adenoma detection rate (ADR), polyp detection rate (PDR), and adenomas per colonoscopy (APC), while reducing the adenoma miss rate (AMR). Furthermore, AI demonstrates substantial potential in polyp characterization, upper GI cancer detection, capsule endoscopy interpretation, ulcerative colitis evaluation, and ultrasound-based hepatobiliary diagnosis. Despite these promising results, clinical implementation faces several challenges, including false-positive rates, the need for external validation, dataset biases, performance variability across different populations, high costs, and regulatory hurdles. In conclusion, AI holds immense potential to elevate the quality of diagnostics and patient care in gastroenterology. While highly effective as a clinical support tool, further research and robust regulatory frameworks are essential to ensure its optimal, safe, and sustainable integration into routine clinical practice.

Keywords: *Artificial Intelligence, computer-aided detection, computer-aided diagnosis, endoscopy, gastroenterology*

Artificial intelligence (AI) merupakan teknologi berbasis komputer yang dirancang untuk meniru kemampuan kognitif manusia, termasuk proses pembelajaran, pengenalan pola, serta pengambilan keputusan berbasis data dalam jumlah besar. Dalam bidang medis, perkembangan AI meningkat pesat dalam dua dekade terakhir, terutama didorong oleh ketersediaan *big data* dan kemajuan metode *machine learning* serta *deep learning* yang memungkinkan analisis data kompleks dengan akurasi tinggi.¹

Gastroenterologi merupakan salah satu bidang kedokteran yang sangat potensial dalam penerapan AI karena sangat bergantung pada interpretasi visual, khususnya dalam prosedur endoskopi. Berbagai

aplikasi AI telah dikembangkan untuk membantu deteksi lesi, klasifikasi penyakit, prediksi prognosis, hingga penilaian kualitas prosedur, seperti evaluasi *bowel preparation* dan kualitas pemeriksaan endoskopi. Kemampuan ini menjadikan AI sebagai alat yang menjanjikan dalam meningkatkan akurasi diagnosis sekaligus mengurangi variasi antar operator.¹

Kanker kolorektal masih menjadi salah satu penyebab utama morbiditas dan mortalitas akibat kanker. Kolonoskopi efektif untuk skrining dan pencegahan kanker kolorektal melalui deteksi serta pengangkatan lesi prakanker. Namun, efektivitas kolonoskopi sangat dipengaruhi oleh kualitas pemeriksaan. *Adenoma detection rate (ADR)* merupakan indikator mutu penting karena peningkatan ADR berhubungan dengan penurunan risiko kanker kolorektal interval dan kematian akibat kanker kolorektal.²

* Corresponding Author: KJF/KSM Ilmu Penyakit Dalam, Fakultas Kedokteran Universitas Riau, RSUD Arifin Achmad Provinsi Riau, email: hendraasputra13@gmail.com

¹ KJF/KSM Ilmu Penyakit Dalam, Fakultas Kedokteran Universitas Riau, RSUD Arifin Achmad Provinsi Riau
Fakultas Kedokteran Universitas Riau

Urgensi penggunaan AI semakin kuat karena endoskopi konvensional tidak sepenuhnya bebas dari lesi yang terlewat. Studi *tandem colonoscopy* menunjukkan bahwa *miss rate* adenoma masih cukup tinggi, yaitu sekitar 26%, terutama pada lesi kecil, sehingga diperlukan teknologi yang mampu membantu pengamatan visual secara konsisten.³ Keterbatasan tersebut dapat dipengaruhi oleh kelelahan operator, blind spot, kualitas bowel preparation, morfologi lesi yang datar, dan variasi pengalaman klinis.

Dalam konteks tersebut, sistem *computer-aided detection* (CADe) dikembangkan untuk memberikan tanda visual atau alarm saat sistem mengenali kemungkinan polip atau lesi. Sementara itu, *computer-aided diagnosis* (CADx) dirancang untuk memprediksi karakteristik histologis atau sifat neoplastik lesi sehingga dapat mendukung pengambilan keputusan klinis, misalnya *resect-and-discard* atau *leave-in-situ* pada polip tertentu.¹

Selain kolonoskopi, aplikasi AI juga berkembang pada endoskopi saluran cerna atas, kapsul endoskopi, *inflammatory bowel disease*, ultrasonografi hati, dan diagnosis lesi hepatobilier. Oleh karena itu, *literature review* ini bertujuan untuk mengkaji perkembangan, peran, manfaat, serta keterbatasan AI dalam gastroenterologi, khususnya dalam meningkatkan deteksi dan karakterisasi lesi gastrointestinal, serta menilai dampaknya terhadap kualitas diagnosis dan praktik klinis.

METODE

Penelitian ini menggunakan metode *literature review* dengan sintesis naratif untuk mengkaji peran AI dalam bidang gastroenterologi. Data yang digunakan merupakan data sekunder yang diperoleh dari berbagai literatur ilmiah yang relevan, terutama penelitian primer seperti *randomized controlled trial*, studi prospektif, studi diagnostik, studi multicenter, dan studi cross-sectional yang membahas penerapan AI dalam praktik klinis gastroenterologi.

Literatur ditelusuri melalui database PubMed, Google Scholar, ScienceDirect, dan Wiley Online Library pada periode Maret-April 2026. Strategi pencarian menggunakan kombinasi kata kunci seperti “artificial intelligence”, “gastroenterology”, “endoscopy”, “computer-aided detection”, serta “computer-aided diagnosis”, yang dikombinasikan dengan operator Boolean seperti AND dan OR untuk memperoleh hasil pencarian yang lebih spesifik dan relevan.

Artikel dipilih berdasarkan relevansi topik, keterbaruan publikasi, ketersediaan *full-text*, kesesuaian desain penelitian, serta kontribusinya terhadap pembahasan penerapan AI dalam gastroenterologi. Literatur yang tidak relevan dengan fokus pembahasan, duplikatif, atau tidak memiliki metodologi yang jelas tidak digunakan sebagai sumber utama.

Literatur yang terpilih kemudian dianalisis secara naratif dengan mengelompokkan temuan berdasarkan tema utama, yaitu deteksi lesi kolorektal dengan CADe, karakterisasi lesi dengan CADx, aplikasi AI di luar kolonoskopi, serta tantangan implementasi klinis. Hasil sintesis disajikan dalam bentuk uraian naratif untuk memberikan gambaran mengenai perkembangan, manfaat, dan keterbatasan penggunaan AI dalam bidang gastroenterologi.

HASIL

Berdasarkan hasil penelusuran literatur yang dilakukan melalui berbagai basis data elektronik, diperoleh sejumlah penelitian primer yang relevan dengan topik penggunaan AI pada bidang gastroenterologi. Artikel-artikel tersebut digunakan sebagai dasar sintesis naratif dan dirangkum pada Tabel 1.

Tabel 1. Karakteristik Penelitian Primer yang Dianalisis

Penulis (Tahun)	Negara	Desain dan Sampel	Jenis AI	Hasil	Implikasi
Wang et al. ⁴ (2019)	China	Prospective randomized controlled study; 1.058 pasien dianalisis (522 CAde, 536 kontrol).	Real-time automatic polyp detection berbasis deep learning/CAde.	ADR meningkat signifikan pada kelompok AI dibanding kontrol (29,1% vs 20,3%; $p<0,001$). Mean adenoma per pasien juga meningkat (0,53 vs 0,31; $p<0,001$), terutama karena bertambahnya deteksi adenoma diminutif (185 vs 102).	CAde efektif meningkatkan deteksi adenoma, terutama lesi kecil/diminutif yang mudah terlewat.
Repici et al. ⁵ (2020)	Italia	Multicenter randomized trial di 3 pusat; 685 subjek dianalisis (341 CAde, 344 kontrol).	GI Genius real-time CAde.	ADR lebih tinggi pada CAde dibanding kontrol (54,8% vs 40,4%; RR 1,30; 95% CI 1,14-1,45). Adenoma per colonoscopy juga meningkat (1,07 +/- 1,54 vs 0,71 +/- 1,20), tanpa peningkatan withdrawal time maupun reseksi lesi non-neoplastik.	CAde meningkatkan ADR dan APC secara real-time tanpa memperpanjang prosedur.
Liu et al. ⁶ (2020)	China	Prospective randomized study; 790 pasien dianalisis (393 CAde, 397 kontrol).	Embedded CAde system pada monitor utama/single-monitor.	ADR meningkat pada kelompok CAde (29,01% vs 20,91%; $p=0,009$). APC meningkat (0,48 vs 0,29; $p<0,001$). Peningkatan terutama pada adenoma diminutif non-advanced, serrated adenoma, dan polip hiperplastik; skor fatigue operator tidak berbeda bermakna.	Integrasi CAde pada monitor utama dapat meningkatkan kualitas deteksi tanpa menambah kelelahan operator.
Xu et al. ⁷ (2021)	China	Prospective multicenter randomized controlled trial di 6 pusat; 2.352 pasien.	Real-time AI-assisted colonoscopy untuk deteksi polip.	PDR tidak meningkat bermakna dibanding kontrol (38,8% vs 36,2%; $p=0,183$). Namun AI meningkatkan PPC-Plus (0,5 vs 0,4; $p<0,05$), serta mendeteksi lebih banyak polip diminutif (76,0% vs 68,8%; $p<0,01$) dan polip datar (5,9% vs 3,3%; $p<0,05$).	Manfaat AI dapat tampak pada jumlah/jenis polip tambahan, meskipun luaran utama PDR tidak selalu signifikan.
Glissen Brown et al. ⁸ (2022)	USA	US multicenter randomized tandem colonoscopy study; 223 pasien setelah eksklusi.	EndoScreener deep-learning CAde.	AMR lebih rendah pada kelompok CAde-first dibanding HDWL-first (20,12% [34/169] vs 31,25% [45/144]; $p=0,0247$). SSL miss rate juga lebih rendah (7,14% vs 42,11%; $p=0,0482$), dan first-pass APC lebih tinggi (1,19 vs 0,90; $p=0,0323$).	CAde dapat mengurangi lesi yang terlewat, termasuk sessile serrated lesion, pada desain tandem.

Shaukat et al. ⁹ (2022)	USA	Randomized trial di 5 pusat akademik/komunitas; modified intention-to-treat 1.359 pasien (677 standar, 682 CAdE).	SKOUT CAdE device.	APC meningkat signifikan dengan CAdE (1,05 vs 0,83; p=0,002). ADR meningkat secara numerik tetapi tidak signifikan (47,8% vs 43,9%; p=0,065). THR tetap memenuhi non-inferiority (67,4% vs 71,7%; p non-inferiority <0,001).	CAdE meningkatkan jumlah adenoma yang ditemukan tanpa penurunan bermakna pada proporsi reseksi dengan histologi klinis penting.
Soons et al. ¹⁰ (2022)	Belanda/ Jerman/ Italia	Prospective feasibility and safety study di 3 pusat tersier Eropa; 90 pasien.	DISCOVERY real-time CAdE pada monitor sekunder.	Mean inspection time 10,8 +/- 4,3 menit; PDR 55,6%, ADR 28,9%, dan SDR 11,1%. Estimasi <20 false positive terjadi pada 81/90 prosedur (90%); tidak ada serious adverse event terkait CAdE; user friendliness median 8/10.	Sistem CAdE layak dan aman untuk penggunaan real-time awal, tetapi perlu RCT untuk mengonfirmasi efektivitas deteksi.
Byrne et al. ¹¹ (2019)	Kanada/ USA/ Argentina/ Irlandia	Diagnostic validation study; 125 video polip diminutif, 106 video dianalisis karena 19 video tidak mencapai confidence threshold.	Deep-learning CAdx/CNN pada video kolonoskopi standar.	Pada 106 video yang dapat diprediksi, akurasi 94%, sensitivitas identifikasi adenoma 98%, spesifisitas 83%, NPV 97%, dan PPV 90% untuk membedakan adenoma vs polip hiperplastik.	CAdx berpotensi mendukung optical biopsy dan strategi resect-and-discard, tetapi perlu validasi live clinical trial.
Mori et al. ¹² (2018)	Jepang	Single-group open-label prospective study; 791 pasien, 466 polip diminutif dari 325 pasien dinilai CAD.	Real-time CAdx dengan endocytoscopy mode NBI dan stained mode.	Prediction rate CAD 98,1% (457/466). NPV untuk adenoma rektosigmoid diminutif: stained mode 96,4% best-case dan 93,7% worst-case; NBI 96,5% best-case dan 95,2% worst-case.	CAdx mencapai ambang performa untuk diagnose-and-leave pada polip rektosigmoid diminutif non-neoplastik.
Hassan et al. ¹³ (2022)	Italia	Prospective clinical study; 162 pasien dengan 544 polip yang direseksi.	Approved CAdx module pada unmagnified white-light colonoscopy.	CAdx feasible pada 291/295 (98,6%) polip rektosigmoid <=5 mm yang terverifikasi histologi. NPV CAdx 97,6% (95% CI 94,1%-99,1%); 242/295 (82%) lesi dapat dipertimbangkan untuk leave-in-situ dan 212/544 (39%) untuk resect-and-discard.	CAdx memenuhi benchmark optical diagnosis untuk mengurangi polipektomi/patologi yang tidak perlu pada polip tertentu.
Hirasawa et al. ¹⁴ (2018)	Jepang	Retrospective diagnostic study; pelatihan 13.584 citra kanker lambung, uji 2.296 citra dari 69 pasien dengan 77 lesi kanker lambung.	CNN berbasis Single Shot MultiBox Detector (SSD).	CNN menganalisis 2.296 citra dalam 47 detik dan mendeteksi 71/77 lesi kanker lambung (sensitivitas 92,2%). PPV 30,6%; 70/71 lesi yang terdeteksi berdiameter >=6 mm dan seluruh kanker invasif terdeteksi.	AI dapat mempercepat skrining citra endoskopi lambung, tetapi false positive masih menjadi keterbatasan.

Luo et al. ¹⁵ (2019)	China	Multicenter case-control diagnostic study; 1.036.496 citra endoskopi dari 84.424 individu.	GRAIDS deep-learning system untuk kanker saluran cerna atas.	Akurasi diagnostik GRAIDS 0,955 pada internal validation set dan 0,927 pada prospective set. Sensitivitas GRAIDS serupa endoskopis ahli (0,942 vs 0,945; p=0,692) dan lebih tinggi dibanding endoskopis kompeten maupun trainee.	AI dapat membantu deteksi kanker saluran cerna atas, terutama untuk mendukung layanan non-rujukan.
Ding et al. ¹⁶ (2019)	China	Multicenter study; 113.426.569 citra SB-CE dari 6.970 pasien di 77 pusat; validasi pada 5.000 pasien.	CNN-based auxiliary reading model untuk capsule endoscopy.	Model mengidentifikasi abnormalitas dengan sensitivitas 99,88% per pasien dan 99,90% per lesi; pembacaan konvensional sensitivitas 74,57% per pasien. Waktu baca rata-rata turun dari 96,6 +/- 22,53 menit menjadi 5,9 +/- 2,23 menit per pasien (p<0,001).	AI meningkatkan sensitivitas dan efisiensi pembacaan kapsul endoskopi secara substansial.
Takenaka et al. ¹⁷ (2020)	Jepang	Development-validation study; training 40.758 citra dan 6.885 biopsi dari 2.012 pasien UC; validasi prospektif 875 pasien.	Deep neural network for evaluation of UC (DNUC).	DNUC mengidentifikasi endoscopic remission dengan akurasi 90,1% dan kappa 0,798. Untuk histologic remission, akurasi 92,9% dan kappa 0,859 dibanding hasil biopsi.	AI dapat membantu standarisasi penilaian aktivitas UC dan berpotensi mengurangi kebutuhan biopsi pada konteks tertentu.
de Groof et al. ¹⁸ (2020)	Belanda/ Belgia	Pilot live endoscopy study; 20 pasien (10 NDBE, 10 Barrett neoplasia).	Deep-learning CAD untuk deteksi neoplasia Barrett pada white-light endoscopy.	Pada analisis per-level, akurasi 90%, sensitivitas 91%, dan spesifisitas 89%. Sebanyak 9/10 pasien dengan neoplasia teridentifikasi dengan benar; satu pasien NDBE menghasilkan false-positive prediction.	Studi pilot menunjukkan kelayakan CAD live untuk neoplasia Barrett, tetapi sampel masih kecil.
Lee et al. ¹⁹ (2020)	Korea	Diagnostic development-validation study; 13.608 citra USG dari 3.446 pasien, dengan internal dan external test set.	DCNN untuk klasifikasi fibrosis hati berdasarkan skor METAVIR.	Akurasi model empat kelas 83,5% pada internal test set dan 76,4% pada external test set. AUC untuk klasifikasi sirosis (F4) 0,901 internal dan 0,857 external; pada external test set AUC DCNN lebih tinggi dibanding 5 radiolog (0,656-0,816).	DCNN berbasis USG dapat membantu penilaian fibrosis/sirosis secara non-invasif.
Li et al. ²⁰ (2021)	China	Retrospective diagnostic study; 226 lesi hati fokal (107 atypical HCC, 119 FNH) yang diperiksa dengan CEUS.	Machine-learning ultrasomics berbasis CEUS dengan SVM.	Model ultrasomics memiliki AUC 0,86, sensitivitas 76,6%, dan spesifisitas 80,5%. Kombinasi ultrasomics + skor radiolog meningkatkan AUC menjadi 0,93, lebih tinggi dibanding skor radiolog saja (AUC 0,84; p<0,001).	Ultrasomics meningkatkan diagnosis banding FNH dan atypical HCC ketika digabungkan dengan penilaian radiolog.

Yao et al. ²¹ (2018)	China	Prospective radiomics study; 177 pasien dengan focal liver lesions yang menjalani multi-modal ultrasound.	Radiomics berbasis SRT dan SVM dari BMUS, SWE, dan SWV.	Dalam LOOCV, AUC 0,94 untuk klasifikasi benign vs malignant, 0,97 untuk malignant subtyping, 0,97 untuk prediksi PD-1, 0,94 untuk Ki-67, dan 0,98 untuk MVI; performa membaik ketika modalitas viscosity dimasukkan.	Radiomics USG multimodal berpotensi membantu diagnosis, subtyping, dan prediksi perilaku biologis HCC.
Du et al. ²² (2025)	China	Multicenter retrospective diagnostic study; 1.052 pasien dengan 1.058 lesi hati <=3 cm dari 55 rumah sakit; external validation 312 lesi dari 14 rumah sakit.	Interpretable ML berbasis grayscale US, radiomics, data klinis; XGBoost/SHAP.	Model terbaik (Model URC berbasis XGBoost) memiliki AUC 0,934 pada internal validation dan 0,899 pada external validation. Model mengungguli radiolog, dan bantuan Model URC meningkatkan AUC diagnostik radiolog (p<0,0001).	Model ML interpretable dapat mendukung diagnosis small HCC <=3 cm pada populasi multicenter.
Asputra et al. ²³ (2025)	Indonesia	Cross-sectional study di Pusat Endoskopi Gastrointestinal RSCM; 60 pasien dewasa dengan 100 polip.	CAD Eye/ CADx untuk karakterisasi polip kolorektal.	Performa keseluruhan CAD Eye: sensitivitas 79,17%, spesifisitas 75,00%, PPV 89,06%, NPV 58,33%, dan akurasi 78,00%.	Memberikan bukti lokal Indonesia bahwa performa CADx bervariasi dan masih perlu validasi pada sampel lebih besar.

Catatan: AI = Artificial Intelligence; CADe = Computer-Aided Detection; CADx = Computer-Aided Diagnosis; ML = Machine Learning; DL = Deep Learning; CNN = Convolutional Neural Network; DCNN = Deep Convolutional Neural Network; ADR = Adenoma Detection Rate; PDR = Polyp Detection Rate; AMR = Adenoma Miss Rate; APC = Adenomas per Colonoscopy; PPC = Polyps per Colonoscopy; PPC-Plus = non-first polyps per colonoscopy; THR = True Histology Rate; SSL = Sessile Serrated Lesion; SDR = Sessile Serrated Lesion Detection Rate; HDWL = High-Definition White-Light; NPV = Negative Predictive Value; PPV = Positive Predictive Value; AUC = Area Under the Curve; UC = Ulcerative Colitis; NDBE = Non-Dysplastic Barrett Esophagus; SB-CE = Small-Bowel Capsule Endoscopy; USG = Ultrasonografi; BMUS = B-mode ultrasound; SWE = Shear Wave Elastography; SWV = Shear Wave Viscosity; CEUS = Contrast-Enhanced Ultrasound; FNH = Focal Nodular Hyperplasia; HCC = Hepatocellular Carcinoma; MVI = Microvascular Invasion; SRT = Sparse Representation Theory; SVM = Support Vector Machine; SHAP = SHapley Additive exPlanations; CI = Confidence Interval; RR = Relative Risk; GI = Gastrointestinal; CRC = Colorectal Cancer; RCT = Randomized Controlled Trial; NBI = Narrow Band Imaging; SSD = Single Shot MultiBox Detector; LOOCV = Leave-One-Out Cross-Validation; DNUC = Deep Neural Network for Evaluation of Ulcerative Colitis; PD-1 = Programmed Death-1; Ki-67 = marker proliferasi sel.

DISKUSI

Urgensi AI dalam Gastroenterologi

Hasil sintesis menunjukkan bahwa AI berkembang sebagai respons terhadap keterbatasan pemeriksaan konvensional yang sangat bergantung pada kemampuan visual dan konsentrasi operator. Pada kolonoskopi, keterbatasan ini penting karena ADR berkaitan dengan risiko kanker kolorektal interval dan kematian akibat kanker.² Selain itu, studi *tandem colonoscopy* menunjukkan adanya adenoma yang tetap terlewat meskipun pemeriksaan dilakukan secara cermat.³ Temuan tersebut memperkuat alasan

penggunaan AI sebagai *second observer* yang dapat membantu mempertahankan kewaspadaan visual selama prosedur.

Dalam konteks gastroenterologi modern, AI tidak hanya berfungsi sebagai alat deteksi otomatis, tetapi juga sebagai sistem pendukung keputusan klinis. Aplikasi ini mencakup deteksi lesi kolorektal, prediksi histologi polip, deteksi kanker saluran cerna atas, pembacaan kapsul endoskopi, penilaian aktivitas penyakit inflamasi usus, dan klasifikasi penyakit hati berbasis pencitraan.

Deteksi Lesi Kolorektal dengan CADE

Bukti paling kuat mengenai AI dalam gastroenterologi berasal dari penelitian CADE pada kolonoskopi. Wang et al. menunjukkan bahwa sistem deteksi polip otomatis berbasis deep learning meningkatkan ADR dari 20,3% menjadi 29,1% dan meningkatkan jumlah adenoma per pasien.⁴ Repici et al. juga menunjukkan bahwa CADE real-time meningkatkan deteksi neoplasia kolorektal dalam uji acak multicenter.⁵ Temuan serupa dilaporkan oleh Liu et al., dengan peningkatan ADR dari 20,91% menjadi 29,01% tanpa peningkatan fatigue operator.⁶

Namun, hasil penelitian tidak sepenuhnya homogen. Xu et al. menemukan bahwa AI tidak meningkatkan PDR secara signifikan, meskipun jumlah polip tambahan dan polip diminutif yang terdeteksi lebih tinggi pada kelompok AI.⁷ Perbedaan ini menunjukkan bahwa performa CADE dapat dipengaruhi oleh populasi, pengalaman endoskopis, sistem AI yang digunakan, kualitas persiapan kolon, dan indikator luaran yang dinilai.

Studi tandem oleh Glissen Brown et al. memberikan bukti penting bahwa CADE bukan hanya meningkatkan deteksi, tetapi juga menurunkan *miss rate*. AMR lebih rendah pada kelompok *CADe-first* dibandingkan *high-definition white light-first*, yaitu 20,12% vs 31,25%.⁸ Dengan demikian, CADE memiliki potensi klinis sebagai alat peningkatan mutu kolonoskopi, khususnya untuk mengurangi lesi yang terlewat.

Selain itu, Shaukat et al. menunjukkan bahwa SKOUT CADE meningkatkan *adenoma per colonoscopy* (1,05 vs 0,83; $p=0,002$), meskipun peningkatan ADR hanya bersifat numerik dan belum mencapai signifikansi statistik (47,8% vs 43,9%; $p=0,065$).⁹ Temuan ini melengkapi bukti bahwa efek CADE dapat lebih terlihat pada jumlah adenoma yang ditemukan per prosedur dibandingkan hanya pada proporsi pasien dengan setidaknya satu adenoma. Soons et al. juga memberikan bukti *feasibility dan safety* awal penggunaan *CADe real-time* di pusat tersier Eropa, dengan PDR 55,6%, ADR 28,9%, estimasi *false positive* yang relatif terkendali, tidak adanya *serious adverse event* terkait CADE, serta penilaian *user friendliness* yang baik.¹⁰ Dengan demikian, bukti CADE mencakup tidak hanya efektivitas deteksi, tetapi juga kelayakan integrasi sistem dalam alur kerja endoskopi.

Karakterisasi Lesi dan *Optical Biopsy* dengan CADx

Selain mendeteksi lesi, AI juga digunakan untuk memprediksi karakteristik histologis polip melalui CADx. Byrne et al. menunjukkan bahwa model CNN dapat membedakan polip adenomatosa dan hiperplastik dengan akurasi 94%, sensitivitas 98%, spesifisitas 83%, dan NPV 97%.¹¹ Hasil tersebut mendukung potensi AI untuk *optical biopsy*, terutama pada polip diminutif.

Mori et al. menunjukkan penggunaan AI secara *real-time* dalam identifikasi polip diminutif selama kolonoskopi, sedangkan Hassan et al. melaporkan NPV 97,6% untuk polip rektosigmoid kecil sehingga membuka peluang strategi *leave-in-situ*.^{12,13} Meskipun menjanjikan, penerapan CADx harus dilakukan hati-hati karena keputusan tidak melakukan reseksi atau tidak mengirim jaringan ke patologi membutuhkan akurasi yang tinggi dan validasi luas.

Di Indonesia, Asputra et al. menunjukkan performa CAD Eye dengan sensitivitas 79,17%, spesifisitas 75,00%, PPV 89,06%, NPV 58,33%, dan akurasi 78,00%.²³ Temuan ini penting karena memperlihatkan bahwa performa AI pada praktik klinis nyata dapat berbeda dari studi di pusat rujukan besar, sehingga validasi lokal tetap diperlukan.

Aplikasi AI di Luar Kolonoskopi

Pada saluran cerna atas, Hirasawa et al. mengembangkan CNN untuk deteksi kanker lambung dan melaporkan sensitivitas 92,2% dengan waktu analisis cepat terhadap 2.296 gambar uji.¹⁴ Luo et al. mengembangkan GRAIDS dalam studi diagnostik multicenter dan menemukan akurasi tinggi untuk mendeteksi kanker saluran cerna atas, dengan sensitivitas yang serupa ahli endoskopi dan lebih baik dibandingkan endoskopis yang belum ahli.¹⁵ Hal ini menunjukkan bahwa AI berpotensi membantu pemerataan kualitas diagnosis, terutama pada pusat layanan dengan keterbatasan tenaga ahli.

Masih pada endoskopi saluran cerna atas, de Groof et al. menunjukkan kelayakan penggunaan CAD berbasis deep learning secara live untuk mendeteksi neoplasia Barrett pada *white-light endoscopy*.¹⁸ Pada studi pilot tersebut, sistem mencapai akurasi 90%, sensitivitas 91%, dan

spesifisitas 89% pada analisis per-level. Walaupun sampelnya masih kecil, temuan ini memperluas aplikasi AI dari deteksi kanker lambung dan kanker saluran cerna atas secara umum menuju identifikasi lesi prakanker atau neoplasia spesifik seperti Barrett esophagus.

Pada kapsul endoskopi, Ding et al. menunjukkan bahwa CNN mampu meningkatkan sensitivitas deteksi abnormalitas dan menurunkan waktu baca dari 96,6 menit menjadi 5,9 menit per pasien.¹⁶ Aplikasi ini sangat relevan karena kapsul endoskopi menghasilkan volume gambar yang besar dan membutuhkan waktu interpretasi yang panjang.

Pada *inflammatory bowel disease*, Takenaka et al. mengembangkan deep neural network untuk evaluasi *ulcerative colitis* berdasarkan citra endoskopi dan data histologis.¹⁷ AI berpotensi membantu standarisasi penilaian aktivitas penyakit, mengurangi variasi antar penilai, dan mendukung pemantauan objektif.

Pada hepatologi, AI berbasis ultrasonografi juga berkembang. Lee et al. menggunakan DCNN untuk klasifikasi fibrosis hati berdasarkan skor METAVIR, sedangkan Li et al., Yao et al., dan Du et al. menunjukkan pemanfaatan radiomics atau machine learning untuk diagnosis dan karakterisasi lesi hepatobilier.¹⁹⁻²² Dengan demikian, penerapan AI dalam gastroenterologi semakin bergerak dari deteksi endoskopik menuju integrasi multimodal berbasis citra dan data klinis.

Tantangan Implementasi Klinis

Walaupun hasil berbagai penelitian primer menunjukkan potensi besar, implementasi AI masih menghadapi beberapa tantangan berupa *false positive* yang dapat mengganggu alur kerja endoskopi dan menimbulkan alarm fatigue; hal ini relevan dengan studi feasibility Soons et al. yang mengevaluasi CADE *real-time* serta melaporkan sistem relatif aman, dapat digunakan dalam praktik klinis, dan memiliki estimasi *false positive* yang masih terkendali.¹⁰ Performa AI tidak selalu konsisten pada seluruh indikator klinis; Shaukat et al. menunjukkan peningkatan adenomas per colonoscopy yang signifikan, tetapi peningkatan ADR belum mencapai signifikansi statistik.⁹ Model AI yang dikembangkan pada dataset atau pusat

layanan tertentu belum tentu memberikan performa serupa pada populasi, perangkat endoskopi, kualitas citra, dan setting layanan yang berbeda. Hal ini terlihat dari studi multicenter Glissen Brown et al., studi live endoscopy de Groof et al., serta data Asputra et al. yang menegaskan pentingnya validasi eksternal dan validasi lokal sebelum implementasi luas.^{8,18,23} Selain itu, penerapan CADx untuk strategi klinis seperti *leave-in-situ* dan *resect-and-discard* memerlukan kehati-hatian karena keputusan tersebut membutuhkan nilai prediksi negatif dan akurasi yang sangat tinggi, sebagaimana ditunjukkan dalam studi Hassan et al.¹³

Sebagian besar studi masih menilai luaran antara seperti ADR, PDR, APC, AMR, sensitivitas, dan spesifisitas. Dampak jangka panjang AI terhadap penurunan kanker interval, mortalitas, biaya layanan, serta kepuasan pasien yang masih memerlukan penelitian lanjutan. Aspek regulasi, tanggung jawab klinis, keamanan data, transparansi algoritma, dan pelatihan operator harus disiapkan sebelum implementasi luas.

Dengan demikian, AI sebaiknya diposisikan sebagai alat bantu klinis yang memperkuat performa endoskopis, bukan menggantikan peran klinisi. Keputusan akhir tetap memerlukan integrasi antara hasil AI, gambaran klinis, temuan endoskopi, pengalaman operator, dan preferensi pasien.

SIMPULAN

Artificial intelligence (AI) memiliki potensi besar dalam meningkatkan kualitas diagnosis dan pelayanan gastroenterologi. Berdasarkan sintesis penelitian primer, bukti paling kuat terdapat pada penggunaan CADE dalam kolonoskopi untuk meningkatkan deteksi polip dan adenoma serta menurunkan *miss rate*. CADx juga menjanjikan untuk karakterisasi polip dan *optical biopsy*, tetapi penerapannya masih memerlukan validasi eksternal dan standar klinis yang ketat.

Aplikasi AI selain kolonoskopi, seperti deteksi kanker saluran cerna atas, pembacaan kapsul endoskopi, evaluasi ulcerative colitis, dan diagnosis hepatobilier berbasis ultrasonografi, menunjukkan perkembangan yang positif. Namun, AI saat ini lebih tepat diposisikan sebagai alat bantu klinis, bukan pengganti klinisi. Implementasi luas memerlukan

validasi populasi, regulasi, pelatihan operator, evaluasi biaya-manfaat, serta pengawasan etika dan keamanan data.

DAFTAR PUSTAKA

1. Pannala R, Krishnan K, Melson J, Parsi MA, Schulman AR, Sullivan S, et al. Artificial intelligence in gastrointestinal endoscopy. *VideoGIE*. 2020;5(12):598–613.
2. Kaminski MF, Wieszczy P, Rupinski M, Wojciechowska U, Didkowska J, Kraszewska E, et al. Increased rate of adenoma detection associates with reduced risk of colorectal cancer and death. *Gastroenterology*. 2017;153(1):98–105.
3. Zhao S, Wang S, Pan P, Xia T, Chang X, Yang X, et al. Magnitude, risk factors, and factors associated with adenoma miss rate of tandem colonoscopy: a systematic review and meta-analysis. *Gastroenterology*. 2019;156(6):1661–1674.e11.
4. Wang P, Berzin TM, Glissen Brown JR, Bharadwaj S, Becq A, Xiao X, et al. Real-time automatic detection system increases colonoscopic polyp and adenoma detection rates: a prospective randomised controlled study. *Gut*. 2019;68(10):1813–1819.
5. Repici A, Badalamenti M, Maselli R, Correale L, Radaelli F, Rondonotti E, et al. Efficacy of real-time computer-aided detection of colorectal neoplasia in a randomized trial. *Gastroenterology*. 2020;159(2):512–520.e7.
6. Liu P, Wang P, Glissen Brown JR, Berzin TM, Zhou G, Liu W, et al. The single-monitor trial: an embedded CAdE system increased adenoma detection during colonoscopy: a prospective randomized study. *Therap Adv Gastroenterol*. 2020;13.
7. Xu L, He X, Zhou J, Zhang J, Mao X, Ye G, et al. Artificial intelligence-assisted colonoscopy: a prospective, multicenter, randomized controlled trial of polyp detection. *Cancer Med*. 2021;10(20):7184–7193.
8. Glissen Brown JR, Mansour NM, Wang P, Chuchuca MA, Minchenberg SB, Chandnani M, et al. Deep learning computer-aided polyp detection reduces adenoma miss rate: a United States multi-center randomized tandem colonoscopy study (CAdE-T-CS Trial). *Clin Gastroenterol Hepatol*. 2022;20(7):1499–1507.e4.
9. Shaukat A, Lichtenstein DR, Somers SC, Chung DC, Perdue DG, Gopal M, et al. Computer-aided detection improves adenomas per colonoscopy for screening and surveillance colonoscopy: a randomized trial. *Gastroenterology*. 2022;163(3):732–741.
10. Soons E, Rath T, Hazewinkel Y, van Dop WA, Esposito D, Testoni PA, et al. Real-time colorectal polyp detection using a novel computer-aided detection system (CAdE): a feasibility study. *Int J Colorectal Dis*. 2022;37(10):2219–2228.
11. Byrne MF, Chapados N, Soudan F, Oertel C, Pérez ML, Kelly R, et al. Real-time differentiation of adenomatous and hyperplastic diminutive colorectal polyps during analysis of unaltered videos of standard colonoscopy using a deep learning model. *Gut*. 2019;68(1):94–100.
12. Mori Y, Kudo SE, Misawa M, Saito Y, Ikematsu H, Hotta K, et al. Real-time use of artificial intelligence in identification of diminutive polyps during colonoscopy. *Ann Intern Med*. 2018;169(6):357–366.
13. Hassan C, Balsamo G, Lorenzetti R, Zullo A, Antonelli G. Artificial intelligence allows leaving-in-situ colorectal polyps. *Clin Gastroenterol Hepatol*. 2022;20(11):2505–2513.e4.
14. Hirasawa T, Aoyama K, Tanimoto T, Ishihara S, Shichijo S, Ozawa T, et al. Application of artificial intelligence using a convolutional neural network for detecting gastric cancer in endoscopic images. *Gastric Cancer*. 2018;21(4):653–660.
15. Luo H, Xu G, Li C, He L, Luo L, Wang Z, et al. Real-time artificial intelligence for detection of upper gastrointestinal cancer by endoscopy: a multicentre, case-control, diagnostic study. *Lancet Oncol*. 2019;20(12):1645–1654.
16. Ding Z, Shi H, Zhang H, Meng L, Fan M, Han C, et al. Gastroenterologist-level identification

- of small-bowel diseases and normal variants by capsule endoscopy using a deep-learning model. *Gastroenterology*. 2019;157(4):1044–1054.e5.
17. Takenaka K, Ohtsuka K, Fujii T, Negi M, Suzuki K, Shimizu H, et al. Development and validation of a deep neural network for accurate evaluation of endoscopic images from patients with ulcerative colitis. *Gastroenterology*. 2020;158(8):2150–2157.
18. de Groof AJ, Struyvenberg MR, Fockens KN, van der Putten J, van der Sommen F, Boers TG, et al. Deep learning algorithm detection of Barrett's neoplasia with high accuracy during live endoscopic procedures: a pilot study (with video). *Gastrointest Endosc*. 2020;91(6):1242–1250.
19. Lee JH, Joo I, Kang TW, Paik YH, Sinn DH, Ha SY, et al. Deep learning with ultrasonography: automated classification of liver fibrosis using a deep convolutional neural network. *Eur Radiol*. 2020;30(2):1264–1273.
20. Li W, Lv XZ, Zheng X, Ruan SM, Hu HT, Chen LD, et al. Machine learning-based ultrasomics improves the diagnostic performance in differentiating focal nodular hyperplasia and atypical hepatocellular carcinoma. *Front Oncol*. 2021;11:544979.
21. Yao Z, Dong Y, Wu G, Zhang Q, Yang D, Yu JH, et al. Preoperative diagnosis and prediction of hepatocellular carcinoma: radiomics analysis based on multi-modal ultrasound images. *BMC Cancer*. 2018;18(1):1089.
22. Du Z, Fan F, Ma J, Liu J, Yan X, Chen X, et al. Development and validation of an ultrasound-based interpretable machine learning model for the classification of ≤ 3 cm hepatocellular carcinoma: a multicentre retrospective diagnostic study. *eClinicalMedicine*. 2025;81:103098.
23. Asputra H, Maulahela H, Fauzi A, Rumende CM, Pitarini A, Sari NK, et al. Diagnostic value of artificial intelligence computer-assisted diagnosis (computer assisted-diagnosis eye function) for colorectal polyps. *Artif Intell Gastroenterol*. 2025;6(3).