

Original Research

The association between alcohol consumption and colorectal carcinogenesis is partially mediated by the gut microbiome

Ane Sørлие Kværner^a, Einar Birkeland^b, Ekaterina Avershina^{d,e}, Edoardo Botteri^{a,f}, Cecilie Bucher-Johannessen^{c,d,f}, Markus Dines Knudsen^{g,h}, Anette Hjartåker^g, Christian M. Page^{i,j}, Johannes R. Hov^{k,l,m,n}, Mingyang Song^{h,o,p}, Kristin Ranheim Randel^a, Geir Hoff^{a,q}, Trine B. Rounge^{e,f} and Paula Berstad^a

Author Affiliations

^aSection of Colorectal Cancer Screening, Cancer Registry of Norway, Norwegian Institute of Public Health, Oslo, Norway

^bSection for Pharmacology and Pharmaceutical Biosciences, Department of Pharmacy, University of Oslo, Oslo, Norway

^cCentre for Bioinformatics, Department of Informatics, University of Oslo, Oslo, Norway

^dDepartment of Tumor Biology, Institute of Cancer Research, Oslo University Hospital, Oslo, Norway

^eDepartment of Pharmacy, University of Oslo, Oslo, Norway

^fDepartment of Research, Cancer Registry of Norway, Norwegian Institute of Public Health, Oslo, Norway

^gDepartment of Nutrition, Institute of Basic Medical Sciences, University of Oslo, Oslo, Norway

^hDepartment of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, Massachusetts, USA

ⁱCentre for Fertility and Health, Norwegian Institute of Public Health, Oslo, Norway

^jDepartment of Physical Health and Ageing, Division of Mental and Physical Health, Norwegian Institute of Public Health, Oslo, Norway

^kNorwegian PSC Research Center, Department of Transplantation Medicine, Division of Surgery and Specialized Medicine, Oslo University Hospital, Oslo, Norway

^lResearch Institute of Internal Medicine, Division of Surgery and Specialized Medicine, Inflammatory Diseases and Transplantation, Oslo University Hospital, Oslo, Norway

^mFaculty of Medicine, Institute of Clinical Medicine, University of Oslo, Oslo, Norway

ⁿSection of Gastroenterology, Department of Transplantation Medicine, Division of Surgery and Specialized Medicine, Oslo University Hospital, Oslo, Norway

^oDepartment of Nutrition, Harvard T.H. Chan School of Public Health, Boston, Massachusetts, USA

^pClinical and Translational Epidemiology Unit and Division of Gastroenterology, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts, USA

^qDepartment of Research, Telemark Hospital, Skien, Norway

Authors' last names

Kværner

Birkeland

Avershina

Botteri

Bucher-Johannessen

Knudsen

Hjartåker

Page

Hov

Song

Randel

Hoff

Rounge

Berstad

Corresponding Authors

Ane Sørli Kværner

Tel: +47 22 92 89 60

Mailing address: Cancer Registry of Norway, Norwegian Institute of Public Health, Postboks 5313 Majorstuen, 0304 Oslo, Norway

E-mail address: ane.sorlie.kvaerner@krefregisteret.no

ORCID identifier: 0000-0001-6247-7304

Paula Berstad

Tel: +47 932 932 35

Mailing address: Cancer Registry of Norway, Norwegian Institute of Public Health, Postboks 5313 Majorstuen, 0304 Oslo, Norway

E-mail address: paula.berstad@krefregisteret.no

ORCID identifier: 0000-0002-9025-4455

Sources of Support

This project would not have been possible without funding from the Norwegian Cancer Society (grant nos. 190179 and 198048), the Norwegian Cancer Society's umbrella organization for cancer research ("Kreftforeningens paraplystiftelse for kreftforskning"), the Research Council of Norway (grant no. 280667) and the South-Eastern Norway Regional Health Authority (grant nos. 2022067 and 2020056). The BCSN trial study was funded by the Norwegian Parliament (Norwegian national budget from 2011). The bowel preparation used for colonoscopy was provided free of charge by Ferring Pharmaceuticals. The funders of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

Short running head: Alcohol, gut bacteria and colorectal cancer

Abbreviations

ACME: Average causal mediation effect

ADE: Average direct effect

ADH: Alcohol dehydrogenase

AICR: American Institute for Cancer Research

ALDH: Acetaldehyde dehydrogenase

BCSN: Bowel Cancer Screening in Norway

BMI: Body mass index

CRC: Colorectal cancer

CUP: Continuous Update Project

CYP2E1: Cytochrome P450 2E1 pathway

E: Exposure

E%: Energy percentage

FFQ: Food frequency questionnaire

FIT: Fecal immunochemical test

KBS: Kostberegningssystem

MaAsLin: Microbiome multivariable associations with linear models

M: Mediator

O: Outcome

PCoA: Principal coordinate analysis

PERMANOVA: Permutational multivariate analysis of variance

Q: Quartile

TE: Total effect

TSS: Total sum scaling

WCRF: World Cancer Research Fund

Trial Registration: The BCSN is registered at clinicaltrials.gov (National clinical trial (NCT) no. 01538550).

1 **Abstract**

2 **Background:** Alcohol consumption is one of the major risk factors of colorectal cancer (CRC).
3 However, the mechanisms underlying this relationship are not fully understood, particularly the
4 potential role of gut microbes.

5 **Objective:** To study associations of alcohol intake with the gut microbiome and colorectal
6 lesions among CRC screening participants. Of particular interest was the potential role of gut
7 microbes in mediating the association between alcohol intake and colorectal lesions.

8 **Methods:** Participants included fecal immunochemical test-positive women and men enrolled in
9 the CRCbiome study, aged 55-77 years at inclusion. Intake of alcohol was assessed using a
10 validated, semi-quantitative food frequency questionnaire. Integrating with shotgun metagenome
11 based taxonomic and functional profiles, we studied associations with screen-detected colorectal
12 lesions. The potential role of alcohol-associated gut microbes in mediating the association
13 between alcohol intake and colorectal lesions was examined using causal mediation analysis.

14 **Results:** Of 1,468 participants with dietary data, 414 were diagnosed with advanced lesions
15 (advanced adenoma, advanced serrated lesions or CRC). Alcohol intake was positively
16 associated with advanced lesions in a dose-dependent manner ($p_{trend} = 0.008$), with odds ratio of
17 1.09 (95% confidence interval, 1.00, 1.19) per 10 g/day increase. Compared to non-consumers,
18 those consuming alcohol were characterized by a distinct microbial profile, manifested as
19 modest, but consistent, shifts in α - and β -diversity, and differentially abundant bacteria (Log₂
20 fold change (Log₂FC) >0: *B. finegoldii* and *L. asaccharolyticus*; Log₂FC <0: *S. mutans*, *B.*
21 *dentinum*, *C. symbiosum* and *E. boltae*). A causal mediation analysis showed that 12% of the

22 association between alcohol intake and advanced lesions was mediated by alcohol-associated gut
23 bacteria.

24 **Conclusions:** Alcohol consumption was associated with a distinct microbial profile, which partly
25 explained the association between alcohol intake and advanced colorectal lesions.

26 **Keywords:** Colorectal cancer, advanced colorectal lesions, colorectal carcinogenesis, bowel
27 cancer screening, FIT, alcohol, food frequency questionnaire, gut microbiome, metagenome
28 sequencing, α -diversity, β -diversity, differential abundance, mediation

29 **Introduction**

30 Alcohol increases the risk of cancer at multiple sites, in particular those of the gastrointestinal
31 system¹⁻³. According to recent global estimates, approximately 1 in 20 cancers in 2020 could be
32 attributed to alcohol consumption⁴. Considering the rise in adult per capita consumption seen
33 worldwide, especially in developing countries, these numbers are likely to increase in the years
34 to come⁵. The changing patterns of alcohol consumption among women represent a particular
35 cause of concern⁵.

36 Alcohol (ethanol) is efficiently absorbed by diffusion in the upper gastrointestinal tract, mainly
37 in the stomach and small intestine, before entering the liver via the portal vein⁶. The predominant
38 pathway for alcohol metabolism involves the enzymes alcohol dehydrogenase (ADH) and
39 acetaldehyde dehydrogenase (ALDH), converting ethanol to acetaldehyde and acetate,
40 respectively⁶. Alcohol can also be metabolized to acetaldehyde through the cytochrome P450
41 2E1 (CYP2E1) pathway⁷. Although the liver is the primary site for alcohol metabolism, some
42 alcohol is also catabolized in the gastrointestinal tract, either by mucosal cells lining the gut or
43 bacteria expressing enzymes involved in alcohol metabolism⁷.

44 Alcohol may promote cancer either directly or indirectly through its metabolites (acetaldehyde
45 and acetate) and/or enzymes involved in alcohol metabolism⁷. Their potential carcinogenic
46 effects are multiple, encompassing genomic, biochemical, inflammatory and immune-
47 modulatory mechanisms, among others⁷. More recently, the gut microbiome has emerged as a
48 plausible pathway through which alcohol may promote cancer⁸. However, it remains unclear
49 whether and how alcohol contributes to carcinogenesis through this microbial pathway.

50 To gain deeper insight into the potential role of gut bacteria in alcohol-associated carcinogenesis,
51 colorectal cancer (CRC) represents a particularly relevant malignancy. Not only are incidence
52 rates highly connected to alcohol consumption (with approximately one in ten new cancer cases
53 attributable to alcohol consumption⁴), but substantial data also support gut microbes as key
54 players in the development of the disease⁹.

55 In the current study, we combined data on alcohol consumption with metagenome based
56 taxonomic and functional profiles from participants in a large bowel cancer screening trial in
57 Norway to shed light on this interplay. Specifically, our objectives included: I) investigating
58 associations between alcohol consumption and the occurrence of screening-detected colorectal
59 lesions, II) identifying microbial features linked to alcohol consumption, and III) evaluating
60 whether potential associations between alcohol consumption and colorectal lesions are mediated
61 by the microbiome.

62 **Subjects and Methods**

63 **Bowel Cancer Screening in Norway (BCSN) and the CRCbiome study**

64 The CRCbiome study is nested within the Bowel Cancer Screening in Norway (BCSN)
65 study^{10,11}, a pilot for the Norwegian national screening program. BCSN is a randomized trial
66 comparing once-only sigmoidoscopy with four rounds of biennial fecal immunochemical testing
67 (FIT). The BCSN trial was initiated in 2012 and has invited 139,291 women and men to
68 participate. Of these, 70,096 were included in the FIT arm. FIT-positive participants (≥ 15 μg
69 hemoglobin/g feces) were referred for work-up colonoscopy.

70 The CRCbiome study was initiated in 2017 and has during its four-year recruitment period
71 invited 2,700 participants (starting from the second FIT round). The long-term goal of
72 CRCbiome is to develop a microbiome-based biomarker to improve the current FIT based
73 testing¹¹. FIT-positive participants were invited to CRCbiome in the interval between being
74 informed about their FIT screening result and attending colonoscopy. Besides the invitation
75 letter, participants received two questionnaires to be completed prior to the colonoscopy
76 examination: a food frequency questionnaire (FFQ) and a lifestyle and demographics
77 questionnaire. Returning at least one of these questionnaires was regarded as consent to the
78 study, being fulfilled by 1,640 (61%) participants. The age range at enrollment was 55-77 years.

79 Both the BCSN and the CRCbiome study have been approved by the Regional Committee for
80 Medical Research Ethics in South-East Norway (Approval no.: 2011/1272 and 63148,
81 respectively). The BCSN is also registered at clinicaltrials.gov (Clinical Trial (NCT) no.:
82 01538550).

83 **Study sample**

84 The current study included participants from the CRCbiome study with available dietary
85 information (n=1,616). After excluding participants who had withdrawn from the study after
86 inclusion (n=15), not attended colonoscopy (n=39), had a poor quality FFQ (n=21) or reported
87 too low (<600 and <800 kcal/day for women and men, respectively, n=9) or too high (>3,500
88 and >4,200 kcal/day for women and men, respectively, n=46) energy intake¹², a final number of
89 1,486 participants were eligible for the study, including 947 individuals with available gut
90 metagenome data (see flowchart, Figure 1).

91 **Assessment of alcohol intake**

92 Information on dietary intake, including alcohol, was obtained using a self-administered
93 semiquantitative, 14-page FFQ, designed to capture the habitual diet during the preceding year.
94 The questionnaire is a modified version of an FFQ developed by the Department of Nutrition,
95 University of Oslo¹³⁻¹⁹, which has been validated for a variety of nutrients^{13,15,18,19} and food
96 groups¹⁵⁻¹⁹, including alcohol intake¹³⁻¹⁵. The questionnaire covers a total of 256 food and
97 beverage items, of which eight concern alcoholic beverages (with one additional item covering
98 non-alcoholic drinks). For each beverage type, participants were asked to record frequency of
99 consumption, ranging from never/seldom to several times a week, and amount, typically as
100 standardized alcoholic units. Daily alcohol intake was calculated using the dietary calculation
101 system KBS (short for “**K**ost**b**eregningssystem”), developed at the Department of Nutrition,
102 University of Oslo. The most recent database at the time, AE-18, was used. AE-18 is an extended
103 version of the official Norwegian Food Composition Table, version 2018²⁰. Alcohol intake was
104 quantified both as ethanol (g/day and energy percentage (E%)) and by alcoholic beverage type
105 (g/day). The alcoholic beverages were categorized as follows (with the standardized unit used in
106 calculations given in brackets): ‘Wine’, including both red and white wine (15 cl); ‘beer’,

107 comprising regular and light beer (33 cl); ‘spirits’, encompassing spirits, mulled wines like port
108 and cherry, and liqueurs (4 cl); and ‘drinks’, which included cocktails, alcoholic cider, and
109 alcohol-containing soft drinks (20 cl). Additionally, the ‘non-alcoholic drinks’ category,
110 consisting of non-alcoholic beers (33 cl), was evaluated for comparison purposes. One alcoholic
111 unit was defined as 12 g²¹.

112 Prior to analyses, all questionnaires were reviewed and evaluated by trained personnel according
113 to a standardized framework for quality control assessment developed by the study group¹¹.

114 **Outcome assessment**

115 Outcome data were obtained from the BCSN database, containing detailed clinicopathological
116 information on all colorectal lesions detected at work-up colonoscopy. The information was
117 recorded by the responsible endoscopist using a structured reporting system. Based on the most
118 severe finding at colonoscopy, participants were categorized into the following three diagnostic
119 groups: advanced lesions, comprising CRC (any adenocarcinoma of the colon or rectum),
120 advanced adenomas (any adenoma with villous histology, high-grade dysplasia or diameter ≥ 10
121 mm) and advanced serrated lesions (any serrated lesion with size ≥ 10 mm or dysplasia); non-
122 advanced adenomas; and controls, i.e., no CRC, adenoma nor advanced serrated lesions detected.

123 **Sample collection, library generation and shotgun metagenome sequencing**

124 Protocols for sample collection, library generation and shotgun metagenome sequencing have
125 been described in detail elsewhere¹¹. In brief, DNA was extracted from 500 μ l aliquots, derived
126 from left-over buffer of FIT samples, using the QIA Symphony DSP Virus/Pathogen Midikit
127 (Qiagen, Hilden, Germany), with an offboard lysis protocol based on bead-beating. Purified
128 DNA was further eluted in 60 μ l AVE buffer (Qiagen, Hilden, Germany). Following extraction,
129 DNA concentration was measured on Qubit (Thermo Fisher Scientific, MA, USA). For samples

130 with a DNA concentration <1.5 ng/ μ l, DNA was extracted from a second aliquot. FIT samples
131 with a concentration of 0.7 ng/ μ l or more were considered eligible for shotgun metagenomic
132 library preparation, whereas for samples with DNA extracted from multiple aliquots, the one
133 with the highest DNA concentration was used.

134 Sequencing libraries were generated according to the Nextera DNA Flex Library Prep Reference
135 Guide, except scaling down the reaction volumes to one fourth of the reference. Library pools of
136 240 samples were combined and size selected to a fragment size of 650 – 900 bp. Sequencing
137 was performed on the Illumina NovaSeq system (Illumina Inc., CA, USA) using S4 flow cells
138 with lane divider, with each pool sequenced on a single lane resulting in paired end 2 x 151 bp
139 reads.

140 **Determination of taxonomic and functional profiles**

141 Sequencing reads were processed for removal of adapters and low-quality bases using
142 trimmomatic (v0.36)²² with the following trimming options: leading 20, trailing 20, minlength
143 50. Reads mapping to the human genome (hg38) and PhiX were removed using Bowtie2
144 (v2.3.5.1)²³. Read-based taxonomy and gene content was assessed using MetaPhlAn3 (v3.0.7)
145 and HumanN3 (v3.0.0)²⁴, respectively, with the mpa v30 ChocoPhlAn 201901 pangenome
146 database, using the UniRef90 database to assign gene families to Metacyc pathways. Read-based
147 taxonomic abundance was evaluated at the species level. Pathway abundance was scaled by the
148 number of quality-controlled reads per million.

149 **Assessment of covariates**

150 Information on covariates was obtained using a self-administered, four-page questionnaire on
151 lifestyle and demographic data, which has been described in detail previously¹¹. The questions
152 relevant to the current study concerned demographic factors (national affiliation, education,

153 occupation and marital status), clinical factors (family history of CRC and diagnosis of chronic
154 bowel disorders) and lifestyle factors (smoking and snus habits and physical activity level).
155 Smokers and snusers were defined as self-reported regular or occasional users or those having
156 quit consumption within the last ten years. Total amount of moderate to vigorous physical
157 activity (minutes/week) was calculated by summing the time spent in moderate and vigorous
158 activity, the latter weighted by a factor of two to best match recent guidelines²⁵⁻²⁷. Body mass
159 index (BMI) was calculated based on self-reported weight (kg) and height (cm) obtained from
160 the FFQ.

161 **Statistics**

162 Descriptive statistics are presented as median and interquartile range (Q1, Q3) for continuous
163 variables and numbers and percentages for categorical variables.

164 Pairwise correlations between continuous measures were computed using Spearman's correlation
165 coefficients (r).

166 To study the association between alcohol intake and colorectal lesions, multinomial logistic
167 regression analysis was conducted. Colonoscopy findings were categorized into advanced
168 lesions, non-advanced adenomas and controls, according to the outcome definition given above.
169 Alcohol (as ethanol in g/day) was categorized by consumption level (0 g/day, >0-10 g/day, \geq 10-
170 20 g/day and \geq 20 g/day) and by adherence to national guidelines (full adherence: 0 g/day, partial
171 adherence: <10 and 20 g/day for women and men, respectively, and non-adherence: \geq 10 and 20
172 g/day for women and men, respectively^{21,25,28}). Linear (per 10 g/increase/day) and exponential
173 (per 2-fold increase) consumption was evaluated, the latter based on a log₂-transformation of the
174 continuous alcohol variable in g/day plus 1 g/day, to enable inclusion of non-consumers in the
175 analysis. For the alcohol subtypes, participants were categorized as consumers/non-consumers.

176 The selection of covariates was based on *a priori* knowledge on the relationship between alcohol
177 intake and colorectal lesions²⁹⁻³¹, with all multinomial logistic regression analyses being adjusted
178 for the following covariates: age (continuous), sex (women, men), national affiliation
179 (Norwegian, non-Norwegian, missing), screening center (center 1, center 2), education level
180 (primary school, high school, college/university, missing), family history of CRC (yes, no,
181 unknown/missing), smoking status (non-smoker, smoker, missing), BMI (continuous, with
182 missing set to median) and level of physical activity (continuous, with missing set to median).
183 Additional adjustment for potential dietary risk or protective factors, such as red and processed
184 meat, whole grain, and dairy products – as well as alternative ways of categorizing covariates,
185 including a more refined smoking variable, were also evaluated, but not included in the final
186 model as they did not influence the interpretation of results.

187 To study potential differential influence of alcohol intake on colorectal lesions by sex, separate
188 analyses were conducted in women and men. Potential interactions were examined using the
189 Wald test. Subgroup analyses were also conducted by precursor lesion subtype (advanced
190 adenoma or advanced serrated lesion) and location (advanced proximal or advanced distal
191 lesion).

192 As sensitivity analyses, the main association analyses were run with alcohol intake calculated as
193 E% rather than g/day, use of a multiple imputation approach for handling of missing data, or the
194 exclusion of participants without metagenome data (n=539). The potential influence of leaving
195 out participants with self-reported bowel disorders (n=216) was also evaluated.

196 To identify microbial features linked to higher alcohol intake, associations of alcohol with the
197 following three measures were examined: α -diversity by means of the Shannon and inverse
198 Simpson indices, β -diversity based on the Bray-Curtis dissimilarity metric, and bacterial species

199 and pathway abundance. Associations with α -diversity were examined using linear regression
200 models with the diversity indices as the dependent variable. To improve normality and ease
201 interpretation, diversity indices were log-transformed prior to analysis. β -diversity was evaluated
202 by principal coordinate analysis (PCoA) and permutational multivariate analysis of variance
203 (PERMANOVA) with 999 permutations. Effect sizes were determined by calculating the partial
204 omega-squared (Ω^2) values. Differential abundance analyses were performed using microbiome
205 multivariable associations with linear models (MaAsLin) ²³² with the following settings: Min.
206 abundance: 0.0; min. prevalence: 0.1; normalization: none and total sum scaling (TSS) for
207 bacteria and pathways, respectively; transformation: log2-transformation with a pseudo-count of
208 half the minimum value; and analyzing method set to a linear model. Benjamini-Hochberg
209 corrected p-values were used as basis for interpretation of results. All analyses on microbial data
210 were conducted in the study group as a whole and stratified by sex. The same set of covariates
211 were adjusted for as in the association analyses with lesions as outcome variable, but with the
212 addition of sequencing depth as a continuous variable. Other combinations of covariates were
213 also evaluated (i.e. the addition of dietary risk or protective factors (as elaborated on above),
214 presence of bowel disorders, as an indicator of bowel movement pattern, and antibiotic usage),
215 but not included in subsequent analyses as they only marginally altered the results.

216 To evaluate whether the gut microbiome mediated the association between alcohol intake and
217 advanced lesions, a causal mediation analysis was applied. An alcohol-associated microbial
218 score was developed as an indicator of the effect of alcohol consumption on the gut microbiome.
219 This score was constructed based on the output of adjusted differential abundance analyses,
220 following the formula proposed by Gevers, *et al.*³³:

221
$$\text{score} = \log \left(\frac{\sum_i s_i^{\text{pos}}}{\sum_i s_i^{\text{neg}}} \right) \quad (1)$$

222 where s_i^{pos} and s_i^{neg} were the abundance of species positively and negatively associated with any
223 level of alcohol consumption, respectively. Zeros in either the numerator or denominator were
224 avoided by substituting zeros with a pseudo-count of half the minimum value of the numerator or
225 denominator across the dataset, respectively. To minimize bias in the estimated effects, a five-
226 fold cross-validation approach was employed for generating the microbial score. Specifically, the
227 dataset was divided into five non-overlapping subsets, where for each subset, a microbial score
228 was calculated using the significantly associated microbial species identified from a fully
229 adjusted differential abundance analysis conducted on the remaining four subsets. The
230 significance threshold for this analysis was set to a Benjamini-Hochberg-adjusted p-value of 0.1.
231 In the causal mediation analysis, 2-fold increases in alcohol intake were treated as the
232 independent variable, advanced lesions as the dependent variable and the alcohol-associated
233 microbial score as the potential mediator. Non-advanced adenomas were grouped with
234 colonoscopy-negatives for these analyses. The mediation analysis was adjusted for the same set
235 of covariates as in the multinomial logistic regression models, as well as the cross-validation
236 partition identity. We used the R package *mediation*³⁴ and the function ‘*mediate*’ based on the
237 following two models: 1) a multivariate linear regression model examining the association
238 between the exposure (E) and the mediator (M), and 2) a multivariate generalized linear model
239 examining the association between the mediator (M) and the outcome (O), controlling for the
240 exposure (E). We report the point estimates along with non-parametric bootstrap confidence
241 intervals using the percentile method. Uncertainty estimates were calculated using 1000
242 simulations. We also employed the *medflex* R package³⁵ as an alternative approach to mediation
243 analysis. In this analysis we used a natural effects model with imputation of nested
244 counterfactuals, ensuring that mediation effects were not dependent on covariate levels.

245 Statistical analyses were performed using R, version 4.1.0 (The R Foundation for Statistical
246 Computing, Vienna, Austria). In addition to packages included in tidyverse (version 1.3.1), main
247 packages, with version number in parentheses, included skimr (2.1.5), corrplot (0.88), nnet (7.3-
248 16), VGAM (1.1-5), mice (3.13.0), vegan (2.5-7), micEco (0.9.15), Maaslin2 (1.12.0), mediation
249 (4.5.0) and medflex (0.6-7).

250

251 **Results**

252 **Study population**

253 **Table 1** shows the key characteristics of the study population overall and by alcohol intake. The
254 median (Q1, Q3) age of participants was 67 (62, 72) years, with a slightly higher representation
255 of men (56%). Compared to non-consumers (13% of participants), those consuming alcohol were
256 more likely to be male and be affiliated to the screening center localized in Bærum (center 2).
257 Higher alcohol intakes were in general positively associated with markers of higher
258 socioeconomic status such as being married or cohabiting, being employed, and holding a
259 university or college degree. The high consumers were also more likely to use snus tobacco and
260 reported having a higher level of physical activity than the non-consumers.

261 **Daily intake of alcohol**

262 Daily intake of alcohol and different types of alcoholic beverages are presented in **Table 2**. The
263 median (Q1, Q3) intake of alcohol (as ethanol in g/day) was 9 (2-19) g/day; 13 (4-25) g/day in
264 men and 5 (1-13) g/day in women. Despite higher intakes in men, the proportion adhering to
265 national guidelines was relatively similar between sexes (68% in men, 65% in women). In terms
266 of beverage types, wine was the most frequently consumed by women (74%), whereas beer was
267 the most frequently consumed by men (80%). There were fewer consumers of spirits (30%) and
268 drinks (17%), particularly among women (17 and 13%, respectively). Alcohol (overall and by
269 subtype) was only modestly correlated with energy intake (with Spearman's r 's ranging from
270 0.06-0.24 in the study group as a whole).

271 **Alcohol intake and colorectal lesions**

272 Associations between alcohol intake (as ethanol in g/day) and colorectal lesions are shown in
273 **Table 3**. Compared to non-consumption, all levels of alcohol intake, and in particular high
274 levels, were positively associated with advanced lesions ($p_{trend} = 0.008$). The probability of
275 advanced lesions increased by 9% per 10 g increase/day and 14% per 2-fold increase/day. Partial
276 and non-adherence to guidelines were also positively associated with advanced lesions relative to
277 not consuming alcohol, with odds ratios (ORs) of 1.91 (95% confidence intervals (CI) 1.22,
278 2.99) and 1.97 (1.23, 3.17), respectively. Notable differences were observed between women and
279 men, with associations being consistently stronger for women. As an example, relative to those
280 not consuming alcohol, women in the highest consumption category (≥ 20 g/day) had an OR for
281 advanced lesions of 4.81 (95% CI 2.18, 10.62) compared to 1.48 (0.74, 2.96) in men ($p_{interaction} =$
282 0.040). No associations between alcohol intake and non-advanced adenomas were detected.
283 Subgroup analyses by lesion subtype and location implied associations for both groups of
284 precursor lesions, regardless of lesion location (**Supplementary Table 1**).
285 Sensitivity analyses with alcohol intake as E%, use of a multiple imputation approach for
286 handling of missing values and restricting the study population to those with metagenome data
287 only (**Supplementary Tables 2-4**) produced similar results as in the main analyses. This was
288 also the case for an analysis excluding participants with a self-reported bowel disorder (data not
289 shown).
290 In terms of beverage types, positive associations with advanced lesions were seen for consumers
291 of wine (OR 1.41; 95%CI 1.03, 1.92) and beer (OR 1.31; 0.97, 1.77; **Figure 2**). Stratifying the
292 analyses by sex, the association for wine was stronger (and only statistically significant) in
293 women, whereas the association for beer tended to be stronger for men (although not statistically
294 significant). No associations with non-advanced adenomas were observed.

295 **Alcohol intake and gut microbial features**

296 Metagenome shotgun sequencing data derived from leftover buffer containing fecal matter
297 collected in screening FIT cartridges was available for 947 participants (mean = 12 million
298 paired end reads, sd = 3.8 million). Taxonomic classification resulted in identification of a total
299 of 787 microbial species (mean = 88, sd = 15.5 per sample).

300 α -diversity

301 Associations of alcohol intake with the α -diversity indices Shannon and Inverse Simpson are
302 presented in **Figure 3a-c** and **Supplementary Table 5**. In general, there was a weak, but
303 statistically significant positive association between alcohol intake (as ethanol in g/day) and both
304 diversity indices. The shift in diversity became noticeable already at low intake levels. Compared
305 to the non-consumers, those consuming alcohol had a 2.7 and 10.2% higher Shannon and Inverse
306 Simpson index, respectively. No dose-response associations were observed. Stratification by sex
307 showed the associations between total alcohol intake and α -diversity to be statistically significant
308 in women only. Still, looking at alcoholic beverage subtypes, statistically significant positive
309 associations were observed for consumers of beer (overall and in men), wine (women), spirits
310 (overall and in women) and non-alcoholic drinks (women). As for ethanol in g/day, there were
311 no clear dose-response relationships.

312 β -diversity

313 Associations of alcohol intake with the Bray-Curtis β -diversity index are presented in Figure 3d-
314 g and Supplementary Table 5. Irrespective of the approach to quantifying alcohol intake (as
315 ethanol in g/day), the microbial composition differed by intake level (Ω values of 0.0005-0.0010,
316 PERMANOVA-derived p-values <0.05, Supplementary Table 5). Consumers of alcohol

317 displayed a small shift in microbial composition (Figure 3d), and were less heterogeneous
318 (Figure 3e) than the non-consumers. There was, however, no difference in microbial composition
319 by intake level (data not shown). Consuming alcohol was associated with microbial composition
320 regardless of sex.

321 Consumption of wine, and to a lesser extent, beer, seemed to be related to microbial composition
322 (Figure 3f).

323 Differentially abundant bacteria and pathways

324 Associations between alcohol intake and bacteria and pathway abundance are presented in
325 **Figures 4a** and **5a**. In total, 6 bacteria (2 positively and 4 negatively) and 7 pathways (5
326 positively and 2 negatively) were statistically significantly associated with at least one of the
327 alcohol consumption categories (Figures 4a and 5a). Of those predictive of the highest
328 consumption category, 4 out of 5 bacteria (i.e. *L. asaccharolyticus*, *B. fingoldii*, *S. mutans* and
329 *C. symbiosum*) and 3 out of 4 pathways (i.e. the tricarboxylic acid (TCA) cycle II, superpathway
330 of sulfur oxidation and L-lysine biosynthesis II), remained statistically significant after mutual
331 adjustment for the other bacteria and pathways, respectively (**Supplementary Figure 2a**).
332 Several of the identified bacteria and pathways were inter-correlated (Supplementary Figure 2b).
333 None of the identified bacteria or pathways showed signs of interaction with sex (separate
334 analyses for women and men can be found in **Supplementary Figure 3**).

335 With regard to consumption of different alcoholic beverages (Figure 4b and 5b), beer was
336 statistically significantly associated with 4 pathways (the superpathway of sulfur oxidation, L-
337 glutamate degradation via hydroxyglutarate, ppGpp biosynthesis and L-lysine biosynthesis),
338 wine with 3 bacteria (*L. asaccharolyticus*, *S. mutans* and *B. dentium*) and 2 pathways (L-lysine

339 biosynthesis and succinate fermentation to butanoate), and spirits with 1 bacterium (*F. bacterium*
340 *GAG 95*) and 1 pathway (L-glutamate and L-glutamine biosynthesis). For the other beverage
341 types, no statistical differences were detected.

342 **The gut microbiome as a potential mediator**

343 Based on output from differential abundance analysis, an alcohol-associated microbial score was
344 developed to examine the potential mediating role of alcohol-related bacteria in colorectal
345 carcinogenesis (see **Supplementary Tables 6-7** for bacterial species significantly associated
346 with alcohol intake overall and using the previously described five-fold cross-validation
347 approach). The causal mediation analysis showed that the alcohol-associated microbial score
348 partially mediated the association between alcohol intake and advanced colorectal lesions
349 (**Figure 6a**). The total effect (95% CI) of alcohol on advanced lesions was 0.033 (0.017, 0.047),
350 the average direct effect (ADE) was 0.029 (0.012, 0.043) and the average causal mediation effect
351 (ACME) was 0.004 (0.001, 0.008). Overall, the proportion mediated was 12.1% (3.1, 32.2;
352 **Figure 6b**), confirmed using an approach based on nested counterfactuals (here the proportion
353 mediated was 12.2%, **Supplementary Table 8**). To assess potential differences in mediation
354 effects between women and men while maintaining power, we conducted a mediation analysis
355 omitting sex as a covariate. This did, however, only marginally affect the results.

356

357 **Discussion**

358 This study provides new evidence on the detrimental role of alcohol in the development of
359 colorectal lesions. We found a positive association between alcohol consumption and advanced
360 colorectal lesions detected at screening even at modest intake levels. This association was strong
361 in the overall population and in women, while no clear association was observed in men.
362 Consumers of alcohol had a distinct gut microbial profile, characterized by higher species
363 diversity, altered microbial composition and differentially abundant bacteria and pathways. This
364 distinctive microbial profile partly explained the association between alcohol intake and
365 advanced lesions observed. Our results support a role of the microbiome in alcohol-induced
366 colorectal carcinogenesis.

367 In the present study, every 10 g per day increase in alcohol consumption was associated with a
368 9% increased probability of advanced lesions being detected at colonoscopy, consistent across
369 lesion subtype and location. Our results are in line with the literature on precancerous colorectal
370 lesions. Meta-analyses have demonstrated a 27% increased risk of adenoma per 25 g alcohol
371 consumed per day³⁶, along with comparable increases in the risk of serrated polyps^{37,38}.
372 Together, these results coincide with the literature on alcohol and CRC^{29,39,40}.

373 In our study, consumption levels even below 10 g/day were associated with advanced lesions. As
374 such, our findings reinforce cancer prevention guidelines of complete abstinence to achieve the
375 lowest possible risk⁴¹.

376 The stronger association observed in women compared to men contrasts with prior literature. In
377 the CUP 2016 meta-analysis from WCRF/AICR²⁹, as well as two pooled analyses (from UK and
378 Japan, respectively)^{42,43}, no particular heterogeneity was found between sexes. Women from

379 Europe and Australia have the highest per-capita alcohol consumption worldwide, surpassing the
380 global average by a factor of about two⁵. It is noteworthy that these regions also bear the highest
381 incidence rates of CRC among women, Norway ranking at the top in 2020⁴⁴. The consumption
382 pattern of participants included in the present study seems to mirror those of the general
383 Norwegian population^{45,46}. Whether country-specific trends in alcohol consumption may account
384 for divergence in findings remains a matter of speculation.

385 Our study demonstrated notable differences in the gut microbiome of participants consuming
386 relative to not consuming alcohol, potentially of relevance to tumorigenesis. A link between
387 alcohol intake and gut microbiome perturbations has been documented by others⁴⁷⁻⁵¹. However,
388 this has typically been studied in the context of chronic alcoholism^{47,48,50} and/or presence of
389 severe liver pathologies^{48,49} that could confound the alcohol-gut microbiome relationship. We
390 observed that even modest alcohol consumption was associated with an increase in α - and a
391 convergence of β -diversity. The diversity change has also been reported in studies of other
392 Western populations⁵²⁻⁵⁵, revealing alcohol as a strong source of gut microbiome variation.

393 We identified four bacterial species to be independently associated with alcohol consumption: *L.*
394 *asaccharolyticus* and *B. finegoldii* were positively associated, whereas *C. symbiosum* and *S.*
395 *mutans* were negatively associated. These results reproduce associations reported in the
396 PREDICT study, which assessed microbial relationships with habitual diet in a large population,
397 also finding *L. asaccharolyticus* and *C. symbiosum* to be strong determinants of alcohol
398 consumption⁵⁵. Among the identified bacteria, *C. symbiosum* has earlier been suggested as a
399 potential biomarker for CRC^{56,57}.

400 Interestingly, we found the association between alcohol consumption and detection of advanced
401 lesions to be partially mediated by the gut microbiome. The mediated proportion was a modest,

402 but conservative, 12% of the total effect, but nonetheless suggests that microbial changes caused
403 by alcohol consumption could contribute to CRC development. How the relatively small
404 concentrations of ethanol reaching the large intestine can induce microbial changes of relevance
405 to carcinogenic development is however unclear. A conventional belief has been that the
406 intestinal bacteria play an important role in metabolizing the remaining amounts of ethanol,
407 through bacterial catalases and ADH activity, leading to accumulation of acetaldehyde and thus
408 causing local damage^{7,8,58}. However, this was recently questioned by Martino, *et al.*⁶ who used a
409 mouse model to demonstrate that rather than metabolizing ethanol directly, gut bacteria
410 responded to ethanol-feeding by activating acetate dissimilation. In line with these findings, we
411 observed that TCA cycle gene abundance was positively associated with alcohol consumption.
412 Acetate has recently received renewed attention in the context of CRC⁷. Although historically
413 being regarded as protective, recent evidence suggests that acetate may contribute to cancer cell
414 growth by serving as a substrate for the synthesis of acetyl-CoA⁷. We also found alcohol
415 consumption to be associated with increased abundance of sulfur oxidation superpathway genes,
416 regardless of consumption levels. Sulfur metabolism is characteristic of a wide range of
417 bacteria⁵⁹ and has been associated with CRC⁵⁹ and other gut disorders⁶⁰. In combination, these
418 results suggest a potential for alcohol consumption to increase risk of CRC via bacterial acetate
419 and sulfur metabolism.

420 A major strength of our study includes its large set of microbiome samples obtained through
421 state-of-the-art methodology, coupled with validated exposure information. Access to clinically
422 verified outcome data facilitated thorough investigations of screening-relevant outcomes, with
423 minimal risk of misclassification. With a study population solely consisting of FIT positive
424 participants, the proportion of detected lesions was high (63%). Nonetheless, this selective

425 inclusion of participants may have restricted the generalizability of our findings. Other
426 limitations must also be considered. First, the cross-sectional design limits causal interpretations,
427 and the results must be viewed as hypothesis-generating only. Thus, although access to
428 comprehensive data on lifestyle and demography allowed for detailed covariate adjustment
429 (being particularly unique for the microbiome analyses), we cannot exclude the possibility of
430 residual or reverse confounding. Second, selective inclusion of participants with colon bleeding
431 at specimen collection, could have introduced bias. The proportion of participants with
432 gastrointestinal morbidity may have been unevenly distributed between alcohol intake and
433 outcome categories. This may be of particular concern for the present study, as some over-
434 representation of former drinkers and “sick quitters” among the non-consumers, is to be
435 expected⁶¹. However, excluding participants with self-reported gastrointestinal morbidity did not
436 alter the observed associations.

437 To conclude, our study confirms the role of alcohol in the etiology of CRC. Consistent and
438 positive associations were observed between alcohol consumption and advanced lesions even at
439 moderate consumption levels, and particularly in women. Consuming alcohol was associated
440 with a distinct microbial profile in the gut, manifested as increased species diversity, altered
441 microbial composition and differentially abundant bacteria and pathways. Collectively, alcohol-
442 associated bacteria mediated 12% of the association between alcohol intake and advanced
443 colorectal lesions. The potential role of alcohol-associated microbial alterations in cancer
444 development should be further examined in prospective cohort studies with long-term follow-up.
445 Such studies should investigate potential sex differences and ideally expand the repertoire of
446 biological mechanisms by evaluating metabolic, inflammatory, and immune-modulatory
447 pathways.

448 **Acknowledgements**

449 We would like to thank the devoted healthcare personnel, technicians and administrative staff at
450 the two screening centers (Moss center and Bærum center) and university hospitals (Oslo
451 University Hospital and Akershus University Hospital) for their important contributions to the
452 BSCN and CRCbiome study. We offer special thanks to Jan Inge Nordby and Vahid Bemanian
453 for their instrumental role in sample preparation and laboratory work. Library preparation and
454 sequencing were carried out at the FIMM Technology Centre supported by HiLIFE and
455 Biocenter Finland. We would particularly express our gratitude to Harri A. Kangas and Pekka J.
456 Ellonen for their good cooperation, service and support. All processing of dietary data, including
457 food and nutrient calculations, was conducted at the Department of Nutrition, University of Oslo.
458 We would particularly like to thank Anne Marte Wetting Johansen for her dedicated work and
459 attention to detail. Lastly, this article is the outcome of collective teamwork and scientific
460 discussions among both former and current colleagues within our research group. Therefore, we
461 would like to acknowledge Elina Vinberg, Even Sannes Riiser, Erik Natvig, Paula Istvan and
462 Maja Jacobsen for their valuable contributions.

463

464 **Conflict of Interest (COI) Statement:**

465 Dr. Mingyang Song serves as a consultant for Etiome Inc. The remaining authors declare no
466 conflicts of interest.

467

468

469

470 **Authors' contributions**

471 ASK, EB, EB, CBJ, TBR and PB designed the research (project conception, development of
472 overall research plan, and study oversight).

473 ASK, EB, EA, CBJ, KRR, TBR and PB conducted the research (hands-on conduct of the
474 experiments and data collection).

475 ASK, EB, CBJ, AH, KRR, GH, TBR and PB provided essential materials (applies to authors
476 who contributed by providing animals, constructs, databases, etc, necessary for the research).

477 ASK, EB, CBJ and EB analyzed data or performed statistical analysis.

478 ASK wrote the paper (only authors who made a major contribution).

479 ASK and PB had the primary responsibility for the final content.

480 ASK, EB, EA, EB, CBJ, MDK, AH, CMP, JRH, MS, KRR, GH, TNR and PB have read and
481 approved the final manuscript.

482 **Data availability**

483 DNA sequencing data analyzed in this study are deposited in the database Federated EGA under
484 accession code EGAS50000000170 (<https://ega-archive.org/studies/EGAS50000000170>). Per
485 participant consent, submitted FASTQ files exclude reads mapping to the human genome. The
486 data are available under restricted access due to the sensitive nature of data derived from human
487 subjects. Processing of data from this study must comply with the General Data Protection
488 Regulation (GDPR). Access can be obtained by following the procedure described here:
489 <https://www.mn.uio.no/sbi/english/groups/rounegrp/crcbiome/>. Requests for data access can
490 also be directed to Trine B Rounge, trinro@uio.no.

491 **Declaration of generative AI and AI-assisted technologies in the writing process**

492 During the preparation of this work the first author(s) used Chat-GTP in order to improve
493 language and readability. After using this tool/service, the author(s) reviewed and edited the
494 content as needed and take(s) full responsibility for the content of the publication.

References

1. World Cancer Research Fund/American Institute for Cancer Research. Alcoholic drinks and the risk of cancer. *Continous Updat. Proj.* (2018).
2. LoConte, N. K., Brewster, A. M., Kaur, J. S., Merrill, J. K. & Alberg, A. J. Alcohol and cancer: A statement of the American society of clinical oncology. *J. Clin. Oncol.* **36**, 83–93 (2018).
3. Scherübl, H. Alcohol Use and Gastrointestinal Cancer Risk. *Visc. Med.* **36**, 175–181 (2020).
4. Rumgay, H. *et al.* Global burden of cancer in 2020 attributable to alcohol consumption: a population-based study. *Lancet Oncol.* **22**, 1071–1080 (2021).
5. Manthey, J. *et al.* Global alcohol exposure between 1990 and 2017 and forecasts until 2030: a modelling study. *Lancet* **393**, 2493–2502 (2019).
6. Martino, C. *et al.* Acetate reprograms gut microbiota during alcohol consumption. *Nat. Commun.* **13**, (2022).
7. Johnson, C. H. *et al.* Molecular mechanisms of alcohol-induced colorectal carcinogenesis. *Cancers* vol. 13 at <https://doi.org/10.3390/cancers13174404> (2021).
8. Song, M. & Chan, A. T. Environmental Factors, Gut Microbiota, and Colorectal Cancer Prevention. *Clinical Gastroenterology and Hepatology* vol. 17 275–289 at <https://doi.org/10.1016/j.cgh.2018.07.012> (2019).
9. Tilg, H., Adolph, T. E., Gerner, R. R. & Moschen, A. R. The Intestinal Microbiota in Colorectal Cancer. *Cancer Cell* **33**, 954–964 (2018).
10. Randel, K. R. *et al.* Colorectal cancer screening with repeated fecal immunochemical test versus sigmoidoscopy: baseline results from a randomized trial. *Gastroenterology* (2020) doi:10.1053/j.gastro.2020.11.037.
11. Kværner, A. S. *et al.* The CRCbiome study: A large prospective cohort study examining the role of lifestyle and the gut microbiome in colorectal cancer screening participants. *medRxiv* (2020) doi:10.1101/2020.12.22.20248658.
12. Willett, W. *Nutritional Epidemiology*. (Oxford University Press, Oxford; New York, 2013).
13. Andersen, L. F. *et al.* Evaluation of a food frequency questionnaire with weighed records, fatty acids, and alpha-tocopherol in adipose tissue and serum. *Am. J. Epidemiol.* **150**, 75–87 (1999).
14. Andersen, L. F., Tomten, H., Haggarty, P., Løvø, A. & Hustvedt, B. E. Validation of energy intake estimated from a food frequency questionnaire: A doubly labelled water study. *Eur. J. Clin. Nutr.* **57**, 279–284 (2003).
15. Carlsen, M. H. *et al.* Evaluation of energy and dietary intake estimates from a food frequency questionnaire using independent energy expenditure measurement and weighed food records. *Nutr. J.* **9**, 1–9 (2010).

16. Andersen, L. F. *et al.* Evaluation of three dietary assessment methods and serum biomarkers as measures of fruit and vegetable intake, using the method of triads. *Br. J. Nutr.* **93**, 519–527 (2005).
17. Carlsen, M. H. *et al.* Relative validity of fruit and vegetable intake estimated from an FFQ, using carotenoid and flavonoid biomarkers and the method of triads. *Br. J. Nutr.* **105**, 1530–1538 (2011).
18. Brunvoll, S. H. *et al.* Validation of repeated self-reported n-3 PUFA intake using serum phospholipid fatty acids as a biomarker in breast cancer patients during treatment. *Nutr. J.* **17**, 1–12 (2018).
19. Nes, M. *et al.* Accuracy of a quantitative food frequency questionnaire applied in elderly Norwegian women. *Eur. J. Clin. Nutr.* (1992).
20. Norwegian Food Safety Authority. Norwegian Food Composition Database 2019. www.matvaretabellen.no.
21. Nordic Council of Ministers. *Nordic Nutrition Recommendations 2012 - Integrating Nutrition and Physical Activity. 5th Ed.* (2012) doi:10.1080/1102680410003794.
22. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
23. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
24. Beghini, F. *et al.* Integrating taxonomic, functional, and strain-level profiling of diverse microbial communities with biobakery 3. *Elife* **10**, 1–42 (2021).
25. Helsedirektoratet. *Anbefalinger Om Kosthold , Ernæring Og Fysisk Aktivitet.* <https://helsedirektoratet.no/Lists/Publikasjoner/Attachments/806/Anbefalinger-om-kosthold-ernæring-og-fysisk-aktivitet-IS-2170.pdf> (2014).
26. World Health Organization (WHO). *Global Recommendations on Physical Activity for Health.* (2010).
27. U.S. Department of Health and Human Services. *Physical Activity Guidelines for Americans.* (2018).
28. Blomhoff, R., Andersen, R., Arnesen, E.K., Christensen, J.J., Eneroth, H., Erkkola, M., Gudaviciene, I., Halldorsson, T.I., Høyer-Lund, A., Lemming, E.W., Meltzer, H.M., Pitsi, T., Schwab, U., Siksna, I., Thorsdottir, I. and Trolle, E. *Nordic Nutrition Recommendations 2023. Integrating Environmental Aspects. 6th Ed.* (2023).
29. World Cancer Research Fund/American Institute for Cancer Research. *Diet, Nutrition, Physical Activity and Colorectal Cancer. Continuous Update Project* <https://www.wcrf.org/sites/default/files/Colorectal-cancer-report.pdf> (2018).
30. Dekker, E., Tanis, P. J., Vleugels, J. L. A., Kasi, P. M. & Wallace, M. B. Colorectal cancer. *Lancet (London, England)* **394**, 1467–1480 (2019).
31. Kuipers, E. J. *et al.* Colorectal cancer. *Nat. Rev. Dis. Prim.* **1**, (2015).

32. Mallick, H. *et al.* Multivariable association discovery in population-scale meta-omics studies. *PLoS Comput. Biol.* **17**, 1–27 (2021).
33. Gevers, D. *et al.* The treatment-naïve microbiome in new-onset Crohn’s disease. *Cell Host Microbe* **15**, 382–392 (2014).
34. Tingley, D., Yamamoto, T., Hirose, K., Keele, L. & Imai, K. Mediation: R package for causal mediation analysis. *J. Stat. Softw.* **59**, 1–38 (2014).
35. Steen, J. & Loeys, T. Medflex: An R Package for Flexible Mediation Analysis using Natural Effect Models. (2010).
36. Ben, Q. *et al.* Alcohol drinking and the risk of colorectal adenoma: A dose-response meta-analysis. *Eur. J. Cancer Prev.* **24**, 286–295 (2015).
37. Bailie, L., Loughrey, M. B. & Coleman, H. G. Lifestyle Risk Factors for Serrated Colorectal Polyps: A Systematic Review and Meta-analysis. *Gastroenterology* **152**, 92–104 (2017).
38. Wang, Y.-M. *et al.* Systematic Review with Meta-Analysis: Alcohol Consumption and Risk of Colorectal Serrated Polyp. *Dig. Dis. Sci.* **60**, 1889–1902 (2015).
39. Lyon, F. *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans VOLUME 96 Alcohol Consumption and Ethyl Carbamate.* (2010).
40. Weiderpass, E. & Stewart, B. W. *World Cancer Report.*
41. World Cancer Research Fund/American Institute for Cancer Research. *Diet, Nutrition, Physical Activity and Cancer: A Global Perspective. Continuous Update Project Expert Report 2018* (2018). doi:10.1016/j.scienta.2014.02.005.
42. Park, J. Y. *et al.* Alcohol intake and risk of colorectal cancer: Results from the UK Dietary Cohort Consortium. *Br. J. Cancer* **103**, 747–756 (2010).
43. Mizoue, T. *et al.* Alcohol drinking and colorectal cancer in Japanese: a pooled analysis of results from five cohort studies. *Am. J. Epidemiol.* **167**, 1397–1406 (2008).
44. Ferlay, J. *et al.* Global Cancer Observatory: Cancer Today. vol. 2024 <https://gco.iarc.who.int/today> (2024).
45. Folkehelseinstituttet. Alkohol i Norge. <https://www.fhi.no/le/alkohol/alkoholinorge/?term=> (2018).
46. Helsedirektoratet. *Norkost 3. En Landsomfattende Kostholdsundersøkelse Blant Menn Og Kvinner i Norge i Alderen 18-70 År, 2010-11.* www.helsedirektoratet.no (2012).
47. Tsuruya, A. *et al.* Ecophysiological consequences of alcoholism on human gut microbiota: Implications for ethanol-related pathogenesis of colon cancer. *Sci. Rep.* **6**, 1–12 (2016).
48. Dubinkina, V. B. *et al.* Links of gut microbiota composition with alcohol dependence syndrome and alcoholic liver disease. *Microbiome* **5**, (2017).
49. Chen, Y. *et al.* Characterization of fecal microbial communities in patients with liver cirrhosis. *Hepatology* **54**, 562–572 (2011).

50. Day, A. W. & Kumamoto, C. A. Gut Microbiome Dysbiosis in Alcoholism: Consequences for Health and Recovery. *Front. Cell. Infect. Microbiol.* **12**, 1–10 (2022).
51. Pohl, K., Moodley, P. & Dhanda, A. D. Alcohol's impact on the gut and liver. *Nutrients* **13**, (2021).
52. Zhernakova, A. *et al.* Population-based metagenomics analysis reveals markers for gut microbiome composition and diversity. *Science (80-.)*. **352**, 565–569 (2016).
53. Falony, G. *et al.* Population-level analysis of gut microbiome variation. *Science (80-.)*. **352**, 560–564 (2016).
54. Vujkovic-cvijin, I. *et al.* Host variables confound gut microbiota studies of human disease. *Nature* **587**, (2020).
55. Asnicar, F. *et al.* Microbiome connections with host metabolism and habitual diet from 1,098 deeply phenotyped individuals. *Nat. Med.* **27**, 321–332 (2021).
56. Thomas, A. M. *et al.* Metagenomic analysis of colorectal cancer datasets identifies cross-cohort microbial diagnostic signatures and a link with choline degradation. *Nat Med* **25**, 667–678 (2019).
57. Xie, Y. H. *et al.* Fecal *Clostridium symbiosum* for Noninvasive Detection of Early and Advanced Colorectal Cancer: Test and Validation Studies. *EBioMedicine* **25**, 32–40 (2017).
58. Rossi, M., Anwar, M. J., Usman, A., Keshavarzian, A. & Bishehsari, F. Colorectal cancer and alcohol consumption—populations to molecules. *Cancers (Basel)*. **10**, (2018).
59. Wolf, P. G. *et al.* Diversity and distribution of sulfur metabolic genes in the human gut microbiome and their association with colorectal cancer. *Microbiome* **10**, 1–16 (2022).
60. Khorsand, B. *et al.* Overrepresentation of Enterobacteriaceae and *Escherichia coli* is the major gut microbiome signature in Crohn's disease and ulcerative colitis; a comprehensive metagenomic analysis of IBDMDB datasets. *Front. Cell. Infect. Microbiol.* **12**, 1–11 (2022).
61. Zhao, J. *et al.* Association between Daily Alcohol Intake and Risk of All-Cause Mortality: A Systematic Review and Meta-analyses. *JAMA Netw. Open* **6**, E236185 (2023).
62. Nasjonalt råd for ernæring. *Kostråd for å Fremme Folkehelsen Og Forebygge Kroniske Sykdommer Metodologi Og Vitenskapelig Kunnskapsgrunnlag.* (2011).

Tables

Table 1. Key characteristics of the study population overall and by alcohol intake (n=1,486)¹.

Variables	n	Level of alcohol intake				
		Overall (n=1,486)	0 g/day (n=187)	>0-10 g/day (n=592)	≥10-20 g/day (n=361)	≥20 g/day (n=346)
Age, years	1,486	67 (62, 72)	67 (62, 73)	67 (62, 72)	66 (62, 72)	68 (62, 72)
Male sex, n (%)	1,486	826 (55.6)	70 (37.4)	281 (47.5)	212 (58.7)	263 (76.0)
Screening center, n (%)	1,486					
Center 1 (Moss)		774 (52.1)	135 (72.2)	323 (54.6)	166 (46.0)	150 (43.4)
Center 2 (Bærum)		712 (47.9)	52 (27.8)	269 (45.4)	195 (54.0)	196 (56.6)
National affiliation, n (%)	1,427					
Norwegian		1345 (94.3)	160 (89.4)	540 (94.7)	333 (96.0)	312 (94.3)
Non-Norwegian		82 (5.7)	19 (10.6)	30 (5.3)	14 (4.0)	19 (5.7)
Family history of CRC, n (%)	1,347					
No		1,092 (81.1)	133 (83.1)	442 (81.1)	272 (83.2)	245 (77.8)
Yes		255 (18.9)	27 (16.9)	103 (18.9)	55 (16.8)	70 (22.2)
Marital status, n (%)	1,465					
Married/cohabiting		1,171 (79.9)	123 (67.6)	460 (79.0)	308 (85.8)	280 (81.9)
Not married/non-cohabiting		294 (20.1)	59 (32.4)	122 (21.0)	51 (14.2)	62 (18.1)
Education, n (%)	1,462					
Primary school		251 (17.2)	54 (29.8)	102 (17.5)	50 (14.0)	45 (13.2)
High school		580 (39.7)	75 (41.4)	243 (41.8)	144 (40.2)	118 (34.6)
University/college		631 (43.2)	52 (28.7)	237 (40.7)	164 (45.8)	178 (52.2)
Working status, n (%)	1,464					
Employed		498 (34.0)	40 (22.1)	190 (32.6)	140 (39.1)	128 (37.4)
Retired/unemployed		966 (66.0)	141 (77.9)	393 (67.4)	218 (60.9)	214 (62.6)
Bowel disorder, n(%)	1,451					
No bowel disease		1,235 (85.1)	147 (82.6)	482 (84.0)	300 (83.8)	306 (89.7)
IBS		81 (5.6)	10 (5.6)	30 (5.2)	31 (8.7)	10 (2.9)
Celiac disease		18 (1.2)	3 (1.7)	8 (1.4)	4 (1.1)	3 (0.9)
IBD		24 (1.7)	5 (2.8)	11 (1.9)	7 (2.0)	1 (0.3)
Other		93 (6.4)	13 (7.3)	43 (7.5)	16 (4.5)	21 (6.2)
Smoking status ² , n (%)	1,462					
Non-smoker		1,082 (74.0)	127 (70.2)	438 (75.3)	270 (75.4)	247 (72.4)
Smoker		380 (26.0)	54 (29.8)	144 (24.7)	88 (24.6)	94 (27.6)
Snus status ³ , n(%)	1,405					
Non-snuser		1,308 (93.1)	171 (97.2)	530 (95.3)	318 (92.4)	289 (87.8)
Snuser		97 (6.9)	5 (2.8)	26 (4.7)	26 (7.6)	40 (12.2)
BMI, kg/m ²	1,480	26 (24, 29)	27 (25, 30)	26 (24, 29)	26 (24, 29)	27 (24, 29)
Physical activity, min/week	1,466	135 (0, 300)	45 (0, 195)	135 (0, 300)	180 (15, 390)	180 (45, 352)
Questionnaires completed prior to colonoscopy, n(%)	1,486					
No		126 (8.5)	15 (8.0)	42 (7.1)	30 (8.3)	39 (11.3)
Yes		1,360 (91.5)	172 (92.0)	550 (92.9)	331 (91.7)	307 (88.7)

¹Values are median (Q1, Q3) for continuous variables and n (%) for categorical variables.

^{2,3}To be defined as a smoker or snuser one had to be a regular or occasional user or having quit consumption within the last ten years.

Abbreviations: BMI; Body mass index, CRC; colorectal cancer, g; gram, n; number

Table 2. Alcohol consumption in the study population as a whole (n=1,486) and by sex (660 women, 826 men).

	Percentiles				Adherent to guidelines ¹	Zero consumers	Correlation with energy intake
	25	50	75	100	n (%)	n (%)	r_s^2
Overall							
Alcohol, g/day	2.2	9.0	19	195	991 (67)	187 (13)	0.22**
Alcohol, E%	0.7	2.9	6.0	45	1003 (68)	187 (13)	-0.04
Alcohol units ³ /day	0.2	0.7	1.6	16	1083 (73)	187 (13)	0.22**
Alcohol subtypes, g/day							
Beer	0.0	33	140	3743	-	535 (36)	0.24**
Wine	0.0	36	108	837	-	417 (28)	0.06*
Spirits	0.0	0.0	1.2	223	-	1040 (70)	0.11**
Drinks, cider, etc.	0.0	0.0	0.0	768	-	1240 (83)	0.10**
Non-alcoholic drinks, g/day	0.0	0.0	14	2000	-	1100 (74)	0.13**
Men							
Alcohol, g/day	3.8	13	25	195	563 (68)	70 (8)	0.17**
Alcohol, E%	1.2	3.6	7.2	45	503 (61)	70 (8)	-0.10*
Alcohol units ³ /day	0.3	1.1	2.1	16	610 (74)	70 (8)	0.17**
Alcohol subtypes, g/day							
Beer	16	80	203	3743	-	167 (20)	0.19**
Wine	0	35	108	837	-	247 (30)	0.04
Spirits	0	0	4,8	223	-	489 (59)	0.07
Drinks, cider, etc.	0	0	0	768	-	667 (81)	0.07*
Non-alcoholic drinks, g/day	0	0	14	2000	-	591 (72)	0.11*
Women							
Alcohol, g/day	1.1	5.4	13	87	428 (65)	117 (18)	0.13*
Alcohol, E%	0.4	2.0	4.8	34	500 (76)	117 (18)	-0.12*
Alcohol units ³ /day	0.1	0.5	1.1	7.23	473 (72)	117 (18)	0.13*
Alcohol subtypes, g/day							
Beer	0.0	0.0	45	1280	-	368 (56)	0.11*
Wine	0.0	41	108	837	-	170 (26)	0.09*
Spirits	0.0	0.0	0.0	78	-	549 (83)	0.04
Drinks, cider, etc.	0.0	0.0	0.0	288	-	571 (87)	0.08
Non-alcoholic drinks, g/day	0	0	0	500	-	511 (77)	0.19**

¹<10 g/day for women and <20 g/day for men according to Nordic and national nutritional recommendations^{21,25}; <5 energy percentage (E%) for both sexes according to Nordic and national nutritional recommendations^{21,25}; <1 unit for women and <2 units for men according to national food-based dietary guidelines⁶².

²**<0.001, *<0.05.

³1 unit set to 12 grams in line with national guidelines^{21,25}.

Abbreviations: E%; energy percentage, n; number, r_s ; Spearman's correlation coefficient.

Table 3. Odds ratios (ORs) and 95% confidence intervals (CIs) for presence of non-advanced adenoma¹ and advanced lesions² relative to controls by level of alcohol consumption in the study population as a whole (n=1,486) and by sex (660 women, 826 men)^{3,4}.

	Control (n=548)	Non-advanced adenoma (n=524)		Advanced lesions (n=414)	
	n	n	OR (95% CI)	n	OR (95% CI)
Overall					
Level of intake					
0 g/day	88	65	Ref.	34	Ref.
>0-10 g/day	226	202	1.16 (0.79, 1.70)	164	1.81 (1.15, 2.85)
≥10-20 g/day	126	131	1.31 (0.86, 2.00)	104	1.99 (1.22, 3.27)
≥20 g/day	108	126	1.30 (0.84, 2.02)	112	2.19 (1.32, 3.63)
<i>P_{trend}</i>			0.20		0.008
Per 10 g increase/day			1.03 (0.94, 1.12)		1.09 (1.00, 1.19)
<i>P_{cont}</i>			0.52		0.047
Per 2-fold increase/day			1.03 (0.96, 1.11)		1.14 (1.05, 1.23)
<i>P_{cont}</i>			0.42		0.002
Adherence to guidelines					
Fully adhering (0 g/day)	88	65	Ref.	34	Ref.
Partially adhering (<10/20 g/day)	291	275	1.16 (0.80, 1.70)	238	1.91 (1.22, 2.99)
Not adhering (≥10/20 g/day)	169	184	1.34 (0.89, 2.00)	142	1.97 (1.23, 3.17)
Men					
Level of intake					
0 g/day	31	22	Ref.	17	Ref.
>0-10 g/day	87	111	1.58 (0.84, 2.97)	83	1.55 (0.78, 3.06)
≥10-20 g/day	65	73	1.37 (0.71, 2.67)	74	1.92 (0.95, 3.88)
≥20 g/day	84	99	1.32 (0.69, 2.51)	80	1.48 (0.74, 2.96)
<i>P_{trend}</i>			0.99		0.48
Per 10 g increase/day			1.00 (0.91, 1.11)		1.04 (0.95, 1.15)
<i>P_{cont}</i>			0.92		0.37
Per 2-fold increase/day			0.99 (0.89, 1.09)		1.08 (0.98, 1.21)
<i>P_{cont}</i>			0.81		0.13
Adherence to guidelines					
Fully adhering (0 g/day)	31	22	Ref.	17	Ref.
Partially adhering (<20 g/day)	152	184	1.49 (0.81, 2.74)	157	1.69 (0.88, 3.25)
Not adhering (≥20 g/day)	84	99	1.32 (0.69, 2.52)	80	1.47 (0.74, 2.93)
Women					
Level of intake					
0 g/day	57	43	Ref.	17	Ref.
>0-10 g/day	139	91	0.96 (0.58, 1.58)	81	2.07 (1.10, 3.90)
≥10-20 g/day	61	58	1.46 (0.82, 2.57)	30	1.80 (0.87, 3.73)
≥20 g/day	24	27	1.65 (0.81, 3.35)	32	4.81 (2.18, 10.62)
<i>P_{trend}</i>	57	43	0.05	17	0.001
Per 10 g increase/day			1.09 (0.91, 1.30)		1.27 (1.06, 1.52)
<i>P_{cont}</i>			0.36		0.010
Per 2-fold increase/day			1.09 (0.97, 1.22)		1.22 (1.08, 1.39)
<i>P_{cont}</i>			0.13		0.002
Adherence to guidelines					
Fully adhering (0 g/day)	57	43	Ref.	17	Ref.
Partially adhering (<10 g/day)	139	91	0.96 (0.58, 1.58)	81	2.09 (1.11, 3.93)
Not adhering (≥10 g/day)	85	85	1.51 (0.89, 2.57)	62	2.66 (1.37, 5.19)

¹Includes any adenoma (adenomatous polyp) not fulfilling the criteria of being advanced.

²Includes advanced adenoma, defined as any adenoma with either villous histology ($\geq 25\%$ villous components), high-grade dysplasia or polyp size greater than or equal to 10 mm; advanced serrated lesions, defined as any serrated lesions with size ≥ 10 mm or dysplasia; and colorectal cancer, defined as presence of adenocarcinoma arising from the colon or rectum.

³Odds ratios (ORs) and 95% confidence intervals (CIs) are obtained from multinomial logistic regression analyses adjusting for the following covariates: age (continuous), sex (except in the sex-specific analyses), national affiliation (Norwegian affiliation, non-Norwegian affiliation, missing), screening center (center 1, center 2), education level (primary school, high school, college/university, missing), family history of CRC (yes, no, unknown/missing), smoking status (non-smoker, smoker, missing), BMI (continuous with missing set to median) and level of physical activity (continuous with missing set to median).

⁴Potential interactions of alcohol intake with sex was examined using the Wald test, resulting in the following p-values: 'Level of intake' (p-values in sequential order (2-4): 0.57, 0.82, 0.040), 'Per 10 g increase/day' (p-value=0.10), 'Per 2-fold increase/day' (p-value=0.24), 'Adherence to guidelines' (p-values for partial and non-adherence of 0.69 and 0.28, respectively).

Figure legends

Figure 1. Flowchart of study participants.

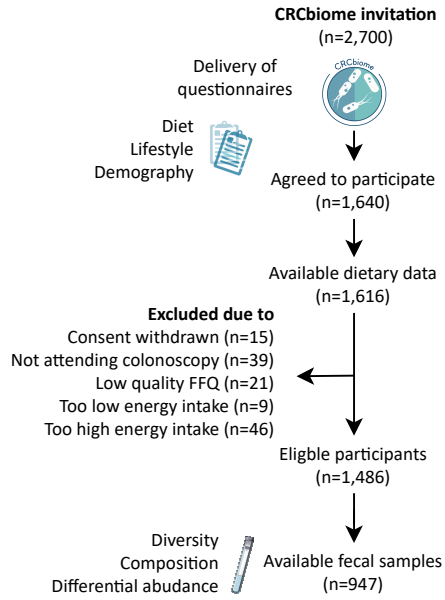
Figure 2. Odds ratios (ORs) and 95% confidence intervals (CIs) for presence of non-advanced adenoma and advanced lesions relative to controls for participants consuming *vs.* not consuming alcohol, as well as the different types of alcoholic beverages in the study overall (n=1,486) and by sex (660 women, 826 men).

Figure 3. Associations of alcohol intake with the α -diversity indices Shannon and Inverse Simpson (a-c) and β -diversity based on the Bray-Curtis dissimilarity metric (d-g).

Figure 4. Differential abundance analyses of bacterial species by total alcohol intake and alcoholic beverage types.

Figure 5. Differential abundance analyses of pathways by total alcohol intake and alcoholic beverage types.

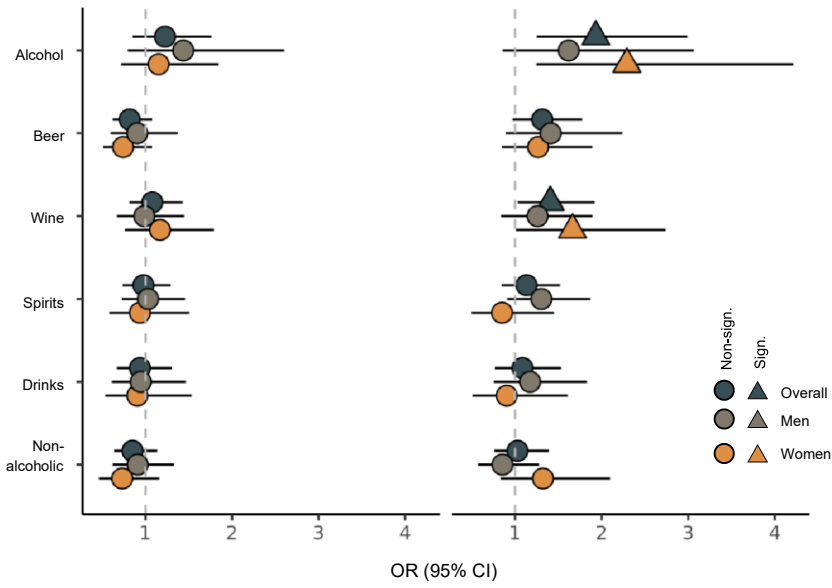
Figure 6. Causal mediation analysis.



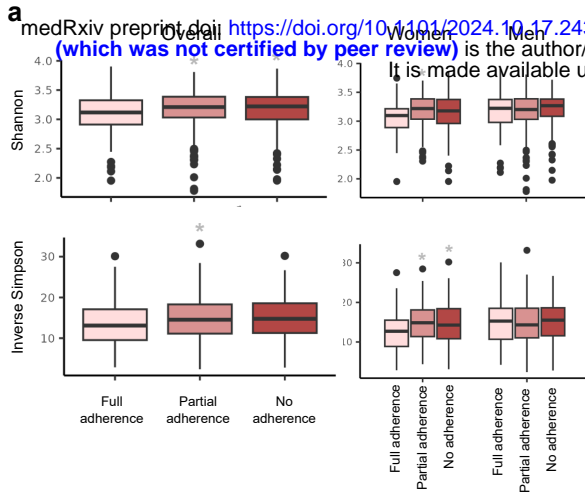
Consumption vs. non-consumption

Non-advanced adenoma

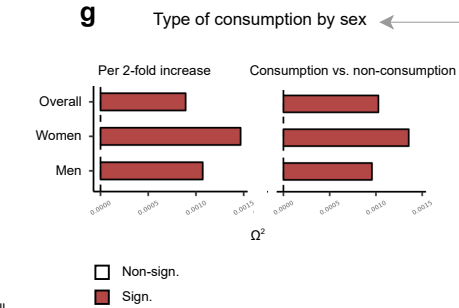
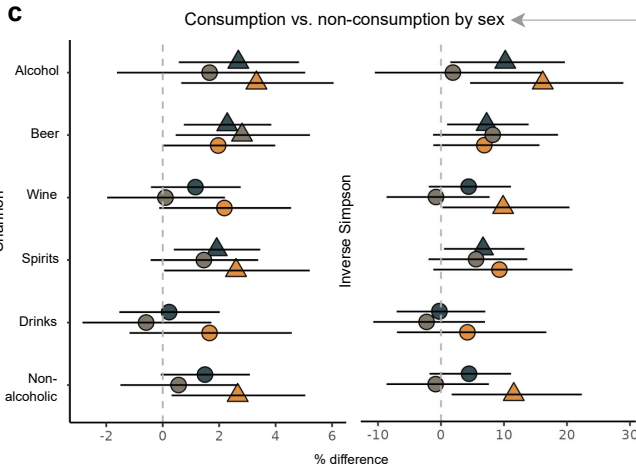
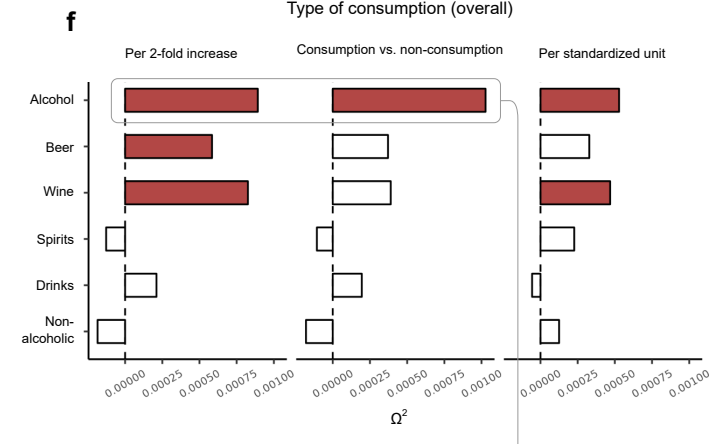
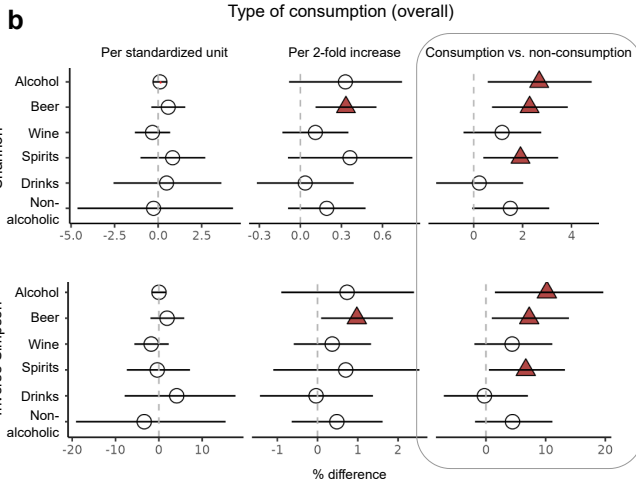
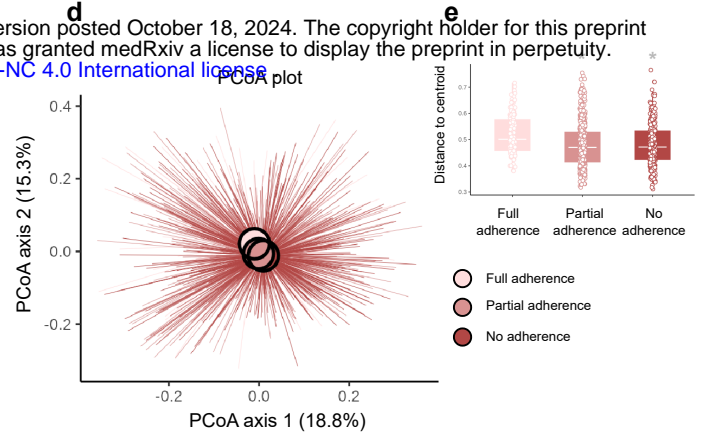
Advanced lesions



α -diversity



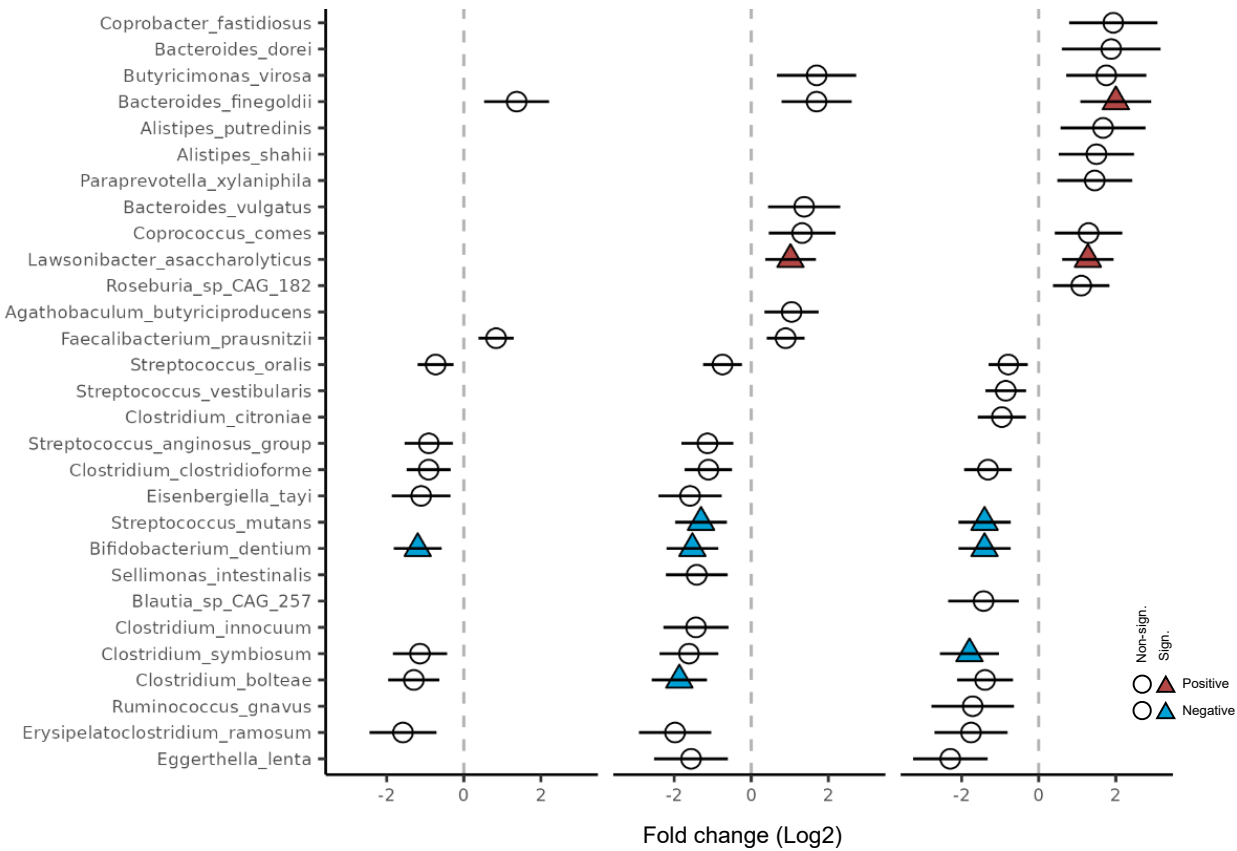
β -diversity



Differentially abundant bacteria

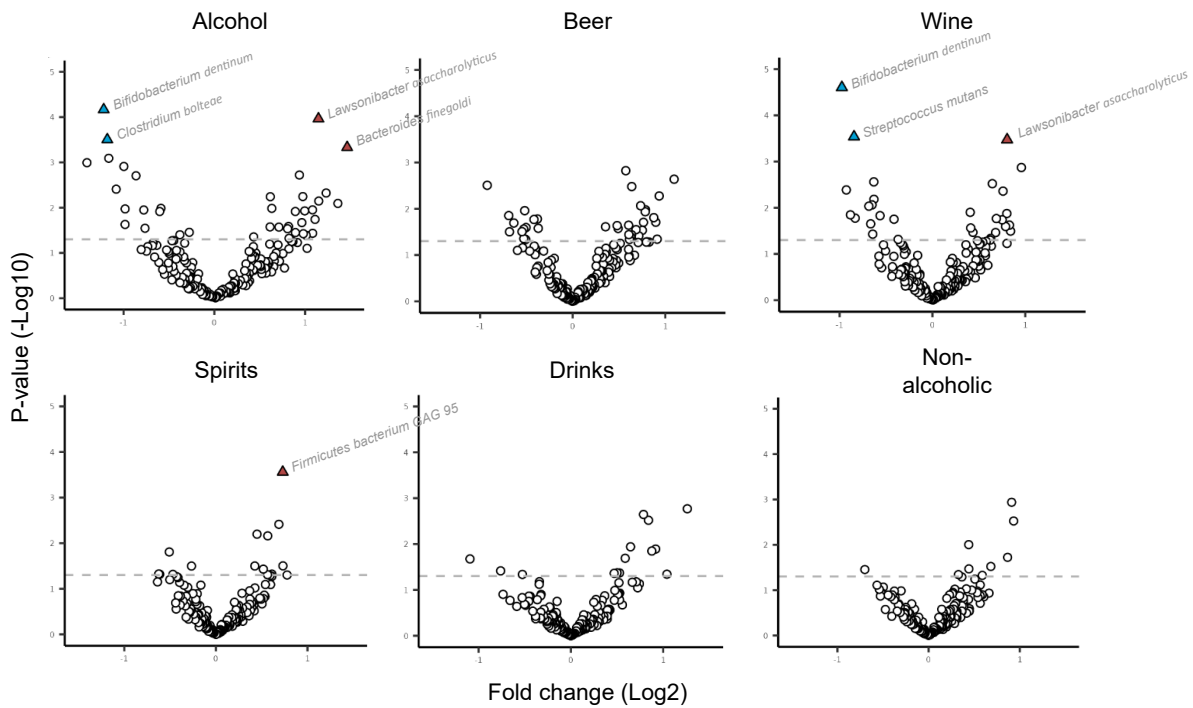
Level of alcohol intake

medRxiv preprint doi: <https://doi.org/10.1101/2024.10.17.24315656>; this version posted October 18, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license.



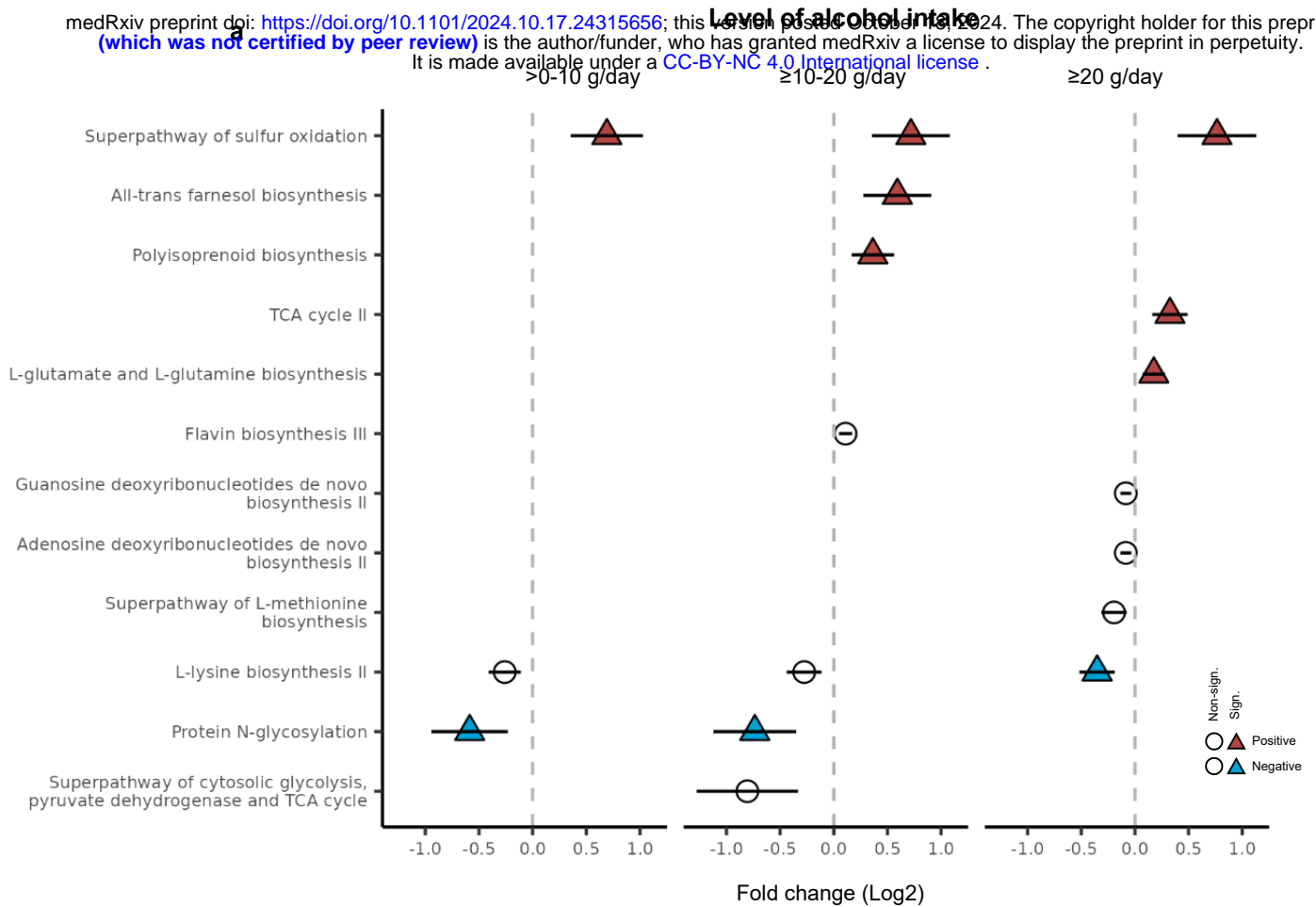
b

Consumption/non consumption



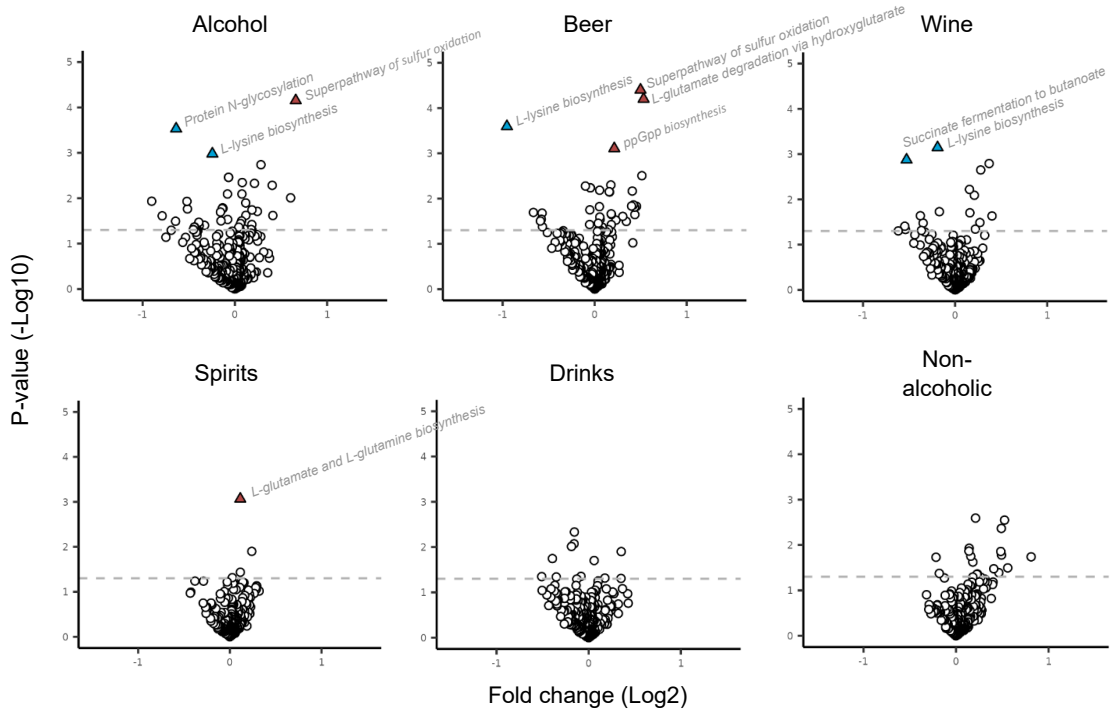
Differentially abundant pathways

medRxiv preprint doi: <https://doi.org/10.1101/2024.10.17.24315656>; this version posted October 24, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a [CC-BY-NC 4.0 International license](https://creativecommons.org/licenses/by-nc/4.0/).

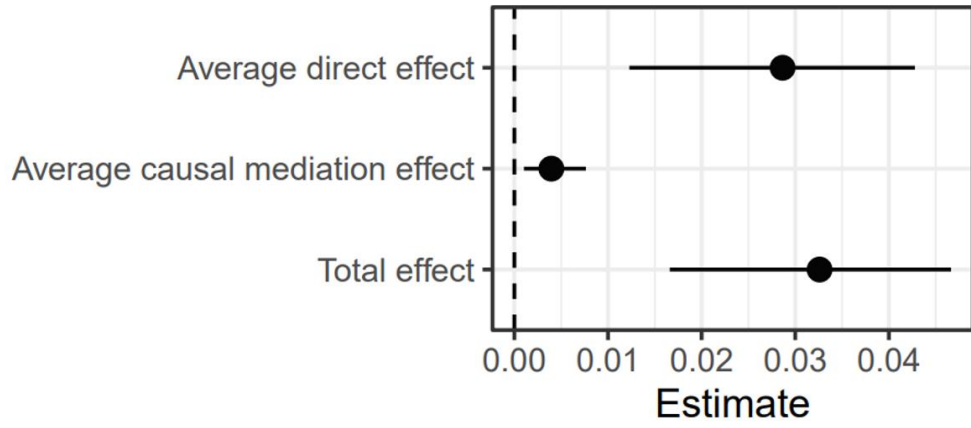


b

Consumption/non consumption



a



b

