

Incidence and Disparities in Colorectal Cancer After Endoscopic Mucosal Resection: A Population-Based Cohort Study

Mohamed H. Eldesouki

meldesouki@nymc.edu

New York Medical College

Khaled Elfert

West Virginia University

Mohammed Y. Youssef

Hunt Regional Medical center

Aasma Shaukat

New York University

Shyam Thakkar

West Virginia University

Research Article

Keywords: endoscopic mucosal resection, colorectal cancer, incidence, racial disparities, surveillance, TriNetX, real-world data

Posted Date: April 23rd, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-9228195/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Background

Endoscopic mucosal resection (EMR) is the standard treatment for complex colorectal lesions, offering an effective minimally invasive alternative to surgery. However, the incidence of colorectal cancer (CRC) arising after EMR is incompletely characterized across diverse real-world populations.

Methods

We conducted a retrospective cohort study using the TriNetX US Collaborative Network. Adult patients (≥ 18 years) who underwent colonoscopy or sigmoidoscopy with EMR between January 2016 and January 2017 were included. Patients with a prior diagnosis of inflammatory bowel disease (IBD) or familial adenomatous polyposis (FAP) were excluded. The primary outcome was incident CRC, observed annually from 2018 through 2021. Incidence rates were calculated as cases per 1,000 person-years (PYO), stratified by age group, sex, race, and ethnicity. Temporal trends were assessed using annual percentage change (APC) analysis.

Results

A total of 4,187 patients were included. The overall post-EMR CRC incidence was 11.85 per 1,000 person-years. Incidence was higher in men (13.61) than women (10.28). Marked racial and ethnic disparities were observed: Hispanic/Latino patients had the highest incidence (28.30), compared to non-Hispanic/non-Latino individuals (12.10). By race, Asian patients had the highest incidence (23.38), followed by Black (11.81) and White patients (10.78). Age-stratified analysis identified the highest incidence in the 40–49-year group (42.2 per 1,000 PYO), with lower rates in the 50–59 (7.8) and 60–69 (9.4) groups, increasing again to 16.8 in the 70–79-year group. Temporal analysis revealed a 27.6% decline in incidence from 30.1 in 2018 to 21.8 per 1,000 PYO in 2021.

Conclusion

Post-EMR CRC incidence varied substantially across racial, ethnic, sex, and age groups, with notable disparities disproportionately affecting Hispanic/Latino and Asian populations, as well as younger adults. These findings underscore the need for risk-stratified, equity-focused surveillance protocols following EMR and call for prospective research integrating procedural, demographic, and molecular data.

Introduction

Colorectal cancer (CRC) remains the second leading cause of cancer-related mortality in the United States, with an estimated 154,270 new cases and 52,180 deaths projected in 2023 [1]. Despite significant advances in screening and treatment, CRC incidence and mortality rates continue to demonstrate persistent racial, ethnic, and socioeconomic disparities [1, 2]. A critical component of the CRC prevention strategy involves endoscopic detection and removal of precancerous and early cancerous lesions, with colonoscopy-based polypectomy reducing CRC incidence by as much as 76–90% in national cohort studies [3].

Endoscopic mucosal resection (EMR) has emerged as the preferred technique for the management of complex colorectal lesions, including large (typically ≥ 20 mm), flat, sessile, or laterally spreading tumors that cannot be safely removed by conventional polypectomy [4]. EMR provides a minimally invasive, organ-sparing alternative to surgical segmental resection, substantially reducing procedural morbidity and healthcare costs [4]. Prior studies have established the efficacy of EMR for large non-pedunculated colorectal polyps (LNPCPs), demonstrating high technical success rates and acceptable recurrence profiles with structured surveillance [5, 6].

Despite this established efficacy, a clinically meaningful subset of patients develops colorectal cancer after EMR, termed post-EMR CRC, which we defined as CRC diagnosed after EMR but before the next recommended surveillance colonoscopy. Post-EMR CRC is a composite quality indicator that may reflect missed synchronous lesions at the time of the index procedure, incomplete resection with residual or locally recurrent neoplasia, or delayed, non-adherent, or inadequate post-procedural surveillance [7, 8]. Understanding the incidence and risk factors of post-EMR CRC is therefore essential for optimizing EMR technique, defining appropriate surveillance intervals, and identifying patients at highest risk who may benefit from intensified follow-up.

The incidence of interval CRC (iCRC) following negative colonoscopy is increasingly characterized in large population-based studies. A prior analysis demonstrated that Black Americans had a 31% higher risk of interval CRC compared with White Americans [2]. A recent population-based study reported a CRC incidence rate of approximately 2.9 per 1,000 person-years within five years after a negative colonoscopy [9]. However, data specific to the post-EMR setting is considerably more limited. Unlike post-polypectomy populations in general, patients undergoing EMR represent a distinctly higher-risk cohort with large, morphologically complex lesions harboring advanced neoplasia, where incomplete resection and local recurrence carry heightened oncologic significance.

Existing evidence on post-EMR outcomes and surveillance has been predominantly derived from small, single- or multi-center cohort studies conducted in tertiary academic endoscopy centers with highly selected, protocol-adherent patient populations [10, 11]. Two recent real-world studies examining post-EMR surveillance adherence in diverse populations reported adherence rates of only 36.6–49.8%, driven by socioeconomic and geographic barriers including transportation limitations, lower educational attainment, language barriers, and inadequate care navigation [10, 11]. Importantly, neither study evaluated the downstream incidence of post-EMR CRC in their populations. In this study, we leveraged

the TriNetX US Collaborative Network to evaluate the real-world incidence of post-EMR CRC over a five-year follow-up period in a large, multicenter population. We further examined the extent to which incidence varied by race, ethnicity, sex, and age group.

Methods

Study Design and Data Source

We conducted a retrospective population-based cohort study using the TriNetX US Collaborative Network, a federated real-world data (RWD) and analytics platform operated by TriNetX, Inc. (Cambridge, MA, USA). As of the study period, the TriNetX US Collaborative Network aggregated de-identified electronic health record (EHR) data from 69 U.S. healthcare organizations (HCOs), including both academic medical centers and community-based hospitals and outpatient practices, encompassing over 80 million patients. The platform integrates longitudinal patient data from inpatient and outpatient encounters, including structured data elements such as diagnoses, procedures, medications, laboratory results, and demographics [14–16]. Data from participating HCOs are coded using standardized terminologies and are harmonized within the federated architecture without transferring raw individual patient records, thereby maintaining Health Insurance Portability and Accountability Act (HIPAA) compliance [14–16].

The TriNetX platform enables cohort construction via a query-builder interface applying Boolean inclusion and exclusion criteria referencing ICD-10-CM diagnostic codes, Current Procedural Terminology (CPT) procedural codes, and demographic variables. Outcome ascertainment is performed by querying for the occurrence of specified diagnosis codes within user-defined time windows following the index event. Because the platform provides only aggregated, de-identified summary statistics rather than individual-level records, institutional review board (IRB) exemption was granted per the TriNetX data use agreement and applicable federal regulations governing research with de-identified data [14–16].

Patient Identification and Cohort Construction

Adult patients aged 18 years or older who underwent colonoscopy or sigmoidoscopy with EMR were identified using the following CPT codes: 44403 (colonoscopy with EMR), 45390 (colonoscopy through stoma with EMR), and 45349 (sigmoidoscopy with EMR). The index EMR procedure date was required to fall between January 1, 2016 and January 31, 2017 to ensure a minimum of five years of potential follow-up through December 31, 2021. This index window was selected to capture a contemporary, post-ICD-10 transition cohort (ICD-10-CM implementation occurred in October 2015) and to avoid data inconsistencies from the ICD-9 to ICD-10 conversion period. Patients were excluded from the cohort if they had a documented diagnosis of inflammatory bowel disease (IBD), including Crohn's disease (ICD-10: K50) or ulcerative colitis (K51.x), or familial adenomatous polyposis (FAP; ICD-10: D12.6, Z84.81), given the substantially elevated CRC risk associated with these conditions and their distinct surveillance

protocols [8]. Patients with a CRC diagnosis (ICD-10-CM: C18, C19, C20) recorded prior to the index EMR date were also excluded to eliminate prevalent CRC cases and to ensure that the cohort consisted of patients undergoing EMR for precancerous or early-stage lesions rather than known invasive cancer.

Outcome Definition and Follow-Up Period

The primary outcome was incident CRC, defined as a new ICD-10-CM diagnosis of CRC arising in the colon (C18), rectosigmoid junction (C19), or rectum (C20) recorded in the EHR during the follow-up period. Incident CRC diagnoses were queried annually from January 1, 2018 through December 31, 2021, corresponding to years two through five after the EMR index window. A one-year washout period was applied between the index EMR (2016–2017) and the start of outcome ascertainment (2018) to exclude any CRC diagnoses that may represent synchronous, pre-existing, or rapidly developing (peri-procedural) cancers that were present at or immediately following the EMR. This approach is consistent with methodological standards applied in interval CRC literature, where a minimum latency of six to twelve months is typically used to separate synchronous from true post-procedural CRC [2, 9]. Metachronous CRC, arising at a site anatomically distinct from the EMR resection site, could not be differentiated from locally recurrent or residual neoplasia at the EMR site due to the absence of lesion-specific anatomic data within the TriNetX platform.

Incidence Rate Calculation

The incidence rate of post-EMR CRC was calculated as the number of new CRC cases per 1,000 person-years of observation (PYO), consistent with standard epidemiologic methods for estimating disease incidence in prospectively followed cohorts. Person-time at risk was computed as the sum of observation time from the start of follow-up (January 1, 2018) to the date of CRC diagnosis, death, loss to follow-up, or end of the observation window (December 31, 2021), whichever occurred first. Within the TriNetX platform, person-years were derived from available encounter data by summing the time each patient remained in the active EHR system during the observation window, with censoring applied at the last available healthcare encounter for patients without a subsequent CRC diagnosis. Annual incidence rates were calculated for each calendar year (2018, 2019, 2020, 2021) to characterize temporal trends.

Subgroup Analyses and Stratification

Incidence rates were stratified by the following prespecified demographic variables: (1) sex (male, female); (2) race (White, Black or African American, Asian); (3) ethnicity (Hispanic/Latino, non-Hispanic/non-Latino); and (4) age group at the time of index EMR, categorized as 40–49 years, 50–59 years, 60–69 years, 70–79 years, and ≥ 80 years. Subgroup analyses were performed using the TriNetX platform's built-in demographic filtering capabilities applied to the identified cohort.

Temporal Trend Analysis

Annual incidence rates were calculated for each calendar year from 2018 through 2021 to characterize the temporal pattern of post-EMR CRC risk over the follow-up period. The annual percentage change (APC) was estimated as $[(\text{Final incidence} - \text{Initial incidence}) / \text{Initial incidence}] \times 100$, representing the relative change in incidence across the observation period.

Statistical Analysis

Descriptive statistics were used to characterize the patient cohort, including proportions for categorical variables. Incidence rates with 95% confidence intervals were reported where data allowed. Statistical significance was defined as a two-tailed p-value of ≤ 0.05 . All analyses were performed within the TriNetX analytics platform, which applies aggregation and rounding protocols, including reporting any cell count between 1 and 10 as “10” to protect patient privacy, consistent with HIPAA de-identification standards [14–16]. This rounding convention was taken into account when interpreting subgroup counts and rate estimates.

Results

Patient Characteristics

A total of 4,187 patients who underwent EMR between January 2016 and January 2017 and met all eligibility criteria were identified from the TriNetX US Collaborative Network. The mean age of the cohort was 73 ± 10 years. Age distribution was as follows: 40–49 years ($n = 302$, 6.3%), 50–59 years ($n = 474$, 9.9%), 60–69 years ($n = 1,658$, 34.6%), 70–79 years ($n = 1,577$, 32.9%), and ≥ 80 years ($n = 785$, 16.4%). Most patients undergoing EMR were in their sixth and seventh decades of life, aligning with the typical age range for complex colorectal adenomas requiring EMR. Females accounted for 51.2% of the cohort, while males comprised 48.8%. The racial distribution included 77.3% White patients, followed by 12% Black, 4.3% Asian, 1.2% Native Hawaiian, and 0.7% American Indian individuals. Metabolic syndrome was common, affecting 74% of patients, and obesity was present in 34%. Diabetes mellitus was reported in 30% of the cohort, hepatic steatosis in 10%, and nicotine dependence in 24%.

Overall and Sex-Stratified Incidence

The overall incidence of post-EMR CRC was 11.85 cases per 1,000 person-years over the five-year observation period. When stratified by sex, the incidence was higher among men (13.61 per 1,000 PYO) than women (10.28 per 1,000 PYO), reflecting a pattern consistent with the known male predominance in CRC incidence across most population studies.

Racial and Ethnic Disparities

Substantial variation in post-EMR CRC incidence was observed across racial and ethnic groups. By ethnicity, Hispanic/Latino patients had the highest incidence at 28.30 per 1,000 PYO, representing more than twice the rate observed among non-Hispanic/non-Latino patients (12.10 per 1,000 PYO). By race,

the highest incidence was observed among Asian patients (23.38 per 1,000 PYO), followed by Black or African American patients (11.81 per 1,000 PYO) and White patients (10.78 per 1,000 PYO).

Age-Stratified Incidence

Age-stratified analysis revealed a non-linear relationship between age and post-EMR CRC incidence. The highest incidence was observed among the youngest cohort, individuals aged 18–49 years, at 42.2 per 1,000 PYO – more than fourfold the overall cohort rate. Rates declined substantially in the 50–59-year group (7.8 per 1,000 PYO) and the 60–69-year group (9.4 per 1,000 PYO). Incidence increased again in the 70–79-year group, reaching 16.8 per 1,000 PYO. This age-related pattern is depicted in Fig. 1.

The markedly elevated incidence in the 40–49-year cohort likely reflects a combination of factors, including the smaller absolute number of EMRs performed in this age group during the study period ($n = 302$), the selection of younger patients for EMR based on symptomatic presentation or hereditary risk rather than average-risk screening, and the biologically aggressive features increasingly recognized in early-onset CRC.

Temporal Trends

Temporal analysis demonstrated a progressive decline in annual post-EMR CRC incidence across the observation period. Incidence decreased from 30.1 cases per 1,000 PYO in 2018 to 21.8 per 1,000 PYO in 2021, representing an approximately 27.6% relative reduction over four years. This declining trend, illustrated in Fig. 2, is consistent with the biological premise that post-procedural CRC risk is highest in the early post-EMR period, likely reflecting residual or incompletely resected neoplasia, and diminishes with increasing time and successful surveillance.

Discussion

In this large, multicenter retrospective cohort study, we characterized the incidence of CRC arising within five years of EMR and systematically evaluated disparities by race, ethnicity, sex, and age. The overall post-EMR CRC incidence of 11.85 per 1,000 person-years demonstrates a clinically meaningful cancer risk following EMR that is substantially higher than the 2.9 per 1,000 PYO reported after negative colonoscopy in population-based cohorts [9]. This excess risk underscores the importance of the post-EMR period as a distinct, high-risk surveillance window requiring dedicated management protocols. To our knowledge, this is the first study to characterize racial and ethnic disparities in post-EMR CRC incidence using real-world, multicenter data from a geographically diverse U.S. population.

The racial and ethnic gradient in post-EMR CRC incidence observed in our study is striking and differs from the pattern reported in iCRC studies after negative colonoscopy, where Black Americans have historically demonstrated the highest risk [2]. In our cohort, the highest incidence was observed in Hispanic/Latino patients (28.30 per 1,000 PYO) and Asian patients (23.38 per 1,000 PYO), while Black and White patients had more comparable rates (11.81 and 10.78, respectively). This divergence

suggests that the drivers of post-EMR CRC risk are distinct from those operating in the general post-colonoscopy interval CRC setting.

Among Hispanic/Latino patients, lower rates of timely endoscopic evaluation after abnormal stool-based screening and broader inequities across the CRC care continuum may contribute to delays in appropriate surveillance following treatment of high-risk lesions [17]. In addition, metabolic risk factors associated with advanced adenomas in Hispanic patients may increase the burden of high-risk precursor lesions in some individuals [18]. In Asian populations, barriers related to access, language, patient-provider communication, and health literacy may similarly delay detection and surveillance pathways [19]. These disparities likely reflect differences across the screening, treatment, and surveillance continuum rather than a single biologic mechanism. In contrast, prior interval CRC literature after negative colonoscopy has shown higher risk among Black individuals than White individuals [2], highlighting that the drivers of post-EMR CRC may differ from those in the general interval CRC setting.

The higher incidence of post-EMR CRC in men (13.61 per 1,000 PYO) compared to women (10.28 per 1,000 PYO) is broadly consistent with the well-documented male sex predominance in CRC risk and incidence across the general population [1]. While the sex-based differential in our cohort is modest, it provides further corroboration of the biological plausibility of our overall findings.

The markedly elevated post-EMR CRC incidence among individuals aged 40–49 years (42.2 per 1,000 PYO) represents one of the most clinically important findings of our study. Several explanations are likely to contribute to this observation. First, during most of the study period, average-risk CRC screening was recommended beginning at age 50 and was later lowered to age 45 in 2021 [20, 21]. Consequently, younger patients undergoing colonoscopy and EMR were not identified through routine average-risk screening but rather through symptomatic evaluation or high-risk surveillance. This selection effect biases the younger cohort toward higher intrinsic lesion risk and potentially toward lesions with more advanced histology at the time of EMR. Second, early-onset CRC (eoCRC) is increasingly recognized as a biologically and molecularly distinct entity. The incidence of eoCRC has risen dramatically since the 1990s, with projections suggesting that by 2030, eoCRC will account for 11% of colon cancers and 23% of rectal cancers in the United States [12]. These tumors demonstrate a predilection for distal colon and rectal locations, higher rates of advanced pathologic staging at diagnosis, and distinct molecular profiles including chromosomal instability, relative microsatellite stability, and alterations in CTNNB1 (beta-catenin) and ATM pathways that differ from late-onset CRC [13]. These biological features may confer a higher propensity for residual neoplasia following EMR and more rapid progression to cancer. Third, younger patients may be less likely to adhere to post-EMR surveillance colonoscopy schedules due to competing professional and personal priorities, underestimation of personal cancer risk, and lower baseline engagement with the healthcare system compared to older adults. This behavioral factor, combined with the biological aggressiveness of eoCRC, could substantially amplify post-EMR CRC risk in this age group. Our finding reinforces the growing consensus that EMR surveillance protocols and cancer risk counseling should be tailored to age and should incorporate the distinct risk profile of early-onset disease.

The progressive decline in annual post-EMR CRC incidence from 30.1 per 1,000 PYO in 2018 to 21.8 per 1,000 PYO in 2021, a 27.6% relative reduction, provides important insight into the time-course of post-procedural CRC risk. This pattern is consistent with the well-established biology of adenoma-to-carcinoma progression and with specific data from the EMR literature. In the ACE study, early recurrence at the EMR site was detected in 16.0% of lesions at 4-month surveillance, whereas late recurrence at 16 months occurred in only 4.0%, supporting the concept that residual or incompletely resected neoplasia is predominantly identified at early surveillance rather than representing de novo interval development [5]. Similarly, cohorts of large piecemeal EMR have demonstrated very low rates of post-polypectomy CRC during long-term follow-up once early surveillance has been completed [5, 6]. The declining incidence in our cohort may also reflect the incremental benefit of any surveillance that did occur during the follow-up period, even if adherence was suboptimal. Patients who did undergo surveillance colonoscopy at one to two years post-EMR were likely to have residual adenomatous tissue identified and removed, thereby interrupting adenoma-to-carcinoma progression and contributing to reduced CRC incidence in subsequent years. Notably, the decline in 2020–2021 may partially reflect the disruption of elective endoscopy and cancer screening services during the COVID-19 pandemic, which would tend to reduce detected CRC diagnoses through reduced diagnostic ascertainment rather than true incidence reduction.

The findings of this study carry direct implications for EMR practice and post-EMR surveillance policy. Current guidelines from the U.S. Multi-Society Task Force on CRC recommend first surveillance colonoscopy at three to six months after piecemeal resection of lesions ≥ 20 mm and individualized surveillance intervals thereafter based on resection quality and pathology [8]. While these recommendations provide a framework, our data suggest that surveillance strategies should additionally incorporate demographic risk factors, particularly race, ethnicity, and age, to ensure that high-risk subgroups receive appropriately intensified follow-up. Health systems seeking to reduce post-EMR CRC disparities should invest in culturally tailored patient navigation programs, bilingual education materials, and telehealth-based surveillance reminders for minority populations, as navigator-based interventions have improved CRC screening uptake in underserved community settings [22]. Recent real-world studies have shown that surveillance adherence after endoscopic resection is strongly influenced by travel distance, healthcare access, and socioeconomic factors [10, 11]. Standardizing and documenting EMR quality indicators, including completeness of resection, margin ablation status, lesion size and morphology, and histologic findings, within the EHR would facilitate more granular risk stratification and enable future retrospective research beyond what the current administrative coding-based approach allows. Emerging technologies hold promises for further reducing post-EMR CRC risk at the procedural level. Artificial intelligence-assisted polyp detection and characterization have been shown in randomized trials and meta-analyses to significantly improve adenoma detection rates and may reduce the likelihood of missed synchronous lesions at the index EMR examination [23]. Optical coherence technologies and advanced imaging modalities may similarly improve real-time assessment of resection margins and residual neoplasia. Adoption of these tools, particularly in lower-volume and community-based endoscopy centers where the populations most vulnerable to post-EMR CRC disparities are often treated, may have meaningful public health impact.

Our study has several important strengths. It represents the largest multicenter real-world cohort to evaluate the incidence and racial/ethnic disparities of post-EMR CRC in the United States to date. The use of the TriNetX US Collaborative Network, which encompasses data from both academic medical centers and community healthcare organizations with variable adherence rates, enhances generalizability and external validity compared to prior single-center or tertiary-center cohort studies. The five-year longitudinal follow-up enabled robust temporal characterization of incidence trends, and the stratified analyses provide actionable demographic data for surveillance protocol development. However, several limitations must be acknowledged. First, the TriNetX platform relies on administrative codes (ICD-10-CM and CPT) for cohort identification and outcome ascertainment, which are subject to miscoding, undercoding, and coding variability across participating institutions [14–16]. Diagnoses identified by ICD code alone may not reflect clinician-confirmed diagnoses with uniform accuracy [16]. Second, the platform lacks granular procedural data, including EMR technique (en-bloc versus piecemeal), submucosal injection agent, use of margin thermal ablation, lesion morphology, histologic subtype (conventional adenoma versus sessile serrated lesion), and resection adequacy assessment, all of which are critical determinants of post-EMR CRC risk and recurrence [5, 8]. The inability to adjust for these procedural variables represents a significant limitation. Third, the TriNetX platform does not permit reliable determination of the anatomic relationship between the post-EMR CRC and the index resection site, precluding differentiation between locally recurrent or residual neoplasia at the EMR site (true post-EMR CRC) and new metachronous cancers elsewhere in the colon (unrelated metachronous CRC). This limitation is inherent to administrative claims-based and EHR-based analyses and represents a fundamental challenge in interval CRC research [2]. Fourth, unmeasured confounders including insurance status, socioeconomic status, educational attainment, provider adenoma detection rate, and post-EMR surveillance adherence could not be controlled for in the current analysis.

Conclusion

In this large, real-world multicenter cohort study, we identified a clinically significant incidence of post-EMR CRC of 11.85 per 1,000 person-years, with marked racial, ethnic, sex-based, and age-related disparities. Hispanic/Latino and Asian patients demonstrated the highest post-EMR CRC incidence, while individuals aged 40–49 years bore a disproportionate burden suggesting a distinct risk profile for early-onset post-procedural CRC. Incidence declined progressively over the follow-up period, consistent with front-loaded risk concentrated in the early post-EMR period. These findings underscore the need for risk-stratified, equity-focused post-EMR surveillance strategies that extend beyond standard timing recommendations to incorporate patient-level demographic and biological risk factors.

Declarations

Funding:

This study did not receive any external funding. No industry, governmental, or institutional funding sources were involved in the design, conduct, analysis, interpretation, or reporting of this study.

Conflicts of Interest:

The authors declare no conflicts of interest relevant to this study. None of the authors have financial relationships, advisory roles, consultancies, stock ownership, honoraria, or intellectual property interests related to the subject matter of this manuscript.

Author Contributions:

Mohamed H. Eldesouki, MD, and Khaled Elfert, MD contributed equally to the conception and design of the study, data acquisition, analysis, and drafting of the manuscript. Mohammed Y. Youssef, MD writing the manuscript. Aasma Shaukat, MD, MPH: Study conception, critical interpretation of data, and revision of the manuscript. Shyam Thakkar, MD: Study supervision, methodological oversight, and critical revision of the manuscript.

Data Availability Statement

The data used in this study were obtained from the TriNetX U.S. Collaborative Network. Access to these data is subject to TriNetX data use agreements and cannot be publicly shared. Researchers seeking access should apply directly through the TriNetX platform (www.trinetx.com).

Competing Interests:

The authors declare that they have no competing interests.

Ethics Approval and Consent to Participate:

The study used de-identified data from the TriNetX Research Network. All data within TriNetX are de-identified in accordance with the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule (§164.514[b]). Because the data are de-identified, this study was exempt from Institutional Review Board (IRB) review and approval.

Consent to Participate:

Not applicable. This study used de-identified data and did not involve direct contact with human participants.

Consent for Publication:

Not applicable.

Institutional Review Board Statement:

This study utilized de-identified data from the TriNetX Research Network and was therefore exempt from Institutional Review Board approval in accordance with HIPAA Privacy Rule (§164.514[b]).

References

1. Siegel RL, Miller KD, Wagle NS, Jemal A, Cancer statistics. 2023. *CA Cancer J Clin.* 2023;73(1):17–48. doi: 10.3322/caac.21763. PMID: 36633525.
2. Fedewa SA, Flanders WD, Ward KC, Lin CC, Jemal A, Goding Sauer A, Doubeni CA, Goodman M. Racial and Ethnic Disparities in Interval Colorectal Cancer Incidence: A Population-Based Cohort Study. *Ann Intern Med.* 2017;166(12):857–866. 10.7326/M16-1154. PMID: 28531909.
3. Winawer SJ, Zauber AG, Ho MN et al. Prevention of colorectal cancer by colonoscopic polypectomy. *N Engl J Med.* 1993;329(27):1977–1981. 10.1056/NEJM199312303292701. PMID: 8247072.
4. Thiruvengadam SS, Fung BM, Barakat MT, Tabibian JH. Endoscopic Mucosal Resection: Best Practices for Gastrointestinal Endoscopists. *Gastroenterol Hepatol (N Y).* 2022;18(3):133–44. PMID: 35506001.
5. Moss A, Williams SJ, Hourigan LF et al. Long-term adenoma recurrence following wide-field endoscopic mucosal resection (WF-EMR) for advanced colonic mucosal neoplasia is infrequent: results and risk factors in 1000 cases from the Australian Colonic EMR (ACE) study. *Gut.* 2015;64(1):57–65. 10.1136/gutjnl-2013-305237. PMID: 24986245.
6. Tate DJ, Desomer L, Klein A, Brown G, Hourigan LF, Lee EY, Moss A et al. Adenoma recurrence after piecemeal colonic EMR is predictable: the Sydney EMR recurrence tool. *Gastrointest Endosc.* 2017;85(3):647–656.e6. doi: 10.1016/j.gie.2016.11.027. PMID: 27908600.
7. Pohl H, Robertson DJ. Colorectal cancers detected after colonoscopy frequently result from missed lesions. *Clin Gastroenterol Hepatol.* 2010;8(10):858–864. doi: 10.1016/j.cgh.2010.06.028. PMID: 20655393.
8. Kaltenbach T, Anderson JC, Burke CA, Dominitz JA, Gupta S, Lieberman D, Robertson DJ, Shaikat A, Syngal S, Rex DK. Endoscopic Removal of Colorectal Lesions: Recommendations by the US Multi-Society Task Force on Colorectal Cancer. *Am J Gastroenterol.* 2020;115(3):435–464. 10.14309/ajg.0000000000000555. PMID: 32058340.
9. Knudsen MD, Wang K, Wang L, Polychronidis G, Berstad P, Hjartåker A, Fang Z, Ogino S, Chan AT, Song M. Colorectal Cancer Incidence and Mortality After Negative Colonoscopy Screening Results. *JAMA Oncol.* 2025;11(1):46–54. 10.1001/jamaoncol.2024.5227. PMID: 39602147.
10. Farooq A, Bani Fawwaz B, Mian A, Singh G, Zhang Y, Gerges P et al. Patient adherence to surveillance colonoscopy after endoscopic resection of colorectal polyps and factors associated with loss to follow-up. *Endosc Int Open.* 2025;13:a24094916. 10.1055/a-2409-4916. PMID: 39965041.
11. Vashi B, Norwood DA, Sullivan R, Hegazy Y, Sánchez-Luna SA, Ajayi-Fox P et al. Social determinants of health influencing the adherence to post-endoscopic mucosal resection surveillance. *Clin Res Hepatol Gastroenterol.* 2024;48(4):102301. 10.1016/j.clinre.2024.102301. PMID: 38355006.
12. Saraiva MR, Rosa I, Claro I. Early-onset colorectal cancer: A review of current knowledge. *World J Gastroenterol.* 2023;29(8):1289–1303. 10.3748/wjg.v29.i8.1289. PMID: 36925459.

13. Zinkeng A, Taylor FL, Cheong SH, Song H, Merchant JL. Early Onset Colorectal Cancer: Molecular Underpinnings Accelerating Occurrence. *Cell Mol Gastroenterol Hepatol*. 2025;19(2):101425. 10.1016/j.jcmgh.2024.101425. PMID: 39510499.
14. Palchuk MB, London JW, Perez-Rey D, Drebert ZJ, Winer-Jones JP, Thompson CN, Esposito J, Claerhout B. A global federated real-world data and analytics platform for research. *JAMIA Open*. 2023;6(2):o0ad035. 10.1093/jamiaopen/o0ad035. PMID: 37193038.
15. Ludwig RJ, Anson M, Zirpel H, Thaci D, Olbrich H, Bieber K, Kridin K, Dempfle A, Curman P, Zhao SS, Alam U. A comprehensive review of methodologies and application to use the real-world data and analytics platform TriNetX. *Front Pharmacol*. 2025;16:1516126. 10.3389/fphar.2025.1516126. PMID: 40129946.
16. Olaker VR, Fry S, Terebuh P, Davis PB, Tisch DJ, Xu R, Miller MG, Dorney I, Palchuk MB, Kaelber DC. With big data comes big responsibility: Strategies for utilizing aggregated, standardized, de-identified electronic health record data for research. *Clin Transl Sci*. 2025;18(1):e70093. 10.1111/cts.70093. PMID: 39740190.
17. Heintzman JD, Ezekiel-Herrera DN, Quiñones AR, Lucas JA, Carroll JE, Gielbultowicz SH, Cottrell EC, Marino M. Disparities in Colorectal Cancer Screening in Latinos and Non-Hispanic Whites. *Am J Prev Med*. 2022;62(2):203–10. 10.1016/j.amepre.2021.07.009. Epub 2021 Oct 11. PMID: 34649735; PMCID: PMC11458137.
18. Tamai R, Dawit L, Lee P, Simmons OL, Ding L, Yuan L. The Association of Metabolic Risk Factors with Advanced Adenomas in Hispanic Patients. *Dig Dis Sci*. 2024;69(4):1403–10. 10.1007/s10620-024-08323-x. Epub 2024 Feb 16. PMID: 38363522; PMCID: PMC11026227.
19. Oh KM, An K, Lee M, Shin C, Steves SL. Colorectal cancer screening disparities in Asian Americans: the influences of patient-provider communication and social media use. *Cancer Causes Control*. 2023;34(9):813–27. 10.1007/s10552-023-01720-z. Epub 2023 Jun 2. PMID: 37266764; PMCID: PMC10237058.
20. US Preventive Services Task Force. Screening for Colorectal Cancer: US Preventive Services Task Force Recommendation Statement. *JAMA*. 2016;315(23):2564–2575. 10.1001/jama.2016.5989. PMID: 27304597.
21. ; US Preventive Services Task Force, Davidson KW, Barry MJ, Mangione CM et al. Screening for Colorectal Cancer: US Preventive Services Task Force Recommendation Statement. *JAMA*. 2021;325(19):1965–1977. 10.1001/jama.2021.6238. PMID: 34003218.
22. Percac-Lima S, Grant RW, Green AR, Ashburner JM, Gamba G, Oo S, Richter JM, Atlas SJ. A culturally tailored navigator program for colorectal cancer screening in a community health center: a randomized, controlled trial. *J Gen Intern Med*. 2009;24(2):211–7. PMID: 19067085; PMCID: PMC2628981.
23. Makar J, Abdelmalak J, Con D, Hafeez B, Garg M. Use of artificial intelligence improves colonoscopy performance in adenoma detection: a systematic review and meta-analysis. *Gastrointest Endosc*. 2025;101(1):68–81.e8. doi: 10.1016/j.gie.2024.08.033. PMID: 39216648.

Figures

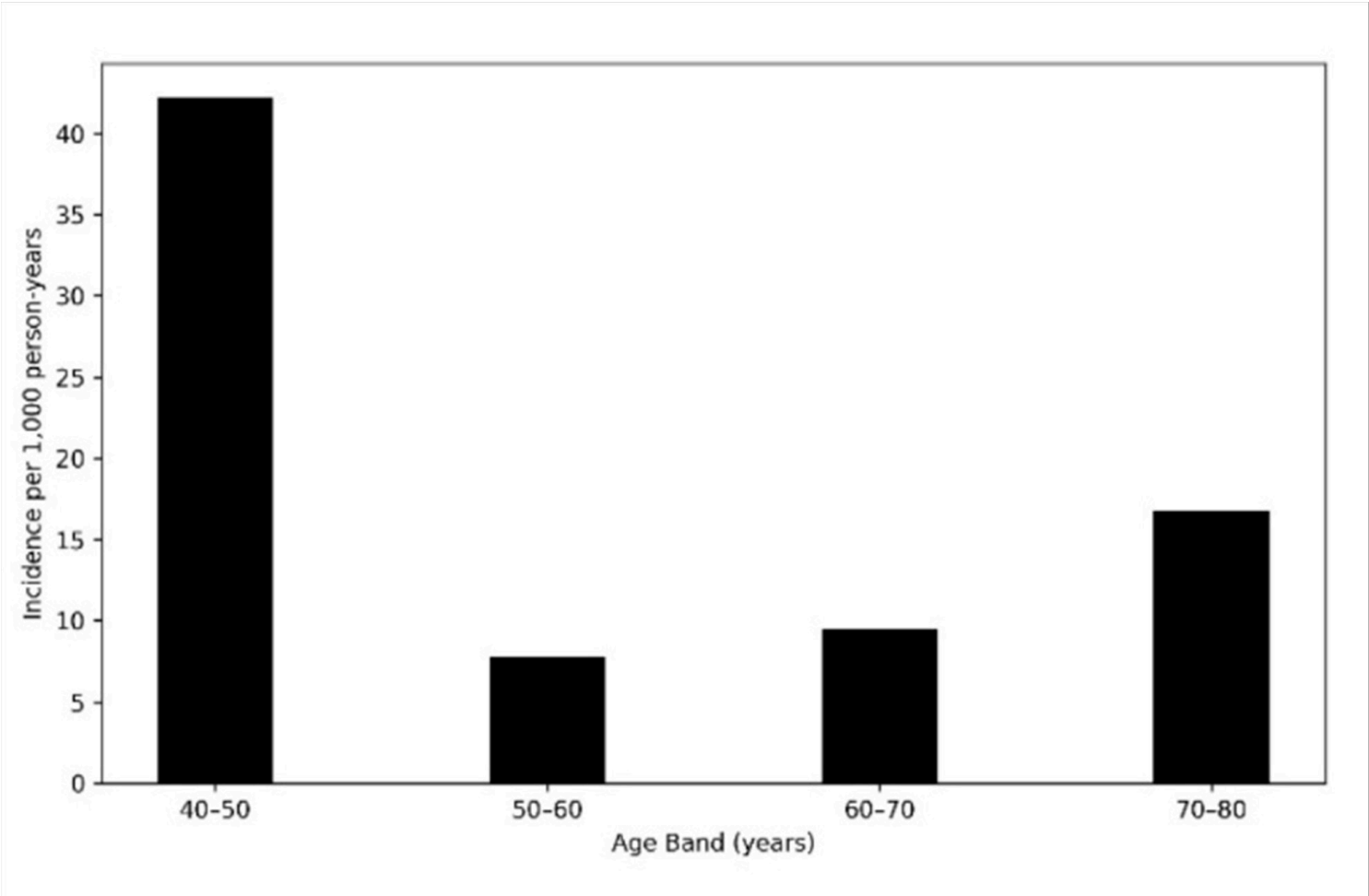


Figure 1

Post-EMR colorectal cancer incidence per 1,000 person-years stratified by age group. Incidence was highest in the 40–49-year group (42.2 per 1,000 PYO), declined in the 50–59- and 60–69-year groups (7.8 and 9.4, respectively), and increased again in the 70–79-year group (16.8 per 1,000 PYO). EMR, endoscopic mucosal resection; PYO, person-years of observation.

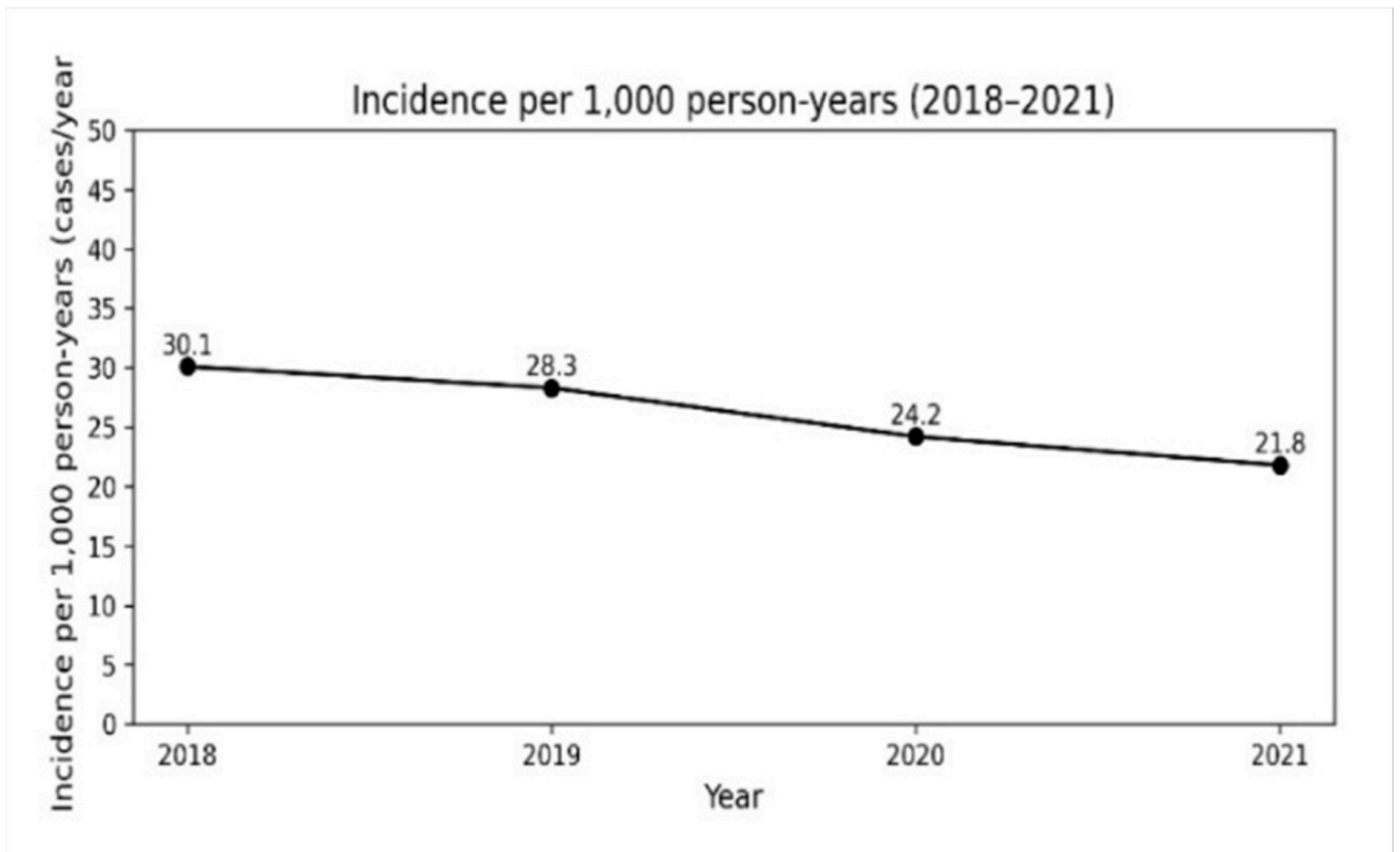


Figure 2

Annual post-EMR colorectal cancer incidence per 1,000 person-years from 2018 to 2021. A progressive decline in incidence from 30.1 cases per 1,000 PYO in 2018 to 21.8 per 1,000 PYO in 2021 (27.6% relative reduction) is observed, consistent with a front-loaded risk pattern concentrated in the early post-resection period. EMR, endoscopic mucosal resection; PYO, person-years of observation.