1 Using genetic variants to evaluate the causal effect of cholesterol

2 lowering on head and neck cancer risk: a Mendelian randomization

3 study

4 Mark Gormley^{1,2,3}*, James Yarmolinsky^{1,3}, Tom Dudding^{1,2,3}, Kimberley Burrows^{1,3}, Richard M

5 Martin^{1,3,4}, Steven Thomas^{2,4}, Jessica Tyrrell⁵, Paul Brennan⁶, Miranda Pring², Stefania

6 Boccia^{7,8}, Andrew F Olshan⁹, Brenda Diergaarde¹⁰, Rayjean J. Hung^{11,12}, Geoffrey Liu^{12,13},

7 Danny Legge¹⁴, Eloiza H Tajara¹⁵, Patricia Severino¹⁶, Martin Lacko¹⁷, Andrew R Ness⁴,

8 George Davey Smith^{1,3}, Emma E Vincent^{1,3,14&}, Rebecca C Richmond^{1,3&}

9 ¹ MRC Integrative Epidemiology Unit, Population Health Sciences, Bristol Medical School,

10 University of Bristol, Bristol, UK.

² Bristol Dental Hospital and School, University of Bristol, Bristol, UK.

- ³ Department of Population Health Sciences, Bristol Medical School, University of Bristol,
 Bristol, UK.
- ⁴National Institute for Health Research Bristol Biomedical Research Centre at the University
- 15 Hospitals Bristol and Weston NHS Foundation Trust, University of Bristol, Bristol, UK.
- ⁵ University of Exeter Medical School, RILD Building, RD&E Hospital, Exeter, UK.
- ⁶ Genetic Epidemiology Group, World Health Organization, International Agency for
- 18 Research on Cancer, Lyon, France.

19 ⁷ Sezione di Igiene, Istituto di Sanità Pubblica, Università Cattolica del Sacro Cuore, Roma,

20 Italy.

- ⁸Department of Woman and Child Health and Public Health, Public Health Area, Fondazione
- 22 Policlinico Universitario A. Gemelli IRCCS, Roma, Italy.
- ⁹ Department of Epidemiology, Gillings School of Global Public Health, University of North
- 24 Carolina, US.
- 25 ¹⁰ Department of Human Genetics, Graduate School of Public Health, University of
- 26 Pittsburgh, and UPMC Hillman Cancer Center, Pittsburgh, US.
- 27 ¹¹ Prosserman Centre for Population Health Research, Lunenfeld-Tanenbaum Research
- 28 Institute, Sinai Health System, Toronto, Canada.
- 29
- 30 ¹² Dalla Lana School of Public Health, University of Toronto, Toronto, Canada.
- 31 ¹³ Princess Margaret Cancer Centre, Toronto Canada.
- 32 ¹⁴ School of Cellular and Molecular Medicine, University of Bristol, Bristol, UK.
- 33 ¹⁵ School of Medicine of São José do Rio Preto, São Paulo, Brazil.
- ¹⁶ Albert Einstein Research and Education Institute, Hospital Israelita Albert Einstein, São
 Paulo, Brazil.
- 36 17 Department of Otorhinolaryngology and Head and Neck Surgery, Research Institute
- 37 GROW, Maastricht University Medical Center, Maastricht, The Netherlands.

38 Corresponding author

39 * Email: <u>mark.gormley@bristol.ac.uk</u>

40

41 **Author contributions**

42 [&] These authors also contributed equally to this work.

43	E.E.V. and R.C.R. are joint senior authors. M.G., R.C.R. and E.E.V. conceived the study
44	and M.G. carried out data curation and formal analysis. K.B. performed the GWAS of head
45	and neck cancer in UK Biobank. Head and neck cancer summary genetic data was obtained
46	through multiple collaborations from studies lead and curated by A.R.N., S.T., A.F.O., B.D.,
47	R.J.H., G.L., M.L., E.T., S.B. and P.B. Expert guidance on MR methodology was provided by
48	R.C.R., J.Y., G.D.S. and R.M.M., R.C.R., G.D.S, J.T., S.T., P.B., R.M.M. and A.R.N. provided
49	expertise on both observational and genetic epidemiological methodology and analysis. The
50	initial manuscript was drafted by M.G., R.C.R. and E.E.V. All authors (M.G., J.Y., T.D., K.B.,
51	R.M.M., S.T., J.T., A.R.N., P.B., M.P., S.B., A.F.O., B.D, R.J.H., G.L., D.L., E.H.T., P.S., M.L.,
52	A.R.N., G.D.S., E.E.V. and R.C.R.) contributed to the interpretation of the results, editing and
53	critical revision of the manuscript. M.G.'s PhD supervisory team includes R.C.R., E.E.V., J.T.,
54	A.R.N and G.D.S.

55

57	Abbreviations: HNSCC, head and neck squamous cell carcinoma; OC, oral cancer; OPC,
58	oropharyngeal cancer; HPV, human papilloma virus; MR, Mendelian randomization; GWAS,
59	genome-wide association study; SNP, single nucleotide polymorphism; PCA, principal
60	component analysis; NOME, no measurement error; SIMEX, simulation extrapolation; MR-
61	PRESSO, Mendelian Randomization Pleiotropy RESidual Sum and Outlier; GAME-ON, Genetic
62	Associations and Mechanisms in Oncology; HMGCR, 3-hydroxy-3-methylglutaryl-CoA
63	reductase; NPC1L1, NPC1 like intracellular cholesterol transporter 1; CETP, cholesteryl ester
64	transfer protein; PCSK9, proprotein convertase subtilisin/kexin type 9; LDLR, low-density
65	lipoprotein receptor; LDL-C, low-density lipoprotein cholesterol; HDL-C, high-density
67	inverse variance weighted: CL confidence intervals: P. <i>p</i> -value
07	

76 Abstract

77 Introduction

78 Head and neck squamous cell carcinoma (HNSCC), which includes cancers of the oral 79 cavity and oropharynx, is a cause of substantial global morbidity and mortality. Strategies to 80 reduce disease burden include discovery of novel therapies and repurposing of existing 81 drugs. Statins are commonly prescribed for lowering circulating cholesterol by inhibiting 82 HMG-CoA reductase (HMGCR). Results from some observational studies suggest that statin 83 use may reduce HNSCC risk. We appraised the relationship of genetically-proxied 84 cholesterol-lowering drug targets and other circulating lipid traits with oral (OC) and 85 oropharyngeal (OPC) cancer risk.

86

87 Methods and findings

We conducted two-sample Mendelian randomization (MR). For the primary analysis, germline genetic variants in *HMGCR*, *NPC1L1*, *CETP*, *PCSK9* and *LDLR* were used to proxy the effect of low-density lipoprotein cholesterol (LDL-C) lowering therapies. In secondary analyses, variants were used to proxy circulating levels of other lipid traits in a genome-wide association study (GWAS) meta-analysis of 188,578 individuals. Both primary and secondary analyses aimed to estimate the downstream causal effect of cholesterol lowering therapies on OC and OPC risk.

95

96 The second sample for MR was taken from a GWAS of 6,034 OC and OPC cases and
97 6,585 controls (GAME-ON). Analyses were replicated in UK Biobank, using 839 OC and OPC

cases and 372,016 controls and the results of the GAME-ON and UK Biobank analyses
combined in a fixed-effects meta-analysis.

100

101 We found limited evidence of a causal effect of genetically-proxied LDL-C lowering 102 using HMGCR, NPC1L1, CETP or other circulating lipid traits on either OC or OPC risk. 103 Genetically-proxied PCSK9 inhibition equivalent to a 1 mmol/L (38.7 mg/dL) reduction in 104 LDL-C was associated with an increased risk of OC and OPC combined (OR 1.8 95%Cl 1.2, 2.8, $p=9.31 \times 10^{-05}$), with good concordance between GAME-ON and UK Biobank ($l^2=22\%$). 105 106 Effects for PCSK9 appeared stronger in relation to OPC (OR 2.6 95%Cl 1.4, 4.9) than OC (OR 107 1.4 95%CI 0.8, 2.4). LDLR variants, resulting in genetically-proxied reduction in LDL-C equivalent to a 1 mmol/L (38.7 mg/dL), reduced the risk of OC and OPC combined (OR 0.7, 108 109 95%CI 0.5, 1.0, p= 0.006). A series of pleiotropy-robust and outlier detection methods 110 showed that pleiotropy did not bias our findings.

111

112 Conclusion

We found limited evidence for a role of cholesterol-lowering in OC and OPC risk, suggesting previous observational results may have been confounded. There was some evidence that genetically-proxied inhibition of PCSK9 increased risk, while lipid-lowering variants in LDLR, reduced risk of combined OC and OPC. This result suggests that the mechanisms of action of PCSK9 on OC and OPC risk may be independent of its cholesterol lowering effects, but further replication of this finding is required.

119 Author summary

120 Why was this study done?

121	•	To determine if genetically-proxied cholesterol-lowering drugs (such as statins which								
122		target HMGCR) reduce oral and oropharyngeal cancer risk.								
123	•	To determine if genetically-proxied circulating lipid traits (e.g. low-density								
124		lipoprotein cholesterol) have a causal effect on oral and oropharyngeal cancer risk.								
125										
126	What	did the researchers do and find?								
127	•	There was little evidence that genetically-proxied inhibition of HMGCR (target of								
128		statins), NPC1L1 (target of ezetimibe) and CETP (target of CETP inhibitors) influences								
129		oral or oropharyngeal cancer risk.								
130	•	There was little evidence of an effect of circulating lipid traits on oral or								
131		oropharyngeal cancer risk.								
132	•	There was some evidence that genetically-proxied inhibition of PCSK9 increases,								
133		while lipid-lowering variants in LDLR reduces oral and oropharyngeal cancer risk .								
134										
135	What	do these findings mean?								
136	•	These findings suggest that the results of previous observational studies examining								
137		the effect of statins on oral and oropharyngeal risk may have been confounded.								
138	•	Given we found little evidence of an effect of other cholesterol lowering therapies,								
139		the mechanism of action of PCSK9 may be independent of cholesterol-lowering.								
140		Further replication of this finding in other head and neck cancer datasets is required.								
141										
142										

143 Introduction

144

145 Head and neck squamous cell carcinoma (HNSCC), which includes cancers of the oral 146 cavity (OC) and oropharynx (OPC), is the sixth most common cancer in the world, with over 147 550,000 new cases and 300,000 deaths annually [1, 2]. Despite some modest improvements 148 in the treatment of HNSCC, survival ranges between 19 - 59% at 10 years [3] and recurrence 149 rates remain high [4]. Patients often undergo a combination of surgery, radiotherapy and 150 chemotherapy which can result in significant morbidity [5]. Established risk factors include 151 smoking, alcohol and human papilloma virus (HPV), the latter mainly linked with 152 oropharyngeal cancer [6, 7]. Given that in contemporary cohorts, around 70% of OPC cases 153 (versus <5% of OC cases) are HPV driven and often present in younger populations, oral and 154 oropharyngeal tumours are considered distinct disease entities, with different risk factor 155 profiles [8]. Despite smoking cessation, alcohol reduction measures and the implementation 156 of HPV vaccination in some areas, HNSCC remains a major global health problem [2]. Novel 157 strategies for prevention of HNSCC are required, in particular for those at high risk and one 158 approach is to identify novel risk factors which can be easily modified, for example by 159 repurposing existing drugs [9].

160

Statins are one of the most commonly prescribed medications worldwide. They are prescribed to reduce levels of circulating total and low-density lipoprotein cholesterol (LDL-C), with proven preventative and therapeutic effects in cardiovascular disease and a good safety profile [10, 11]. Statins inhibit 3-hydroxy-3-methylglutaryl-coenzyme A reductase (HMG-CoA reductase (HMGCR)), the rate-limiting enzyme in the mevalonate pathway, responsible for cholesterol and steroid hormone synthesis [12]. Other clinically approved

167 drugs that target cholesterol metabolism via different mechanisms include ezetimibe 168 Niemann-Pick C1-like protein (NPC1L1)) and (targeting proprotein convertase 169 subtilisin/kexin type 9 (PCSK9) inhibitors such as evolocumab or alirocumab. These agents 170 act by reducing the intestinal absorption of cholesterol or by enhancing uptake of LDL-C 171 through increased cellular membrane expression of the LDL-receptor (LDLR), respectively. 172 Conversely, cholesteryl ester transfer protein inhibitors (CETP) substantially increase levels 173 of high-density lipoprotein cholesterol (HDL-C), lower levels of LDL-C and enhance reverse 174 cholesterol transport [13].

175

176 Cholesterol is vital for a variety of key cellular functions, including membrane 177 integrity, signalling, protein synthesis and cell cycle progression. Therefore, modulation of 178 cholesterol synthesis has the potential to influence several hallmarks of tumourigenesis 179 including cell migration and proliferation [14]. In an experimental study, mice given oral 180 daily doses of simvastatin two weeks prior to subcutaneous injection of FaDu (HPV-negative 181 hypopharyngeal tumour cells), showed a significant reduction in tumour growth. This study 182 was designed to mimic a clinical scenario where patients who present with a tumour may 183 have been taking the medication prior to tumour initiation [15], and suggested that 184 simvastatin antagonises tumour metabolic reprogramming, another important hallmark of 185 cancer [15]. Mechanistic support for the role of LDL-C lowering in cancer development 186 comes largely from the fact that lipids are themselves major cell membrane components 187 essential for cell division and maintaining tissue integrity. Changes in lipid levels have been 188 reported associated with tumour development [16, 17].

189

190 However, the evidence that cholesterol-lowering drugs may reduce the risk of head 191 and neck cancer is limited. Some observational studies report an inverse association of 192 taking statins with both head and neck cancer risk and survival [14, 18], but others indicate 193 little evidence of any effect [19]. Observational studies are not randomised and are 194 susceptible to reverse causality and/or confounding [19]. Mendelian randomization (MR) is 195 an approach that uses germline genetic proxies (referred to as instruments) to help appraise 196 causal effects of potentially modifiable extrinsic exposures or intrinsic traits with disease 197 [20-22]. It has also been used to estimate therapeutic potential by investigating genetic 198 variation at drug targets [23]. Numerous single nucleotide polymorphisms (SNPs) are 199 associated with lower levels of circulating LDL-C [24, 25] and inheriting an LDL-C lowering 200 allele has been proposed to be analogous to being assigned life-long treatment with a 201 cholesterol lowering drug [26]. In this way, germline genetic variants may serve as proxies 202 for exposure to potential pharmacological agents which are less likely than observational 203 measures to be subject to reverse causation or confounding. Genetic proxies can therefore 204 be used to predict both the likely beneficial and adverse effects of long-term modulation of 205 the drug targets on disease.

206

207 MR has previously demonstrated the protective effect of cholesterol-lowering drugs 208 on cardiovascular disease risk [27], but also that inhibition of HMGCR and PCSK9 may have 209 an adverse effect on diabetes risk [28]. Of relevance to cancer, some recent MR studies have 210 shown that genetically-proxied inhibition of HMGCR may be protective against overall 211 cancer [29] and epithelial ovarian cancer [30] risk. Our aim was to use MR to appraise the 212 causal nature and mechanistic basis of the relationship between cholesterol-lowering and

risk of oral and oropharyngeal cancer by investigating germline variation in HMGCR,
NPC1L1, CETP, PCSK9 and LDLR, and other related lipid traits such as circulating LDL-C.

215

216 Methods

217

218 Identifying cholesterol-lowering genetic instruments. For the primary analysis, SNPs in 219 HMGCR, NPC1L1, CETP, PCSK9 and LDLR were used to proxy the effect of lipid-lowering 220 therapies and to estimate the downstream effect of manipulating these targets on OC and 221 OPC risk. SNPs were identified within 100 kb on either side of the target gene (HMGCR, 222 NPC1L1, CETP, PCSK9 and LDLR) that were associated with LDL-C levels. Variants were 223 robustly associated with LDL-C in a meta-analysis of genome-wide association studies 224 (GWAS) involving 188,578 individuals primarily (96%) of European ancestry in the Global 225 Lipids Genetic Consortium (GLGC) [25]. Linkage disequilibrium (LD) (r^2) was evaluated to 226 check the overlap between variants using LDmatrix (4.0 Release) [31]. Proxy SNPs were permitted to be in weak LD (r^2 < 0.2) with each other, to increase the proportion of variance 227 228 explained in each respective drug target therefore maximising instrument strength. Multiple 229 papers have used these genetic instruments at this threshold, to demonstrate causal effects 230 in cardiovascular disease [13, 27, 32], diabetes [28] and ovarian cancer [30]. In secondary 231 analyses, SNPs were used to proxy circulating levels of LDL-C, HDL-C, total triglyceride, total 232 cholesterol, apolipoprotein A and B. Betas represented the change in lipid trait levels per copy of the effect allele. These SNPs were already independently ($r^2 < 0.001$) associated with 233 234 the respective traits in large GWAS which have been described previously [25, 33].

235

236 Summary level genetic data on oral and oropharyngeal cancer from GAME-ON.

237 We estimated the effects of the cholesterol-lowering genetic variants on risk of OC and OPC 238 using GWAS performed on 6,034 cases and 6,585 controls from 12 studies which were part 239 of the Genetic Associations and Mechanisms in Oncology (GAME-ON) Network [34]. The 240 study population included participants from Europe (45.3%), North America (43.9%) and 241 South America (10.8%). Cancer cases comprised the following the International 242 Classification of Diseases (ICD) codes: oral cavity (C02.0-C02.9, C03.0-C03.9, C04.0-C04.9, 243 C05.0-C06.9) oropharynx (C01.9, C02.4, C09.0-C10.9), hypopharynx (C13.0-C13.9), 244 overlapping (C14 and combination of other sites) and 25 cases with unknown ICD code 245 (other). A total of 954 individuals with cancers of hypopharynx, unknown code or 246 overlapping cancers were excluded. Genomic DNA isolated from blood or buccal cells was 247 genotyped at the Center for Inherited Disease Research (CIDR) using an Illumina OncoArray 248 custom designed for cancer studies by the OncoArray Consortium [35]. In GAME-ON, all SNPs with a call rate of <95% were excluded. Given the ethnic heterogeneity of the study 249 250 population, the dataset was divided by geographical region and SNPs within each region 251 that showed deviation from Hardy-Weinberg Equilibrium (HWE) in controls ($p = <1 \times 10^{-7}$) 252 were excluded. Principal component analysis (PCA) was performed using approximately 10,000 common markers in low LD (r^2 <0.004), minor allele frequency (MAF) >0.05 and 139 253 254 population outliers were removed. Full details of the included studies, as well as the 255 genotyping and imputation performed, have been described previously [34, 36].

256

Two-sample Mendelian randomization. Two-sample MR was conducted using the "TwoSampleMR" package in R (version 3.5.3), by integrating SNP associations for cholesterol-lowering (sample 1) with those for OC and OPC in GAME-ON (sample 2). For

260 those SNPs instrumenting LDL-C lowering, we first extracted summary statistics for the 261 associations with OC and OPC from GAME-ON. We next performed harmonisation of the 262 direction of effects between the cholesterol-lowering exposures and outcome (OC or OPC) 263 where, for each variant, the allele designated the 'exposure allele' was associated with 264 lower LDL-C levels and palindromic SNPs were aligned when MAFs were <0.3 or were 265 otherwise excluded. In our primary analysis, four palindromic SNPs, one in HMGCR 266 (rs2006760), the other in PCSK9 (rs2149041) and two in CETP (rs5880, rs9929488) were 267 removed. In the secondary analysis with other lipid traits, 11 palindromic SNPs (rs1936800, 268 rs2288912, rs7112577, rs964184, rs150617279, rs1883711, rs4722043, rs2156552, 269 rs2954029, rs581080, rs7534572) were removed.

270

271 Individual effect-estimates for each SNP were calculated using the Wald ratio, by 272 dividing the SNP-outcome association by the SNP-exposure association. Multiple SNPs were 273 then combined into multi-allelic instruments using the random-effects inverse-variance 274 weighted (IVW) meta-analysis method, for each of the genes HMGCR, NPC1L1, CETP, PCSK9 275 and LDLR. This meta-analysis was undertaken to increase the proportion of variance in drug 276 targets and LDL-C lowering explained by each instrument, and thus improve statistical 277 power and the precision of our estimates [37]. The analysis produced an estimate of the 278 effect of the risk factor on OC and OPC risk. As the instruments for HMGCR, NPC1L1, CETP, *PCSK9* and *LDLR* were in weak LD ($r^2 < 0.2$), we accounted for this correlation between SNPs 279 280 in the primary analysis using LDlink (4.0 Release) which employs Phase 3 (Version 5) of the 281 1000 Genomes Project and variant rs numbers based on dbSNP. Correlation matrices were 282 inserted as an MRInput object [31], resulting in MR methods which altered the weightings for correlated SNPs [38-40]. For circulating lipid traits, a more stringent r² <0.001 was 283

284 already applied in the initial GWAS [25], so we did not account further for correlation in the 285 secondary analysis. We computed odds ratios (OR) which represent the change in odds of 286 oral and oropharyngeal squamous cell carcinoma per genetically-proxied inhibition of the 287 drug target, equivalent to a 1 mg/dl decrease in LDL-C. The OR was scaled to be per 1 288 mmol/L decrease by dividing the LDL-C lowering effect (beta) and standard error (se) 289 measured in mg/dL by 38.7 [30]. The betas and standard errors for HMGCR, NPC1L1, CETP, 290 PCSK9 and LDLR SNPs were also converted to reflect the cholesterol-lowering effect, given 291 the "TwoSampleMR" package preference to automatically change these into a positive 292 direction of effect (i.e. lipid increasing). The correct direction of effect was checked using a 293 positive control of coronary heart disease from the CARDIoGRAM GWAS data [41]. MR was 294 also used to examine the effect of circulating levels of LDL-C, HDL-C, total triglyceride, total 295 cholesterol, apolipoprotein A and B levels directly with OC and OPC cancer risk. The ORs in 296 this analysis represent the change in odds of oral or oropharyngeal squamous cell 297 carcinoma, per SD unit increase in lipid trait.

298

299 The IVW method can provide an unbiased effect estimate in the absence of 300 horizontal pleiotropy or when horizontal pleiotropy is balanced [42]. We therefore 301 performed sensitivity analyses to evaluate the potential for unbalanced horizontal 302 pleiotropy, where genetic variants influence two or more traits through independent 303 biological pathways. To ensure the genetic instrument was associated with the instrument it 304 was proxying, estimates of the proportion of variance in each risk factor explained by the 305 instrument (R^2) and F-statistics were generated. An F-statistic of <10 is indicative of a weak 306 instrument which may be subject to weak instrument bias. To account for directional 307 pleiotropy, we compared the IVW results with three MR sensitivity analyses, which each

308 make different assumptions: MR Egger [43], weighted median [44] and weighted mode [45]. 309 While these three methods are best used when genetic instruments consist of a large numbers of independent SNPs, since r² values between SNPs in our instruments were low 310 $(r^2 < 0.2)$, we a-priori decided to include them, with a further sensitivity analysis to account 311 312 for correlation in the MR Egger analysis. The weighted median stipulates that at least 50% of 313 the weight in the analysis stems from variants that are valid instruments [44], while the 314 weighted mode requires that the largest subset of instruments which identify the same 315 causal effect to be valid instruments [45]. MR-Egger can provide unbiased estimates even 316 when all SNPs in an instrument violate the exclusion restriction assumption (i.e. affect the 317 outcome by means other than via the risk factor of interest). However, there must be 318 negligible measurement error (NOME) [46] in the genetic instrument and the InSIDE 319 (Instrument Strength Independent of Direct Effect) assumption must be satisfied [43]. 320 Where there was evidence of violation of the NOME assumption, this was assessed using the I² statistic and MR-Egger was performed with simulation extrapolation (SIMEX) 321 322 correction [46]. To further assess the robustness of findings, we examined evidence of heterogeneity in the individual SNP estimates using the Cochran Q-statistic, which may 323 324 indicate the presence of invalid instruments (e.g. due to horizontal pleiotropy) [47]. Scatter 325 and leave-one-out plots were produced to evaluate influential outliers and MR-PRESSO 326 (Mendelian Randomization Pleiotropy RESidual Sum and Outlier) was used to detect and 327 correct for potential outliers (where Q-statistic p < 0.05) [48].

328

Stratification by cancer subsite. Given the difference in established aetiology (i.e. smoking, alcohol and HPV) at each HNSCC subsite [6], we performed MR analyses with stratification by cancer subsite to evaluate potential heterogeneity in effects. For this, we used GWAS summary data on a subset of 2,641 OPC cases and 2,990 OC cases from the 6,034 HNSCC cases and the 6,585 common controls in the GAME-ON GWAS [34].

334

335 **Replication in UK Biobank.** UK Biobank data was used as a replication dataset for primary 336 analyses. A GWAS was performed on 839 combined OC and OPC cases and 372,016 controls, 337 with two further stratified GWAS for OC (n= 357) and OPC (n= 494). UK Biobank is a large 338 population-based cohort study that recruited over 500,000 men and women aged between 339 37 and 73 years between 2006 and 2010 throughout the UK. It received ethical approval 340 from the National Health Service North West Centre for Research Ethics Committee 341 (reference: 11/NW/0382). Details of genotyping quality control, phasing and imputation are 342 described elsewhere [49]. Participant records are linked to cancer registry data and HNSCC 343 was grouped using the same ICD-codes as described above. Squamous cell carcinoma cases 344 were identified using histology codes 8070 - 8078. UK Biobank GWAS analyses were 345 adjusted for sex and genotyping array and performed in BOLT-LMM, a mixed model that 346 accounts for population stratification and relatedness [50, 51]. Primary MR analyses as 347 described above in GAME-ON were repeated in UK Biobank data.

348

Meta-analysis of results. We performed both fixed-effects and random-effects metaanalysis of the MR estimates in GAME-ON and UK Biobank using the R package 'meta'. However, we focus more on the fixed-effects estimates since we assume that the causal

effect is constant between the studies [52]. Heterogeneity between study populations was assessed using l^2 statistic [53].

354

355 Results

356

357	Primary analysis in GAME-ON. In total, 5 SNPs in HMGCR (rs12916, rs17238484, rs5909,
358	rs2303152, rs10066707) were used to proxy HMG-CoA reductase inhibition (statins); 5 SNPs
359	in NPC1L1 (rs217386, rs2073547, rs7791240, rs10234070, rs2300414) proxied NPC1L1
360	inhibition (ezetimibe); 6 SNPs in CETP (rs9989419, rs12708967, rs3764261, rs1800775,
361	rs1864163, rs289714) proxied CETP inhibition; 6 SNPs in <i>PCSK9</i> (rs11206510, rs2479409,
362	rs2479394, rs10888897, rs7552841, rs562556) proxied PCSK9 inhibition; 3 SNPs proxied
363	cholesterol-lowering in LDLR (rs6511720, rs1122608, rs688) (LDL-receptor inhibition).
364	Further details of these SNP effects are given in Table 1 .
365	

366 Table 1. Detailed summary of LDL-C lowering genetic variants in HMGCR, NPC1L1, CETP,

367 *PCSK9* and *LDLR* variants from the in Global Lipids Genetics Consortium (GLGC)

Target	SNP	Pathway	EA	OA	EAF	Beta	se	<i>P</i> -value
HMGCR	rs12916	LDL-C	т	С	0.57	-0.06061	0.003	7.79E-78
	rs17238484	LDL-C	G	Т	0.75	-0.05184	0.005	1.35E-21
	rs5909	LDL-C	G	А	0.90	-0.05102	0.007	4.93E-13
	rs2303152	LDL-C	G	А	0.88	-0.03498	0.005	1.04E-09
	rs10066707	LDL-C	G	А	0.58	-0.0411	0.005	2.97E-19
	rs2006760*	LDL-C	С	G	0.81	-0.04407	0.006	1.67E-13
NPC1L1	rs217386	LDL-C	А	G	0.41	-0.02908	0.003	1.20E-19
	rs2073547	LDL-C	А	G	0.81	-0.03885	0.004	1.92E-21
	rs7791240	LDL-C	Т	С	0.91	-0.03404	0.005	1.84E-10
	rs10234070	LDL-C	С	Т	0.90	-0.02363	0.005	1.52E-06

	rs2300414	LDL-C	G	А	0.93	-0.02828	0.006	5.45E-06
CETP	rs3764261	LDL-C	А	С	0.29	-0.04471	0.004	2.22E-34
	rs1800775	LDL-C	А	с	0.48	-0.03487	0.003	8.54E-24
	rs1864163	LDL-C	G	А	0.73	-0.03698	0.004	7.97E-21
	rs9929488*	LDL-C	G	С	0.70	-0.03159	0.004	8.15E-13
	rs9989419	LDL-C	G	А	0.60	-0.02344	0.004	2.49E-12
	rs12708967	LDL-C	Т	С	0.80	-0.02963	0.004	3.47E-11
	rs289714	LDL-C	А	G	0.79	-0.03032	0.005	2.85E-10
	rs5880*	LDL-C	G	С	0.94	-0.03979	0.008	1.59E-06
PCSK9	rs11206510	LDL-C	С	Т	0.15	-0.06871	0.001	2.38E-53
	rs2479409	LDL-C	А	G	0.67	-0.05309	0.001	2.52E-50
	rs2149041*	LDL-C	С	G	0.84	-0.05259	0.001	1.44E-35
	rs2479394	LDL-C	А	G	0.72	-0.03192	0.001	1.58E-19
	rs10888897	LDL-C	Т	С	0.40	-0.04192	0.001	8.43E-31
	rs7552841	LDL-C	С	Т	0.63	-0.04589	0.001	5.40E-15
	rs562556	LDL-C	G	А	0.19	-0.05292	0.002	6.16E-21
LDLR	rs6511720	LDL-C	Т	G	0.11	-0.17482	0.004	3.69E-54
	rs1122608	LDL-C	Т	G	0.23	-0.05473	0.003	2.02E-86
	rs688	LDL-C	С	Т	0.56	-0.04465	0.003	3.04E-48

369

370 EA, effect allele or low-density lipoprotein-cholesterol (LDL-C) lowering allele; OA = other or non-

371 effect allele; EAF, effect allele frequency; se= standard error. * Palindromic SNPs removed

Beta represents the change in LDL-C levels per copy of the effect allele. For SI conversion of mmol/L

to mg/dL, multiply by 38.7.

375	There was limited evidence of an effect of genetically-proxied inhibition of HMGCR
376	and NPC1L1 on combined OC and OPC risk (OR IVW 1.1; 95%CI 0.6, 1.9, p= 0.82 and 1.0; 95%
377	CI 0.4, 2.7, p= 0.99), respectively. (Table 2 and S1 Fig). A similar result was found for
378	genetically-proxied LDL-C lowering inhibition of CETP on OC and OPC combined (OR IVW 1.3;
379	95%CI 0.6, 2.6, p= 0.49) (Table 2 and S1 Fig). However, higher risk of combined OC and OPC
380	was found in relation to genetically-proxied PCSK9 inhibition, equivalent to a 1 mmol/L (38.7
381	mg/dL) reduction in LDL-C (OR IVW 2.1; 95%Cl 1.2, 3.4, p= 0.01; Table 2 and Fig 1). This is in
382	contrast to the reduction in odds seen in relation to cardiovascular disease using the same
383	instrument (OR IVW 0.6; 95%Cl 0.4, 0.8, p <1 x10 ⁻⁰³ ; Fig 2) in 60,801 cases and 123,504

384	control subjects enrolled in the CARDIoGRAM consortia studies [41]. There was also some
385	evidence that LDLR variants, resulting in genetically-proxied reduction in LDL-C equivalent to
386	a 1 mmol/L (38.7 mg/dL), reduced the risk of combined OC and OPC (OR IVW; 0.7; 95%CI
387	0.4, 1.0, p= 0.05; Table 2, Fig 1).
388	
389	Fig 1. Forest plot showing the causal effects of HMGCR, NPC1L1, CETP, PCSK9 and LDLR
390	variants on the oral and oropharyngeal cancer subsites in GAME-ON. Effect estimates on
391	oral and oropharyngeal cancer are reported on the log odds scale.
392	
393	Fig 2. Forest plots showing the causal effects of cholesterol-lowering PCSK9 and LDLR
394	single nucleotide polymorphisms on coronary heart disease and combined oral/
395	oropharyngeal cancer in GAME-ON. Effect estimates on oral and oropharyngeal cancer are
396	reported on the log odds scale.

397 **Table 2.** Mendelian randomization results of genetically-proxied inhibition of HMGCR, NPC1L1, CETP, PCSK9 and LDLR with risk of oral and

398 oropharyngeal cancer including sensitivity analyses in GAME-ON

					IVW		Weighted median		Weighted mode		MR-Egger	
	Outcome	Exposure/ Outcome dataset	Outcome N	Number of SNPs	OR (95%Cl)	Р	OR (95%Cl)	Р	OR (95%Cl)	Р	OR (95%Cl)	Р
		GAME-ON/ GLGC										
	Oral/ Oropharyngeal cancer		6,034	5	1.07 (0.62, 1.84)	0.82	1.20 (0.63, 2.28)	0.58	1.20 (0.57, 2.50)	0.66	1.33 (0.07, 26.58)	0.86
		GAME-ON/ GLGC										
HMGCR	Oral cancer		2,990	5	1.49 (0.75, 2.96)	0.25	1.66 (0.73, 3.78)	0.23	1.69 (0.65, 4.42)	0.35	1.25 (0.03, 55.80)	0.91
		GAME-ON/ GLGC										
	Oropharyngeal cancer		2,641	5	0.90 (0.43, 1.85)	0.77	1.01 (0.43, 2.34)	0.99	1.08 (0.44, 2.64)	0.88	0.77 (0.01, 45.07)	0.91
		GAME-ON/ GLGC										
	Oral/ Oropharyngeal cancer		6,034	5	1.01 (0.38, 2.69)	0.99	0.90 (0.29, 2.82)	0.86	0.86 (0.24, 3.08)	0.83	0.22 (0.00, 102.42)	0.66
NPC1L1		GAME-ON/ GLGC										
	Oral cancer		2,990	5	1.02 (0.30, 3.41)	0.98	1.20 (0.29, 5.05)	0.80	1.29 (0.22, 7.41)	0.79	0.09 (0.00, 160.66)	0.57
		GAME-ON/ GLGC										
	Oropharyngea cancer		2,641	5	0.60 (0.16, 2.25)	0.45	0.60 (0.12, 3.04)	0.53	0.60 (0.10, 3.73)	0.62	0.20 (0.01, 688.64)	0.72
		GAME-ON/ GLGC										
CETP	Oral/ Oropharyngeal cancer		6,034	6	1.28 (0.64, 2.55)	0.49	1.24 (0.54, 2.81)	0.61	1.25 (0.43, 3.64)	0.71	0.82 (0.02, 27.60)	0.92
		GAME-ON/ GLGC										
	Oral cancer		2,990	6	1.65 (0.70, 3.88)	0.25	1.69 (0.61, 4.70)	0.31	1.63 (0.43, 5.87)	0.49	0.42 (0.01, 32.58)	0.72
		GAME-ON/ GLGC										
	Oropharyngea cancer		2,641	6	1.12 (0.45, 2.77)	0.81	1.05 (0.35, 3.20)	0.93	0.73 (0.17, 3.07)	0.69	1.50 (0.02, 142.22)	0.87

	GAME-ON/										
	GLGC										
	Oral/ Oropharyngeal cancer	6,034	6	2.05 (1.24, 3.38)	0.01	2.21 (1.12, 4.08)	0.01	2.19 (0.81, 5.92)	0.18	1.83 (0.17, 20.00)	0.65
	GAME-ON/										
	GLGC										
PCSK9	Oral cancer	2,990	6	1.62 (0.89, 2.94)	0.11	1.80 (0.90, 3.62)	0.10	2.00 (0.73, 5.48)	0.23	2.17 (0.15, 32.13)	0.60
	GAME-ON/										
	GLGC				2.00						
	Oropharyngeal cancer	2,641	6	3.49 (1.58, 7.68)	E-03	3.19 (1.40, 7.26)	0.01	2.73 (0.86, 8.70)	0.15	1.99 (0.04, 92.11)	0.74
	GAME-ON/										
	GLGC										
	Oral/ Oropharyngeal cancer	6,034	3	0.65 (0.42, 1.00)	0.05	0.71 (0.47, 1.08)	0.11	0.74 (0.47, 1.15)	0.31	0.90 (0.41, 1.95)	0.83
LDLR	GAME-ON/										
	GLGC										
	Oral cancer	2,990	3	0.53 (0.32, 0.87)	0.01	0.56 (0.32, 0.98)	0.04	0.57 (0.32, 1.04)	0.21	0.56 (0.18, 1.76)	0.50
	GAME-ON/										
	GLGC										
	Oropharyngeal cancer	2 64 1	3	0 90 (0 53 1 52)	0.69	0 91 (0 55 1 53)	0.73	1 08 (0 62 1 89)	0.81	1 56 (0 64 3 82)	0.51

400

401 Abbreviations: IVW, inverse variance weighted; OR, odds ratio; CI, confidence intervals; P, *p*-value.

402

403 OR represents the exponential change in odds of oral/ oropharyngeal squamous cell carcinoma per genetically-proxied inhibition of drug target equivalent

404 to a 1 mmol/L decrease in LDL-C.

405	Stratification by cancer subsite in GAME-ON. When stratified by subsite, the
406	adverse effect of PCSK9 appeared to be mainly in the oropharynx, (OR IVW 3.5; 95%Cl 1.6,
407	7.7, p= 2 $\times 10^{-03}$), with limited evidence in the oral cavity (OR IVW 1.6; 0.9, 2.94, p= 0.11)
408	(Table 2 & Fig 1). The effects appeared stronger in OPC versus OC, but with overlapping
409	confidence intervals. LDLR was associated with a reduction in risk of OC (OR IVW 0.5; 95%Cl
410	0.3, 0.9, $p=$ 0.01), but there was little evidence of an association with OPC (OR IVW 0.9;
411	95%CI 0.5, 1.5, p= 0.69) (Table 2 and Fig 1). For both PCSK9 and LDLR associations, the
412	direction of effect was generally consistent across the four MR methods tested (Table 2).
413	
414	Secondary analysis in GAME-ON. SNPs were also used to proxy circulating levels of LDL-C
415	(77 SNPs), high-density lipoprotein cholesterol (HDL-C) (85 SNPs), total triglyceride (54
416	SNPs), total cholesterol (82 SNPs), apolipoprotein A (9 SNPs) and apolipoprotein B (14 SNPs)
417	(S1 Table). There was limited evidence of an effect of any of these other lipid traits on either
418	OC or OPC (Table 3 and S2 Fig).
 419 420 421 422 423 424 	
425 426 427	
428 429 430	
430 431 432	
433 434 435	

438 **Table 3.** Mendelian randomization results of circulating lipid traits with risk of oral and

439 oropharyngeal cancer in GAME-ON

Target	N SNPs	Outcome	IVW OR (95% CI)	<i>P</i> -value
LDL-C	77	Oral/ Oropharyngeal cancer	0.98 (0.87, 1.11)	0.79
		Oral cancer	0.99 (0.85, 1.15)	0.88
		Oropharyngeal cancer	1.03 (0.87, 1.21)	0.76
HDL-C	85	Oral/ Oropharyngeal cancer	0.98 (0.83, 1.16)	0.79
		Oral cancer	1.08 (0.88, 1.32)	0.45
		Oropharyngeal cancer	0.87 (0.71, 1.05)	0.15
Total triglycerides	54	Oral/ Oropharyngeal cancer	1.19 (1.01, 1.04)	0.04
		Oral cancer	1.19 (0.96, 1.46)	0.11
		Oropharyngeal cancer	1.19 (0.96, 1.47)	0.12
Total cholesterol	82	Oral/ Oropharyngeal cancer	1.04 (0.91, 1.18)	0.55
		Oral cancer	1.09 (0.93, 1.28)	0.30
		Oropharyngeal cancer	1.00 (0.84, 1.20)	0.96
Apolipoprotein A	9	Oral/ Oropharyngeal cancer	0.87 (0.73, 1.03)	0.11
		Oral cancer	0.90 (0.73, 1.12)	0.34
		Oropharyngeal cancer	0.82 (0.65, 1.04)	0.10
Apolipoprotein B	14	Oral/ Oropharyngeal cancer	1.04 (0.88, 1.22)	0.67
		Oral cancer	1.15 (0.89, 1.49)	0.29
		Oropharyngeal cancer	1.00 (0.83, 1.21)	0.98

440

441 Abbreviations: IVW, inverse variance weighted; OR, odds ratio; CI, confidence intervals.

442 IVW OR represents the exponential change in odds of oral/ oropharyngeal squamous cell carcinoma

443 per SD increase in the circulating lipid trait (one SD for LDL-C = 38.7 mg/dL, HDL-C = 15.5 mg/dL,

444 Apolipoprotein A = 0.32 g/L, Apolipoprotein B = 0.52 g/L, Total triglycerides = 90.7 mg/dL, Total

445 cholesterol = 41.8 mg/dL).

446	Sensitivity analyses. IVW, MR Egger, weighted median, simple and weighted mode were
447	carried out, in addition to IVW analysis accounting for LD structure (S2.1-2.5 Tables). The
448	results adjusting for SNP correlation followed the same pattern as the main results (S3
449	Table). There was limited evidence of weak instrument bias being present (F-statistic >10)
450	and the proportion of variance in the phenotype (R^2) explained by the genetic instruments
451	ranged from 0.1 to 6% (S4 Table). In both primary and secondary analyses there was limited
452	evidence of heterogeneity in the SNP effect estimates for IVW and MR Egger regression,
453	except for in HDL-C (Q IVW 115.7, p= 0.01; Q MR Egger 115.6, p= 0.01) (S5 and S6 Tables).
454	
455	MR Egger intercepts also indicated limited evidence of directional pleiotropy (S7 and
456	S8 Tables). For the primary analysis there were no clear outliers in both scatter and leave-
457	one-out plots (Fig 3 and S3-6 Figs) and MR-PRESSO detected no individual outliers (S9
458	Table). Where there was evidence of violation of the NOME assumption for the HMGCR,
459	NPC1L1 and CETP instruments (i.e. I ² statistic <0.90) (S10 Table), MR-Egger was performed
460	with SIMEX correction and effects were still consistent with the null (S11 Table).
461	
462	Fig 3. Scatter plots for LDLR and PCSK9 single nucleotide polymorphisms effect on
463	combined oral/ oropharyngeal cancer in GAME-ON.
464	
465	Replication in UK Biobank and meta-analysis of results. Primary MR analyses as described
466	above were replicated in UK Biobank, showing limited evidence of an effect of genetically-
467	proxied inhibition of HMGCR, NPC1L1, CETP, PCSK9 and LDLR on risk of OC and OPC (S12
468	Table). Following IVW fixed-effects meta-analysis of GAME-ON and UK Biobank MR results,

469 there was a consistently strong effect of genetically-proxied PSCK9 inhibition on combined

470	OC and OPC (OR IVW 1.8; 95%Cl 1.2, 2.8), with good concordance between studies (l^2 = 22%)
471	and methods used (Fig 4 and S7 Fig). Effects for PCSK9 appeared stronger in relation to OPC
472	(OR IVW 2.6; 95%CI 1.4, 4.9) than OC (OR IVW 1.4; 95%CI 0.8, 2.4), but with moderate
473	heterogeneity between studies (l^2 = 41%) (Fig 4). Conversely, the protective effect for LDLR
474	on OC and OPC combined was also consistent in the meta-analysis (OR IVW 0.7; 95%Cl 0.5,
475	1.0), with good concordance between studies (I^2 = 0%) (Fig 5 and S7 Fig). However, the
476	protective effect seen specifically in relation to OC in GAME-ON (OR IVW 0.5; 95%CI 0.3, 0.9)
477	was not replicated in UK Biobank (OR IVW 1.6; 95%CI 0.5, 4.8), with strong evidence of
478	heterogeneity between the studies (I^2 = 66%).
479	
480	Fig 4. Forest plots showing inverse variance weighted meta-analysis effects of cholesterol-
481	lowering <i>PCSK9</i> single nucleotide polymorphisms on head and neck cancer subsites.
482	
482 483	Fig 5. Forest plots showing inverse variance weighted meta-analysis effects of cholesterol-
482 483 484	Fig 5. Forest plots showing inverse variance weighted meta-analysis effects of cholesterol- lowering <i>LDLR</i> single nucleotide polymorphisms on head and neck cancer subsites.
482 483 484 485	Fig 5. Forest plots showing inverse variance weighted meta-analysis effects of cholesterol- lowering <i>LDLR</i> single nucleotide polymorphisms on head and neck cancer subsites.
482 483 484 485 486	Fig 5. Forest plots showing inverse variance weighted meta-analysis effects of cholesterol- lowering <i>LDLR</i> single nucleotide polymorphisms on head and neck cancer subsites. Discussion
482 483 484 485 486 487	Fig 5. Forest plots showing inverse variance weighted meta-analysis effects of cholesterol- lowering <i>LDLR</i> single nucleotide polymorphisms on head and neck cancer subsites. Discussion
482 483 484 485 486 487 488	Fig 5. Forest plots showing inverse variance weighted meta-analysis effects of cholesterol- lowering LDLR single nucleotide polymorphisms on head and neck cancer subsites. Discussion We found limited evidence for a role of cholesterol-lowering in OC and OPC risk. This
482 483 484 485 486 487 488 489	Fig 5. Forest plots showing inverse variance weighted meta-analysis effects of cholesterol-lowering LDLR single nucleotide polymorphisms on head and neck cancer subsites. Discussion We found limited evidence for a role of cholesterol-lowering in OC and OPC risk. This included the absence of a protective effect of genetically-proxied inhibition of HMGCR,
482 483 484 485 486 487 488 489 490	Fig 5. Forest plots showing inverse variance weighted meta-analysis effects of cholesterol- lowering LDLR single nucleotide polymorphisms on head and neck cancer subsites. Discussion We found limited evidence for a role of cholesterol-lowering in OC and OPC risk. This included the absence of a protective effect of genetically-proxied inhibition of HMGCR, suggesting previous observational studies investigating the relationship between statins and
482 483 484 485 486 487 488 489 490 491	Fig 5. Forest plots showing inverse variance weighted meta-analysis effects of cholesterol-lowering LDLR single nucleotide polymorphisms on head and neck cancer subsites. Discussion We found limited evidence for a role of cholesterol-lowering in OC and OPC risk. This included the absence of a protective effect of genetically-proxied inhibition of HMGCR, suggesting previous observational studies investigating the relationship between statins and head and neck cancer risk may be subject to residual confounding or bias. However, we did

magnitude to the protective effect seen in relation to cardiovascular disease using the same

494 genetic instrument (**Fig 2**). This PCSK9 effect was evident in both the GAME-ON (n= 6,034 OC 495 and OPC cases and n= 6,585 controls) and UK Biobank datasets (n= 839 OC and OPC cases 496 and n= 372,016 controls). We also found some evidence for a protective effect of 497 cholesterol-lowering variants in *LDLR* on OC and OPC risk in both studies. The IVW analysis 498 for HMGCR, NPC1L1, CETP, PCSK9 and LDLR accounting for LD structure ($r^2 < 0.2$) followed 499 the same pattern as the main results (**S3 Table**).

500

501 Association between PCSK9, LDLR and cancer risk. Despite the lack of consistent evidence 502 for a role of cholesterol-lowering on risk of OC or OPC in this study, individual effects of both 503 PCSK9 and LDLR were demonstrated which may implicate a role for these drug targets in the 504 development of OC or OPC via other mechanisms. The effects identified in the present study 505 are directionally consistent with a recent MR analysis of 1,615 combined head and neck 506 cancer cases from UK Biobank, which also found that a 1 SD unit increase in LDL-C proxied 507 by PCSK9 and LDLR was associated with a reduction (OR 0.7 95%Cl 0.4, 1.4, p= 0.35) and 508 increase in odds (OR 1.6 95%CI 1.0, 2.4, p= 0.05) of head and neck cancer, respectively [29]. 509 The opposing effects of PCSK9 and LDLR in this study also suggests that cholesterol lowering 510 is unlikely to be the main mechanism of action. However, this previous analysis is limited by 511 the relatively small number of cases and heterogeneity of head and neck cancer subtypes, 512 with no selection for the histological subtype of squamous cell carcinoma. In the present 513 study, we focused specifically on oral and oropharyngeal subtypes of HNSCC.

514

515 Beyond the established role of *PCSK9* in cholesterol homeostasis, other potential 516 pleiotropic effects are not well understood. Variants in *PCSK9* have been associated with an 517 increased risk of diabetes (OR 1.1, 95%CI 1.0, 1.2 for each 10 mg per decilitre decrease in

518 LDL-C) [28]. However, a recent phenome-wide association study (PheWAS) did not 519 find *PCSK9* or *LDLR* to be correlated with any non-lipid-related phenotypes, including 520 diabetes [54]. There is limited *in vivo* and *in vitro* evidence that *PCSK9* might be involved in 521 both cell proliferation and apoptosis. The gene was initially designated as NARC1 (neural 522 apoptosis-regulated convertase 1), involved in apoptosis of cerebellar neurons [55] and 523 *PCSK9* has since been found to be upregulated in some cancers [56, 57].

524

525 One suggested mechanism for a link with cancer progression is that the increased 526 expression of PCSK9 prevents LDL-receptor (LDLR) recycling, leading to 527 hypercholesterolaemia and more exogenous lipid to support the proliferation of the tumour 528 [55]. Our study suggests the opposite, that genetically-proxied inhibition of *PCSK9* results in 529 an increased risk of OC or OPC. We hypothesise that access to intracellular LDL-C could in 530 fact be pro-tumourigenic, providing a favourable environment for a developing tumour cell, 531 maintaining membrane integrity and promoting cell division. Therapies such as statins, 532 ezetimibe, and PCSK9 inhibitors may all lower LDL-C level through the upregulation of LDL-533 receptors, resulting in elevated intracellular cholesterol. However, cancer mechanisms are 534 often context dependent and perhaps only the expression of *PCSK9* and *LDLR* is relevant in 535 head and neck cancer. CETP inhibitors instead block the transfer of cholesteryl ester from 536 HDL-C to LDL-C, thereby raising HDL-C and lowering LDL-C (and apolipoprotein B), as well as 537 enhancing reverse cholesterol transport [58]. Unlike statins, CETP inhibitors do not appear 538 to increase the risk of type 2 diabetes, thought to be as a result of pancreatic islet cell 539 cholesterol accumulation with use of other cholesterol-lowering drugs [59]. Therefore, the 540 absence of effect when proxying CETP inhibition in this study, supports the possible 541 mechanism of LDL-C uptake via the LDL-receptor in OC or OPC.

542

543	While both PCSK9 [60] and LDLR [61] are expressed in head and neck tumours, this was not
544	evident in normal oral or oropharyngeal tissue, and sufficient tissue is currently not
545	available in expression datasets [62]. Recent studies have associated elevated PCSK9 with
546	alcohol use disorder, including the interesting possibility of using anti-PSCK9 monoclonal
547	antibodies for the treatment of alcoholic liver disease [63, 64]. Given that alcohol is a well-
548	known risk factor for HNSCC, this pleiotropy could have partially explained the effect seen in
549	our study, however we proxied the inhibition of <i>PCSK9</i> and so would have expected to see a
550	protective effect of this gene in HNSCC cases who may have been heavy alcohol drinkers.
551	Further investigation is required to untangle the relationship between PCSK9, alcohol and
552	head and neck cancer.

553

554 **Comparison with previous studies.** It is believed that statins could play a potential role in 555 cancer chemoprevention which may reduce the risk of some site-specific cancers such as 556 prostate [65] and ovarian [30], but not all. Some of these studies have reported cancer risk 557 reductions by as much as 50 – 65% [66-68]. However, meta-analyses and clinical trials have 558 contradicted these findings [69, 70]. In addition to confounding, immortal time bias likely 559 inflated observational results because, to be classified as a long-term statin user 560 necessitates that users survived without cancer over a long period [19, 71]. Dickerman et al. 561 used electronic records from 733,804 adults with 10-year follow-up to emulate a trial 562 design. To achieve this a pre-specified protocol was set, including eligibility criteria and 563 checks were made to ensure effect estimates for statins on cancer were comparable 564 between the large observational dataset and trial. The authors found little indication that 565 statin therapy influences cancer incidence, which was consistent with the analyses of

566 randomised trials (with a 10-year cancer-free survival difference of -0.3% 95%Cl -1.5%, 567 0.5%) [19]. Nonetheless, recent MR studies [29, 30] have identified an association between 568 variants in HMGCR with cancer risk, but not alternative cholesterol-lowering treatments or 569 genetically-predicted LDL-C, suggesting that statins may reduce cancer risk through a 570 cholesterol independent pathway. A recent case-control study of over 11,000 participants 571 found an inverse association between statin use and the occurrence of HNSCC (OR 0.86, 572 95%Cl 0.77, 0.95, p = < 0.01, of prior statin exposure for cases compared to propensity 573 score-matched controls) [18]. However, a wider systematic review found that the evidence 574 for the role of statins in the prevention of HNSCC was limited [72]. As HNSCC incidence is a 575 rare outcome, randomised controls trials are not feasible, so we must be cautious 576 interpreting the available observational findings given the potential for bias and 577 confounding as discussed previously.

578

579 In contrast to previous MR studies assessing overall cancer risk [29] and ovarian 580 specific risk [30], the MR carried out here in relation to OC and OPC showed no effect using 581 genetic instruments for HMGCR (statins). There was also limited evidence for a causal effect 582 of NPC1L1 (ezetimibe), CETP (CETP inhibitors) as well as a number of other circulating lipid 583 traits on OC or OPC. Therefore, it remains unclear as to whether the effects observed with 584 PCSK9 and LDLR are via LDL-C lowering or another less well-established pathway [73, 74], 585 such as receptor regulation for viral entry, synthesis of sex hormones and resultant 586 dysregulated metabolism [75-77].

587

588 **Strengths and limitations of this study.** Protective associations between cholesterol-589 lowering therapies such as statins and head and neck cancer risk seen in previous

590 observational studies could be a result of reverse causation, immortal time bias, lack of 591 randomisation or confounding by socioeconomic status, smoking or HPV infection, for 592 example. Our study applied MR in an attempt to overcome these issues, using the largest 593 number of SNPs identified from the latest GWAS for both cholesterol-lowering and head 594 and neck cancer that could be identified in the literature [25, 34]. A series of pleiotropy-595 robust MR methods and outlier detection were applied to rigorously explore the possibility 596 that findings were not biased as a result of pleiotropy. However, there was no HPV data 597 available in these summary results to enable more detailed stratified analysis of OPC. We 598 did replicate findings for PCSK9 and LDLR, but the number of cases are low. Finally, most 599 participants in the GAME-ON network [34] were of European or North American decent, 600 with only around 11% from South America, and participants included in the UK Biobank. 601 analysis were exclusively of European descent, so more work is required to determine if our 602 results translate to other ancestry groups.

603

604 In conclusion, our MR analyses provided little evidence for a role of cholesterol 605 lowering in OC or OPC risk although effects of genetically-proxied inhibition of PCSK9 and 606 cholesterol-lowering variants in LDLR were observed in relation to OPC and OC risk. Given 607 the lack of a common pathway to carcinogenesis in OC or OPC, metabolic targets that may 608 be common to all tumours, regardless of the activated molecular pathway, could help 609 simplify a preventative or therapeutic approach [78]. Replication of our findings in other 610 head and neck cancer datasets and use of individual-level follow-up data with HPV status 611 could provide further insight into the effect of these genetic instruments on risk, treatment outcomes and survival in head and neck cancer. With more evidence, future randomised 612

613	control trials could focus on treating high risk OC or OPC groups, such as those with severe
614	dysplasia or who are HPV-16 E6 positive, with cholesterol-lowering therapies.
615	
616	Data availability statement
617	
618	Summary-level analysis was conducted using publicly available GWAS data [25, 33].
619	Full summary statistics for the GAME-ON GWAS can be accessed via dbGAP (OncoArray:
620	Oral and Pharynx Cancer; study accession number: phs001202.v1.p1) [34]. Access to UK
621	Biobank (<u>https://www.ukbiobank.ac.uk/</u>) data is available to researchers through
622	application.
623	
624	Code availability statement
625	
626	Two-sample MR analyses were conducted using the "TwoSampleMR" package in R
627	(version 3.5.3). A copy of the code used in this analysis will be made available at:
628	https://github.com/rcrichmond/cholesterol_lowering_headandneckcancer
629	
630	Acknowledgements
631	
632	M.G. was a National Institute for Health Research (NIHR) academic clinical fellow and
633	is currently supported by a Wellcome Trust GW4-Clinical Academic Training PhD Fellowship.
634	This research was funded in part, by the Wellcome Trust [Grant number 220530/Z/20/Z].

635 For the purpose of open access, the author has applied a CC BY public copyright licence to

636 any Author Accepted Manuscript version arising from this submission. UK Biobank approval 637 was given for this project (ID 40644 "Investigating aetiology, associations and causality in 638 diseases of the head and neck") and UK Biobank GWAS data was also accessed under the 639 application (ID 15825 "MR-Base: an online resource for Mendelian randomization using 640 summary data"- Dr. Philip Haycock). R.C.R. is a de Pass VC research fellow at the University 641 of Bristol. J.T. is supported by an Academy of Medical Sciences (AMS) Springboard award, 642 which is supported by the AMS, the Wellcome Trust, Global Challenges Research Fund 643 (GCRF), the Government Department of Business, Energy and Industrial strategy, the British 644 Heart Foundation and Diabetes UK (SBF004\1079). R.M.M. was supported by a Cancer 645 Research UK (C18281/A19169) programme grant (the Integrative Cancer Epidemiology 646 Programme) and is part of the Medical Research Council Integrative Epidemiology Unit at 647 the University of Bristol supported by the Medical Research Council (MC UU 12013/1, 648 MC UU 12013/2, and MC UU 12013/3) and the University of Bristol. JY is supported by a 649 Cancer Research UK Population Research Postdoctoral Fellowship (C68933/A28534). R.M.M. 650 and A.R.N. are supported by the National Institute for Health Research (NIHR) Bristol 651 Biomedical Research Centre which is funded by the National Institute for Health Research 652 (NIHR) and is a partnership between University Hospitals Bristol and Weston NHS 653 Foundation Trust and the University of Bristol. This publication presents data from the Head and Neck 5000 study. The study was a component of independent research funded by the 654 655 NIHR under its Programme Grants for Applied Research scheme (RP-PG-0707-10034). The 656 views expressed in this publication are those of the author(s) and not necessarily those of 657 the NHS, the NIHR or the Department of Health. Core funding was also provided through 658 awards from Above and Beyond, University Hospitals Bristol Research Capability Funding 659 and the NIHR Senior Investigator award to A.R.N. Human papillomavirus (HPV) serology was

660 supported by a Cancer Research UK Programme Grant, the Integrative Cancer Epidemiology 661 Programme (grant number: C18281/A19169). B.D. and the University of Pittsburgh head and 662 neck cancer case-control study are supported by US National Institutes of Health (NIH) 663 grants: P50 CA097190, P30 CA047904 and R01 DE025712. The genotyping of the HNSCC 664 cases and controls was performed at the Center for Inherited Disease Research (CIDR) and 665 funded by the US National Institute of Dental and Craniofacial Research (NIDCR; 1X01HG007780-0). The University of North Caroline (UNC) CHANCE study was supported in 666 667 part by the National Cancer Institute (R01-CA90731). E.E.V and D.L are supported by 668 Diabetes UK (17/0005587). E.E.V is also supported by the World Cancer Research Fund 669 (WCRF UK), as part of the World Cancer Research Fund International grant programme 670 (IIG 2019 2009). E.H.T and P.S. were supported by the GENCAPO/FAPESP grant (10/51168-671 0).

672

673 Competing Interests

The authors declare no conflicts of interest. The funders had no role in the design of the study; the collection, analysis, and interpretation of the data; the writing of the manuscript; and the decision to submit the manuscript for publication. The authors alone are responsible for the views expressed in this article.

678

679 **References**

- 681 1. Warnakulasuriya S. Global epidemiology of oral and oropharyngeal cancer. Oral
- 682 Oncol. 2009;45(4-5):309-16.
- 683 2. Saba NF, Goodman M, Ward K, Flowers C, Ramalingam S, Owonikoko T, et al. Gender
- and ethnic disparities in incidence and survival of squamous cell carcinoma of the oral
- tongue, base of tongue, and tonsils: a surveillance, epidemiology and end results program-
- 686 based analysis. Oncology. 2011;81(1):12-20.
- 687 3. Cancer Research UK (CRUK). Head and neck cancer statistics 2019 [cited 2019
- 688 11/04/2019]. Available from: https://www.cancerresearchuk.org/health-
- 689 professional/cancer-statistics/statistics-by-cancer-type/head-and-neck-cancers#heading-
- 690 <u>Two</u>.
- 691 4. Mehanna H, Kong A, Ahmed SK. Recurrent head and neck cancer: United Kingdom

692 National Multidisciplinary Guidelines. J Laryngol Otol. 2016;130(S2):S181-S90.

- 5. De Boer MF, McCormick LK, Pruyn JF, Ryckman RM, van den Borne BW. Physical and
- 694 psychosocial correlates of head and neck cancer: a review of the literature. Otolaryngol
- 695 Head Neck Surg. 1999;120(3):427-36.
- 696 6. Thomas SJ, Penfold CM, Waylen A, Ness AR. The changing aetiology of head and neck
- 697 squamous cell cancer: A tale of three cancers? Clin Otolaryngol. 2018;43(4):999-1003.
- 698 7. Ang KK, Harris J, Wheeler R, Weber R, Rosenthal DI, Nguyen-Tan PF, et al. Human
- 699 papillomavirus and survival of patients with oropharyngeal cancer. N Engl J Med.
- 700 2010;363(1):24-35.

- 8. Elrefaey S, Massaro MA, Chiocca S, Chiesa F, Ansarin M. HPV in oropharyngeal
- cancer: the basics to know in clinical practice. Acta Otorhinolaryngol Ital. 2014;34(5):299-
- 703 309.
- 9. Sleire L, Forde HE, Netland IA, Leiss L, Skeie BS, Enger PO. Drug repurposing in
- 705 cancer. Pharmacol Res. 2017;124:74-91.
- 10. Hu M, Cheung BM, Tomlinson B. Safety of statins: an update. Ther Adv Drug Saf.
 2012;3(3):133-44.
- 708 11. Audi S, Burrage DR, Lonsdale DO, Pontefract S, Coleman JJ, Hitchings AW, et al. The
- 709 'top 100' drugs and classes in England: an updated 'starter formulary' for trainee
- 710 prescribers. British Journal of Clinical Pharmacology. 2018;84(11):2562-71.
- 711 12. Istvan ES, Deisenhofer J. Structural mechanism for statin inhibition of HMG-CoA
- 712 reductase. Science. 2001;292(5519):1160-4.
- 713 13. Ference BA, Kastelein JJP, Ginsberg HN, Chapman MJ, Nicholls SJ, Ray KK, et al.
- 714 Association of Genetic Variants Related to CETP Inhibitors and Statins With Lipoprotein
- T15 Levels and Cardiovascular Risk. JAMA. 2017;318(10):947-56.
- 716 14. Nielsen SF, Nordestgaard BG, Bojesen SE. Statin Use and Reduced Cancer-Related
- 717 Mortality. New Engl J Med. 2012;367(19):1792-802.
- 718 15. Mehibel M, Ortiz-Martinez F, Voelxen N, Boyers A, Chadwick A, Telfer BA, et al.
- 719 Statin-induced metabolic reprogramming in head and neck cancer: a biomarker for targeting
- 720 monocarboxylate transporters. Sci Rep. 2018;8(1):16804.

- 16. Chawda JG, Jain SS, Patel HR, Chaduvula N, Patel K. The relationship between serum
- 722 lipid levels and the risk of oral cancer. Indian J Med Paediatr Oncol. 2011;32(1):34-7.
- 17. Mehta R, Gurudath S, Dayansoor S, Pai A, Ganapathy KS. Serum lipid profile in
- patients with oral cancer and oral precancerous conditions. Dent Res J (Isfahan).

725 2014;11(3):345-50.

- 18. Kao LT, Hung SH, Kao PF, Liu JC, Lin HC. Inverse association between statin use and
- 727 head and neck cancer: Population-based case-control study in Han population. Head Neck-J
- 728 Sci Spec. 2019;41(5):1193-8.
- 19. Dickerman BA, Garcia-Albeniz X, Logan RW, Denaxas S, Hernan MA. Avoidable flaws
- in observational analyses: an application to statins and cancer. Nat Med. 2019;25(10):1601-
- 731 6.
- 732 20. Smith GD, Ebrahim S. 'Mendelian randomization': can genetic epidemiology
- 733 contribute to understanding environmental determinants of disease? Int J Epidemiol.
- 734 2003;32(1):1-22.
- 735 21. Davey Smith G, Hemani G. Mendelian randomization: genetic anchors for causal
- inference in epidemiological studies. Hum Mol Genet. 2014;23(R1):R89-98.
- 737 22. Davies NM, Holmes MV, Davey Smith G. Reading Mendelian randomisation studies: a
- guide, glossary, and checklist for clinicians. BMJ. 2018;362:k601.
- 739 23. Walker VM, Smith GD, Davies NM, Martin RM. Mendelian randomization: a novel
- 740 approach for the prediction of adverse drug events and drug repurposing opportunities. Int J
- 741 Epidemiol. 2017;46(6):2078-89.

742 24. Teslovich TM, Musunuru K, Smith AV, Edmondson AC, Stylianou IM, Koseki M, et al.

- 743 Biological, clinical and population relevance of 95 loci for blood lipids. Nature.
- 744 2010;466(7307):707-13.
- 745 25. Willer CJ, Schmidt EM, Sengupta S, Peloso GM, Gustafsson S, Kanoni S, et al.
- 746 Discovery and refinement of loci associated with lipid levels. Nat Genet. 2013;45(11):1274-+.
- 747 26. Ference BA, Ginsberg HN, Graham I, Ray KK, Packard CJ, Bruckert E, et al. Low-
- 748 density lipoproteins cause atherosclerotic cardiovascular disease. 1. Evidence from genetic,
- 749 epidemiologic, and clinical studies. A consensus statement from the European
- 750 Atherosclerosis Society Consensus Panel. Eur Heart J. 2017;38(32):2459-72.
- 751 27. Ference BA, Majeed F, Penumetcha R, Flack JM, Brook RD. Effect of naturally random
- allocation to lower low-density lipoprotein cholesterol on the risk of coronary heart disease
- 753 mediated by polymorphisms in NPC1L1, HMGCR, or both: a 2 x 2 factorial Mendelian
- randomization study. J Am Coll Cardiol. 2015;65(15):1552-61.
- 755 28. Ference BA, Robinson JG, Brook RD, Catapano AL, Chapman MJ, Neff DR, et al.
- 756 Variation in PCSK9 and HMGCR and Risk of Cardiovascular Disease and Diabetes. N Engl J
- 757 Med. 2016;375(22):2144-53.
- Carter P, Vithayathil M, Kar S, Potluri R, Mason AM, Larsson SC, et al. Predicting the
 effect of statins on cancer risk using genetic variants: a Mendelian randomization study in
 UK Biobank. medRxiv. 2020:2020.02.28.20028902.

761 30. Yarmolinsky J, Bull CJ, Vincent EE, Robinson J, Walther A, Smith GD, et al. Association

- 762 Between Genetically Proxied Inhibition of HMG-CoA Reductase and Epithelial Ovarian
- 763 Cancer. JAMA. 2020;323(7):646-55.
- 764 31. Machiela MJ, Chanock SJ. LDlink: a web-based application for exploring population-
- 765 specific haplotype structure and linking correlated alleles of possible functional variants.
- 766 Bioinformatics. 2015;31(21):3555-7.
- 767 32. Ference BA, Ray KK, Catapano AL, Ference TB, Burgess S, Neff DR, et al. Mendelian
- 768 Randomization Study of ACLY and Cardiovascular Disease. N Engl J Med. 2019;380(11):1033-

769 42.

- 770 33. Kettunen J, Demirkan A, Wurtz P, Draisma HH, Haller T, Rawal R, et al. Genome-wide
- study for circulating metabolites identifies 62 loci and reveals novel systemic effects of LPA.
- 772 Nat Commun. 2016;7:11122.
- 773 34. Lesseur C, Diergaarde B, Olshan AF, Wunsch V, Ness AR, Liu G, et al. Genome-wide
- association analyses identify new susceptibility loci for oral cavity and pharyngeal cancer.
- 775 Nat Genet. 2016;48(12):1544-50.
- 776 35. Consortium O. Consortium launches genotyping effort. Cancer Discov.
- 777 2013;3(12):1321-2.
- 778 36. Dudding T, Johansson M, Thomas SJ, Brennan P, Martin RM, Timpson NJ. Assessing
- the causal association between 25-hydroxyvitamin D and the risk of oral and oropharyngeal
- 780 cancer using Mendelian randomization. Int J Cancer. 2018;143(5):1029-36.

- 781 37. Brion MJ, Shakhbazov K, Visscher PM. Calculating statistical power in Mendelian
- randomization studies. Int J Epidemiol. 2013;42(5):1497-501.
- 783 38. Yavorska OO, Burgess S. MendelianRandomization: an R package for performing
- 784 Mendelian randomization analyses using summarized data. Int J Epidemiol.
- 785 2017;46(6):1734-9.
- 786 39. Burgess S, Scott RA, Timpson NJ, Davey Smith G, Thompson SG. Using published data
- 787 in Mendelian randomization: a blueprint for efficient identification of causal risk factors. Eur
- 788 J Epidemiol. 2015;30(7):543-52.
- 789 40. Burgess S, Dudbridge F, Thompson SG. Combining information on multiple
- instrumental variables in Mendelian randomization: comparison of allele score and

791 summarized data methods. Stat Med. 2016;35(11):1880-906.

- 792 41. Nikpay M, Goel A, Won HH, Hall LM, Willenborg C, Kanoni S, et al. A comprehensive
- 793 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease.
- 794 Nat Genet. 2015;47(10):1121-30.
- 42. Hemani G, Bowden J, Smith GD. Evaluating the potential role of pleiotropy in
- 796 Mendelian randomization studies. Hum Mol Genet. 2018;27(R2):R195-R208.
- 43. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid
- instruments: effect estimation and bias detection through Egger regression. Int J Epidemiol.

799 2015;44(2):512-25.

800 44. Bowden J, Smith GD, Haycock PC, Burgess S. Consistent Estimation in Mendelian
801 Randomization with Some Invalid Instruments Using a Weighted Median Estimator. Genet
802 Epidemiol. 2016;40(4):304-14.

45. Hartwig FP, Smith GD, Bowden J. Robust inference in summary data Mendelian
randomization via the zero modal pleiotropy assumption. Int J Epidemiol. 2017;46(6):1985-

805 98.

806 46. Bowden J, Del Greco MF, Minelli C, Smith GD, Sheehan NA, Thompson JR. Assessing

807 the suitability of summary data for two-sample Mendelian randomization analyses using

808 MR-Egger regression: the role of the I-2 statistic. Int J Epidemiol. 2016;45(6):1961-74.

809 47. Bowden J, Spiller W, Del Greco F, Sheehan N, Thompson J, Minelli C, et al. Improving

810 the visualization, interpretation and analysis of two-sample summary data Mendelian

randomization via the Radial plot and Radial regression. Int J Epidemiol. 2018;47(4):1264-78.

812 48. Verbanck M, Chen CY, Neale B, Do R. Publisher Correction: Detection of widespread

813 horizontal pleiotropy in causal relationships inferred from Mendelian randomization

between complex traits and diseases. Nat Genet. 2018;50(8):1196.

815 49. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, et al. Genome-wide

816 genetic data on ~500,000 UK Biobank participants. bioRxiv. 2017:166298.

817 50. Mitchell R, Hemani, G, Dudding, T, Corbin, L, Harrison, S, Paternoster, L. UK Biobank

818 Genetic Data: MRC-IEU Quality Control, version 2.

819 <u>https://doi.org/10.5523/bris.1ovaau5sxunp2cv8rcy88688v</u>. 2019.

- 820 51. Mitchell R, Elsworth BL, Mitchell R, Raistrick CA, Paternoster L, Hemani G, Gaunt TR
- 821 MRC IEU UK Biobank GWAS pipeline version 2.
- 822 https://doi.org/10.5523/bris.pnoat8cxo0u52p6ynfaekeigi. 2019.
- 52. Schwarzer G. meta: an R package for metaanalysis. R News. 2007;7:40-5.
- 824 53. Higgins SG. Measuring inconsistency in metaanalyses. BMJ Br Med J. 2003;327:557-
- 825 60.
- 826 54. Safarova MS, Satterfield BA, Fan X, Austin EE, Ye Z, Bastarache L, et al. A phenome-
- 827 wide association study to discover pleiotropic effects of PCSK9, APOB, and LDLR. NPJ Genom
- 828 Med. 2019;4:3.
- 829 55. Stoekenbroek RM, Lambert G, Cariou B, Hovingh GK. Inhibiting PCSK9 biology
- 830 beyond LDL control. Nat Rev Endocrinol. 2018;15(1):52-62.
- 831 56. Lee S, Zhang C, Liu Z, Klevstig M, Mukhopadhyay B, Bergentall M, et al. Network
- analyses identify liver-specific targets for treating liver diseases. Mol Syst Biol.
- 833 2017;13(8):938.
- 834 57. Marimuthu A, Subbannayya Y, Sahasrabuddhe NA, Balakrishnan L, Syed N, Sekhar
- 835 NR, et al. SILAC-based quantitative proteomic analysis of gastric cancer secretome.
- 836 Proteomics Clin Appl. 2013;7(5-6):355-66.
- 837 58. Dong B, Singh AB, Fung C, Kan K, Liu J. CETP inhibitors downregulate hepatic LDL
- 838 receptor and PCSK9 expression in vitro and in vivo through a SREBP2 dependent
- 839 mechanism. Atherosclerosis. 2014;235(2):449-62.

- 840 59. Barter PJ, Cochran BJ, Rye KA. CETP inhibition, statins and diabetes. Atherosclerosis.
- 841 2018;278:143-6.
- 842 60. The Human Protein Atlas. PCKS9 2020 [Available from:
- 843 https://www.proteinatlas.org/ENSG00000169174-PCSK9/pathology.
- 844 61. The Human Protein Atlas. LDLR 2020 [Available from:
- 845 <u>https://www.proteinatlas.org/ENSG00000130164-LDLR/pathology</u>.
- 846 62. The Human Protein Atlas. LDLR Tissue 2020 [Available from:
- 847 https://www.proteinatlas.org/ENSG00000130164-LDLR/tissue.
- 848 63. Lee JS, Mukhopadhyay P, Matyas C, Trojnar E, Paloczi J, Yang YR, et al. PCSK9
- 849 inhibition as a novel therapeutic target for alcoholic liver disease. Scientific Reports.
- 850 2019;9(1):17167.
- 851 64. Lee JS, Rosoff D, Luo A, Longley M, Phillips M, Charlet K, et al. PCSK9 is Increased in
- 852 Cerebrospinal Fluid of Individuals With Alcohol Use Disorder. Alcoholism: Clinical and
- 853 Experimental Research. 2019;43(6):1163-9.
- 854 65. Wang K, Gerke TA, Chen X, Prosperi M. Association of statin use with risk of Gleason
- 855 score-specific prostate cancer: A hospital-based cohort study. Cancer Medicine.
- 856 2019;8(17):7399-407.
- 857 66. Poynter JN, Gruber SB, Higgins PD, Almog R, Bonner JD, Rennert HS, et al. Statins and
- the risk of colorectal cancer. N Engl J Med. 2005;352(21):2184-92.
- 859 67. Khurana V, Bejjanki HR, Caldito G, Owens MW. Statins reduce the risk of lung cancer
- in humans: a large case-control study of US veterans. Chest. 2007;131(5):1282-8.

- 861 68. Shannon J, Tewoderos S, Garzotto M, Beer TM, Derenick R, Palma A, et al. Statins
- and prostate cancer risk: a case-control study. Am J Epidemiol. 2005;162(4):318-25.
- 863 69. Cholesterol Treatment Trialists C, Emberson JR, Kearney PM, Blackwell L, Newman C,
- 864 Reith C, et al. Lack of effect of lowering LDL cholesterol on cancer: meta-analysis of
- 865 individual data from 175,000 people in 27 randomised trials of statin therapy. PLoS One.
- 866 2012;7(1):e29849.
- 867 70. Dale KM, Coleman CI, Henyan NN, Kluger J, White CM. Statins and cancer risk: a
- 868 meta-analysis. JAMA. 2006;295(1):74-80.
- 869 71. Hernan MA, Sauer BC, Hernandez-Diaz S, Platt R, Shrier I. Specifying a target trial
- 870 prevents immortal time bias and other self-inflicted injuries in observational analyses. J Clin
- 871 Epidemiol. 2016;79:70-5.
- 872 72. Saka Herran C, Jane-Salas E, Estrugo Devesa A, Lopez-Lopez J. Protective effects of
- 873 metformin, statins and anti-inflammatory drugs on head and neck cancer: A systematic
- 874 review. Oral Oncol. 2018;85:68-81.
- 875 73. Dimitroulakos J, Marhin WH, Tokunaga J, Irish J, Gullane P, Penn LZ, et al. Microarray
- and biochemical analysis of lovastatin-induced apoptosis of squamous cell carcinomas.
- 877 Neoplasia. 2002;4(4):337-46.
- 878 74. Osmak M. Statins and cancer: current and future prospects. Cancer Lett.
- 879 2012;324(1):1-12.
- 880 75. Gallagher EJ, LeRoith D. Obesity and Diabetes: The Increased Risk of Cancer and
 881 Cancer-Related Mortality. Physiol Rev. 2015;95(3):727-48.

882	76.	Joost HG. Diabetes and cancer: Epidemiology and potential mechanisms. Diabetes		
883	Vasc Dis Re. 2014;11(6):390-4.			
884	77.	Hursting SD, DiGiovanni J, Dannenberg AJ, Azrad M, LeRoith D, Demark-Wahnefried		
885	W, et a	al. Obesity, Energy Balance, and Cancer: New Opportunities for Prevention. Cancer		
886	Prev Res. 2012;5(11):1260-72.			
887	78.	Bhat M, Skill N, Marcus V, Deschenes M, Tan X, Bouteaud J, et al. Decreased PCSK9		
888	expres	sion in human hepatocellular carcinoma. BMC Gastroenterol. 2015;15:176.		
889				
890				
891				
892				
893	Fig 1. F	Forest plot showing the causal effects of HMGCR, NPC1L1, CETP, PCSK9 and LDLR		
894	variant	ts on the oral and oropharyngeal cancer subsites in GAME-ON.		
895				
896				
	HMG	Cancer ∳ OPC+OC ∲ OPC ∳ OC Cancer ∲ OPC+OC ∲ OPC ∳ OC OCR NPC1L1		
Inverse variance	e weighted -			



- 912 Effect estimates on oral and oropharyngeal cancer are reported on the log odds scale.



Effect estimates on oral and oropharyngeal cancer are reported on the log odds scale.

Fig 2. Forest plots showing the causal effects of cholesterol-lowering PCSK9 and LDLR single

nucleotide polymorphisms on coronary heart disease and combined oral/ oropharyngeal

cancer in GAME-ON.



Effect estimates on oral and oropharyngeal cancer are reported on the log odds scale.



Fig 3. Scatter plots for LDLR and PCSK9 single nucleotide polymorphisms effect on combined oral/oropharyngeal cancer in GAME-ON.

Fig 4. Forest plots showing inverse variance weighted meta-analysis effects of cholesterol-

lowering PCSK9 single nucleotide polymorphisms on head and neck cancer subsites .

Study	Odds Ratio	OR	95%-CI
	All sites		
GAMEON UKBiobank Fixed effect model Random effects model Heterogeneity: $I^2 = 22\%$, $\tau^2 = 0.035$ Heterogeneity: $I^2 = 0\%$, $\tau^2 = 0$, $p = 0.01$	55. p = 0.26 0.78 0.1 1 10 100	2.05 1.17 1.77 1.71	[1.24; 3.38] [0.50; 2.69] [1.15; 2.71] [1.02; 2.87]
	Oropharyngeal		
GAMEON UKBiobank Fixed effect model Random effects model Heterogeneity: $I^2 = 41\%$, $\tau^2 = 0.160$ Heterogeneity: $I^2 = 0\%$, $\tau^2 = 0$, $\rho = 0.001$	06. p = 0.19 0.70 ¹ 0.1 1 10 100	3.49 1.43 2.57 2.43	[1.58; 7.68] [0.48; 4.26] [1.35; 4.87] [1.03; 5.72]
	Oral		
GAMEON UKBiobank Fixed effect model Random effects model	¢	1.62 0.70 1.40 1.30	[0.89; 2.94] [0.19; 2.53] [0.81; 2.40] [0.64; 2.68]

Heterogeneity: $l^2 = 26\%$, $\tau^2 = 0.0891$, p = 0.25Heterogeneity: $l^2 = 0\%$, $\tau^2 = 0$, p = 0.670.01 0.1 1 10 100

Fig 5. Forest plots showing inverse variance weighted meta-analysis effects of cholesterol-lowering

LDLR single nucleotide polymorphisms on head and neck cancer subsites

Study	Odds Ratio	OR	95%-CI
	All sites		
GAMEON UKBiobank Fixed effect model Random effects model Heterogeneity: $l^2 = 0\%$, $\tau^2 = 0$, μ Heterogeneity: $l^2 = 0\%$, $\tau^2 = 0$, μ	0.5 1 2	0.65 0.91 0.71 0.71	[0.42; 1.00] [0.44; 1.90] [0.49; 1.03] [0.49; 1.03]

Oropharyngeal



Oral

GAMEON		0.53	[0.32; 0.87]
UKBiobank		1.55	[0.50; 4.78]
Fixed effect model	\Leftrightarrow	0.63	[0.40; 0.99]
Random effects model	-	0.80	[0.29; 2.23]
Heterogeneity: $J^2 = 66\%$, $\tau^2 = 0.37$	797, p = 0.09		
Heterogeneity: $I^2 = 24\%$, $\tau^2 = 0.05$	560, p = 0.22		
0.	1 0.5 1 2 10		