Original research

Lifestyle factors, genetic susceptibility and risk of incident diverticulitis: an integrated analysis of four prospective cohort studies and electronic health records-linked biobank

Wenjie Ma ,[1,2,3] Jane Ha,[1,2,3] Christopher J Neylan,[4] Heather Munro,[5] Duncan Skerrett,[6] Jonathan M Downie,[1,2,3] Magdalena Sevilla-González,[1,2,7] Mark Steinwandel,[5] Michael Mumma,[5] Wei Zheng,[8,9] Lillias H Maguire,[4] Edward L Giovannucci ,[10] Lisa L Strate,[11] Andrew T Chan [1,2,3,12,13]

► Additional supplemental material is published online only. To view, please visit the journal online (https://doi.org/10.1136/gutjnl-2025-335364).

For numbered affiliations see end of article.

Correspondence to
Dr Andrew T Chan;
achan@mgh.harvard.edu

WM and JH contributed equally.

Received 19 March 2025
Accepted 22 May 2025

To cite: Ma W, Ha J, Neylan CJ, *et al*. *Gut* Epub ahead of print: [*please include* Day Month Year]. doi:10.1136/gutjnl-2025-335364

ABSTRACT

Background Both lifestyle factors and genetic predisposition contribute to the development of diverticulitis.

Objective To examine whether lifestyle modification can reduce the genetic risk of diverticulitis.

Design We derived an overall healthy lifestyle score for diverticulitis based on smoking, body mass index (BMI), physical activity, fibre and red meat among 179 564 participants in three prospective cohorts-the Nurses' Health Study (NHS), NHSII and the Health Professionals Follow-Up Study. The association between the healthy lifestyle score and incident diverticulitis was confirmed among 30 750 participants in the Southern Community Cohort Study (SCCS). We assessed genetic risk using a polygenic risk score among 36 077 individuals with genotype data available. We further validated our findings in the Mass General Brigham Biobank (MGBB).

Results A healthy lifestyle score was associated with a decreased risk of diverticulitis. Compared with a score of 0, the multivariable-adjusted HR for a score of 5 was 0.50 (95% CI, 0.44 to 0.57; p trend<0.0001). This association was consistent across the SCCS in both non-Hispanic black and white populations. Each unit increase in the healthy lifestyle score was associated with a reduced diverticulitis risk similarly across genetic risk categories, with HRs of 0.89 (95% CI, 0.83 to 0.95) for low, 0.86 (0.81 to 0.92) for mid and 0.87 (0.83 to 0.91) for high genetic risk. In the MGBB cohort, a higher BMI was associated with an increased diverticulitis risk across genetic risk categories.

Conclusion Maintaining a healthy lifestyle was associated with a reduced risk of developing diverticulitis, regardless of population differences and genetic susceptibilities.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Prospective cohort studies have linked lifestyle factors to risk of diverticulitis and diverticular disease. Genome-wide association studies have identified up to 150 genetic variants associated with diverticular disease. However, no study has examined whether lifestyle changes can mitigate the genetic risk of diverticulitis or how genetic predisposition and lifestyle factors interact to influence disease risk.

WHAT THIS STUDY ADDS

⇒ We provide consistent evidence from four prospective cohort studies that a healthy lifestyle score incorporating obesity, physical activity, smoking, dietary fibre and red meat consumption was associated with a lower risk of diverticulitis across diverse populations. Moreover, regardless of genetic risk, adherence to a healthy lifestyle was associated with a reduced risk of diverticulitis in both prospective cohort studies and an electronic health records-linked biobank study.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ Maintaining a healthy lifestyle was associated with a reduced risk of developing diverticulitis, regardless of population differences and genetic susceptibilities.

INTRODUCTION

Diverticulitis, inflammation of diverticula that are small sacs forming through weak areas of the colon wall, is the sixth most common gastrointestinal indication for hospitalisation[1] and is a major reason for emergent colectomy. Patients with diverticulitis suffer from acute symptoms of pain and fever, and some may develop other complications, including peritonitis, obstruction, fistula or abscess, chronic gastrointestinal problems and recurrent episodes.[2 3]

Both genetic and lifestyle factors have been identified as critical drivers for diverticulitis and more broadly, diverticular disease. Earlier twin studies estimated that 40–53% of the heritability of diverticular disease can be attributed to genetic factors.[4 5] The candidate gene approach identified the *TNFSF15* gene encoding a cytokine of the tumour necrosis factor family associated with diverticulitis.[6] Multiple genome-wide association studies

(GWAS) have been conducted and associated 150 genetic variants with diverticular disease, pointing to a role for altered structure of the colon, gut motility, gastrointestinal mucus and ionic homeostasis that may all contribute to the development of this condition.[7–11] Polygenic risk scores (PRS) combining the influences of specific risk alleles have demonstrated predictive ability for diverticulitis overall and its severity and recurrences.[11–13]

Accumulating evidence has also linked lifestyle factors to the development of diverticulitis, including smoking, obesity, lack of physical activity and a Western dietary pattern of high intake of red and processed meat and low dietary fibre.[14 15] Adherence to a healthy lifestyle incorporating the individual components could potentially reduce the incidence of diverticulitis by up to 50%.[16]

However, with the need for data sets containing both genetic inflammation and comprehensive lifestyle data, it remains largely unknown the extent to which genetic risk of diverticulitis can be mitigated by lifestyle changes, or any potential interplay between genetic predisposition and lifestyle in influencing diverticulitis risk. In the current study, we analysed genetic and detailed lifestyle data from participants in three prospective cohorts, Nurses' Health Study (NHS), NHSII and the Health Professionals Follow-up Study (HPFS). We separately validated the healthy lifestyle score in the Southern Community Cohort Study (SCCS), a cohort with two-thirds non-Hispanic black individuals. Gene-lifestyle associations were further validated in Mass General Brigham Biobank (MGBB).

METHODS

Study population

A schema of study population is shown in online supplemental figure 1. The NHS, NHSII and HPFS are three ongoing prospective cohort studies of US health professionals. The NHS recruited 121 700 female registered nurses aged 30–55 in 1976, while the NHSII recruited 116 429 younger female nurses between 25 and 42 years old in 1989.[17] The HPFS enrolled 51 520 male health professionals aged 40–75 in 1986.[18] The participants were followed up every 2 years by questionnaires querying about health status, medication use and lifestyle behaviours. Detailed dietary information was obtained through validated food-frequency questionnaires (FFQs) every 4 years.[19 20] A subset of NHS, NHSII and HPFS participants of European ancestry had genomic data available from previous GWAS.[21–23]

The SCCS is a prospective cohort study recruiting 84 507 men and women aged 40–79 across 12 US southeastern states between March 2002 and September 2009.[24] Two-thirds of the SCCS cohort were self-reported non-Hispanic black participants; over half had annual household income <US$15 000. Baseline interviews were conducted to collect information on sociodemographics, lifestyle and medical history. Habitual dietary intake was assessed using a validated FFQ.[25]

The MGBB is a comprehensive repository that houses stored biospecimens linked to clinical information from electronic health records (EHRs) and lifestyle surveys for more than 150 000 consented patients within the MGB healthcare network in Boston, Massachusetts, USA. Genomic data is available for over 65 000 participants.[26]

Genotyping and polygenic risk score

The GWAS samples of NHS, NHSII and HPFS were genotyped in single-nucleotide polymorphism (SNP) arrays including Illumina HumanHap, Affymetrix, OmniExpress, OncoArray, HumanCoreExome2 and Global Screening Array (GSA). The MGBB samples were genotyped on three arrays offered by Illumina, including the Multi-Ethnic Genotyping Array, Multi-Ethnic Global BeadChip and the GSA. For both datasets, we used data imputed using the TopMed imputation server based on the GRCh38 human reference genome.

A PRS for diverticulitis was calculated using The Polygenic Score Catalog Calculator (pgsc_calc) with weights derived based on meta-analysed summary statistics of GWAS for diverticular disease in UK Biobank and Million Veteran Program.[27] This included all SNPs, not just the significant ones. In brief, a genome-wide association meta-analysis was performed using METAL and was comprised of two independent GWASs for diverticular disease (phecode 562), one in the Million Veteran Program cohort[28] and the other performed by the PanUKBB initiative (https://pan.ukbb.broadinstitute.org/phenotypes). The Department of Veteran Affairs Million Veteran Program is a longitudinal cohort of over 1 million veterans and collects precision and genomic information, as previously described in detail elsewhere.[29] The PanUKBB initiative undertook pan-ancestry genetic analysis of the UK biobank, a publicly available database that recruited approximately 500 000 volunteers and has been described elsewhere.[30] Polygenic risk weights were calculated using PRS-CSx based on the combined European and African population summary statistics from the GWAS meta-analysis. PRS-CSx is a Bayesian polygenic modelling framework[31] and in this study produced a weight file with >1.2 million SNPs.

Lifestyle factors and healthy lifestyle score

We derived the healthy lifestyle score by considering five major lifestyle factors that have been associated with diverticulitis,[15] including smoking, body mass index (BMI), physical activity, dietary fibre intake and total red meat intake (red and processed meat). A summary of the current evidence on the five lifestyle risk factors for diverticulitis is shown in online supplemental table 1. Participants received a score of 1 for each component if they met the low-risk lifestyle criteria and a score of 0 otherwise. Specifically, those who have never smoked were considered low-risk, as both past and current smoking have been linked to an increased risk of diverticulitis.[32 33] Individuals with a BMI of less than 25 kg/m^2 were classified as low-risk since being overweight or obese was associated with a greater risk.[34] The thresholds for low-risk levels of physical activity (top 40%), fibre intake (top 40%) and total red meat intake (lowest 40%) were determined based on cohort-specific and questionnaire cycle-specific distributions, consistent with lifestyle score definitions for major chronic diseases and guideline-recommended levels.[16 35] The overall healthy lifestyle score was calculated by summing up the scores for each individual component (ranging from 0 to 5).

Ascertainment of diverticulitis

Participants of NHS and NHSII were asked if they had diverticulitis requiring antibiotic therapy or hospitalisation on biennial questionnaires (2008, 2012, 2014, 2016 and 2020 for NHS; 2015, 2017 and 2019 for NHSII). Those who reported having diverticulitis were subsequently asked the year of the episode, dating back to 1990 for NHS and 2003 for NHSII. For HPFS, participants were asked biennially, starting in 1990 until 2020, if they were newly diagnosed with diverticulitis. If they reported diverticulitis, supplementary questionnaires were sent for the date of diagnosis, presenting symptoms, methods of diagnosis and relevant treatments. Participants were defined as having diverticulitis if they had abdominal pain related to diverticular disease along with one of the following: (1) treated with antibiotics; surgery or hospitalisation; (2) complications such

Table 1 Multivariable-adjusted HR (95% CI) of diverticulitis according to lifestyle risk factors in NHS, NHSII and HPFS

	Pooled	NHS 1990–2020 (n=56 158)	NHSII 2003–2019 (n=79 055)	HPFS 1986–2020 (n=44 351)
Cases/person-years	10 309/3 690 578	5832/1 440 078	3316/1 197 075	1151/1 051 628
BMI, kg/m^2				
<25	1 (ref)	1 (ref)	1 (ref)	1 (ref)
25.0–29.9	1.32 (1.26 to 1.38)	1.29 (1.22 to 1.37)	1.48 (1.36 to 1.61)	1.13 (0.99 to 1.29)
≥30	1.44 (1.37 to 1.52)	1.41 (1.31 to 1.51)	1.61 (1.46 to 1.76)	1.24 (1.03 to 1.50)
Smoking status				
Never smoker	1 (ref)	1 (ref)	1 (ref)	1 (ref)
Past smoker	1.17 (1.12 to 1.22)	1.16 (1.10 to 1.22)	1.20 (1.12 to 1.30)	1.13 (1.00 to 1.28)
Current smoker	1.13 (1.04 to 1.23)	1.09 (0.97 to 1.21)	1.18 (1.02 to 1.36)	1.31 (1.04 to 1.67)
Physical activity				
Q1 (lowest)	1 (ref)	1 (ref)	1 (ref)	1 (ref)
Q2	0.96 (0.91 to 1.02)	0.96 (0.89 to 1.04)	0.95 (0.86 to 1.05)	1.01 (0.84 to 1.20)
Q3	0.96 (0.90 to 1.02)	0.97 (0.90 to 1.04)	0.92 (0.83 to 1.02)	0.98 (0.82 to 1.18)
Q4	0.92 (0.87 to 0.98)	0.91 (0.84 to 0.99)	0.91 (0.82 to 1.02)	0.97 (0.81 to 1.17)
Q5 (highest)	0.84 (0.79 to 0.90)	0.82 (0.75 to 0.89)	0.84 (0.74 to 0.94)	0.94 (0.78 to 1.14)
Fibre intake				
Q1 (lowest)	1 (ref)	1 (ref)	1 (ref)	1 (ref)
Q2	0.94 (0.88 to 1.00)	0.97 (0.90 to 1.05)	0.93 (0.83 to 1.03)	0.83 (0.70 to 0.99)
Q3	0.95 (0.89 to 1.01)	1.00 (0.92 to 1.08)	0.92 (0.82 to 1.02)	0.85 (0.71 to 1.01)
Q4	0.91 (0.95 to 0.97)	0.93 (0.86 to 1.02)	0.92 (0.92 to 1.03)	0.80 (0.66 to 0.96)
Q5 (highest)	0.86 (0.80 to 0.92)	0.93 (0.85 to 1.01)	0.84 (0.75 to 0.95)	0.64 (0.51 to 0.79)
Total red meat intake (unprocessed and processed)				
Q1 (lowest)	1 (ref)	1 (ref)	1 (ref)	1 (ref)
Q2	1.09 (1.02 to 1.17)	1.07 (0.98 to 1.16)	1.08 (0.96 to 1.21)	1.28 (1.04 to 1.57)
Q3	1.13 (1.06 to 1.21)	1.09 (1.00 to 1.19)	1.14 (1.01 to 1.28)	1.39 (1.13 to 1.71)
Q4	1.10 (1.03 to 1.18)	1.06 (0.97 to 1.16)	1.15 (1.02 to 1.29)	1.21 (0.97 to 1.50)
Q5 (highest)	1.09 (1.01 to 1.17)	1.00 (0.91 to 1.10)	1.16 (1.02 to 1.31)	1.42 (1.13 to 1.79)
Healthy lifestyle score, out of 5 (components: BMI<25 kg/m^2, never smoker, PA Q4–5, fibre intake Q4–5 and red meat intake Q1–2)				
0 (unhealthy)	1 (ref)	1 (ref)	1 (ref)	1 (ref)
1	0.86 (0.81 to 0.92)	0.85 (0.78 to 0.93)	0.86 (0.76 to 0.98)	0.76 (0.63 to 0.91)
2	0.79 (0.74 to 0.84)	0.78 (0.72 to 0.85)	0.76 (0.67 to 0.86)	0.69 (0.57 to 0.83)
3	0.70 (0.65 to 0.75)	0.73 (0.66 to 0.80)	0.60 (0.53 to 0.69)	0.58 (0.48 to 0.71)
4	0.56 (0.51 to 0.61)	0.57 (0.51 to 0.64)	0.49 (0.42 to 0.60)	0.46 (0.36 to 0.60)
5 (healthy)	0.50 (0.44 to 0.57)	0.49 (0.41 to 0.60)	0.42 (0.33 to 0.53)	0.54 (0.37 to 0.78)
Per 1 increase	0.88 (0.86 to 0.89)	0.89 (0.87 to 0.90)	0.84 (0.81 to 0.86)	0.86 (0.82 to 0.90)
P for trend	<0.0001	<0.0001	<0.0001	<0.0001

Models were adjusted for alcohol intake, total calorie intake, regular aspirin, non-steroidal anti-inflammatory drug and acetaminophen use, physical examination in the past 2 years and menopausal hormone therapy (only for NHS and NHSII), and stratified by age, questionnaire cycle and cohort (in the pooled analyses).
BMI, body mass index; HPFS, Health Professionals Follow-up Study; NHS, Nurses' Health Study; PA, physical activity.

as perforation, abscess, fistula or obstruction; (3) presentation with fever, requiring medical therapy or evaluated with CT. The accuracy of ascertaining diverticulitis within the cohorts has been validated,[33 34 36] with 84–92% of self-reported cases being confirmed through medical records review.

In the SCCS cohort, diagnosis of diverticulitis was identified through linkage to the Centers for Medicare and Medicaid Services (CMS) and defined as any inpatient, emergency room or outpatient encounter with a colonic diverticulitis diagnosis based on the International Classification of Diseases (ICD) codes (ICD9: 562.11, 562.13; ICD10: K57.2, K57.32, K57.33, K57.4, K57.52, K57.53, K57.8 (K57.80, K57.81), K57.92, K57.93) listed as the primary diagnosis.

In the MGBB data set, we used the same ICD codes (ICD9: 562.11, 562.13; ICD10: K57.2, K57.32, K57.33, K57.4, K57.52, K57.53, K57.8 (K57.80, K57.81), K57.92, K57.93) to identify diverticulitis cases among participants with genomic data and matched each case with five controls according to age, sex and the last encounter year. We assessed the validity of the identification of patients with diverticulitis using ICD codes by reviewing the medical records of a randomly selected sample of 106 patients, with 91.5% of diverticulitis cases being confirmed.

Statistical analysis

We evaluated the associations of lifestyle factors and the overall healthy lifestyle score with incident diverticulitis in a combined cohort of NHS, NHSII and HPFS participants. We excluded participants with a history of diverticulitis, cancer or inflammatory bowel disease, as well as those with missing lifestyle information at baseline. Person-time was calculated from baseline until the date of diagnosis of diverticulitis, death, last follow-up

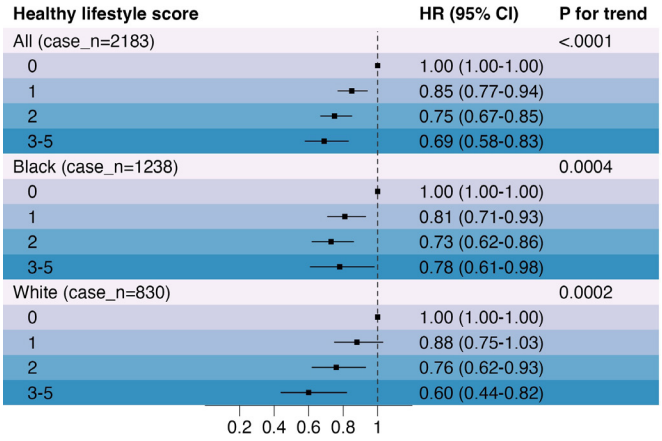

Figure 1 Associations between healthy lifestyle score and risk of incident diverticulitis among 30 750 participants in the Southern Community Cohort Study. Results from Cox proportional hazards regression models adjusted for age at enrolment, sex, race, educational attainment, annual household income, marital status, enrolment source and alcohol consumption. Missing data was imputed for race, educational attainment, annual household income, marital status and alcohol consumption using multiple imputation with five imputations. Healthy lifestyle score was derived by summing up scores for individual lifestyle factors. Participants got a score of 1 if they met the low-risk lifestyle criteria (body mass index <25 kg/m^2, never smoker, physically active (≥150 min of moderate or ≥75 vigorous leisure-time physical activity per week), fibre intake in the highest quintile, red/processed meat intake in the lowest quintile) and 0 otherwise. Scores 3, 4 and 5 were combined due to small numbers in each group. No statistically significant interaction was observed between a healthy lifestyle score and race on the risk of diverticulitis, based on the Wald test of a product term of healthy lifestyle score and race (p for interaction=0.33).

questionnaire or the end of the study period (1990–2020 for NHS, 2003–2019 for NHSII and 1986–2020 for HPFS), whichever came first. We used time-varying Cox proportional hazards regression models stratified by age, questionnaire cycle and cohort and adjusted for other demographical and covariates, including alcohol intake, total calorie intake, regular medication use (aspirin, non-steroidal anti-inflammatory drug or acetaminophen), physical examination in the past 2 years and menopausal hormone therapy (only for women). Lifestyle factors and covariates were updated by using the most recent information prior to the questionnaire cycle of interest to account for changes over time.

Among participants with genomic data, we examined the association of healthy lifestyle score and PRS individually with diverticulitis, as well as the association between healthy lifestyle score and diverticulitis across PRS categories. We assessed multiplicative interaction using the Wald test by including a product term between continuous measures of lifestyle score and PRS. We also estimated additive interaction between continuous measures of PRS and unhealthy lifestyle score (evaluated in reverse from healthy lifestyle score so effect estimate was positive) by calculating the relative excess risk due to interactions (RERIs), with RERI >0 indicating the presence of significant additive interaction between PRS and unhealthy lifestyle.[37] We also calculated population-attributable risk (PAR) for lifestyle score overall and within each PRS category.

In the SCCS, we excluded any participants with prevalent diverticulitis or missing lifestyle information and restricted the remaining SCCS participants to those individuals aged ≥65 years at cohort enrolment, or persons<65 years at enrolment who: (1) reported being covered by Medicaid (provides medical benefits to low-income adults and uninsured persons) or Medicare (the primary health insurance programme for persons aged ≥65 or those with disability under age 65) on the baseline questionnaire; or (2) did not report Medicare or Medicaid on the baseline questionnaire but had a CMS claim within 1 year of being enrolled in SCCS. A total of 30 750 participants were included in the analysis of lifestyle and incident diverticulitis using Cox proportional hazard models. Participants were censored on the date of death, diagnosis of diverticulitis, loss to follow-up or the end of claims data linkage in 2021. We similarly examined the association between healthy lifestyle score and incident diverticulitis, as well as potential effect modification by race, with significance evaluated using the Wald test of a product term of healthy lifestyle score and race.

We conducted a case-control study within the MGBB data set, matching each diverticulitis case with five controls according to age, sex and year of last encounter. Logistic regression models were used to evaluate the associations of PRS and diverticulitis. To examine potential interaction with lifestyle, we examined BMI in relation to PRS since data on other lifestyle factors were not consistently collected in this cohort.

Patient involvement

No patients were involved in setting the research question or the outcome measures, nor were they involved in the design and implementation of the study. There are no plans to involve patients in dissemination.

Table 2 Association of PRS with diverticulitis risk and its interaction with age in NHS, NHSII and HPFS

		Age subgroups			P value for interaction*
	All participants	<60 years	60–69 years	≥70 years	
Cases/person-years	2357/808 408	651/291 530	832/252 267	874/264 611	–
Age-adjusted HR (95% CI) per 1 SD increase	1.59 (1.54 to 1.66)	1.83 (1.69 to 1.98)	1.62 (1.51 to 1.73)	1.42 (1.33 to 1.52)	<0.0001
Multivariable-adjusted HR (95% CI) per 1 SD increase†	1.58 (1.52 to 1.65)	1.81 (1.67 to 1.96)	1.61 (1.50 to 1.72)	1.42 (1.33 to 1.52)	<0.0001

Cox proportional hazards regression models were stratified by age, questionnaire cycle and cohort.
*P value for interaction was assessed using the Wald test by including a product term between continuous age and continuous PRS.
†Multivariable-adjusted models were adjusted for body mass index, aspirin use, non-steroidal anti-inflammatory drug use, acetaminophen use, physical examination, smoking status, physical activity, menopausal hormone use, fibre, red meat, alcohol consumption, total calorie intake.
HPFS, Health Professionals Follow-Up Study; NHS, Nurses' Health Study; PRS, polygenic risk score.

Table 3 Association of healthy lifestyle score with incident diverticulitis according to PRS tertile in NHS, NHSII and HPFS

	All participants	Low PRS	Middle PRS	High PRS	P value for multiplicative interaction*	P value for additive interaction†
Cases/person-years	2357/808 408	470/276 342	726/272 013	1161/26 005	–	
Healthy lifestyle score, multivariable-adjusted HR (95% CI)‡						
0 (unhealthy)	1 (ref)	1 (ref)	1 (ref)	1 (ref)	0.67	<0.0001
1	0.81 (0.71 to 0.93)	1.24 (0.88 to 1.74)	0.74 (0.58 to 0.94)	0.71 (0.59 to 0.86)		
2	0.68 (0.59 to 0.78)	0.88 (0.62 to 1.25)	0.67 (0.52 to 0.85)	0.62 (0.51 to 0.75)		
3	0.66 (0.57 to 0.76)	0.99 (0.70 to 1.41)	0.58 (0.45 to 0.75)	0.60 (0.49 to 0.74)		
4,5 (healthy)	0.50 (0.43 to 0.59)	0.63 (0.42 to 0.94)	0.48 (0.36 to 0.64)	0.50 (0.39 to 0.63)		
Per 1 increase	0.87 (0.84 to 0.90)	0.89 (0.83 to 0.95)	0.86 (0.81 to 0.92)	0.87 (0.83 to 0.91)		
P for trend	<0.0001	0.001	<0.0001	<0.0001		
PAR of healthy lifestyle score§	26% (18–34%)	42% (37–46%)	31% (13–47%)	23% (11–35%)		

*P value for multiplicative interaction was assessed using the Wald test by including a product term between continuous healthy lifestyle score and continuous PRS.
†P value for additive interaction was assessed between continuous unhealthy lifestyle score (evaluated in reverse from healthy lifestyle score so effect estimate was positive) and continuous PRS. Relative excess risk due to interaction (RERI) was 0.09 (95% CI, 0.06 to 0.13).
‡Models were adjusted for alcohol intake, total calorie intake, regular aspirin, non-steroidal anti-inflammatory drug and acetaminophen use, physical examination in the past 2 years and menopausal hormone therapy, and were stratified by age, questionnaire cycle and cohort.
§PAR was calculated with healthy lifestyle score of 4 or 5 as reference, indicating the percentage of diverticulitis that can be prevented if all individuals were in this group.
HPFS, Health Professionals Follow-Up Study; NHS, Nurses' Health Study; PAR, population attributable risk; PRS, polygenic risk score.

RESULTS

In a combined cohort of 179 564 participants from NHS, NHSII and HPFS, we documented a total of 10 299 incident diverticulitis cases over an average follow-up of 20 years. Individual lifestyle factors were each significantly associated with the incidence of diverticulitis (table 1). For example, compared with participants with a BMI <25 kg/m^2, the multivariable-adjusted HR was 1.32 (95% CI, 1.26 to 1.38) for those who were overweight and 1.44 (95% CI, 1.37 to 1.52) for those who were obese. Both past smokers (HR, 1.17; 95% CI, 1.12 to 1.22) and current smokers (HR, 1.13; 95% CI, 1.04 to 1.23) had an increased risk of diverticulitis when compared with those who never smoked. Higher levels of physical activity were associated with a reduced risk, with an HR of 0.84 (95% CI, 0.79 to 0.90) comparing participants in the highest to the lowest quintile. Finally, a greater intake of fibre was associated with a reduced risk (extreme-quintile HR, 0.86; 95% CI, 0.80 to 0.92), whereas a greater intake of red meat was associated with an increased risk (extreme-quintile HR, 1.09; 95% CI, 1.01 to 1.17). When all five lifestyle factors were considered together, a higher score indicating a healthier lifestyle was linearly linked to a lower risk of diverticulitis (HR per 1-point increase: 0.88; 95% CI, 0.86 to 0.89). Participants with a score of 5 had an HR of 0.50 (95% CI, 0.44 to 0.57) when compared with those with a score of 0.

The association between the healthy lifestyle score and incident diverticulitis was consistently observed across NHS, NHSII and HPFS (table 1). In the SCCS cohort, a total of 2183 incident diverticulitis cases were ascertained during an average follow-up period of 11.9 years. A healthy lifestyle score similarly showed a significant association with diverticulitis (figure 1). Those with a healthy lifestyle score of 3–5 had a substantially lower risk (HR: 0.69; 95% CI: 0.58 to 0.83) compared with those with a score of 0. No notable differences were found comparing the white (HR, 0.60; 95% CI, 0.44 to 0.82) and black (HR, 0.78; 95% CI, 0.61 to 0.98; p interaction=0.33) populations.

Among the subset of NHS, NHSII and HPFS population with available genomic data, there were no notable differences in lifestyle factors across PRS categories (online supplemental table 2). PRS was significantly associated with incident diverticulitis (table 2). The adjusted HR was 1.58 (95% CI, 1.52 to 1.65) for each SD increase in PRS. The link between PRS and diverticulitis appeared to be stronger among younger participants. The HR for diverticulitis with each SD increase in PRS was 1.81 (95% CI, 1.67 to 1.96) for those younger than 60 years, 1.61 (95% CI, 1.50 to 1.72) for those aged 60–69 years and 1.42 (95% CI, 1.33 to 1.52) for those over 70 years (p interaction <0.0001).

The association between the healthy lifestyle score and diverticulitis was similar across categories of PRS (table 3). For instance, compared with participants with a score of 0, those with a score of 4–5 had an HR (95% CI) of 0.63 (0.42 to 0.94) for individuals in the lowest tertile of PRS, 0.48 (0.36 to 0.64) in the middle tertile and 0.50 (0.39 to 0.63) in the highest tertile (p for multiplicative interaction=0.67). PAR analysis showed that adopting a healthy lifestyle could prevent 23–42% of diverticulitis cases across PRS categories.

Table 4 Risk of diverticulitis according to joint categories of BMI and PRS in MGBB

	BMI			P value for additive interaction*
PRS	<25 kg/m^2	25–29.9 kg/m^2	≥30 kg/m^2	
Low	1 (ref)	1.27 (1.00 to 1.61)	1.73 (1.37 to 2.18)	<0.0001
Middle	1.46 (1.14 to 1.87)	2.38 (1.92 to 2.96)	2.96 (2.38 to 3.68)	
High	3.12 (2.50 to 3.89)	4.36 (3.53 to 5.36)	5.55 (4.51 to 6.82)	

*P value for additive interaction was assessed between continuous BMI and continuous PRS. Relative excess risk due to interaction (RERI) was 0.13 (95% CI, 0.09 to 0.17).
BMI, body mass index; MGBB, Mass General Brigham Biobank; PRS, polygenic risk score.

We also examined the combined impact of PRS and lifestyle on the risk of diverticulitis. Individuals in the highest PRS group with a healthy lifestyle score of 0 or 1 were 4.8-fold more likely to develop diverticulitis compared with those in the lowest PRS group with a score of 4 or 5 (online supplemental figure 2). There was a statistically significant additive interaction between lifestyle score and PRS when both were evaluated as continuous variables (RERI: 0.09; 95% CI, 0.06 to 0.13; p for additive interaction <0.0001), indicating that with each 1-point decrease in healthy lifestyle score and each SD increase in PRS, the HR of developing diverticulitis was 0.09 higher than it would be without the interaction.

In our validation MGBB cohort, an association between PRS and diverticulitis risk was confirmed. We found that for each SD increase in PRS, the odds of developing diverticulitis increased by 1.67 times (95% CI, 1.60 to 1.75). A higher BMI was associated with an increased risk of diverticulitis regardless of the genetic risk level (table 4). Individuals in the highest PRS tertile with a BMI ≥30 kg/m^2 were 5.55 times (95% CI, 4.51 to 6.82) more likely to develop diverticulitis compared with those in the lowest PRS tertile with a BMI <25 kg/m^2. Furthermore, a similar additive interaction between BMI and PRS on diverticulitis risk was also observed (RERI, 0.13; 95% CI, 0.09 to 0.17; p for additive interaction <0.0001).

DISCUSSION

In the current study, we showed that a healthy lifestyle score, comprising five components, including no smoking, normal BMI, adequate physical activity, high fibre intake and low red meat intake, was associated with a lower risk of diverticulitis across multiple cohorts, including a cohort comprised of predominantly non-Hispanic black participants. Conversely, individuals with a higher PRS were shown to be at increased risk of developing diverticulitis. However, regardless of the genetic risk, adherence to a healthy lifestyle was associated with a significantly decreased risk of the disease in both prospective cohort studies and an EHR-linked biobank study.

Our results support the benefits of adherence to a healthy lifestyle in reducing the risk of developing diverticulitis among various racial groups. Existing evidence on the associations of lifestyle, including obesity,[36 38] physical activity,[38 39] smoking[32 33] and diet,[40 41] with the risk of incident diverticulitis was largely from prospective cohorts of predominantly non-Hispanic white populations, whereas the role lifestyle plays in diverticulitis risk among other minority groups remains poorly understood.[42] The risk of diverticulitis or other diverticular disease may vary by race or ethnicity. For example, non-Hispanic black individuals were reported to have the highest prevalence of proximal diverticulosis among those who underwent colonoscopies,[43] while analyses using National Inpatient or Emergency Department Visits data have found that the prevalence of diverticular bleeding was highest in blacks, whereas that of diverticulitis was highest in whites.[44 45] By leveraging data from the SCCS cohort including 65% non-Hispanic black individuals, we were able to enhance the generalisability of our findings and offer insight into the impact of lifestyle on diverticulitis risk in diverse populations.

While the association between overall lifestyle score and diverticulitis was similar across cohorts, we observed differences for individual lifestyle factors. For example, the inverse association of physical activity with diverticulitis in women (NHS/NHSII) was linear. In men (HPFS), however, only vigorous physical activity showed a significant association.[39] Overweight and obesity appeared to be a more important risk factor in women, whereas diet showed a stronger association with diverticulitis in men. These differences may reflect variation in the relative contribution of potential mechanisms, such as metabolic dysfunction or inflammation, in the development of diverticulitis across populations.

It should be noted that existing GWASs have focused on diverticular disease as the primary outcome due to limitations in the diagnostic coding within biobank studies, making it challenging to distinguish diverticulitis from diverticulosis. Despite this, recent studies have demonstrated that using a PRS combining the subtle effects of individual genetic variants associated with diverticular disease can successfully predict the risk of diverticulitis.[11–13] A higher PRS was associated with a greater risk of diverticulitis as well as severe diverticulitis and recurrent diverticulitis.[12 13] We here validated the association between PRS and incident diverticulitis in both population-based cohort and EHR-linked biobank. Additionally, the association on the relative scale appeared to be more pronounced in individuals under the age of 60, indicating that PRS might be more predictive for diverticulitis in younger populations.

There was no statistically significant multiplicative interaction between the healthy lifestyle score and PRS in relation to diverticulitis, suggesting that the impact of genetic or lifestyle factors on diverticulitis does not depend on each other. These results suggest that most individuals, regardless of genetic risk, could achieve similar relative benefits in reducing the risk of diverticulitis by adopting a healthy lifestyle. However, we observed a modest yet significant interaction on the additive scale, indicating the combined impact of genetic susceptibility and lifestyle resulted in a slightly greater risk than the sum of the individual effects when considered independently. This may suggest that the absolute risk of diverticulitis associated with a poor lifestyle varies according to genetic risk, possibly depending on the baseline risk at the population level. These findings have direct implications for both clinical practice and public health.

The notable strengths of this study include the incorporation of well-established prospective cohort studies including diverse racial and ethnic populations leading to greater generalisability, detailed and repeated collection of lifestyle factors to minimise the potential for measurement error, the utilisation of PRS based on results from the latest and largest GWAS, as well as the validation of genetic-lifestyle associations with diverticulitis in an external biobank study. Nevertheless, it should be acknowledged that we were unable to examine the interaction between PRS-lifestyle derived from the NHS, NHSII and HPFS cohorts with the full healthy lifestyle score in the MGB Biobank given the available data on lifestyle. Further, the ascertainment of diverticulitis was based on different approaches (ie, self-reports, ICD codes) in different cohorts. However, the comparative validity of these methods has been demonstrated to be high. Moreover, variation in the ascertainment of diverticulitis within each individual cohort would be expected to underestimate the consistency of the associations that we observed between PRS, lifestyle score and diverticulitis across the cohorts.

In conclusion, our data provide consistent evidence from multiple data sets indicating that adherence to a healthy lifestyle is linked to a reduced risk of developing diverticulitis, irrespective of one's genetic predisposition.

Author affiliations
[1]Clinical and Translational Epidemiology Unit, Massachusetts General Hospital, Boston, Massachusetts, USA
[2]Department of Medicine, Harvard Medical School, Boston, Massachusetts, USA
[3]Division of Gastroenterology, Massachusetts General Hospital, Boston, Massachusetts, USA

[4]Department of Surgery, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania, USA
[5]Vanderbilt Institute for Clinical and Translational Research, Vanderbilt University Medical Center, Nashville, Tennessee, USA
[6]Department of Medicine, University of Washington School of Medicine, Seattle, Washington, USA
[7]Programs in Metabolism and Medical and Population Genetics, The Broad Institute of MIT and Harvard, Cambridge, Massachusetts, USA
[8]Division of Epidemiology, Department of Medicine, Vanderbilt University Medical Center, Nashville, Tennessee, USA
[9]Vanderbilt-Ingram Cancer Center, Vanderbilt University Medical Center, Nashville, Tennessee, USA
[10]Departments of Nutrition and Epidemiology, Harvard T H Chan School of Public Health, Boston, Massachusetts, USA
[11]Division of Gastroenterology and Hepatology, University of Wisconsin-Madison School of Medicine and Public Health, Madison, Wisconsin, USA
[12]Department of Immunology and Infectious Diseases, Harvard T H Chan School of Public Health, Boston, Massachusetts, USA
[13]Cancer Center, Massachusetts General Hospital, Boston, MA, USA

Acknowledgements We thank the participants and staff of Nurses' Health Studies, Health Professionals Follow-Up Study, the Southern Community Cohort Study and the Mass General Brigham Biobank.

Contributors Study concept and design: WM, JH, CJN, WZ, LHM, LLS, AC; acquisition of data: WM, JH, CJN, DS, JMD, MS-G, MS, MM, WZ, EG, LLS, AC; statistical analysis: WM, JH, HM; interpretation of data: all authors; writing—original draft: WM, JH, AC; writing—review and editing: all authors. AC is the guarantor for this work and accepts full responsibility for the conduct of the study, the integrity of the data and the decision to publish.

Funding This work is supported by the National Institutes of Health (NIH) grants UM1 CA186107, R01 CA49449, U01 CA176726, R01 CA67262, U01 CA167552, U01 CA202979, R01 DK101495 (EG, LLS, AC), R01 DK131694 (LLS), the American Gastroenterological Association Research Foundation's Research Scholar Awards AGA2021-13-01 (WM), NIH 1K01DK135854-01A1 (WM) and Massachusetts General Hospital (MGH) Claflin Distinguished Scholar Award (WM). AC is an American Cancer Society Clinical Research Professor. MS-G is supported by the American Diabetes Association grant 9-22-PDFPM-04 and by the NIH grant 5U24DK132733-02. JMD is supported by the MGH Research Training in Digestive Diseases training grant (NIH grant T32 DK007191) and the NIH Loan Repayment Program (L30DK137289). LHM is supported by NIH grant 1K08DK124687. CJN is supported by the American Society of Colon and Rectal Surgeons Surgery Resident Research Initiation Grant.

Competing interests AC served as a consultant for Pfizer, Bayer Pharma AG and Boehringer Ingelheim and received grants from Pfizer, Zoe and Freenome for work unrelated to the topic. LLS is on Data Monitoring Committee for Medtronic for work unrelated to the topic.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Ethics approval The study was approved by the Institutional Review Boards of Mass General Brigham and Harvard T H Chan School of Public Health (2014P000837; 2023P001404). Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request. Because of participant confidentiality and privacy concerns, data cannot be shared publicly and requests to data access must be submitted in writing. According to standard controlled access procedures, applications to use NHS/NHSII/HPFS resources will be reviewed by our External Collaborations Committee to verify that the proposed use maintains the protection of the privacy of participants and the confidentiality of the data. Investigators wishing to use NHS/NHSII/HPFS data are asked to submit a brief description of the proposed project (go to https://www.nurseshealthstudy.org/researchers (contact email: nhsaccess@channing.harvard.edu) and https://sites.sph.harvard.edu/hpfs/for-collaborators/ for details). For SCCS cohort data, a Request for Data and Biospecimen Use must be submitted through the SCCS online request system (ORS) at https://ors.southerncommunitystudy.org/, or by following the 'Information for Researchers' link at the SCCS website (www.southerncommunitystudy.org). Any questions concerning the electronic submission and review process can be emailed to datause@southerncommunitystudy.org. For MGB Biobank data, requests need to be sent to biobank@mgb.org.

ORCID iDs
Wenjie Ma http://orcid.org/0000-0001-9894-7072
Edward L Giovannucci http://orcid.org/0000-0002-6123-0219
Andrew T Chan http://orcid.org/0000-0001-7284-6767

REFERENCES

1 Peery AF, Crockett SD, Murphy CC, *et al*. Burden and Cost of Gastrointestinal, Liver, and Pancreatic Diseases in the United States: Update 2021. *Gastroenterology* 2022;162:621–44.
2 Andeweg CS, Berg R, Staal JB, *et al*. Patient-reported Outcomes After Conservative or Surgical Management of Recurrent and Chronic Complaints of Diverticulitis: Systematic Review and Meta-analysis. *Clin Gastroenterol Hepatol* 2016;14:183–90.
3 Bharucha AE, Parthasarathy G, Ditah I, *et al*. Temporal Trends in the Incidence and Natural History of Diverticulitis: A Population-Based Study. *Am J Gastroenterol* 2015;110:1589–96.
4 Granlund J, Svensson T, Olén O, *et al*. The genetic influence on diverticular disease--a twin study. *Aliment Pharmacol Ther* 2012;35:1103–7.
5 Strate LL, Erichsen R, Baron JA, *et al*. Heritability and familial aggregation of diverticular disease: a population-based study of twins and siblings. *Gastroenterology* 2013;144:736–42.
6 Connelly TM, Berg AS, Hegarty JP, *et al*. The TNFSF15 gene single nucleotide polymorphism rs7848647 is associated with surgical diverticulitis. *Ann Surg* 2014;259:1132–7.
7 Sigurdsson S, Alexandersson KF, Sulem P, *et al*. Sequence variants in ARHGAP15, COLQ and FAM155A associate with diverticular disease and diverticulitis. *Nat Commun* 2017;8:15789.
8 Maguire LH, Handelman SK, Du X, *et al*. Genome-wide association analyses identify 39 new susceptibility loci for diverticular disease. *Nat Genet* 2018;50:1359–65.
9 Schafmayer C, Harrison JW, Buch S, *et al*. Genome-wide association analysis of diverticular disease points towards neuromuscular, connective tissue and epithelial pathomechanisms. *Gut* 2019;68:854–65.
10 Joo YY, Pacheco JA, Thompson WK, *et al*. Multi-ancestry genome- and phenome-wide association studies of diverticular disease in electronic health records with natural language processing enriched phenotyping algorithm. *PLoS One* 2023;18:e0283553.
11 Wu Y, Goleva SB, Breidenbach LB, *et al*. 150 risk variants for diverticular disease of intestine prioritize cell types and enable polygenic prediction of disease susceptibility. *Cell Genom* 2023;3:100326.
12 De Roo AC, Chen Y, Du X, *et al*. Polygenic Risk Prediction in Diverticulitis. *Ann Surg* 2023;277:e1262–8.
13 Schaeffer HD, Smelser DT, Rao HS, *et al*. Development of a Polygenic Risk Score to Predict Diverticulitis. *Dis Colon Rectum* 2024;67:254–63.
14 Strate LL. Lifestyle factors and the course of diverticular disease. *Dig Dis* 2012;30:35–45.
15 Strate LL, Morris AM. Epidemiology, Pathophysiology, and Treatment of Diverticulitis. *Gastroenterology* 2019;156:1282–98.
16 Liu P-H, Cao Y, Keeley BR, *et al*. Adherence to a Healthy Lifestyle is Associated With a Lower Risk of Diverticulitis among Men. *Am J Gastroenterol* 2017;112:1868–76.
17 Bao Y, Bertoia ML, Lenart EB, *et al*. Origin, Methods, and Evolution of the Three Nurses' Health Studies. *Am J Public Health* 2016;106:1573–81.
18 Rimm EB, Giovannucci EL, Stampfer MJ, *et al*. Reproducibility and validity of an expanded self-administered semiquantitative food frequency questionnaire among male health professionals. *Am J Epidemiol* 1992;135:1114–26.
19 Yuan C, Spiegelman D, Rimm EB, *et al*. Validity of a Dietary Questionnaire Assessed by Comparison With Multiple Weighed Dietary Records or 24-Hour Recalls. *Am J Epidemiol* 2017;185:570–84.
20 Al-Shaar L, Yuan C, Rosner B, *et al*. Reproducibility and Validity of a Semiquantitative Food Frequency Questionnaire in Men Assessed by Multiple Methods. *Am J Epidemiol* 2021;190:1122–32.
21 Qi Q, Chu AY, Kang JH, *et al*. Sugar-sweetened beverages and genetic risk of obesity. *N Engl J Med* 2012;367:1387–96.
22 Ding M, Ahmad S, Qi L, *et al*. Additive and Multiplicative Interactions Between Genetic Risk Score and Family History and Lifestyle in Relation to Risk of Type 2 Diabetes. *Am J Epidemiol* 2020;189:445–60.
23 Thomas G, Jacobs KB, Kraft P, *et al*. A multistage genome-wide association study in breast cancer identifies two new risk alleles at 1p11.2 and 14q24.1 (RAD51L1). *Nat Genet* 2009;41:579–84.
24 Signorello LB, Hargreaves MK, Steinwandel MD, *et al*. Southern community cohort study: establishing a cohort to investigate health disparities. *J Natl Med Assoc* 2005;97:972–9.

25 Signorello LB, Munro HM, Buchowski MS, *et al*. Estimating nutrient intake from a food frequency questionnaire: incorporating the elements of race and geographic region. *Am J Epidemiol* 2009;170:104–11.
26 Boutin NT, Schecter SB, Perez EF, *et al*. The Evolution of a Large Biobank at Mass General Brigham. *J Pers Med* 2022;12:1323.
27 Neylan CJ, Levin MG, Hartmann K, *et al*. Genome-wide association meta-analysis identifies 126 novel loci for diverticular disease and implicates connective tissue and colonic motility. *medRxiv* 2025:2025.03.27.25324777.
28 Verma A, Huffman JE, Rodriguez A, *et al*. Diversity and scale: Genetic architecture of 2068 traits in the VA Million Veteran Program. *Science* 2024;385:eadj1182.
29 Gaziano JM, Concato J, Brophy M, *et al*. Million Veteran Program: A mega-biobank to study genetic influences on health and disease. *J Clin Epidemiol* 2016;70:214–23.
30 Sudlow C, Gallacher J, Allen N, *et al*. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med* 2015;12:e1001779.
31 Ruan Y, Lin Y-F, Feng Y-CA, *et al*. Improving polygenic prediction in ancestrally diverse populations. *Nat Genet* 2022;54:573–80.
32 Hjern F, Wolk A, Håkansson N. Smoking and the risk of diverticular disease in women. *Br J Surg* 2011;98:997–1002.
33 Gunby SA, Ma W, Levy MJ, *et al*. Smoking and Alcohol Consumption and Risk of Incident Diverticulitis in Women. *Clin Gastroenterol Hepatol* 2024;22:1108–16.
34 Ma W, Jovani M, Liu P-H, *et al*. Association Between Obesity and Weight Change and Risk of Diverticulitis in Women. *Gastroenterology* 2018;155:58–66.
35 Chomistek AK, Chiuve SE, Eliassen AH, *et al*. Healthy lifestyle in the primordial prevention of cardiovascular disease among young women. *J Am Coll Cardiol* 2015;65:43–51.
36 Strate LL, Liu YL, Aldoori WH, *et al*. Obesity increases the risks of diverticulitis and diverticular bleeding. *Gastroenterology* 2009;136:115–22.
37 Knol MJ, van der Tweel I, Grobbee DE, *et al*. Estimating interaction on an additive scale between continuous determinants in a logistic regression model. *Int J Epidemiol* 2007;36:1111–8.
38 Hjern F, Wolk A, Håkansson N. Obesity, physical inactivity, and colonic diverticular disease requiring hospitalization in women: a prospective cohort study. *Am J Gastroenterol* 2012;107:296–302.
39 Strate LL, Liu YL, Aldoori WH, *et al*. Physical activity decreases diverticular complications. *Am J Gastroenterol* 2009;104:1221–30.
40 Crowe FL, Appleby PN, Allen NE, *et al*. Diet and risk of diverticular disease in Oxford cohort of European Prospective Investigation into Cancer and Nutrition (EPIC): prospective study of British vegetarians and non-vegetarians. *BMJ* 2011;343:d4131.
41 Crowe FL, Balkwill A, Cairns BJ, *et al*. Source of dietary fibre and diverticular disease incidence: a prospective study of UK women. *Gut* 2014;63:1450–6.
42 Ma W, Hua S, Giovannucci EL, *et al*. Association of Obesity With Diverticulitis in the Hispanic Community Health Study/Study of Latinos. *JAMA Surg* 2023;158:666–9.
43 Peery AF, Keku TO, Galanko JA, *et al*. Sex and Race Disparities in Diverticulosis Prevalence. *Clin Gastroenterol Hepatol* 2020;18:1980–6.
44 Wheat CL, Strate LL. Trends in Hospitalization for Diverticulitis and Diverticular Bleeding in the United States From 2000 to 2010. *Clin Gastroenterol Hepatol* 2016;14:96–103.
45 Zheng NS, Ma W, Shung DL, *et al*. Sex, Race, and Ethnicity Differences in Patients Presenting With Diverticular Disease at Emergency Departments in the United States: A National Cross-Sectional Study. *Gastro Hep Adv* 2024;3:178–80.