Demographic and Clinicopathologic Risk Factors for Colorectal Adenoma Recurrence: A Large-Scale Surveillance Cohort Study of 59,667 Adults

Usman Ayub Awan, PhD[1,2†], Qingyuan Song, PhD[3,4†], Kristen K. Ciombor, MD[5], Adetunji T. Toriol, MD, PhD[6], Jungyoon Choi, MD, PhD[7], Timothy Su, MD, PhD[1], Xiao-ou Shu, PhD[1], Kamran Idrees, MD[8], Kay M. Washington[9], Wei Zheng, MD, PhD[1], Wanqing Wen, MD[1*], Zhijun Yin, PhD[3,4,10*], Xingyi Guo, PhD[1,3*]

Author Affiliations:

[1] Division of Epidemiology, Department of Medicine, Vanderbilt-Ingram Cancer Center, Vanderbilt University Medical Center, Nashville, TN, USA
[2] Department of Medical Laboratory Technology, The University of Haripur, Haripur, Khyber Pakhtunkhwa, Pakistan
[3] Department of Biomedical Informatics, Vanderbilt University Medical Center, Nashville, TN, USA
[4] Department of Department of Computer Science, Vanderbilt University, Nashville, TN, USA
[5] Division of Hematology/Oncology, Department of Medicine, Vanderbilt University Medical Center, Vanderbilt-Ingram Cancer Center, Nashville, TN, USA
[6] Division of Public Health Sciences, Department of Surgery, and Siteman Cancer Center, Washington University School of Medicine St. Louis, MO, USA
[7] Division of Oncology and Hematology, Department of Internal Medicine, Korea University Ansan Hospital, Korea University College of Medicine, Ansan, Korea
[8] Division of Surgical Oncology and Endocrine Surgery, Department of Surgery, Vanderbilt University Medical Center, Vanderbilt-Ingram Cancer Center Nashville, TN, USA.
[9] Department of Pathology, Vanderbilt University Medical Center, Nashville, Tennessee
[10] Department of Electrical and Computer Engineering, Vanderbilt University, Nashville, TN, USA

Abstract: 350 / 350

Figures & Tables: 05

Supplementary Online Figures & Tables: 04

[†]These authors contributed equally to this work.

*** Corresponding Authors:**
Xingyi Guo, PhD
Vanderbilt University Medical Center
2525 West End Ave, Suite 330
Nashville, Tennessee 37203
Office: 615-936-3471
xingyi.guo@vumc.org

Zhijun Yin, PhD
Vanderbilt University Medical Center
2525 West End Ave. Suite 1475, Nashville, TN 37203
Tel: +1 615-936-3690; Fax: +1 615-322-0102
Email: zhijun.yin.1@vumc.org

Wanqing Wen, MD
Vanderbilt University Medical Center
2525 West End Ave, Suite 334-G, Nashville, 37203
Office: 615-322-0105
wanqing.wen@vumc.org

Text Word Count: 2992

Running Headline: Investigation of Disparities in Adenoma Recurrence Risk

Research in context

Evidence before this study

We conducted a PubMed search for publications dated before June 2024 using combinations of keywords such as “colonoscopic polypectomy,” “Demographic and Clinicopathologic Risk Factors,” “Vannderbilt,” and “electronic health records.” We found no studies that comprehensively evaluated the associations of demographic characteristics, obesity, and adenoma features with recurrence risk over time in a large, longitudinal surveillance cohort.

Added Value of This Study

Using a longitudinal cohort of 59,667 patients, our study reveals substantial temporal heterogeneity in adenoma recurrence. Non-Hispanic Black and Asian or Pacific Islander individuals exhibited a lower risk of recurrence within the first 5 years but experienced increased risk at 5–10 and >10 years post-polypectomy. Females showed heightened early recurrence risk, particularly when initial adenomas were tubulovillous, mixed-type, or large. Early recurrence was predominantly driven by high-grade dysplasia, high-risk adenomas, villous or tubulovillous histology, and multiplicity.

Implications of All the Available Evidence

These findings highlight the critical need to recognize and address temporal, racial, and sex-specific heterogeneity in adenoma recurrence risk. The observed variability in histopathologic and demographic factors over time underscores the importance of personalized, adaptive surveillance strategies to reduce adenoma recurrence and enhance colorectal cancer prevention.

Summary

Background: Current colorectal surveillance guidelines emphasize adenoma characteristics but overlook temporal, racial, and sex-based heterogeneity in recurrence risk— an gap that limits equitable and personalized care. To evaluate the associations of demographic factors, obesity, and adenoma features with recurrence risk over time in a large longitudinal surveillance cohort.

Methods: This retrospective cohort study included 59,667 adults who underwent their first colonoscopic polypectomy between January 1990 and July 2024 at a tertiary medical center. Median follow-up was 4 years. Demographic variables included race and ethnicity (non-Hispanic White [NHW], non-Hispanic Black [NHB], Hispanic, Asian or Pacific Islander [API]), sex, obesity (BMI >30), family history of colorectal cancer (CRC) or polyps, and age at adenoma onset (<50 vs ≥50 years). Adenoma features included histology, size, number, and dysplasia. The primary outcome was recurrence-free survival, defined as time from initial polypectomy to histologically confirmed recurrence. Cox proportional hazards models estimated associations adjusted for confounders, with stratified analyses over 5-, 10-, and >10-year follow-up intervals.

Findings: Among 59,667 patients, 17,596 (29.5%) experienced recurrence within 5 years, revealing substantial temporal heterogeneity. Early recurrence was associated with male sex (adjusted hazard ratio [aHR], 1.10; 95% CI, 1.06–1.14), obesity (aHR, 1.18; 95% CI, 1.13–1.23), early-onset adenomas (aHR, 1.17; 95% CI, 1.11–1.23), and family history of CRC (aHR, 1.24; 95% CI, 1.18–1.31). Compared with NHW patients, NHB individuals had lower early recurrence risk (aHR, 0.89; 95% CI, 0.83–0.96) but higher late recurrence (>10 years; aHR, 1.26; 95% CI, 1.06–1.50). API patients had a similar shift, with lower early risk (aHR, 0.80; 95% CI, 0.67–0.96) and elevated mid-term risk (5–10 years; aHR, 1.40; 95% CI, 1.08–1.81). High-grade dysplasia (aHR 2.86; 95% CI, 2.54–3.22) and villous histology (aHR 2.55; 95% CI, 2.31–2.81showed the largest effect sizes for early recurrence. Females had stronger associations with tubulovillous histology, mixed adenomas, and large lesions.

Interpretation: Temporal, demographic, and histologic differences in adenoma recurrence highlight the need for surveillance strategies that incorporate population- and time-specific risk profiles to enhance colorectal cancer prevention.

Funding: This work was supported by the National Cancer Institute (Grant No. R37CA227130 to Xingyi Guo).

Keywords: Colorectal Adenoma Recurrence; Demographic Disparities; Surveillance; Poly Characteristics; Cohort Study

Introduction

Colorectal cancer (CRC) persists as a leading cause of cancer-related mortality globally, with adenomatous polyps representing the primary precursor lesions for 60–90% of cases [1,2]. Current post-polypectomy surveillance guidelines, such as those from the U.S. Multi-Society Task Force (USMSTF) and the European Society of Gastrointestinal Endoscopy (ESGE), stratify recurrence risk primarily based on polyp characteristics, including size, histology, and multiplicity [1-3]. Recent updates to these guidelines, including the USMSTF recommendations of 7–10 years for 1–2 nonadvanced adenomas (NAAs) and the 2020 ESGE's omission of surveillance for 1–4 NAAs, reflect evolving evidence but reveal persistent inconsistencies. For instance, the ESGE and British Society of Gastroenterology/Association of Coloproctology of Great Britain and Ireland/Public Health England (BSG/ACPGBI/PHE) guidelines recommend a return to routine screening for 1–4 NAAs (tubular adenomas <10 mm with low-grade dysplasia), whereas the USMSTF advocates surveillance intervals ranging from 3–10 years, depending on adenoma number [1,3,4].

These discrepancies underscore a broader limitation: guidelines universally neglect demographic variables—such as race, sex, family history of polyps, and family history of CRC, and age—and modifiable risk factors like obesity, potentially perpetuating inequities in colorectal adenoma recurrence and CRC prevention [5-7]. Of note, racial disparities in CRC outcomes are well-documented, with non-Hispanic Black (NHB) individuals experiencing higher CRC incidence and mortality compared to White populations; however, limited studies have explored the association between demographic disparities and adenoma recurrence [8,9]. Despite differential recurrence patterns across ethnic groups, current surveillance protocols lack race-specific risk stratification, while central obesity (waist-to-hip ratio) independently increases CRC risk by 15–18% per standard deviation, surpassing BMI's effect [6,7].

Histopathological features, such as high-grade dysplasia and villous architecture, are recognized as having significant effect sizes in their association with colorectal adenoma recurrence [9-11]. However, their prioritization varies across current clinical guidelines,

contributing to inconsistent surveillance recommendations. USMSTF classifies villous histology or high-grade dysplasia as high-risk for 3-year surveillance, but ESGE prioritizes size and dysplasia, potentially underestimating villous histology's long-term impact, with neither addressing racial or sex-specific recurrence variations [9,12]. Compounding these issues, long-term surveillance data from racially and ethnically diverse populations remain sparse, restricting insights into how demographic and histopathological factors interact to influence post-polypectomy outcomes over time [9,13,14]. Investigating recurrent colorectal adenoma demands longitudinal follow-up with serial colonoscopies to confirm polyp recurrence and polyp-free status in controls.

Current literature provides limited insight into temporal recurrence dynamics—namely, how risk patterns evolve over extended follow-up—thus constraining the optimization of long-term surveillance strategies [1,2,15,16]. Although guidelines assume constant risk, obesity and male sex drive recurrence within 3–5 years post-polypectomy, while villous histology, lesion multiplicity, and high-grade dysplasia show persistent or re-emerging risk after 10 years, necessitating temporally adaptive surveillance frameworks[17-19].

To overcome these barriers, we analyzed Vanderbilt electronic health records (EHRs), covering diagnoses, procedures, and pathology reports from over 3.5 million patients, using a dataset of 59,667 Polypectomy Cohort individuals to clarify post-polypectomy surveillance determinants, focusing on obesity, demographic disparities, and polyp characteristics, thereby providing a foundation for refining polyp-centric risk models and establishing targeted surveillance strategies for high-risk populations.

Methods

Study Design and Setting

The Vanderbilt Polypectomy Cohort was recently established to investigate adenoma recurrence patterns across demographic subgroups using clinical data from Vanderbilt University Medical Center (VUMC) spanning January 1990 through July 2024. To ensure consistency and accuracy despite changes in clinical practice, diagnostic standards, and coding systems over the

years—including the transition from ICD-9 to ICD-10—harmonized data extraction protocols were meticulously implemented [20].

Data Source

Data was extracted from VUMC EHR system (Epic Systems Corporation's eStar), encompassing colonoscopy reports, pathology findings, demographics, and comorbidities. To ensure robustness, a multistep validation process was implemented. First, polyp cases and procedures were identified using standardized diagnostic and procedural codes (ICD-9, ICD-10, CPT, HCPCS) to address discrepancies across coding systems. Second, all polyp diagnoses were cross verified against original histopathology reports to confirm neoplastic status, histology, and dysplasia grade. Third, a secure institutionally managed large language model (LLM) was deployed to extract adenoma characteristics—including polyp histology, size, and number—from unstructured clinical notes and pathology reports, with a verification rate > 95% based on our comparison between procedural codes and LLM from randomly selected 100 patient records [21].

Vanderbilt Polypectomy Cohort Identification

The study population comprised adults aged ≥18 years with histopathologically confirmed colorectal adenoma who underwent polypectomy at VUMC. Exclusion criteria were prior CRC and other cancers, non-neoplastic polyps, colectomy, or incomplete colonoscopies. Colorectal adenomas were categorized as high-risk if they met one or more criteria: ≥3 number, size ≥10 mm, villous/tubulovillous histology, or high-grade dysplasia. Low-risk adenomas were defined as <3 adenomas, <10 mm in size, and low-grade dysplasia. A multidisciplinary team (U.A.A., Q.S., Z.Y., X.G.) independently reviewed cohort eligibility to ensure adherence to clinical and histopathologic standards. Additional details on data extraction and variable definitions are provided in the Supplemental Methods.

Follow-Up and Censoring:

Patients were followed from the index polypectomy until the earliest occurrence of adenoma recurrence, death, loss to follow-up, or the study's end in July 2024. Longitudinal recurrence patterns were primarily evaluated using a 5-year follow-up interval, with events after this period censored, while additional intervals of <5 years, 5-10 years, and >10 years were

assessed to explore trends, as detailed in the supplementary files (eTable 1, 2), censoring events after 5 years for the <5-year analysis, after 10 years for the 5-10-year analysis, and analyzing events beyond 10 years for the >10-year interval. Patients undergoing colectomy (complete or partial) or lost to follow-up were censored at their last documented clinical contact.

The primary outcome, recurrence-free survival, was defined as the time from initial polypectomy to histologically confirmed recurrence on subsequent colonoscopy (≥6 months post-procedure). This was analyzed both as a binary endpoint and as time-to-event data to estimate recurrence-free survival. Recurrence timing was evaluated across pre-defined intervals (<5, 5–10, and >10 years), and incorporated into multivariable Cox regression models to assess associations with demographic, clinical, and histopathological factors. Censoring rules and follow-up windows were applied as detailed in the corresponding analytical framework.

Statistical Analysis

Baseline characteristics were summarized using descriptive statistics, with categorical variables reported as numbers and percentages (%). Cox proportional hazard models were used to estimate hazard ratios (HRs) and 95% confidence intervals (CIs), adjusting for potential confounders including race/ethnicity, sex, obesity, family history of polyp, family history of CRC, NSAID, Aspirin use (during the post-polypectomy period) and polyp characteristics (number, size, and type), and baseline adenoma risk (high vs low). Recurrence patterns were primarily assessed using a 5-year follow-up interval, with extended intervals up to and beyond 10 years evaluated to explore long-term trends, censoring events after each interval as appropriate. The proportional hazards assumption was verified using Schoenfeld residuals, and survival curves were generated via the Kaplan-Meier method to compare recurrence-free probabilities across demographic and clinical subgroups. Stratified analyses were conducted by race/ethnicity, sex, recurrence status, and follow-up interval (< 5 Years, 5-10 years and >10-year windows). Statistical significance was defined as a two-tailed $P < 0.05$.

Results

The Vanderbilt Polypectomy Cohort included 59,667 patients who underwent colon polypectomy, with a median follow-up of 4 years (Figure 1). The cohort consisted of 29,401

females (49.3%) and 30,266 males (50.7%), with a racial/ethnic distribution of 52,042 (NHW, 87.2%), 5,972 (NHB, 10.0%), 1,007 (API, 1.7%), and 646 Hispanic (1.1%) individuals. Early-onset adenoma was identified in 11,018 patients (18.5%), while late-onset adenoma was observed in 48,649 patients (81.5%). Adenoma characteristics varied significantly by sex and race/ethnicity. Males exhibited higher prevalence of tubulovillous (4.3% vs. 3.6% in females) and mixed adenomas (27.9% vs. 27.2%), while females had more serrated (5.9% vs. 3.4%) and villous adenomas (3.5% vs. 2.2%; $P = 1.4 \times 10^{\square\square\square}$). Among racial groups, NHW patients showed elevated mixed (27.7%) and serrated adenoma rates (4.8%), NHB had higher tubular (36.3%) and tubulovillous adenomas (5.4%), and API had the highest tubular adenoma prevalence (40.0%; $P = 1.6 \times 10^{\square\square\square}$). Early-onset adenomas were more common in Hispanic (30.5%) and API (20.3%) than NHW patients (18.2%; $P = 2.6 \times 10^{\square 1\square}$). Family history of CRC was more frequent in females (15.7% vs. 11.9% in males; $P = 4.9 \times 10^{\square\square 2}$) and NHW (14.0%; $P = 1.7 \times 10^{\square\square}$), while family history of polyps was higher in females (3.2% vs. 2.1%; $P = 2.5 \times 10^{\square 1\square}$) and NHW (2.8%; $P = 5.5 \times 10^{\square 1\square}$). High-risk adenomas were slightly more prevalent in males (25.3% vs. 23.8%; $P = 5.6 \times 10^{\square\square}$) (Table 1).

Recurrence Rates and Demographic-Clinical Associations

Adenoma recurrence occurred in 17,596 patients (29.5%) within 5 years, with significant associations observed across demographic and clinical variables (Table 2). Males showed a higher recurrence rate than females (30.9% vs. 28.1%; adjusted hazard ratio [aHR] 1.1, 95% CI 1.06–1.14; $P = 3.46 \times 10^{\square\square}$). By race/ethnicity, NHW patients had a 30.0% recurrence rate, compared to 26.6% in NHB (aHR 0.89, 95% CI 0.83–0.96; $P = 1.97 \times 10^{\square 3}$), 25.8% in API (aHR 0.8, 95% CI 0.67–0.96; $P = 1.36 \times 10^{\square 2}$), and 17.2% in Hispanic patients (aHR 0.89, 95% CI 0.70–1.14; $P = 3.63 \times 10^{\square 1}$). Obesity, present in 19,641 patients (32.9%), was associated with elevated recurrence risk (32.3% vs. 67.7% non-obese; aHR 1.18, 95% CI 1.13–1.23; $P = 1.84 \times 10^{\square 1\square}$). Patients with early-onset adenomas had a 26.4% recurrence rate, demonstrating a higher risk than those with late-onset (30.2%; aHR 1.17, 95% CI 1.11–1.23; $P = 6.77 \times 10^{\square\square}$). Family history of CRC exhibited a strong association with recurrence (38.7%; aHR 1.24, 95% CI 1.18–1.31; $P = 4.44 \times 10^{\square 1\square}$), particularly in females and NHW patients. Family history of polyps showed a weaker association (38.5%; aHR 1.04, 95% CI 0.93–1.17; $P = 4.64 \times 10^{\square 1}$).

Adenoma Characteristics and Recurrence

Adenoma characteristics demonstrated significant associations with recurrence, with histopathological features exhibiting the largest effect sizes (Table 3). Villous histology had the highest recurrence rate (57.7%; aHR 2.55, 95% CI 2.31–2.81; $P = 4.55 \times 10^{-}$), followed by tubulovillous (32.6%; aHR 2.18, 95% CI 1.99–2.39; $P = 1.58 \times 10^{-1}$) and mixed adenomas (31.2%; aHR 1.37, 95% CI 1.30–1.45; $P = 6.99 \times 10^{-32}$), compared to tubular adenomas (28.2%). Serrated adenomas showed a moderate association (28.1%; aHR 1.3, 95% CI 1.18–1.44; $P = 1.10 \times 10^{-}$). High-grade dysplasia was strongly associated with recurrence (30.9%; aHR 2.86, 95% CI 2.54–3.22; $P = 4.13 \times 10^{-}$) relative to low-grade dysplasia (29.5%). Multiple adenomas (≥3) had a recurrence rate of 33.5% (aHR 1.69, 95% CI 1.60–1.79; $P = 5.95 \times 10^{-}$), higher than fewer than 3 (28.3%). Adenomas ≥10 mm showed increased risk (30.8%; aHR 1.63, 95% CI 1.53–1.73; $P = 5.72 \times 10^{-}$), though with a smaller effect size than histopathological features. High-risk adenomas, observed in 24.6% of patients, had a 34.5% recurrence rate (aHR 2.79, 95% CI 2.68–2.91; $P = 2.22 \times 10^{-1}$), significantly exceeding low-risk adenomas (27.9%). Adjusted survival curves for polyp type, dysplasia grade, number, and size over 5 years are illustrated in Figure 2.

Long-Term Recurrence Patterns and Temporal Heterogeneity

Colorectal adenoma recurrence demonstrated pronounced temporal heterogeneity over the 25-year surveillance period (eTables 1–2). Among racial groups, NHB patients exhibited a temporal shift from reduced early recurrence risk (<5 years; aHR 0.89, 95% CI 0.83–0.96; $P = 1.97 \times 10^{-3}$) to elevated late-phase hazard (>10 years; aHR 1.26, 95% CI 1.06–1.50; $P = 9.22 \times 10^{-3}$). Similarly, API patients experienced early protection (aHR 0.80, 95% CI 0.67–0.96; $P = 1.36 \times 10^{-2}$) followed by heightened mid-term recurrence risk (5–10 years; aHR 1.40, 95% CI 1.08–1.81; $P = 1.04 \times 10^{-2}$). Temporal recurrence risk was most strongly influenced by histopathologic features during the early phase, including high-grade dysplasia (aHR 2.86, 95% CI 2.54–3.22; $P = 4.13 \times 10^{-}$), high-risk adenomas (aHR 2.79, 95% CI 2.68–2.91; $P = 2.22 \times 10^{-1}$), and villous histology (aHR 2.55, 95% CI 2.31–2.81; $P = 4.55 \times 10^{-}$), with attenuation of these associations over long-term follow-up. Tubulovillous histology (aHR 2.18, 95% CI 1.99–2.39; $P = 1.58 \times 10^{-1}$) also conferred elevated early recurrence hazard, reinforcing the importance of temporally adaptive, pathology-specific surveillance strategies.

Heterogeneity of Effects by Race and Sex

Five-year stratified analyses revealed significant sex- and race-specific heterogeneity in recurrence risk (eTables 3–4), reinforcing the rationale for personalized surveillance. Females exhibited significantly greater early recurrence hazards for tubulovillous histology (aHR 2.21, 95% CI 1.93–2.53; $P = 2.67 \times 10^{-3}$), mixed adenomas (aHR 1.43, 95% CI 1.32–1.55; $P = 3.19 \times 10^{-1}$), and large lesions (≥10 mm; aHR 1.60, 95% CI 1.46–1.74; $P = 2.77 \times 10^{-2}$), with stronger associations than those observed in males relative to tubular or smaller (<10 mm) lesions. Among racial groups, NHB patients exhibited significantly higher recurrence risk for multiple adenomas (aHR 1.67, 95% CI 1.42–1.98; $P = 1.85 \times 10^{-}$), indicating a greater burden of lesion multiplicity. These findings underscore the imperative for demographically stratified surveillance paradigms to optimize post-polypectomy care.

Discussion

Current Vanderbilt Polypectomy Cohort, encompassing 59,667 patients who underwent polypectomy, provides critical insights into the demographic and clinicopathological factors driving colorectal adenoma recurrence and long-term outcomes. This study, with its 25-year follow-up, reveals distinct recurrence patterns across sex, age, race, and histopathological subtypes, highlighting opportunities to refine current post-polypectomy surveillance guidelines (e.g., USMSTF and ESGE) through personalized, data-driven risk stratification approaches that enhance precision and equity in CRC prevention.

Males exhibited higher early recurrence risk, driven by tubulovillous and mixed adenomas, consistent with studies linking male sex to advanced neoplasia [22-25]. Females showed increased risk with villous architecture and larger adenoma, independent of adenoma burden, suggesting a sex-histology-size interaction. This challenges USMSTF's uniform 3–10-year surveillance intervals[9], which overlook sex as a risk modifier. Shorter intervals for females with villous morphology could enhance precision prevention.

Recurrence patterns varied by race/ethnicity. NHB patients displayed delayed but elevated recurrence, contrasting with early peaks in API and Hispanic populations. Despite

higher tubulovillous adenoma burden, NHB patients had lower early recurrence, possibly due to lower adenoma detection rates and delayed follow-up, exacerbating disparities [26-28,45,49]. Current guidelines (USMSTF, ESGE, BSG/ACPGBI/PHE) inadequately address this heterogeneity, with ESGE omitting surveillance for villous adenomas <10 mm [2,15,16,36,37]. Race-stratified surveillance—extended for NHB patients and tighter early intervals for those with familial CRC risk—could bridge equity gaps. Prior studies report inconsistent racial differences in advanced neoplasia [29-31], suggesting adenoma characteristics and surveillance intensity influence outcomes. Structural inequities, including unequal healthcare access and lower screening rates among NHB patients [29-31,45,49], necessitate standardized care and prospective multi-institutional studies.

Early-onset adenomas, prevalent among Hispanic and API patients, were associated with higher early recurrence risk, aligning with rising early-onset CRC trends [32,33]. Despite USMSTF's 2022 recommendation for screening at age 45 [34,35], guidelines do not account for age at adenoma detection as a recurrence modifier. Family history of CRC was associated with early recurrence risk (aHR 1.24), while polyp family history showed a stronger effect size for recurrence at 5–10 years, indicating temporally distinct trajectories [36-40]. Integrating onset age and familial subtype into risk models could refine surveillance intervals. Obesity (32.9% prevalence) was associated with recurrence (aHR 1.18), with stronger effects in females, possibly due to visceral adiposity and endocrine signaling [41-43].

Villous architecture and high-grade dysplasia exhibited the largest effect sizes for recurrence, surpassing adenoma number (≥3: aHR 1.69) and size (≥10 mm: aHR 1.63). Villous adenomas conferred sustained risk beyond 10 years, while high-grade dysplasia elevated early risk. Females with villous features or larger adenomas faced higher recurrence, supporting sex-histology-size-specific stratification. USMSTF flags villous histology as high-risk, advocating 3-year surveillance [35,44], but ESGE's omission of villous features [2] may underestimate risk. High-risk adenomas exhibited persistent hazard, with late-phase resurgence (>10 years), underscoring the need for prolonged, individualized follow-up prioritizing histologic severity.

Adenoma recurrence risk is heterogeneous and dynamic, inadequately captured by static guideline frameworks. Time-stratified Cox models revealed evolving risk trajectories: obesity's influence dissipated after 5 years, early-onset lesions shifted from excess to deficit risk, and NHB patients transitioned from early protection to late excess risk. High-risk adenomas showed a late-phase resurgence (>10 years), unlike the time-limited effects of dysplasia and obesity. Villous morphology maintained a persistent hazard, even in diminutive adenomas, challenging ESGE's exclusion of villous features <10 mm [2]. A recent cohort noted higher adenoma burden in NHB at older ages, suggesting biology-care interactions [45], while a study found no race effect with uniform colonoscopy quality [31]. These findings reconcile discordant results, emphasizing dynamic, race-specific hazards. The 2020 USMSTF lengthened surveillance for low-risk adenomas but retained uniform high-risk intervals[35], and ESGE's simplified high-risk definitions [2,55] overlook sustained villous risk. Integrating temporal coefficients and heterogeneity metrics could align surveillance with evolving patient risk, enhancing precision and equity in CRC prevention.

This study's cohort of 59,667 patients, followed up to 25 years, robustly examines colorectal adenoma recurrence across diverse subgroups using validated natural language processing for histologic and clinical feature extraction, enabling stratified analyses by sex, race/ethnicity, onset age, and adenoma subtype. Time-dependent Cox models reveal dynamic recurrence trajectories over <5, 5–10, and >10-year intervals, uncovering late-phase risks often missed, with real-world EHR data enhancing generalizability. However, this single-center study, predominantly non-Hispanic White, may not fully generalize to diverse populations, and its retrospective design risks selection bias and confounding from unmeasured factors (e.g., diet, smoking, socioeconomic status, healthcare access, colonoscopy quality). Variable surveillance frequency across groups may affect recurrence detection, and reliance on VUMC network colonoscopies risks misclassification if patients were surveilled elsewhere.

Conclusion

In this cohort of 59,667 patients, villous histology, high-grade dysplasia, and adenoma multiplicity emerged as dominant recurrence drivers, though their hazards varied by time and population. Risk peaked early among males, obese patients, and those with early-onset lesions,

but shifted later among non-Hispanic Black individuals. Notably, villous features conferred a stronger recurrence hazard in females. Incorporating these temporally dynamic, demography-specific risks into surveillance algorithms may enhance precision, reduce disparities, and help current guidelines to enhance precision and equity in post-polypectomy surveillance.

Contributors

X.G. and Y.Z. conceived and designed the study. U.A.A., Q.S., W.W., Z.Y., and X.G. collected the data and developed the analytical plan. U.A.A., Q.S., Z.Y., and X.G. accessed, verified, and processed the raw cohort data used in this study. U.A.A., together with Q.S., W.W., Z.Y., and X.G., drafted the manuscript with input from all co-authors. All authors contributed to data interpretation, manuscript revision, and approved the final version for submission. All authors had full access to all the data in the study and shared final responsibility for the decision to submit the manuscript for publication.

Declaration of interest

None of the authors have any conflicts of interest to disclose.

Data sharing:

The datasets used and analyzed in the current study are available from the corresponding authors upon reasonable request.

References:

1. Jung YS. Summary and comparison of recently updated post-polypectomy surveillance guidelines. *Intestinal research* 2023; **21**(4): 443-51.

2. Hassan C, Antonelli G, Dumonceau JM, et al. Post-polypectomy colonoscopy surveillance: European Society of Gastrointestinal Endoscopy (ESGE) Guideline - Update 2020. *Endoscopy* 2020; **52**(8): 687-700.

3. Abu-Freha N, Katz LH, Kariv R, et al. Post-polypectomy surveillance colonoscopy: Comparison of the updated guidelines. *United European gastroenterology journal* 2021; **9**(6): 681-7.

4. He X, Hang D, Wu K, et al. Long-term risk of colorectal cancer after removal of conventional adenomas and serrated polyps. *Gastroenterology* 2020; **158**(4): 852-61. e4.

5. Kumar R, Brown A, Okano S, et al. Overweight and obesity are associated with colorectal neoplasia in an Australian outpatient population. *Scientific Reports* 2024; **14**(1): 23501.

6. Safizadeh F, Mandic M, Schöttker B, Hoffmeister M, Brenner H. Central obesity may account for most of the colorectal cancer risk linked to obesity: evidence from the UK Biobank prospective cohort. *International Journal of Obesity* 2024: 1-8.

7. Zhang X, Xu Z, Shang L, et al. Global burden of colorectal cancer attributable to metabolic risks from 1990 to 2021, with predictions to 2046. *BMC cancer* 2025; **25**(1): 228.

8. McLeod MR, Galoosian A, May FP. Racial and Ethnic Disparities in Colorectal Cancer Screening and Outcomes. *Hematology/oncology clinics of North America* 2022; **36**(3): 415-28.

9. Gupta S, Lieberman D, Anderson JC, et al. Recommendations for follow-up after colonoscopy and polypectomy: a consensus update by the US Multi-Society Task Force on Colorectal Cancer. *Official journal of the American College of Gastroenterology| ACG* 2020; **115**(3): 415-34.

10. Hassan C, Antonelli G, Dumonceau J-M, et al. Post-polypectomy colonoscopy surveillance: European society of gastrointestinal endoscopy (ESGE) guideline–update 2020. *Endoscopy* 2020; **52**(08): 687-700.

11. Rutter MD, East J, Rees CJ, et al. British Society of Gastroenterology/Association of Coloproctology of Great Britain and Ireland/Public Health England post-polypectomy and post-colorectal cancer resection surveillance guidelines. *Gut* 2020; **69**(2): 201-23.

12. Rutter MD, East J, Rees C, Cripps N, Docherty J, Dolwani S. BSG/ACPGBI/PHE Post-polypectomy and post-colorectal cancer resection surveillance guidelines. *Gut* 2020; **2020**: 201-23.

13. Kamal Y, Idos G. Incidental Young-Onset Adenomas: Sporadic Findings or Harbingers of Increased Colon Cancer Risk? *Current Treatment Options in Gastroenterology* 2022; **20**(2): 122-32.

14. Issaka RB, Chan AT, Gupta S. AGA clinical practice update on risk stratification for colorectal cancer screening and post-polypectomy surveillance: expert review. *Gastroenterology* 2023.

15. Rutter MD, East J, Rees CJ, et al. British Society of Gastroenterology/Association of Coloproctology of Great Britain and Ireland/Public Health England post-polypectomy and post-colorectal cancer resection surveillance guidelines. *Gut* 2020; **69**(2): 201-23.

16. Kim SY, Kwak MS, Yoon SM, et al. Korean Guidelines for Postpolypectomy Colonoscopic Surveillance: 2022 revised edition. *Intestinal research* 2023; **21**(1): 20-42.
17. Jacobs ET, Martínez ME, Alberts DS, et al. Association between body size and colorectal adenoma recurrence. *Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association* 2007; **5**(8): 982-90.
18. Hsu CH, Taylor JM, Long Q, Alberts DS. Analysis of colorectal adenoma recurrence data subject to informative censoring. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* 2009; **18**(3): 712-7.
19. Paspatis G, Fragaki M, Arna DE, et al. Long-term adenoma recurrence and development of colorectal cancer following endoscopic mucosal resection in large non-pedunculated colonic polyps ≥4 cm. *Digestive and liver disease : official journal of the Italian Society of Gastroenterology and the Italian Association for the Study of the Liver* 2025; **57**(1): 44-50.
20. Indicators AQ. ICD-9-CM to ICD-10-CM/PCS Conversion of AHRQ Quality Indicators2011. https://qualityindicators.ahrq.gov/Downloads/Resources/Publications/2011/ICD-10_Report_02-08-11_Final.pdf (accessed 02-15-2025).
21. Huang J, Yang DM, Rong R, et al. A critical assessment of using ChatGPT for extracting structured data from clinical notes. *npj Digital Medicine* 2024; **7**(1): 106.
22. Calderwood AH, Tosteson TD, Walter LC, Hua P, Onega T. Colonoscopy utilization and outcomes in older adults: Data from the New Hampshire Colonoscopy Registry. *J Am Geriatr Soc* 2022; **70**(3): 801-11.
23. Penn E, Garrow D, Romagnuolo J. Influence of race and sex on prevalence and recurrence of colon polyps. *Archives of internal medicine* 2010; **170**(13): 1127-32.
24. Ferlitsch M, Reinhart K, Pramhas S, et al. Sex-specific prevalence of adenomas, advanced adenomas, and colorectal cancer in individuals undergoing screening colonoscopy. *Jama* 2011; **306**(12): 1352-8.
25. Hao Y, Wang Y, Qi M, He X, Zhu Y, Hong J. Risk factors for recurrent colorectal polyps. *Gut and Liver* 2019; **14**(4): 399.
26. McLeod MR, Galoosian A, May FP. Racial and ethnic disparities in colorectal cancer screening and outcomes. *Hematology/Oncology Clinics* 2022; **36**(3): 415-28.
27. White PM, Itzkowitz SH. Barriers driving racial disparities in colorectal cancer screening in African Americans. *Current Gastroenterology Reports* 2020; **22**: 1-7.
28. O'Sullivan DE, Sutherland RL, Town S, et al. Risk factors for early-onset colorectal cancer: a systematic review and meta-analysis. *Clinical gastroenterology and hepatology* 2022; **20**(6): 1229-40. e5.
29. Kwah J, Schroy PC, Jacobson BC, Calderwood AH. Whites and blacks have similar risk of metachronous advanced colorectal neoplasia. *Digestive diseases and sciences* 2014; **59**: 2264-71.
30. Martínez ME, Baron JA, Lieberman DA, et al. A pooled analysis of advanced colorectal neoplasia diagnoses after colonoscopic polypectomy. *Gastroenterology* 2009; **136**(3): 832-41.

31. Laiyemo AO, Doubeni C, Brim H, et al. Short- and long-term risk of colorectal adenoma recurrence among whites and blacks. *Gastrointest Endosc* 2013; **77**(3): 447-54.
32. Kim NH, Jung YS, Yang HJ, et al. Prevalence of and Risk Factors for Colorectal Neoplasia in Asymptomatic Young Adults (20-39 Years Old). *Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association* 2019; **17**(1): 115-22.
33. Martínez ME, Baron JA, Lieberman DA, et al. A pooled analysis of advanced colorectal neoplasia diagnoses after colonoscopic polypectomy. *Gastroenterology* 2009; **136**(3): 832-41.
34. Patel SG, May FP, Anderson JC, et al. Updates on Age to Start and Stop Colorectal Cancer Screening: Recommendations From the U.S. Multi-Society Task Force on Colorectal Cancer. *Gastroenterology* 2022; **162**(1): 285-99.
35. Gupta S, Lieberman D, Anderson JC, et al. Recommendations for Follow-Up After Colonoscopy and Polypectomy: A Consensus Update by the US Multi-Society Task Force on Colorectal Cancer. *Gastroenterology* 2020; **158**(4): 1131-53.e5.
36. Song M, Emilsson L, Roelstraete B, Ludvigsson JF. Risk of colorectal cancer in first degree relatives of patients with colorectal polyps: nationwide case-control study in Sweden. *Bmj* 2021; **373**: n877.
37. Krakow M, Rising CJ, Trivedi N, Yoon DC, Vanderpool RC. Prevalence and Correlates of Family Cancer History Knowledge and Communication Among US Adults. *Preventing chronic disease* 2020; **17**: E146.
38. Brenner H, Chang-Claude J, Seiler CM, Rickert A, Hoffmeister M. Protection from colorectal cancer after colonoscopy: a population-based, case-control study. *Ann Intern Med* 2011; **154**(1): 22-30.
39. Cottet V, Pariente A, Nalet B, et al. Colonoscopic screening of first-degree relatives of patients with large adenomas: increased risk of colorectal tumors. *Gastroenterology* 2007; **133**(4): 1086-92.
40. Lieberman DA, Rex DK, Winawer SJ, Giardiello FM, Johnson DA, Levin TR. Guidelines for colonoscopy surveillance after screening and polypectomy: a consensus update by the US Multi-Society Task Force on Colorectal Cancer. *Gastroenterology* 2012; **143**(3): 844-57.
41. Jung YS, Park JH, Park DI, Sohn CI, Choi K. Weight Change and Obesity Are Associated with a Risk of Adenoma Recurrence. *Dig Dis Sci* 2016; **61**(9): 2694-703.
42. Safizadeh F, Mandic M, Schöttker B, Hoffmeister M, Brenner H. Central obesity may account for most of the colorectal cancer risk linked to obesity: evidence from the UK Biobank prospective cohort. *International journal of obesity (2005)* 2024.
43. Kim TJ, Kim JE, Choi YH, et al. Obesity-related parameters and colorectal adenoma development. *Journal of gastroenterology* 2017; **52**(12): 1221-9.
44. Shaukat A, Kaltenbach T, Dominitz JA, et al. Endoscopic Recognition and Management Strategies for Malignant Colorectal Polyps: Recommendations of the US Multi-Society Task Force on Colorectal Cancer. *Gastroenterology* 2020; **159**(5): 1916-34.e2.
45. Wu Y, de Groen PC, Natafgi N, Cai C, Wu D, Xirasagar S. Racial differences in colorectal adenomas at screening colonoscopy in the United States. *Cancer Epidemiology, Biomarkers & Prevention* 2025: OF1-OF7.

Table 1: Baseline Characteristics of the Study Population Stratified by Race/Ethnicity and Sex.

Variable	N = 59,667[1]	Ethnicity					Sex		
		Asian or Pacific Islander	Hispanic	Non-Hispanic Black	Non-Hispanic White	*P*-value[2]	Female	Male	*P*-value[2]
		N = 1,007[1]	N = 646[1]	N = 5,972[1]	N = 52,042[1]		N = 29,401[1]	N = 30,266[1]	
Demographics Characteristics									
Sex						5.9×10^{-1}			
Female	29,401 (49.3%)	495 (49.2%)	351 (54.3%)	3,435 (57.5%)	25,120 (48.3%)		-	-	
Male	30,266 (50.7%)	512 (50.8%)	295 (45.7%)	2,537 (42.5%)	26,922 (51.7%)		-	-	
Obesity	19,641 (32.9%)	124 (12.3%)	212 (32.8%)	2,810 (47.1%)	16,495 (31.7%)	2.1×10^{-1}	10,423 (35.5%)	9,218 (30.5%)	1.6×10^{-3}
Adenoma Onset Age (yrs)						2.6×10^{-1}			2.2×10^{-21}
0-49	11,018 (18.5%)	204 (20.3%)	197 (30.5%)	1,161 (19.4%)	9,456 (18.2%)		5,879 (20.0%)	5,139 (17.0%)	
≥ 50	48,649 (81.5%)	803 (79.7%)	449 (69.5%)	4,811 (80.6%)	42,586 (81.8%)		23,522 (80.0%)	25,127 (83.0%)	
Family History of Polyps	1,587 (2.7%)	10 (1.0%)	9 (1.4%)	96 (1.6%)	1,472 (2.8%)	5.5×10^{-1}	943 (3.2%)	644 (2.1%)	2.5×10^{-1}
Family History of Colorectal Cancer	8,233 (13.8%)	104 (10.3%)	70 (10.8%)	752 (12.6%)	7,307 (14.0%)	$1.7 \times 10^{-}$	4,629 (15.7%)	3,604 (11.9%)	4.9×10^{-2}
Adenoma Characteristics									
Sub-types						$1.6 \times 10^{-}$			$1.4 \times 10^{-}$
Mixed	16,455 (27.6%)	255 (25.3%)	156 (24.1%)	1,644 (27.5%)	14,400 (27.7%)		8,005 (27.2%)	8,450 (27.9%)	
Not defined	18,750 (31.4%)	260 (25.8%)	177 (27.4%)	1,495 (25.0%)	16,818 (32.3%)		9,018 (30.7%)	9,732 (32.2%)	
Serrated	2,744 (4.6%)	42 (4.2%)	26 (4.0%)	191 (3.2%)	2,485 (4.8%)		1,730 (5.9%)	1,014 (3.4%)	
Tubular	17,666 (29.6%)	403 (40.0%)	251 (38.9%)	2,168 (36.3%)	14,844 (28.5%)		8,559 (29.1%)	9,107 (30.1%)	
Tubulovillous	2,349 (3.9%)	28 (2.8%)	24 (3.7%)	325 (5.4%)	1,972 (3.8%)		1,056 (3.6%)	1,293 (4.3%)	
Villous	1,703 (2.9%)	19 (1.9%)	12 (1.9%)	149 (2.5%)	1,523 (2.9%)		1,033 (3.5%)	670 (2.2%)	
Dysplasia						3.7×10^{-2}			2.0×10^{-3}
High grade	1,117 (1.9%)	19 (1.9%)	16 (2.5%)	138 (2.3%)	944 (1.8%)		498 (1.7%)	619 (2.0%)	
Low grade	58,550 (98.1%)	988 (98.1%)	630 (97.5%)	5,834 (97.7%)	51,098 (98.2%)		28,903 (98.3%)	29,647 (98.0%)	
Number of Adenoma						3.8×10^{-1}			2.5×10^{-31}

<3	36,047 (60.4%)	718 (71.3%)	424 (65.6%)	4,245 (71.1%)	30,660 (58.9%)		18,442 (62.7%)	17,605 (58.2%)	
≥3	7,417 (12.4%)	113 (11.2%)	85 (13.2%)	743 (12.4%)	6,476 (12.4%)		3,321 (11.3%)	4,096 (13.5%)	
Not defined	16,203 (27.2%)	176 (17.5%)	137 (21.2%)	984 (16.5%)	14,906 (28.6%)		7,638 (26.0%)	8,565 (28.3%)	
Size of Adenoma (mm)						5.0×10^{-31}			$8.4 \times 10^{-}$
<10	24,670 (41.3%)	490 (48.7%)	293 (45.4%)	2,820 (47.2%)	21,067 (40.5%)		12,269 (41.7%)	12,401 (41.0%)	
≥10	6,669 (11.2%)	104 (10.3%)	75 (11.6%)	713 (11.9%)	5,777 (11.1%)		3,101 (10.5%)	3,568 (11.8%)	
Not defined	28,328 (47.5%)	413 (41.0%)	278 (43.0%)	2,439 (40.8%)	25,198 (48.4%)		14,031 (47.7%)	14,297 (47.2%)	
Adenoma Type						2.9×10^{-2}			$5.6 \times 10^{-}$
High Risk	14,654 (24.6%)	224 (22.2%)	167 (25.9%)	1,542 (25.8%)	12,721 (24.4%)		7,009 (23.8%)	7,645 (25.3%)	
Low Risk	45,013 (75.4%)	783 (77.8%)	479 (74.1%)	4,430 (74.2%)	39,321 (75.6%)		22,392 (76.2%)	22,621 (74.7%)	
Recurrence Status						3.3×10^{-1}			1.3×10^{-13}
Not Recurrence	42,071 (70.5%)	747 (74.2%)	535 (82.8%)	4,381 (73.4%)	36,408 (70.0%)		21,143 (71.9%)	20,928 (69.1%)	
Recurrence	17,596 (29.5%)	260 (25.8%)	111 (17.2%)	1,591 (26.6%)	15,634 (30.0%)		8,258 (28.1%)	9,338 (30.9%)	

(%); Mean ± SD. [2]Pearson's Chi-squared test was used to calculate *P*-values for comparisons of categorical variables between groups

Table 2: Results of Adenoma Recurrence associated with Demographic Variables for 5-Year Survival Analysis

Variable	Not Recurrence	Recurrence	*P*-value[2]	Adjusted Hazard Ratio*		
	N = 42,071[1]	N = 17,596[1]		HR[3]	95% CI[4]	*P*-value
Demographic Characteristics						
Sex[5]			1.3×10^{-13}			
Female	21,143 (71.9%)	8,258 (28.1%)		—	—	
Male	20,928 (69.1%)	9,338 (30.9%)		1.1	1.06, 1.14	$3.46 \times 10^{-}$
Ethnicity[6]			3.3×10^{-1}			
Non-Hispanic White	36,408 (70.0%)	15,634 (30.0%)		—	—	
Asian or Pacific Islander	747 (74.2%)	260 (25.8%)		0.8	0.67, 0.96	1.36×10^{-2}
Hispanic	535 (82.8%)	111 (17.2%)		0.89	0.70, 1.14	3.63×10^{-1}
Non-Hispanic Black	4,381 (73.4%)	1,591 (26.6%)		0.89	0.83, 0.96	1.97×10^{-3}
Obesity[7]	13,302 (67.7%)	6,339 (32.3%)	1.5×10^{-2}	1.18	1.13, 1.23	1.84×10^{-1}
Adenoma Onset Age (yr)[8]			8.7×10^{-1}			
0-49	8,104 (73.6%)	2,914 (26.4%)		1.17	1.11, 1.23	$6.77 \times 10^{-}$
≥ 50	33,967 (69.8%)	14,682 (30.2%)		—	—	
Family History of Polyps[9]	976 (61.5%)	611 (38.5%)	1.5×10^{-1}	1.04	0.93, 1.17	4.64×10^{-1}
Family History of Colorectal Cancer[10]	5,044 (61.3%)	3,189 (38.7%)	$2.4 \times 10^{-}$	1.24	1.18, 1.31	4.44×10^{-1}

[1]n (%); [2]Pearson's Chi-squared test; [3]HR = Hazard Ratio, [4]CI = Confidence Interval.
Model adjustments: [5]adjusted for ethnicity; stratified by age groups; [6]adjusted for sex; stratified by age group; [7]adjusted for sex and ethnicity; stratified by age group; [8]adjusted for sex, ethnicity, use of Non-steroidal anti-inflammatory drugs (NSAIDs), use of Aspirin, family history of colorectal cancer, and family history of polyps; stratified by obesity and age group; [9]adjusted for sex, ethnicity, use of NSAIDs, use of Aspirin, and family history of colorectal cancer; stratified by obesity, Adenoma Onset Stratification, and age group; [10]adjusted for sex, ethnicity, use of NSAIDs, use of Aspirin, and family history of polyps; stratified by obesity, Adenoma Onset Stratification, and age group.

Table 3: Results of Adenoma Characteristics Associated with Recurrence for 5-Year Survival Analysis

Variable	Not Recurrence	Recurrence	*P*-value[2]	Adjusted Hazard Ratio*		
	N = 42,071[1]	N = 17,596[1]		HR[3]	95% CI[4]	*P*-value
Adenoma Characteristics[¥]						
Adenoma Type[5]			3.6×10^{-1}			
Tubular	12,687 (71.8%)	4,979 (28.2%)		—	—	
Villous	721 (42.3%)	982 (57.7%)		2.55	2.31, 2.81	$4.55 \times 10^{-}$
Tubulovillous	1,584 (67.4%)	765 (32.6%)		2.18	1.99, 2.39	1.58×10^{-1}
Serrated	1,974 (71.9%)	770 (28.1%)		1.3	1.18, 1.44	$1.10 \times 10^{-}$
Not defined	13,787 (73.5%)	4,963 (26.5%)		0.53	0.48, 0.57	$1.02 \times 10^{-}$
Mixed	11,318 (68.8%)	5,137 (31.2%)		1.37	1.30, 1.45	6.99×10^{-32}
Dysplasia[6]			0.302			
Low grade	41,299 (70.5%)	17,251 (29.5%)		—	—	
High grade	772 (69.1%)	345 (30.9%)		2.86	2.54, 3.22	$4.13 \times 10^{-}$
Number of Adenoma[7]			1.7×10^{-1}			
<3	25,848 (71.7%)	10,199 (28.3%)		—	—	
≥3	4,936 (66.5%)	2,481 (33.5%)		1.69	1.60, 1.79	$5.95 \times 10^{-}$
Not defined	11,287 (69.7%)	4,916 (30.3%)		1.12	1.04, 1.21	3.81×10^{-3}
Size of Adenoma (mm)[8]			$6.8 \times 10^{-}$			
<10	17,186 (69.7%)	7,484 (30.3%)		—	—	
≥10	4,617 (69.2%)	2,052 (30.8%)		1.63	1.53, 1.73	$5.72 \times 10^{-}$
Not defined	20,268 (71.5%)	8,060 (28.5%)		0.79	0.75, 0.83	6.94×10^{-2}
Adenoma Type			5.7×10^{-3}			
High Risk	9,598 (65.5%)	5,056 (34.5%)		2.79	2.68, 2.91	2.22×10^{-1}
Low Risk	32,473 (72.1%)	12,540 (27.9%)		—	—	

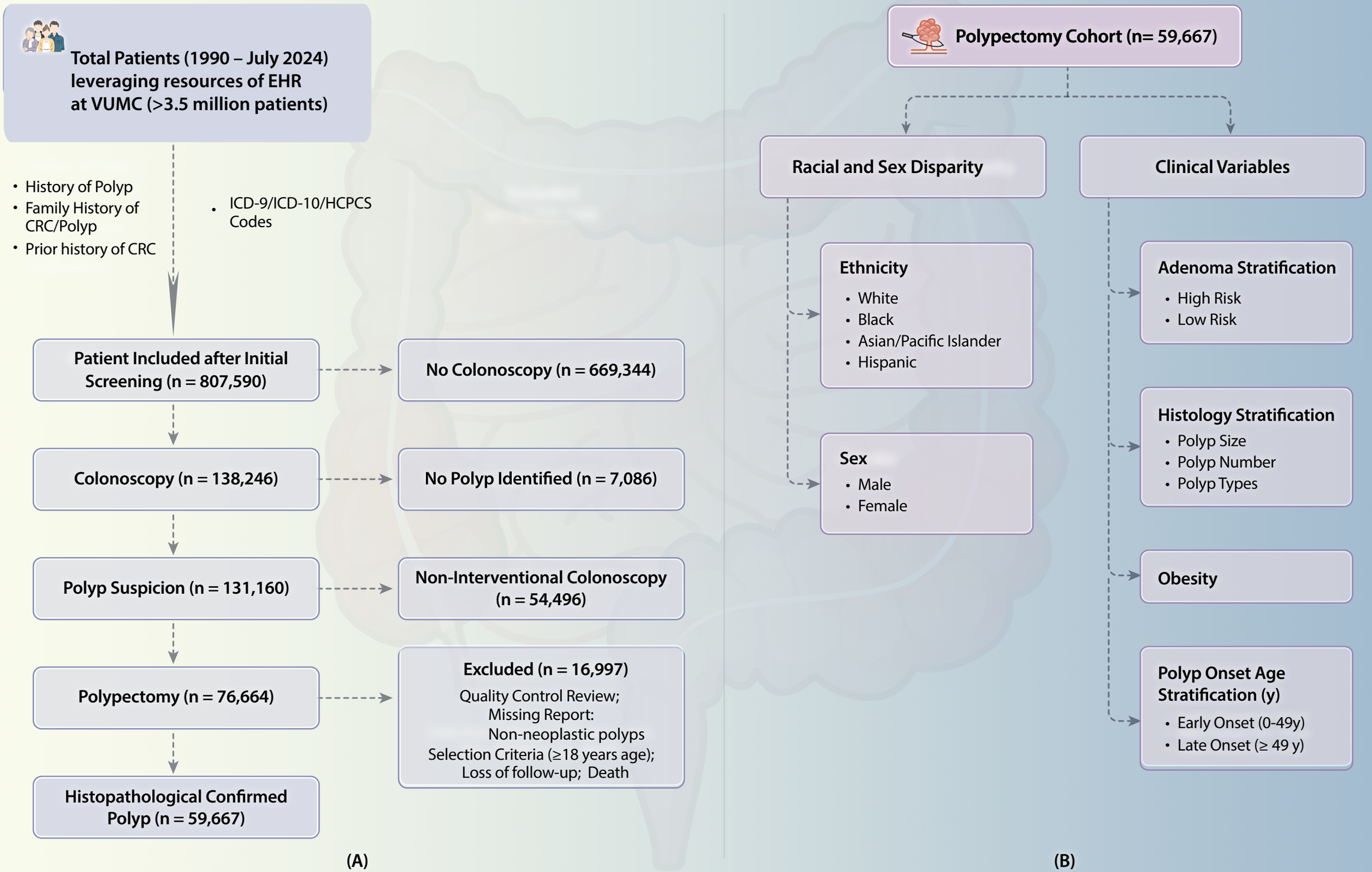
[1]n (%); [2]Pearson's Chi-squared test; [3]HR = Hazard Ratio, [4]CI = Confidence Interval.

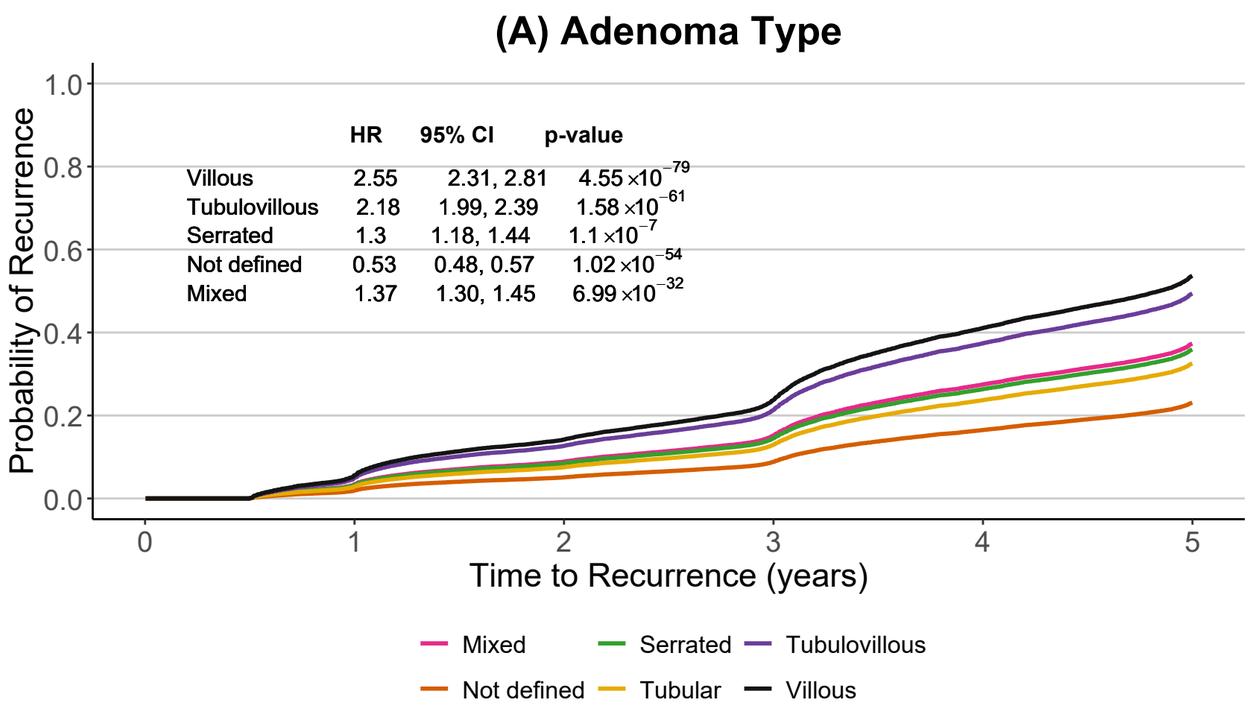
[¥]Cox proportional hazards models with time-dependent effects, adjusted for sex, ethnicity, use of nonsteroidal anti-inflammatory drugs, use of aspirin (NSAIDs), family history of colorectal cancer, and family history of polyps; models 5 to 8 additionally adjusted for other polyp-related variables (polyp type, dysplasia, number of polyps, size of polyps [mm]) when not the primary variable. All models stratified by age group; however, models 5 to 8 are additionally stratified by obesity and polyp onset stratification.

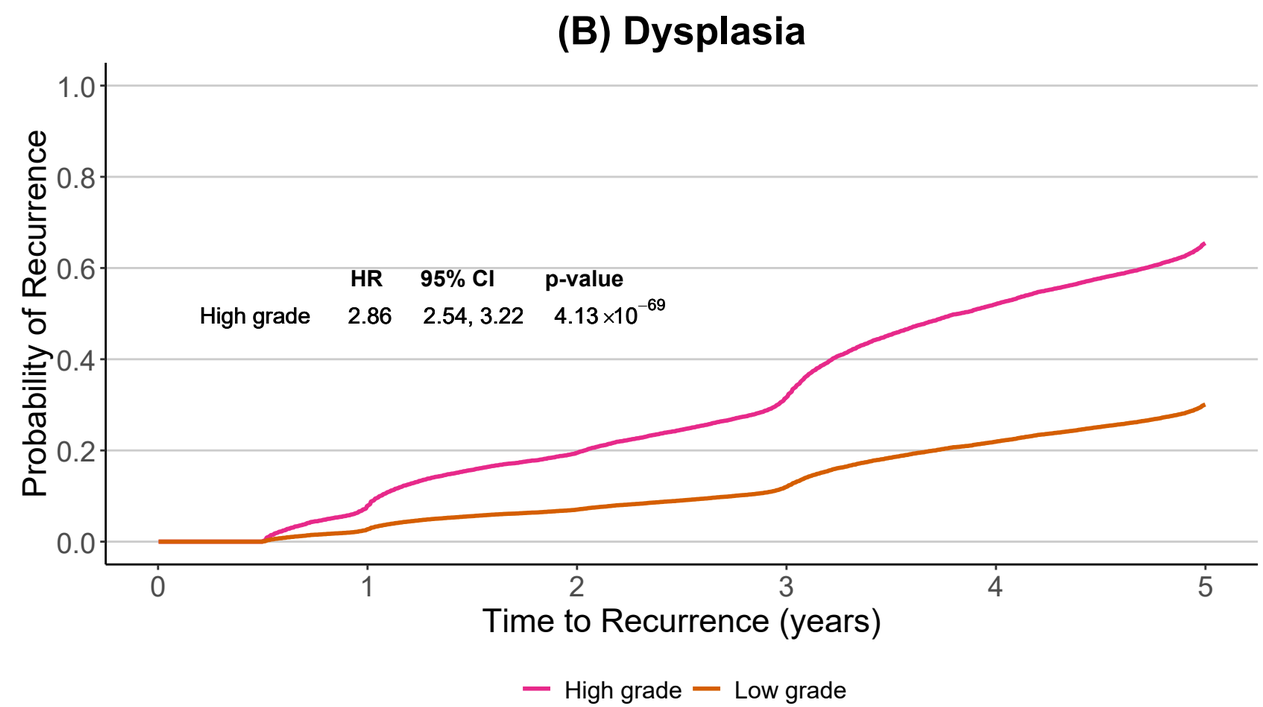
Figure Legends:

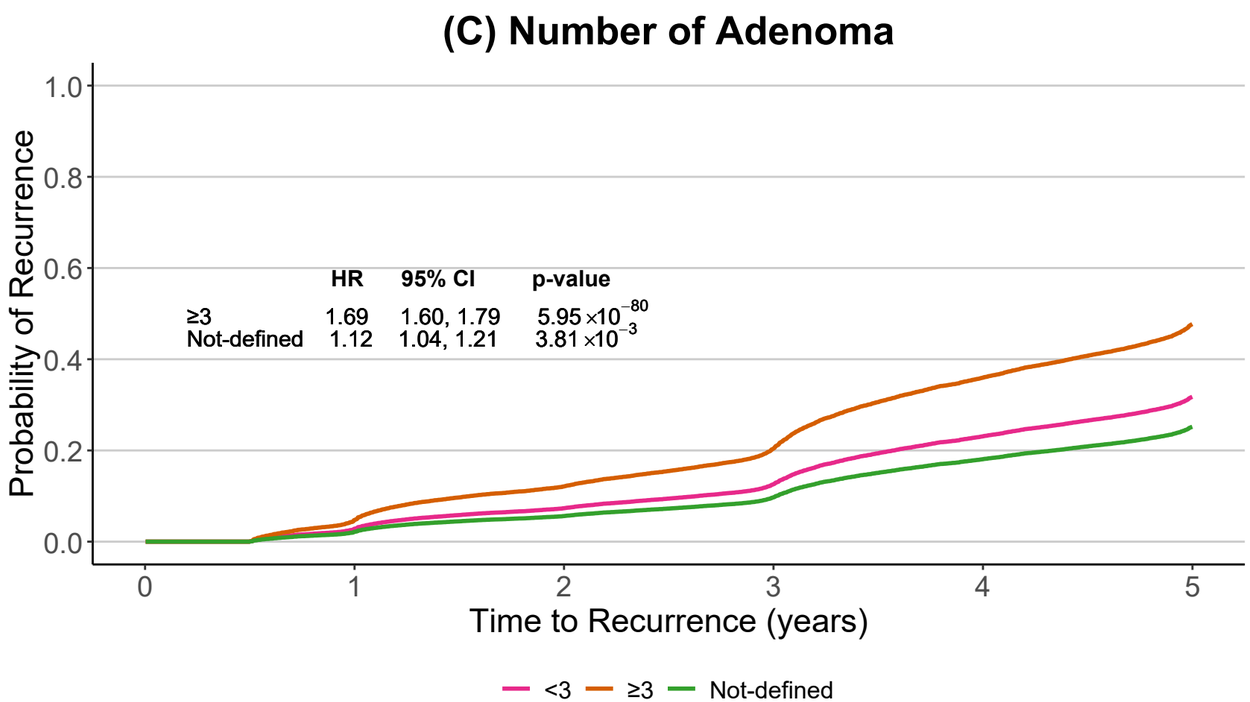
Figure 1: Patient Selection and Cohort Characteristics for the Vanderbilt Polypectomy Cohort. (A): Flowchart illustrates the inclusion and exclusion criteria for patients undergoing colonoscopy and polypectomy from 1990 to July 2024, resulting in a final cohort of 59,667 patients. (B): Stratification of the polypectomy cohort by demographic factors (ethnicity and sex) and clinical variables (adenoma risk, histology, obesity, and age at polyp onset).

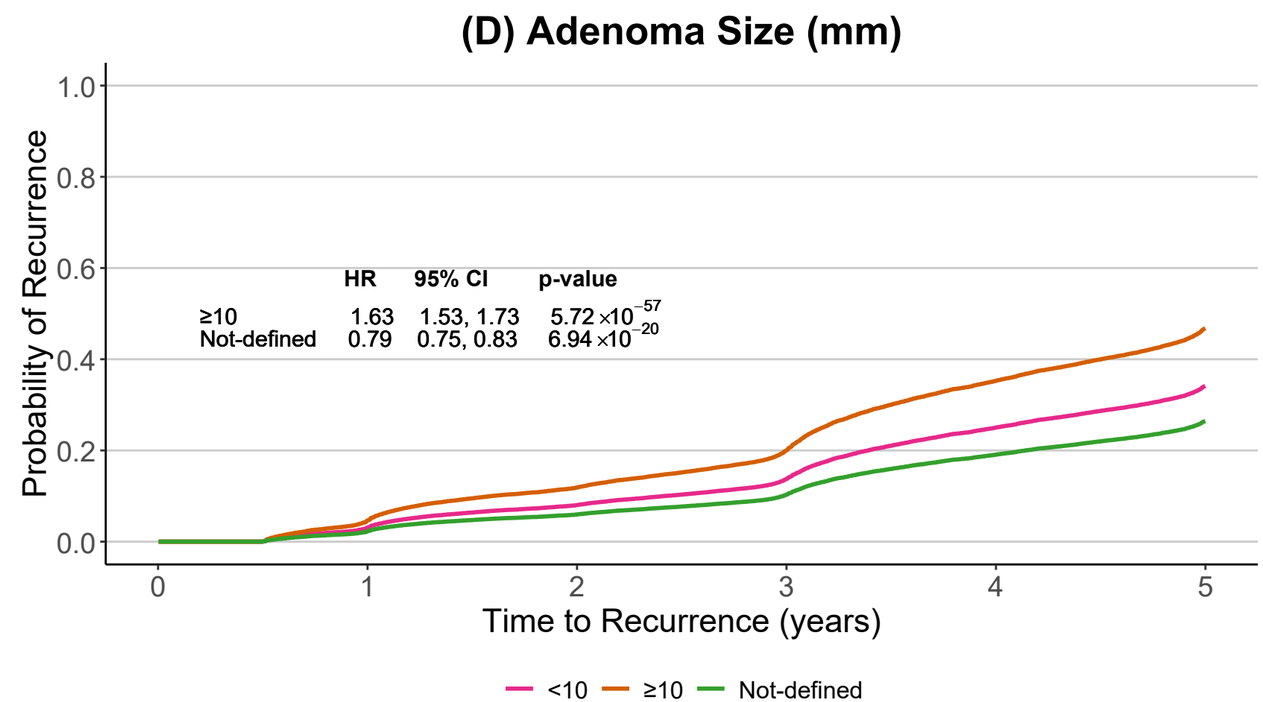
Figure 2: Adjusted Survival Curves for Time to Adenoma Recurrence After Polypectomy in the Vanderbilt Polypectomy Cohort. (A) Survival by polyp type. (B) Survival by dysplasia grade. (C) Survival by number of polyps. (D) Survival by polyp size (mm). Survival curves are derived from Cox proportional hazards models adjusted for ethnicity, sex, polyp onset age, obesity, use of NSAIDs, use of aspirin, family history of colorectal cancer, and family history of polyps, plotted for a 5-year follow-up period. *P*-values are obtained from the adjusted Cox proportional hazards models. aHR, hazard ratio; CI, confidence interval.


Total Patients (1990 – July 2024) leveraging resources of EHR at VUMC (>3.5 million patients)
History of Polyp
Family History of CRC/Polyp
Prior history of CRC
ICD-9/ICD-10/HCPCS Codes
Patient Included after Initial Screening (n = 807,590)
No Colonoscopy (n = 669,344)
Colonoscopy (n = 138,246)
No Polyp Identified (n = 7,086)
Polyp Suspicion (n = 131,160)
Non-Interventional Colonoscopy (n = 54,496)
Polypectomy (n = 76,664)
Excluded (n = 16,997)
Quality Control Review;
Missing Report:
Non-neoplastic polyps
Selection Criteria (≥18 years age);
Loss of follow-up; Death
Histopathological Confirmed Polyp (n = 59,667)
(A)
Polypectomy Cohort (n= 59,667)
Racial and Sex Disparity
Clinical Variables
Ethnicity
White
Black
Asian/Pacific Islander
Hispanic
Sex
Male
Female
Adenoma Stratification
High Risk
Low Risk
Histology Stratification
Polyp Size
Polyp Number
Polyp Types
Obesity
Polyp Onset Age Stratification (y)
Early Onset (0-49y)
Late Onset (≥ 49 y)
(B)


(A) Adenoma Type
HR 95% CI p-value
Villous 2.55 2.31, 2.81 4.55×10^−79
Tubulovillous 2.18 1.99, 2.39 1.58×10^−61
Serrated 1.3 1.18, 1.44 1.1×10^−7
Not defined 0.53 0.48, 0.57 1.02×10^−54
Mixed 1.37 1.30, 1.45 6.99×10^−32
Probability of Recurrence
Time to Recurrence (years)
Mixed
Serrated
Tubulovillous
Not defined
Tubular
Villous


(B) Dysplasia
HR 95% CI p-value
High grade 2.86 2.54, 3.22 4.13×10^−69
Probability of Recurrence
Time to Recurrence (years)
High grade
Low grade


(C) Number of Adenoma
HR 95% CI p-value
≥3 1.69 1.60, 1.79 5.95×10^−80
Not-defined 1.12 1.04, 1.21 3.81×10^−3
Probability of Recurrence
Time to Recurrence (years)
<3
≥3
Not-defined


(D) Adenoma Size (mm)
HR 95% CI p-value
≥10 1.63 1.53, 1.73 5.72×10^−57
Not-defined 0.79 0.75, 0.83 6.94×10^−20
Probability of Recurrence
Time to Recurrence (years)
<10
≥10
Not-defined