SUPPLEMENTAL MATERIALS Artificial intelligence-assisted colonoscopy in real-world clinical practice: A systematic review and meta-analysis

Mike Tzuhen Wei¹*, Shmuel Fay²*, Diana Yung³, Uri Ladabaum¹**, Uri Kopylov²**

1-Division of Gastroenterology and Hepatology, Stanford University School of Medicine, Stanford, California
2- Department of Gastroenterology, Sheba Medical Center, Ramat Gan, Israel, and Tel Aviv University Medical School
3- Gold Coast Hospital and Health Service, Gold Coast, Australia
*-shared first authorship
**-shared last authorship

Supplementary Table 1: Search strategies used for the meta-analysis

Database	Search Strategy
Pubmed	("artificial intelligence"[All Fields] OR "computer-assisted"[All Fields] OR
	"computer-aided"[All Fields] OR "computer-assisted"[All Fields]) AND
	("colonoscopy"[MeSH Terms] OR "colonoscopy"[All Fields] OR
	"colonoscopies"[All Fields]) AND ("polyps"[All Fields] OR "polypous"[All
	Fields] OR "polyps"[MeSH Terms] OR "polyp"[All Fields] OR
	("adenoma"[MeSH Terms] OR "adenoma"[All Fields] OR "adenomas"[All
	Fields]))
EMBASE	('artificial intelligence'/exp OR 'artificial intelligence' OR 'computer-assisted' OR
	'computer-aided') AND ('colonoscopy'/exp OR 'colonoscopy' OR 'colonoscopies')
	AND ('polypous' OR 'polyps'/exp OR 'polyps' OR 'polyp'/exp OR 'polyp' OR
	'adenoma'/exp OR 'adenoma' OR 'adenomas')

Study	Selection	Compatibility	Outcomes	Total Stars *
Quan ⁶	4	2	3	9
Koh ¹²	4	1	3	8
Ishiyama ¹¹	4	1	3	8
Shaukat ¹⁶	4	2	3	9
Richter ¹⁴	4	1	3	8
Nehme ¹³	4	2	3	9
Ahmad ¹⁸	4	1	3	8
Agazzi ¹⁹	4	1	3	8
Wong ¹⁷	4	2	3	9
Schauer ¹⁵	4	1	3	8
Levy ⁸	4	1	3	8
Ladabaum ⁷	4	2	3	9

Supplemental Table 2: Newcastle–Ottawa scale for assessing the quality of non-randomized studies

*Maximum of 9

Supplemental Table 3: Studies implementation strategy details

Study	Year	Publication type	Location	Study Design	Implementation strategy	Control	CADe used (N)	Without CADe (N)	CAD vs without CADe APC (p- value)	CAD vs without CADe ADR (p- value)
Quan	2022	Published full paper	United States	Multicenter prospective	Prospective single-arm, open-label pilot study with a historical control group.	Historical	EndoVigilant (N=300)	300	1.35 vs 1.07 (0.099)	52 vs 46.3 (0.165)
Koh	2022	Published full paper	Singapore	Single center prospective	Investigator-initiated prospective cohort study in a referral center during the trial period of an AI- assisted colonoscopy device. Specialist-grade endoscopists performed colonoscopies.	Historical	GI Genius (N=298)			30.4 vs 24.3 (0.02)
Ishiyama	2021	Published full paper	Japan	Single center prospective	Propensity score-matched, prospective, non- blinded, registered as a clinical trial. CADe system implemented in 4 out of 8 colonoscopy rooms, allocation based on room availability. CADe system activated after cecal intubation in CADe group. Endoscopists performed withdrawal with or without CADe system based on group assignment.	Concurrent	EndoBRAINEYE (N=918)	918	0.42 vs 0.3, (0.003)	26.4 vs 19.9 (0.001)
Shaukat	2022	Published full paper	United States	Single center prospective	Prospective pilot study was conducted for 6 weeks by three experienced providers. Compared to historical cohort. This was the first clinical pilot of this CADe.	Historical	Skout (N=83)	283	1.46 vs 1.01 (0.104)	54.2 vs 40.6 (0.028)

Richter	2023	Published full paper	Germany	Single center retrospective	CAD group data collected forward from the time of AI system inauguration in routine practice. Compared to historical cohort.	Historical	CADEye (N=163)	140		39 vs 41 (>0.05)
Nehme	2021	Published full paper	United States	Single center prospective	Retrospective analysis of a prospectively maintained database of patients undergoing colonoscopy at a tertiary center, before and after a CADe system was made available. The decision to activate the CADe system was at the discretion of the endoscopist.	Historical	GI Genius (N=403)	641	1.27 vs 1.17, (0.45)	50.4 vs 53 (0.41)
Ahmad	2021	Abstract	England	Single center prospective	A prospective study, in one month period, all Historical GI Genius 86 colonoscopies were performed with the addition of (N=82) a CAD. Endoscopists completed an evaluation form after the procedure. Compared to the previous month's cohort.			48.8 vs 46.5 (0.77)		
Agazzi	2022	Abstract	Italy	Single center prospective	Retrospective, single-center cross-sectional study conducted in an open access high volume endoscopy unit.	Historical	CADEye (N=250)	450		46 vs 30.7 (<0.005)
Wong	2022	Abstract	Hong Kong	Single center retrospective	Retrospective study comparing outcomes before and after ENDO-AID implementation. Single experienced endoscopist. Minimalist deployment strategy.	Historical	ENDO-AID (N=119)	115		52.9 vs 37.4 (0.017)
Schauer	2021	Published full paper	New Zealand	Single center retrospective	A single-center retrospective, real-world setting, study. Endoscopists were able to toggle AI on-and- off at their discretion. The Endo-AID equipment from Olympus was loaned free of charge. There was no prior experience or training provided for the software.	Historical	ENDO-AID (N=213)	213		47.9 vs 38.5 (0.03)
Levy	2022	Published full paper	Israel	Single center retrospective	Retrospective observational study evaluating deployment of GI Genius at an academic tertiary medical center. Data collected for 6 months in 2021 with comparison to historical data for corresponding 6 months in 2020.	Historical	GI Genius (N=1,969)	2175	0.6 vs 0.68 (0.001)	30.3 vs 35.2 (0.001)
Ladabaum	2022	Published full letter	United States	Single center retrospective	Real-world setting, retrospective observational study in high-volume tertiary medical center. For 3 months, the CAD was switched on routinely at the start of every colonoscopy withdrawal. Historical and concurrent non-CADe controls. A representative mix of patients and physicians.	Historical and concurrent	GI Genius (N=619)	619	0.78 vs 0.89 (0.63)	40.1 vs 41.8 (0.41)

CADe: computer-assisted detection; ADR: Adenoma detection rate; APC: Adenoma per colonoscopy

Supplemental Figure 1: Pooled ADR risk ratio of only full studies

	CAE)	Contr	ol		Risk Ratio		Risk Ratio
Study or Subgroup	Events	Total	Events	Total	Weight I	M-H, Random, 95% Cl	Year	M-H, Random, 95% Cl
1.1.1 Prospective studi	es							
Shaukat et al 2021	45	83	115	283	8.7%	1.33 [1.05, 1.70]	2021	_
Ishiyama et al 2022	242	918	183	918	10.8%	1.32 [1.12, 1.57]	2022	
Quan et al 2022	156	300	139	300	10.9%	1.12 [0.95, 1.32]	2022	+
Koh et al 2022	91	298	72	298	8.2%	1.26 [0.97, 1.65]	2023	
Nehme et al 2023	203	403	340	641	12.1%	0.95 [0.84, 1.07]	2023	
Subtotal (95% CI)		2002		2440	50.7%	1.17 [1.01, 1.36]		-
Total events	737		849					
Heterogeneity: Tau ² = 0.	.02; Chi ž :	= 13.93	, df = 4 (F	P = 0.00)8); I ² = 719	6		
Test for overall effect: Z	= 2.03 (P	= 0.04))					
1.1.2 Retrospective stu	dies							
Levy et al 2022	597	1969	766	2175	12.9%	0.86 [0.79, 0.94]	2022	
Richter et al 2022	64	163	57	140	7.9%	0.96 [0.73, 1.27]	2022	
Schauer et al 2022	102	213	82	213	9.3%	1.24 [1.00, 1.55]	2022	
Wong et al 2022	63	119	43	115	7.5%	1.42 [1.06, 1.89]	2022	
Ladabaum et al 2023	248	619	259	619	11.8%	0.96 [0.84, 1.09]	2023	
Subtotal (95% CI)		3083		3262	49.3%	1.04 [0.88, 1.23]		
Total events	1074		1207					
Heterogeneity: Tau ² = 0.	.03; Chi ž :	= 17.83	, df = 4 (F	° = 0.00)1); I² = 789	6		
Test for overall effect: Z	= 0.45 (P	= 0.65))					
Total (95% CI)		5085		5702	100.0%	1.11 [0.98, 1.24]		-
Total events	1811		2056					
Heterogeneity: Tau ² = 0.	.03; Chi ² :	= 42.20	, df = 9 (F	° < 0.00	0001); I² = 7	'9%		
Test for overall effect: Z	= 1.67 (P	= 0.10))					Favours conventional endo Favours CAD
Toot for oubgroup differ	onees: Ol	hi z _ 1 ∩	10 df = 1	$(D = 0)^{1}$	201 12 - 5 7	'0/		

Test for subgroup differences: Chi² = 1.06, df = 1 (P = 0.30), l² = 5.7% ADR: Adenoma detection rate; CADe: computer-assisted detection

Supplemental Figure 2: Pooled APC evaluating prospective studies only

				Rate Ratio		Rate Ratio
Study or Subgroup	log[Rate Ratio]	SE	Weight	IV, Random, 95% Cl	Year	IV, Random, 95% Cl
Shaukat et al 2021	0.36640502	0.10844799	19.5%	1.44 [1.17, 1.78]	2021	
Quan et al 2022	0.23244594	0.07472885	25.9%	1.26 [1.09, 1.46]	2022	
Ishiyama et al 2022	0.33906627	0.07891158	25.1%	1.40 [1.20, 1.64]	2022	_
Nehme et al 2023	0.08234431	0.05732764	29.5%	1.09 [0.97, 1.21]	2023	+
Total (95% CI)			100.0%	1.27 [1.11, 1.46]		•
Heterogeneity: Tau ² =	0.01; Chi² = 9.80,	df = 3 (P = 0.0	2); I ² = 69	%		
Test for overall effect:	Z = 3.45 (P = 0.000	J6)				Favours conventional endo Favours CAD

APC: Adenoma per colonoscopy; CADe: computer-assisted detection

Supplemental Figure 3: Pooled APC risk ratio of full studies evaluating only GI Genius

				Rate Ratio		Rate Ratio
Study or Subgroup	log[Rate Ratio]	SE	Weight	IV, Random, 95% Cl	Year	IV, Random, 95% Cl
Levy et al 2022	-0.12550178	0.039024	36.8%	0.88 [0.82, 0.95]	2022	
Ladabaum et al 2023	-0.13171816	0.06233198	30.9%	0.88 [0.78, 0.99]	2023	
Nehme et al 2023	0.08234431	0.05732764	32.2%	1.09 [0.97, 1.21]	2023	+
Total (95% CI)			100.0%	0.94 [0.82, 1.08]		-
Heterogeneity: Tau² = 0 Test for overall effect: Z	.01; Chi² = 10.04, = 0.89 (P = 0.37)	df= 2 (P = 0.0		0.5 0.7 1 1.5 2 Favours conventional endo Favours CAD		

APC: Adenoma per colonoscopy; CADe: computer-assisted detectio