Haycock, P. C., Burgess, S., Nounu, A., Zheng, J., Okoli, G. N., Bowden, J., ... Telomeres Mendelian Randomization Collaboration (2017). Association Between Telomere Length and Risk of Cancer and Non-Neoplastic Diseases: A Mendelian Randomization Study. *JAMA Oncology*. https://doi.org/10.1001/jamaoncol.2016.5945

Mendelian randomization study of the association between telomere length and risk of cancer and non-neoplastic diseases

The Telomeres Mendelian Randomization Collaboration

Correspondence: Philip Haycock

MRC Integrative Epidemiology Unit

University of Bristol

Bristol

UK

philip.haycock@bristol.ac.uk

Tel: +44 1173 310 088

2995 words [word limit 3000]

3 figures, 2 tables, 132 references; 7 supplementary figures / 6 supplementary tables

ABSTRACT 349 WORDS

Importance The causal direction and magnitude of the association between telomere length and incidence of cancer and non-neoplastic diseases is uncertain, due to the susceptibility of observational studies to confounding and reverse causation.

Objective To conduct a Mendelian randomization study, using germline genetic variants as instrumental variables, to appraise the causal relevance of telomere length for risk of cancer and non-neoplastic diseases.

Data Sources Genome-wide association studies (GWAS) published up to January 15 2015.

Study Selection GWAS of non-communicable diseases that assayed germline genetic variation and did not select cohort or control participants on the basis of pre-existing diseases. Of 163 GWAS of non-communicable diseases identified, summary data from 103 were available.

Data Extraction Summary association statistics for single nucleotide polymorphisms (SNPs) that are strongly associated with telomere length in the general population.

Main Outcomes Odds ratios (ORs) for disease per standard deviation (SD) higher telomere length due to germline genetic variation.

Results Summary data were available for 35 cancers and 48 non-neoplastic diseases, corresponding to 420,081 cases (median 2,526 per disease) and 1,093,105 controls (median 6,789 per disease). Increased telomere length due to germline genetic variation was generally associated with increased risk for site-specific cancers. The strongest associations were observed for (ORs per 1-SD change in genetically increased telomere length): glioma 5.27 (3.15-8.81), serous low-malignant-potential ovarian cancer 4.35 (2.39-7.94), lung adenocarcinoma 3.19 (2.40-4.22), neuroblastoma 2.98 (1.92-4.62), bladder cancer 2.19 (1.32-3.66), melanoma 1.87 (1.55-2.26), testicular cancer 1.76 (1.02-3.04), kidney cancer 1.55

(1.08-2.23) and endometrial cancer 1.31 (1.07-1.61). Associations were stronger for rarer cancers and at tissue sites with lower rates of stem cell division ($P<0.05$). There was generally little evidence of association between genetically increased telomere length and risk of psychiatric, autoimmune, inflammatory, diabetic and other non-neoplastic diseases, except for coronary heart disease (0.78 [0.67-0.90]), abdominal aortic aneurysm (0.63 [0.49-0.81]), celiac disease (0.42 [0.28-0.61]) and interstitial lung disease (0.09 [0.05- 0.15]).

Conclusions It is likely that longer telomeres increase risk for several cancers but reduce risk for some non-neoplastic diseases, including cardiovascular diseases.

INTRODUCTION

At the ends of chromosomes, telomeres are DNA-protein structures that protect the genome from damage, shorten progressively over time in most somatic tissues[1] and are proposed physiological markers of ageing.[2,3] Shorter leukocyte telomeres are correlated with older age, male sex and other known risk factors for non-communicable diseases[4–6] and are generally associated with higher risk for cardiovascular diseases[7,8], type 2 diabetes[9] and non-vascular non-neoplastic causes of mortality.[8] Whether these associations are causal, however, is unknown. Telomere length has also been implicated in risk of cancer but the direction and magnitude of the association is uncertain and contradictory across observational studies.[10–14] The uncertainty reflects the considerable difficulty of designing observational studies of telomere length and cancer incidence that are robust to reverse causation, confounding and measurement error.

The aim of the present report was to conduct a Mendelian randomization study, using germline genetic variants as instrumental variables for telomere length, to help clarify the nature of the association between telomere length and risk of cancer and non-neoplastic diseases. The approach, which mimics the random allocation of individuals to the placebo and intervention arms of a randomized controlled trial, allowed us to: (1) estimate the direction and broad magnitude of the association of telomere length with risk of multiple cancer and non-neoplastic diseases; (2) appraise the evidence for causality in the estimated etiological associations; (3) investigate potential sources of heterogeneity in findings for site-specific cancers; and (4) compare genetic estimates to findings based on directly measured telomere length in prospective observational studies.

METHODS

Study design

The design of our study, illustrated in Figure S1, had three key components: 1) the identification of genetic variants to serve as instruments for telomere length; 2) the acquisition of summary data for the genetic instruments from genome wide association studies (GWASs) of diseases and risk factors for non-communicable diseases; and 3) the classification of diseases and risk factors into primary or secondary outcomes based on *a priori* statistical power. As a first step, we searched the GWAS catalog[15,16] on the 15 January 2015, to identify single nucleotide polymorphisms (SNPs) associated with telomere length. To supplement the list with additional potential instruments, we also searched the original study reports curated by the GWAS catalog (using a P-value threshold of $5x10^{-8}$).[17–25] We acquired summary data for all SNPs identified by our search from a meta-analysis of GWASs of telomere length, involving 9,190 participants of European ancestry.[18]

The second key component of our design strategy involved the acquisition of summary data, corresponding to the selected genetic instruments for telomere length, from GWASs of non-communicable diseases and risk factors (Fig. S1). As part of this step, we invited principal investigators of non-communicable disease studies curated by the GWAS catalog[15,26] to share summary data for our study. We also downloaded summary data for diseases and risk factors from publically available sources, including study-specific websites, dbGAP, ImmunoBase and the GWAS catalog (Fig. S1).

The third key component of our design strategy was the classification of diseases and risk factors into either primary or secondary outcomes, which we defined on the basis of *a priori* statistical power to detect associations with telomere length. Primary outcomes were defined

as diseases with sufficient cases and controls for >50% statistical power and secondary outcomes defined as diseases with <50% statistical power to detect odds ratios ≥2.0 per standard deviation (SD) change in genetically increased telomere length (alpha assumed to be 0.01). All risk factors were defined as secondary outcomes. Risk factors with <50% statistical power were excluded.

Further details on our design strategy can be found in the supplement.

Comparison with prospective observational studies

We searched PubMed for prospective observational studies of the association between telomere length and disease (see Tables S3 and S4 for details of the search strategy and inclusion criteria). Study-specific relative risks for disease per unit change or quantile comparison of telomere length were transformed to a SD scale using previously described methods.[27] Hazard ratios, risk ratios and odds ratios were assumed to approximate the same measure of relative risk. Where multiple independent studies of the same disease were identified, these were combined by fixed effects meta-analysis, unless there was strong evidence of between-study heterogeneity ($P_{Cochran's\ Q}<0.001$), in which case they were kept separate.

Statistical analysis

We combined summary data across SNPs into a single instrument, using maximum likelihood to estimate the slope of the relationship between β_{GD} and β_{GP} and a variance-covariance matrix to make allowance for linkage disequilibrium between SNPs,[28] where β_{GD} is the change in disease log odds or risk factor levels per copy of the effect allele and β_{GP} is the SD change in telomere length per copy of the effect allele (see supplementary methods

144 for technical details). The slope from this approach can be interpreted as the log odds ratio for
145 binary outcomes, or the unit change for continuous risk factors, per SD change in genetically
146 increased telomere length. P-values for heterogeneity amongst SNPs, in the estimated
147 associations of genetically increased telomere length with disease and risk factors, were
148 estimated by likelihood ratio tests.[28] Associations between genetically increased telomere
149 length and continuous risk factors were transformed into SD units. For five secondary disease
150 outcomes where only a single SNP was available for analysis, we estimated associations
151 using the Wald ratio: $\beta_{GD}/\beta_{GP,}$ with standard errors approximated by the delta method.[29]

152 Inference of causality in the estimated etiological associations between telomere length and
153 disease depends on satisfaction of Mendelian randomization assumptions (Fig. S7; see Table
154 S6 for a glossary of terms).[30,31] The assumptions are: 1) the selected SNPs are associated with
155 telomere length; 2) the selected SNPs are not associated with confounders; and 3) the selected
156 SNPs are associated with disease exclusively through their effect on telomere length. If these
157 assumptions are satisfied, the selected SNPs are valid instrumental variables and their
158 association with disease can be interpreted as a causal effect of telomere length. We modeled
159 the impact of violations of these assumptions through two sets of sensitivity analyses: a
160 weighted median function[32] and MR-Egger regression[30] (see supplementary methods for
161 technical details). We restricted our sensitivity analyses to diseases showing the strongest
162 evidence of association with genetically increased telomere length (defined as
163 $P_{Bonferroni} \leq 0.05$).
164
165 We used meta-regression to appraise potential sources of heterogeneity in our findings for
166 cancer. The association of genetically increased telomere length with the log odds of cancer
167 was regressed on cancer incidence, survival time and median age-at-diagnosis, downloaded
168 from the National Cancer Institute Surveillance, Epidemiology, and End Results (SEER)

Program,[33] and tissue-specific rates of stem cell division from Tomasetti and Vogelstein.[34] As the downloaded cancer characteristics from SEER correspond to the United States population, 77% of which was of white ancestry in 2015[35], the meta-regression analyses excluded genetic studies conducted in East Asian populations.

All analyses were performed in R version 3.1.2[36] and Stata release 13.1 (StataCorp, College Station, TX). P-values were two-sided and evidence of association was declared at $P<0.05$. Where indicated, Bonferroni corrections were used to make allowance for multiple testing, although this is likely to be overly conservative given the non-independence of many of the outcomes tested.

RESULTS

We selected 16 SNPs as instruments for telomere length (Fig. S1 & Table 1). The selected SNPs correspond to 10 independent genomic regions that collectively account for 2-3% of the variance in leukocyte telomere length, which is equivalent to an F statistic of ~18. This indicates that the genetic instrument, constructed from these 10 independent genomic regions, is strongly associated with telomere length (details in supplementary discussion).[37] Summary data for the genetic instruments were available for 83 non-communicable diseases, corresponding to 420,081 cases (median 2,526 per disease) and 1,093,105 controls (median 6,789 per disease), and 44 risk factors (Fig. S1, Table 2 and Table S1). The median number of SNPs available across diseases was 11 (min=1, max=13) and across risk factors was 12 (min=11, max=13). Of the 83 diseases, 56 were classified as primary outcomes and 27 as secondary outcomes (Table 2, Fig. S1 and Table S1). For 9 of the 83 non-communicable diseases, additional summary data were available from 10 independent studies for replication

194 analyses, corresponding to 40,465 cases (median 1,416 per disease) and 52,306 controls
195 (median 3,537 per disease) (Table S1).

196 The results from primary analyses of non-communicable diseases are presented in Figure 1;
197 results from secondary analyses of risk factors and diseases with low *a priori* power are
198 presented in the supplement (Fig. S2, S5 and S6). Genetically increased telomere length was
199 associated with higher odds of disease for 9 of 22 primary cancers ($P<0.05$), including (odds
200 ratio [95% confidence interval]): glioma (5.27 [3.15-8.81]), endometrial cancer (1.31 [1.07-
201 1.61]), kidney cancer (1.55 [1.08-2.23]), testicular germ cell cancer (1.76 [1.02-3.04]),
202 melanoma (1.87 [1.55-2.26]), bladder cancer (2.19 [1.32-3.66]), neuroblastoma (2.98 [1.92-
203 4.62]), lung adenocarcinoma (3.19 [2.40-4.22]) and serous low-malignancy-potential (LMP)
204 ovarian cancer (4.35 [2.39-7.94]) (Fig. 1). The associations were, however, highly variable
205 across cancer types, varying from an odds ratio of 0.86 (0.50-1.48) for head and neck cancer
206 to 5.27 (3.15-8.81) for glioma. Substantial variability was also observed within tissue sites.
207 For example, the odds ratio for lung adenocarcinoma was 3.19 (2.40-4.22) compared to 1.07
208 (0.82-1.39) for squamous cell lung cancer. For serous LMP ovarian cancer the odds ratio was
209 4.35 (2.39-7.94) compared to odds ratios of 1.21 (0.87-1.68) for endometrioid ovarian cancer,
210 1.12 (0.94-1.34) for serous invasive ovarian cancer, 1.04 (0.66-1.63) for clear cell ovarian
211 cancer and 1.04 (0.73-1.47) for mucinous ovarian cancer. The strongest evidence of
212 association was observed for glioma, lung adenocarcinoma, neuroblastoma and serous LMP
213 ovarian cancer ($P_{Bonferroni}<0.05$). Results for glioma and bladder cancer showed evidence for
214 replication in independent datasets (independent datasets were not available for other
215 cancers) (Fig. S3).

216 Genetically increased telomere length was associated with reduced odds of disease for 6 of 32
217 primary non-neoplastic diseases ($P<0.05$), including coronary heart disease (0.78 [0.67-0.9]),
218 abdominal aortic aneurysm (0.63 [0.49-0.81]), Alzheimer's disease (0.84 [0.71-0.98]), celiac

disease (0.42 [0.28-0.61]), interstitial lung disease (0.09 [0.05-0.15]) and type 1 diabetes (0.71 [0.51-0.98]) ($P<0.05$) (Figure 1). The strongest evidence of association was observed for coronary heart disease ($P_{Bonferroni}=0.05$) and abdominal aortic aneurysm, celiac disease and interstitial lung disease ($P_{Bonferroni}<0.05$). The associations with coronary heart disease and interstitial lung disease showed evidence for replication in independent datasets (Fig. S3).

Our genetic findings were generally similar in direction and magnitude to estimates based on observational prospective studies of leukocyte telomere length and disease (Figure 3). Our genetic estimates for lung adenocarcinoma, melanoma, kidney cancer and glioma, were, however, stronger in comparison to observational estimates.

In sensitivity analyses, we appraised the potential impact of confounding by pleiotropic pathways on our results. Associations estimated by the weighted median and MR-Egger were broadly similar to the main results for glioma, lung adenocarcinoma, serous LMP ovarian cancer, neuroblastoma, abdominal aortic aneurysm, coronary heart disease and interstitial lung disease (Fig. S4). In the second set of sensitivity analyses, implemented by MR-Egger regression, we found little evidence for the presence of pleiotropy ($P_{intercept}\geq0.27$) (Fig. S4). The MR-Egger analyses were, however, generally underpowered, as reflected by the wide confidence intervals in the estimated odds ratios.

In meta-regression analyses, we observed that genetically increased telomere length tended to be more strongly associated with rarer cancers ($P=0.02$) and cancers at tissue-sites with lower rates of stem cell division ($P=0.02$) (Figure 2). The associations showed little evidence of varying by percentage survival five years after diagnosis or median age-at-diagnosis ($P\geq37$).

DISCUSSION

In this report we show that genetically increased telomere length is associated with increased risk of several cancers and with reduced risk of some non-neoplastic diseases. Given the random distribution of genotypes in the general population with respect to lifestyle and other environmental factors, as well as the fixed nature of germline genotypes, these results should be less susceptible to confounding and reverse causation in comparison to observational studies. Our results could, however, reflect violations of Mendelian randomization assumptions, such as confounding by pleiotropy, population stratification or ancestry.[38] Although we cannot entirely rule out this possibility, the majority of our results persisted in sensitivity analyses that made allowance for violations of Mendelian randomization assumptions. Confounding by population stratification or ancestry is also unlikely, given the adjustments made for ancestry in the original disease GWASs (see supplementary discussion). Our results are therefore compatible with causality.

Comparison with previous studies

Our findings for cancer are generally contradictory to those based on retrospective studies, which tend to report increased risk for cancer in individuals with shorter telomeres.[11,12,39–42] The contradictory findings may reflect reverse causation in the retrospective studies, whereby shorter telomeres arise as a result of disease, or of confounding effects, e.g. due to cases being slightly older than controls even in age-matched analyses. Our findings for cancer are generally more consistent with those based on prospective observational studies, which tend to report weak or null associations of longer leukocyte telomeres with overall and site-specific risk of cancer,[10–13,41,43–62] with some exceptions.[63] Our results are also similar to previously reported Mendelian randomization studies of telomere length and risk of

melanoma, lung cancer, chronic lymphocytic leukemia and glioma.[64–67] The shape of the association with cancer may not, however, be linear over the entire telomere length distribution. For example, individuals with dyskeratosis congenita, a disease caused by germline loss-of-function mutations in the telomerase component genes *TERC* and *TERT,* have chronically short telomeres and are at increased risk of some cancers, particularly acute myeloid leukemia and squamous cell carcinomas arising at sites of leukoplakia,[68,69] presumably due to increased susceptibility to genome instability and chromosomal end-to-end fusions.[70] Our results should therefore be interpreted as reflecting the average association at the population level and may not be generalizable to the extreme ends of the telomere length distribution.

Mechanisms of association

Our cancer findings are compatible with known biology.[70] By limiting the proliferative potential of cells, telomere shortening may serve as a tumour suppressor; and individuals with longer telomeres may be more likely to acquire somatic mutations owing to increased proliferative potential.[70] Rates of cell division are, however, highly variable amongst tissues[34] and thus the relative gain in cell proliferative potential, conferred by having longer telomeres, may also be highly variable across tissues. This could explain the ~6-fold variation in odds ratios observed across cancer types in the present study, as well as the tendency of our results to be stronger at tissue sites with lower rates of stem cell division. For example, the association was strongest for glioma (OR=5.27) and comparatively weak for colorectal cancer (OR=1.09) and the rates of stem cell division in the tissues giving rise to these cancers differ by several orders of magnitude. In neural stem cells, which give rise to gliomas, the number of divisions is ~270 million and for colorectal stem cells is ~1.2 trillion over the average lifetime of an individual.[34] The observation that genetically increased telomere

length was more strongly associated with rarer cancers potentially reflects the same mechanism, since rarer cancers also tend to show lower rates of stem cell division.[34] For example, the incidence of glioma is 0.4 and for colorectal cancer is 42.4 per 100,000 per year in the United States.[33]

The inverse associations observed for some non-neoplastic diseases may reflect the impact of telomere shortening on tissue degeneration and an evolutionary trade-off for greater resistance to cancer at the cost of greater susceptibility to degenerative diseases, particularly cardiovascular diseases.[71,72]

Study limitations

Our study is subject to some limitations, in addition to the Mendelian randomization assumptions already considered above. First, our method assumes that the magnitude of the association between SNPs and telomere length is consistent across tissues. Second, our study assumed a linear shape of association between telomere length and disease risk, whereas the shape could be "J" or "U" shaped.[44,57,68] Third, our results assume that the samples used to define the genetic instrument for telomere length[18] and the various samples used to estimate the SNP-disease associations are representative of the same general population, practically defined as being of similar ethnicity, age and sex distribution.[73] This assumption would, for example, not apply in the case of the SNP-disease associations derived from East Asian or pediatric populations. Generally speaking, violation of the aforementioned assumptions could bias the magnitude of the association between genetically increased telomere length and disease; but would be unlikely to increase the likelihood of false positives (i.e. incorrectly inferring an association when none exists).[74] Our results should therefore remain informative for the direction and broad magnitude of the average association at the population level, even

in the presence of such violations. Fourth, we cannot rule out chance in explaining some of the weaker findings. Fifth, our results may not be fully representative of non-communicable diseases (since not all studies shared data and our analyses were underpowered for the secondary disease outcomes). The diseases represented in our primary analyses probably account for >60% of all causes of death in American adults.[75]

Clinical relevance of findings

Our findings suggest that potential clinical applications of telomere length, e.g. as a tool for risk prediction or as an intervention target for disease prevention, may have to consider a trade-off in risk between cancer and non-neoplastic diseases. For example, a number of companies have been established that offer telomere length measurement services to the public (via a requesting physician), under the claim that shorter telomeres are a general indicator of poorer health status and older biological age and that such information can be used to motivate healthy lifestyle choices in individuals. However, the conflicting direction of association between telomere length and risk of cancer and non-neoplastic diseases, indicated by our findings, suggests that such services to the general public may be premature.

Conclusion

It is likely that longer telomeres increase risk for several cancers but reduce risk for some non-neoplastic diseases, including cardiovascular diseases. Further research is required to resolve whether telomere length is a useful predictor of risk that can help guide therapeutic interventions, to clarify the shape of any dose-response relationships and to characterise the nature of the association in population subgroups.

342 The Telomeres Mendelian Randomization Collaboration

343

344 Philip Haycock[1]; Stephen Burgess[2]; Aayah Nounu[1]; Jie Zheng[1]; George N Okoli[3]; Jack
345 Bowden[1,4]; Kaitlin Wade[1]; Nicholas Timpson[1]; David M. Evans[1,5]; Peter Willeit[2,6]; Abraham
346 Aviv[7]; Tom R. Gaunt[1]; Gibran Hemani[1]; Massimo Mangino[8,9]; Hayley Patricia Ellis[10];
347 Kathreena Mary Kurian[10]; Karen Pooley[11] on behalf of the BCAC and OCAC consortia;
348 Rosalind Eeles[12] on behalf of the PRACTICAL consortium; Jeffrey E Lee[13]; Shenying
349 Fang[13]; Wei Chen[13]; Matthew H Law[14], Lisa M Bowdler[15] and Mark M Iles[16] on behalf of the
350 Melanoma meta-analysis consortium; Qiong Yang[17], Bradford B. Worrall[18] and Hugh
351 Stephen Markus[19] on behalf of the METASTROKE project of the ISGC; Rayjean J. Hung[20,21]
352 and Chris I Amos[22] on behalf of the ILCCO consortium; Amanda Spurdle[23], Deborah J
353 Thompson[24] and Tracy O'Mara[23] on behalf of the ECAC consortium; Brian Wolpin[25], Laufey
354 Amundadottir[26] and Rachael Stolzenberg-Solomon[27] on behalf of the PanScan consortium;
355 Antonia Trichopoulou[29,30], Charlotte Onland-Moret[31], Eiliv Lund[32], Eric J Duell[33], Federico
356 Canzian[34], Gianluca Severi[35,36,37,38], Kim Overvad[39], Marc J Gunter[40], Rosario Tumino[41] and
357 Ulrika Svenson[42] on behalf of EPIC; Andre van Rij[43], Annette F Baas[44], Matthew J Bown[45],
358 Nilesh J Samani[45], Paul IW de Bakker[44], Femke NG van t'Hof[44], Gerard Tromp[46,47], Gregory
359 T Jones[43], Helena Kuivaniemi[46,47] and James R Elmore[48] on behalf of the Aneurysm
360 Consortium; Mattias Johansson[49]; James Mckay[49]; Ghislaine Scelo[49]; Robert Carreras-
361 Torres[49]; Valerie Gaborieau[49]; Paul Brennan[49]; Paige M. Bracci[50], Rachel E Neale[15], Sara H
362 Olson[51], Steven Gallinger[20], Donghui Li[52], Gloria M. Petersen[54], Harvey Risch[55], and Alison
363 P. Klein[56] on behalf of PanC[4]; Jiali Han[57,58]; Christian C. Abnet[59]; Neal D. Freedman[59]; Philip
364 R. Taylor[59]; John M Maris[60]; Katja K Aben[61,62]; Lambertus A Kiemeney[61]; Sita H
365 Vermeulen[61]; John K Wiencke[63,64]; Kyle M Walsh[63,64]; Margaret Wrensch[63,64]; Terri Rice[63];
366 Clare Turnbull[65]; Kevin Litchfield[66]; Lavinia Paternoster[1] and Marie Standl[67] on behalf of the
367 EAGLE consortium; Gonçalo R Abecasis[68]; John Paul SanGiovanni[69]; Yong Li[70] and Vladan
368 Mijatovic[71] on behalf of the CKDGen consortium; Yadav Sapkota[15]; Siew-Kee Low[72]; Krina
369 T Zondervan[73,74]; Grant W Montgomery[15]; Dale R. Nyholt[75,15]; David A van Heel[76]; Karen
370 Hunt[76]; Dan E. Arking[77], Foram N. Ashar[77] and Nona Sotoodehnia[78] on behalf of the
371 CHARGE-Sudden Cardiac Arrest Working Group; Daniel Woo[79]; Jonathan Rosand[80]; Mary
372 Comeau[81]; W. Mark Brown[82]; Edwin K. Silverman[83], John E Hokanson[84] and Michael Cho[83]
373 on behalf of COPDGene; Jennie Hui[85,86,87,88], Manuel Ferreira[15] and Philip J. Thompson[89] on
374 behalf of the AAGC consortium; Alanna C. Morrison[90], Janine F Felix[91] and Nicholas L

Smith[92] on behalf of the CHARGE-Heart Failure Working Group; Angela M Christiano[93]; Lynn Petukhova[94]; Regina C. Betz[95]; Xing Fan[96]; Xuejun Zhang[96]; Caihong Zhu[96]; Carl Langefeld[97]; Susan D. Thompson[98]; Feijie Wang[99]; Xu Lin[99,100]; David A. Schwartz[101]; Tasha Fingerlin[102]; Jerome I. Rotter[103,104], Mary Frances Cotch[105] and Richard A Jensen on behalf of the CHARGE-Eye Working Group[106,107]; Matthias Munz[108], Henrik Dommisch[108] and Arne S Schaefer[108] on behalf of the European Periodontitis Genetics Group; Fang Han[109]; Hanna M Ollila[110]; Ryan P. Hillary[110]; Omar Albagha[111]; Stuart H. Ralston[112]; Chenjie Zeng[113]; Wei Zheng[113]; Xiao-Ou Shu[113]; Andre Reis[114]; Steffen Uebe[114]; Ulrike Hüffmeier[114]; Yoshiya Kawamura[115], Takeshi Otowa[116],[117] and Tsukasa Sasaki[118] on behalf of the Japanese Collaboration Team for GWAS of Panic Disorder; Martin Lloyd Hibberd[119]; Michael Levin[120]; Sonia Davila[121]; Gang Xie1[22,20]; Katherine Siminovitch[122,20]; Jin-Xin Bei[123]; Yi-Xin Zeng[123,124]; Asta Försti[125,126]; Bowang Chen[125]; Stefano Landi[127]; Andre Franke[128]; Annegret Fischer[128,129]; David Ellinghaus[128]; Carlos Flores[130,131]; Imre Noth[132]; Shwu-Fan Ma[132]; Jia Nee Foo[133]; Jianjun Liu[133]; Jong-Won Kim[134]; David G. Cox[135]; Olivier Delattre[136]; Olivier Mirabeau[136]; Christine F. Skibola[137]; Clara S. Tang[138]; Merce Garcia-Barcelo[138]; Kai-Ping Chang[139]; Wen-Hui Su[140]; Yu-Sun Chang[141]; Nicholas G Martin[15]; Scott Gordon[15]; Tracey Wade[142]; Chaeyoung Lee[143]; Michiaki Kubo[144]; Pei-Chieng Cha[145]; Yusuke Nakamura[146]; Daniel Levy[147]; Masayuki Kimura[7]; Shih-Jen Hwang[147]; Steven Hunt[148]; Tim Spector[8]; Nicole Soranzo[149]; Ani W Manichaikul[150]; R Graham Barr[151]; Bratati Kahali[152], Elizabeth Speliotes[152] and Laura M Yerges-Armstrong[153] on behalf of the GOLD Consortium; Ching-Yu Cheng[154,155,156], Jost B. Jonas[157,158] and Tien Yin Wong[154,155,156] on behalf of the SEED consortium; Isabella Fogh[159], Kuang Lin[159] and John F. Powell[159] on behalf of the SLAGEN and ALSGEN consortia; Kenneth Rice[160] on behalf of the ICBP; Caroline Relton[1]; Richard M Martin[1,3,161]; George Davey Smith[1]

[1] MRC Integrative Epidemiology Unit, University of Bristol, Bristol, UK
[2] Department of Public Health and Primary Care, University of Cambridge
[3] School of Social and Community Medicine, University of Bristol, Bristol, UK
[4] MRC Biostatistics Unit, Cambridge, UK.
[5] University of Queensland Diamantina Institute, Translational Research Institute, Brisbane, Queensland, Australia
[6] Department of Neurology, Innsbruck Medical University, Austria
[7] Center of Human Development and Aging, Department of Pediatrics, New Jersey Medical School, Rutgers, The State University of New Jersey
[8] Department of Twin Research and Genetic Epidemiology, King's College London, London UK
[9] NIHR Biomedical Research Centre at Guy's and St. Thomas' Foundation Trust, London, UK
[10] Brain Tumour Research Group, Institute of Clinical Neuroscience, Learning and Research Building, Southmead Hospital, University of Bristol
[11] Strangeways Research Laboratory, University of Cambridge, Cambridge, UK
[12] The Institute of Cancer Research, London, UK
[13] Department of Surgical Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX.
[14] Statistical Genetics, QIMR Berghofer Medical Research Institute, Brisbane, Australia
[15] QIMR Berghofer Medical Research Institute, Brisbane, Australia
[16] Section of Epidemiology and Biostatistics, Leeds Institute of Cancer and Pathology, University of Leeds, Leeds, UK
[17] Department of Biostatistics, Boston University School of Public Health, Boston, Massachusetts, United States of America and the Framingham Heart Study, Framingham, Massachusetts, United States of America
[18] Departments of Neurology and Public Health Sciences University of Virginia Charlottesville, Virginia 22908
[19] Department of Clinical Neurosciences, University of Cambridge, UK
[20] Lunenfeld-Tanenbaum Research Institute of Mount Sinai Hospital, Toronto, Ontario, Canada
[21] Division of Epidemiology, Dalla Lana School of Public Health, University of Toronto, 60 Murray St. Rm L5-215, Box 18, Toronto, ON M5T 3L9, Canada
[22] Geisel School of Medicine, Dartmouth College
[23] Genetics and Computational Biology Division, QIMR Berghofer Medical Research Institute, Brisbane, QLD 4006, Australia
[24] Centre for Cancer Genetic Epidemiology, Department of Public Health and Primary Care, University of Cambridge, Cambridge, CB1 8RN, UK.
[25] Dana-Farber Cancer Institute
[26] Laboratory of Translational Genomics, Division of Cancer Epidemiology & Genetics, National Cancer Institute
[27] Division of Cancer Epidemiology and Genetics, National Cancer Institute, NIH, DHHS.

[29] Hellenic Health Foundation, Athens, Greece
[30] WHO Collaborating Center for Nutrition and Health, Unit of Nutritional Epidemiology and Nutrition in Public Health, Dept. of Hygiene, Epidemiology and Medical Statistics, University of Athens Medical School, Greece
[31] Dept of Epidemiology, Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht
[32] Institute of community medicine, UiT The Artcic University of Norway
[33] Unit of Nutrition and Cancer, Cancer Epidemiology Research Program, Bellvitge Biomedical Research Institute (IDIBELL), Catalan Institute of Oncology (ICO), Avda Gran Via 199-203, 08908 L'Hospitalet de Llobregat, Barcelona, Spain
[34] Genomic Epidemiology Group, German Cancer Research Center (DKFZ), Heidelberg, Germany
[35] Université Paris-Saclay, Université Paris-Sud, UVSQ, CESP, INSERM, Villejuif, France
[36] Gustave Roussy, F-94805, Villejuif, France
[37] Human Genetics Foundation (HuGeF), Torino, Italy

[38] Cancer Council Victoria and University of Melbourne, Australia

[39] Department of Public Health, Section for Epidemiology, Aarhus University, Aarhus, Denmark

[40] School of Public Health, Imperial College London, London W2 1PG

[41] Cancer Registry, Azienda Ospedaliera "Civile M.P.Arezzo", via Dante 109, Ragusa, IT

[42] Department of Medical Biosciences, Umea University, Umea, Sweden

[43] Surgery Department, University of Otago, Dunedin, New Zealand

[44] Department of Medical Genetics, Center for Molecular Medicine and Department of Epidemiology, Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, The Netherlands

[45] The Department of Cardiovascular Sciences and the NIHR Leicester Cardiovascular Biomedical Research Unit, University of Leicester, Leicester, LE2 7LX, UK.

[46] Division of Molecular Biology and Human Genetics, Department of Biomedical Sciences, Faculty of Medicine and Health Sciences, Stellenbosch University, Tygerberg, 7505, South Africa

[47] The Sigfried and Janet Weis Center for Research, Geisinger Health System, Danville, PA, USA

[48] Department of Vascular and Endovascular Surgery, Geisinger Health System, Danville, PA 17822, USA

[49] Genetic Epidemiology Group, International Agency for Research on Cancer, Lyon, France

[50] Department of Epidemiology and Biostatistics, University of California San Francisco, San Francisco, California

[51] Department of Epidemiology and Biostatistics, Memorial Sloan Kettering Cancer Center, New York, New York, USA.

[52] Department of Gastrointestinal Medical Oncology, University of Texas MD Anderson Cancer Center, Houston, Texas, USA

[54] Department of Health Sciences Research, Mayo Clinic College of Medicine, Rochester, Minnesota, USA

[55] Yale School of Public Health & Yale School of Medicine & Yale Cancer Center, 60 College St., PO Box 208034, New Haven, CT 06520-8034

[56] Departments of Oncology, Pathology and Epidemiology, Johns Hopkins School of Medicine, Baltimore Maryland 21231

[57] Department of Epidemiology, Fairbanks School of Public Health, Indiana University

[58] Indiana University Melvin and Bren Simon Cancer Center, Indianapolis, IN, USA

[59] Division of Cancer Epidemiology and Genetics, National Cancer Institute, Rockville MD USA

[60] Children's Hospital of Philadelphia, Perelman School of Medicine at the University of Pennsylvania

[61] Radboud university medical center, Radboud Institute for Health Sciences, Nijmegen, The Netherlands

[62] Netherlands Comprehensive Cancer Organization, Utrecht, The Netherlands

[63] Department of Neurological Surgery, University of California, San Francisco, San Francisco, CA

[64] Institute of Human Genetics, University of California, San Francisco, San Francisco, CA

[65] William Harvey Research Institute, Queen Mary University, London, UK

[66] Division of Genetics and Epidemiology, The Institute of Cancer Research, London, UK

[67] Institute of Epidemiology I, Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg, Germany

[69] National Eye Institute, Division of Epidemiology and Clinical Research Clinical Trials Branch National Institutes of Health, Bethesda, USA

[70] Department of Internal Medicine IV, University Hospital Freiburg

[71] Department of Life and Reproduction Sciences, University of Verona

[72] Laboratory of Statistical Analysis, Centre for Integrative Medical Sciences, The Institute of Physical and Chemical Research (RIKEN], Yokohama, Japan

[73] Genetic and Genomic Epidemiology Unit, Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK

[74] Nuffield Department of Obstetrics and Gynecology, University of Oxford, John Radcliffe Hospital, Oxford, UK
[75] Institute of Health and Biomedical Innovation, Queensland University of Technology, Brisbane, Australia
[76] Blizard Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London E1 2AT, UK
[77] McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD USA
[78] Division of Cardiology and Cardiovascular Health Research Unit, Department of Medicine, University of Washington, Seattle, Washington 98101
[79] University of Cincinnati College of Medicin, Department of Neurology, Cincinnati, OH, USA 45267
[80] Massachusetts General Hospital, Neurology, Center for Human Genetic Research, MA, USA
[81] Center for Public Health Genomics, Department of Biostatistical Sciences, Division of Public Health Sciences, Wake Forest School of Medicine, Medical Center Blvd, Winston-Salem, NC 27157
[82] Department of Biostatistical Sciences, Division of Public Health Sciences, Wake Forest School of Medicine, Medical Center Blvd, Winston-Salem, NC 27157
[83] Channing Division of Network Medicine, Brigham and Women's Hospital, Boston, MA 02115
[83] Channing Division of Network Medicine, Brigham and Women's Hospital, Boston, MA, 02115
[84] Department of Epidemiology, University of Colorado Anschutz Medical Campus, Aurora, Colorado, USA
[85] Busselton Population Medical Research Institute Inc, Sir Charles Gairdner Hospital, Perth, Australia
[86] PathWest Laboratory Medicine of Western Australia (WA), Perth, Australia
[87] School of Pathology and Laboratory Medicine, University of WA, Perth, Australia
[88] School of Population Health, University of WA, Perth, Australia
[89] The Lung Health Clinic and Institute for Respiratory Health, University of Western Australia, Perth, Australia
[90] Department of Epidemiology, Human Genetics, and Environmental Sciences, University of Texas Health Science Center at Houston, Houston, TX 77030 USA
[91] Department of Epidemiology, Erasmus MC, University Medical Center Rotterdam, Rotterdam, the Netherlands
[92] Department of Epidemiology, University of Washington, Seattle WA 98101 USA
[93] Departments of Dermatology and Genetics & Development, Columbia University, New York, NY, US
[94] Departments of Dermatology and Epidemiology, Columbia University, New York, NY, US
[95] Institute of Human Genetics, University of Bonn, Bonn D-53127, Germany
[96] Institute of Dermatology & Department of Dermatology, First Affiliated Hospital of Anhui Medical University
[97] Director, Center for Public Health Genomics, Department of Biostatistical Sciences, Division of Public Health Sciences, Wake Forest School of Medicine, Medical Center Blvd, Winston-Salem, NC 27157
[98] Center for Autoimmune Genomics and Etiology, Cincinnati Children's Hospital Medical Center, Department of Pediatrics,University of Cincinnati College of Medicine, Cincinnati, OH, USA
[99] Institute for Nutritional Sciences, SIBS, Chinese Academy of Sciences, Shanghai, 200031, PR China
[100] Key Laboratory of Nutrition and Metabolism, Chinese Academy of Sciences, Shanghai, 200031, PR China
[101] University of Colorado, 12631 East 17th Avenue, B178, Aurora, CO 80045
[102] Department of Biomedical Research, National Jewish Health Hospital
[103] Institute for Translational Genomics and Population Sciences, Los Angeles Biomedical Research Institute at Harbor-UCLA Medical Center
[104] Departments of Pediatrics and Medicine, 1124 W. Carson Street, Harbor-UCLA Medical Center, Torrance, CA 90502
[105] Epidemiology Branch, Division of Epidemiology and Clinical Applications, NIH Intramural Research Program, National Eye Institute, National Institutes of Health, Clinical Research Center 3A2521,

[106] Cardiovascular Health Research Unit, University of Washington, Seattle, Washington, USA
[107] Department of Medicine, University of Washington, Seattle, Washington, USA
[108] Charité – University Medicine Berlin, CC 03, Institute of Dental, Oral and Maxillary Medicine, Dept. of Periodontology and Synoptic Dentistry, Aßmannshauser Str. 4-6, 14197 Berlin, Germany
[109] Department of Pulmonary Medicine, Peking University People's Hospital, 100044 Beijing, China
[110] Stanford University, Center for Sleep Sciences, Palo Alto, CA, USA
[111] Centre for Genomic and Experimental Medicine, Institute of Genetics and Molecular Medicine, University of Edinburgh, Western General Hospital, Edinburgh EH4 2XU, UK
[112] Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, UK
[113] Division of Epidemiology, Department of Medicine, Vanderbilt Epidemiology Center, Vanderbilt-Ingram Cancer Center, Vanderbilt University Medical Center, Nashville, Tennessee
[114] Institute of Human Genetics, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany
[115] Department of Psychiatry, Shonan Kamakura General Hospital, Kanagawa, Japan
[116] Department of Neuropsychiatry, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan
[117] Graduate School of Clinical Psychology, Teikyo Heisei University Major of Professional Clinical Psychology, Tokyo, Japan
[118] Department of Physical and Health Education, Graduate School of Education, The University of Tokyo, Tokyo, Japan
[119] Infectious Diseases, Genome Institute of Singapore, Singapore
[120] Division of Infectious diseases, Department of medicine, Imperial College London, UK
[121] Human genetics, Genome Institute of Singapore, Singapore
[122] Departments of Medicine, Immunology, Molecular Genetics, University of Toronto
[123] Sun Yat-sen University Cancer Center, State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine Guangzhou 510060, P. R. China
[124] Peking Union Medical College, Beijing 100730, P.R. China
[125] Molecular Genetic Epidemiology, German Cancer Research Center (DKFZ), Heidelberg, Germany
[126] Center for Primary Health Care Research, Clinical Research Center, Lund University, Malmö, Sweden
[127] Department of Biology, University of Pisa, Pisa, Italy
[128] Institute of Clinical Molecular Biology, Christian-Albrechts-University of Kiel, Kiel, Germany.
[129] University Hospital Schleswig-Holstein, Kiel, Germany
[130] Research Unit, Hospital Universitario N.S. de Candelaria, Tenerife, Spain
[131] CIBER de Enfermedades Respiratorias, Instituto de Salud Carlos III, Madrid, Spain
[132] Section of Pulmonary and Critical Care Medicine, University of Chicago, 5841 S. Maryland Ave., Chicago IL 60637-6076
[133] Human Genetics, Genome Institute of Singapore, A*STAR, Singapore 138672, Singapore
[134] Dept. of Laboratory Medicine and Genetics, Samsung Medical Center, Sungkyunkwan, University School of Medicine, Ilwon-dong 50, Gangnam-gu, Seoul, Korea, 135-710
[135] Cancer Research Center of Lyon, INSERM U1052, Lyon, France
[136] Inserm U830, Institut Curie, PSL University, 26 rue d'Ulm, 75248 Paris Cedex 05 France, France.
[137] Department of Epidemiology, University of Alabama at Birmingham. 1665 University Boulevard, Birmingham, AL 35294-0022, USA
[138] Department of Surgery, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong, SAR China
[139] Department of Otolaryngology - Head and Neck Surgery, Chang Gung Memorial Hospital at Lin-Kou, Taoyuan, Taiwan,
[140] Department of Biomedical Sciences, Graduate Institute of Biomedical Sciences, College of Medicine, Molecular Medicine Research Center, Chang Gung University, Taoyuan, Taiwan
[141] Molecular Medicine Research Center, Chang Gung University, Taoyuan, Taiwan

[142] School of Psychology, Flinders University
[143] School of Systems Biomedical Science, Soongsil University, 369 Sangdo-ro, Dongjak-gu, Seoul 156-743, Korea
[144] RIKEN Center for Integrative Medical Science, 1-7-22, Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, JAPAN
[145] Division of Molecular Brain Science, Kobe University Graduate School of Medicine, 7-5-1 Kusunoki-chou, Chuo-ku, Kobe 650-0017, Japan
[146] Center for Personalized Therapeutics, The University of Chicago, 900E 57th Street, Chicago IL 60637 USA
[147] The NHLBI's Framingham Heart Study, Framingham, MA, Population Sciences Branch of the National Heart, Lung, and Blood Institute, Bethesda, MD.
[148] Department of Genetic Medicine, Weill Cornell Medicine in Qatar, Doha, Qatar
[149] Human Genetics, Wellcome Trust Sanger Institute, Genome Campus, Hinxton Cambridge
[150] Center for Public Health Genomics, Department of Public Health Sciences, University of Virginia, Charlottesville, VA USA
[151] Department of Medicine and Department of Epidemiology, Columbia University Medical Center, New York, NY 10032, USA
[152] Department of Internal Medicine, Division of Gastroenterology and Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI 48109, USA
[153] Department of Medicine, University of Maryland, Baltimore, MD 21201, USA
[154] Singapore Eye Research Institute, Singapore National Eye Center, Singapore 168751, Singapore
[155] Department of Ophthalmology, National University of Singapore and National University Health System, Singapore
[156] Duke-National University of Singapore Graduate Medical School, Singapore
[157] Beijing Institute of Ophthalmology, Beijing Tongren Eye Center, Beijing Tongren Hospital, Capital Medical University, Beijing Ophthalmology and Visual Science Key Lab, Beijing, China
[158] Department of Ophthalmology, Medical Faculty Mannheim of the Ruprecht-Karls-University Heidelberg, Mannheim, Germany
[159] Department of Basic and Clinical Neuroscience, Maurice Wohl Clinical Neuroscience Institute, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, United Kingdom
[160] Dept of Biostatistics, University of Washington, Seattle, USA
[161] University of Bristol / University Hospitals Bristol NHS Foundation Trust National Institute for Health Research Bristol Nutrition Biomedical Research Unit, Bristol, UK

399

400 Affiliations of the Telomeres Mendelian Randomization Collaboration

401

402 **Acknowledgements**

Access to Data Statement: Dr Philip C Haycock had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Role of the Funder/Sponsor: The funding institutions had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

This work was supported by CRUK grant number C18281/A19169 (the Integrative Cancer Epidemiology Programme). Dr Haycock is supported by CRUK Population Research Postdoctoral Fellowship C52724/A20138. The MRC Integrative Epidemiology Unit is supported by grants MC_UU_12013/1 and MC_UU_12013/2. Dr Martin is supported by the National Institute for Health Research (NIHR), the Bristol Nutritional Biomedical Research Unit and the University of Bristol.

We gratefully acknowledge all the studies and databases that made GWAS summary data available (see supplementary materials for detailed acknowledgements): **AC** (the aneurysm consortium), **ALSGEN (**the International Consortium on Amyotrophic Lateral Sclerosis Genetics), **AMD Gene (**Age-related Macular Degeneration Gene Consortium), **BCAC** (Breast Cancer Association Consortium), **C4D** (Coronary Artery Disease Genetics Consortium), **CARDIoGRAM** (Coronary ARtery DIsease Genome wide Replication and Meta-analysis), **CHARGE-HF** (Cohorts for Heart and Aging Research in Genomic Epidemiology Consortium – Heart Failure Working Group), **COPDGene** (The Genetic Epidemiology of Chronic Obstructive Pulmonary Disease), **CORECT** (ColoRectal Transdisciplinary Study), **CKDGen** (Chronic Kidney Disease Genetics consortium), **dbGAP** (database of Genotypes and Phenotypes), **DIAGRAM** (DIAbetes Genetics Replication And Meta-analysis), **EAGLE** (EArly Genetics & Lifecourse Epidemiology Eczema Consortium, excluding 23andMe), **ECAC** (Endometrial Cancer Association Consortium), **EGG** (Early Growth Genetics Consortium), **EPG** (European Periodontitis Genetics Group), **GABRIEL** (A Multidisciplinary Study to Identify the Genetic and Environmental Causes of Asthma in the European Community), **GCAN** (Genetic Consortium for Anorexia Nervosa), **GECCO** (Genetics and Epidemiology of Colorectal Cancer Consortium), **GIANT** (Genetic Investigation of ANthropometric Traits), **GLGC** (Global Lipids Genetics Consortium),

GUGC (Global Urate and Gout consortium), **ICBP** (International Consortium for Blood Pressure), **IGAP** (International Genomics of Alzheimer's Project), **HPFS** (Health Professionals Follow-Up Study), **JCTGPD** (Japanese Collaboration Team for GWAS of Panic Disorder), **ILCCO** (International Lung Cancer Consortium), **ImmunoBase** (genetic database of immunologically related human diseases), **IMSGC** (International Multiple Sclerosis Genetic Consortium), **IIBDGC** (International Inflammatory Bowel Disease Genetics Consortium); **KIDRISK** (Kidney cancer consortium), **MAGIC** (Meta-Analyses of Glucose and Insulin-related traits Consortium), **MC** (the melanoma meta-analysis consortium), **MESA** (Multi-Ethnic Study of Atherosclerosis), **METASTROKE/ISGC** (METASTROKE project of the International Stroke Genetics Consortium), **NBCS** (Nijmegen Bladder Cancer Study), **NHGRI-EBI GWAS catalog** (National Human Genome Research Institute and European Bioinformatics Institute Catalog of published genome-wide association studies), **NHS** (Nurses' Health Study), **OCAC** (Ovarian Cancer Association Consortium), **PanScan** (Pancreatic Cancer Cohort Consortium), **PGC** (Psychiatric Genomics Consortium), **PRACTICAL** (Prostate Cancer Association Group to Investigate Cancer Associated Alterations in the Genome), **SEEDS** (the Singapore Epidemiology of Eye Diseases Study), **SLAGEN** (Italian Consortium for the Genetics of Ayotrophic Lateral Sclerosis), **SSGAC (**Social Science Genetics Association Consortium), **TAG** (Tobacco and Genetics Consortium), **T1Dbase** (type 1 diabetes database), **TICG** (Tourette International Collaborative-Genetics); **TSAICG** (Tourette Syndrome Association International Consortium for Genetics).

We gratefully acknowledge the assistance and contributions of Dr Julia Gumy, Ms Lisa Wright, Dr Georg B. Ehret (ICBP), Dr Louise V. Wain (ICBP), Dr Caroline Fox (CKDGen), Dr Stephan Ripke (IIBDGC), Dr Jimmy Liu (IIBDGC), Dr Carl Anderson (IIBDGC), Dr

452 Jeremiah Scharf (TSAICG and TICG), Dr Lars Fritsche (AMD Gene), Dr Joanne Elena and

453 Dr Paul KH Tam (Hirschsprung's disease GWAS).

Table 1. Single nucleotide polymorphisms associated with telomere length

SNPs	Chr	Pos	Gene	EA	OA	EAF*	Beta*	SE*	P-value*	Phet*	No. studies*	Sample size*	Discovery p-value	% variance explained	Discovery study
rs11125529	2	54248729	*ACYP2*	A	C	0.16	0.065	0.012	0.000606	0.313	6	9177	8.00E-10	0.080	Codd[21]
rs6772228	3	58390292	*PXK*	T	A	0.87	0.041	0.014	0.049721	0.77	6	8630	3.91E-10	0.200	Pooley[17]
rs12696304	3	169763483	*TERC*	C	G	0.74	0.090	0.011	5.41E-08	0.651	6	9012	4.00E-14	0.319	Codd[22]
rs10936599	3	169774313	*TERC*	C	T	0.76	0.100	0.011	1.76E-09	0.087	6	9190	3.00E-31	0.319	Codd[21]
rs1317082	3	169779797	*TERC*	A	G	0.71	0.097	0.011	4.57E-09	0.029	6	9176	1.00E-08	0.319	Mangino[18]
rs10936601	3	169810661	*TERC*	C	T	0.74	0.087	0.011	8.64E-08	0.433	6	9150	4.00E-15	0.319	Pooley[17]
rs7675998	4	163086668	*NAF1*	G	A	0.80	0.048	0.012	0.008912	0.077	6	9161	4.35E-16	0.190	Codd[21]
rs2736100	5	1286401	*TERT*	C	A	0.52	0.085	0.013	2.14E-05	0.54	4	5756	4.38E-19	0.310	Codd[21]
rs9419958	10	103916188	*OBFC1*	T	C	0.13	0.129	0.013	5.26E-11	0.028	6	9190	9.00E-11	0.171	Mangino[18]
rs9420907	10	103916707	*OBFC1*	C	A	0.14	0.142	0.014	1.14E-11	0.181	6	9190	7.00E-11	0.171	Codd[21]
rs4387287	10	103918139	*OBFC1*	A	C	0.14	0.120	0.013	1.40E-09	0.044	6	8541	2.00E-11	0.171	Levy[25]
rs3027234	17	8232774	*CTC1*	C	T	0.83	0.103	0.012	2.75E-08	0.266	6	9108	2.00E-08	0.292	Mangino[18]
rs8105767	19	22032639	*ZNF208*	G	A	0.25	0.064	0.011	0.000169	0.412	6	9096	1.11E-09	0.090	Codd[21]
rs412658	19	22176638	*ZNF676*	T	C	0.35	0.086	0.010	1.83E-08	0.568	6	9156	1.00E-08	0.484	Mangino[18]
rs6028466	20	39500359	*DHX35*	A	G	0.17	0.058	0.013	0.003972	0.533	6	9190	2.57E-08†	0.041	Mangino[18] & Gu
rs755017	20	63790269	*ZBTB46*	G	A	0.17	0.019	0.0129	0.339611	0.757	5	8026	6.71E-09	0.090	Codd[21]

*Summary data from Mangino et al[18]; Chr, chromosome; pos, base-pair position (GRCh38.p3); EA, effect allele, OA, other allele, Beta, standard deviation change in telomere length per copy of the effect allele; SE, standard error; EAF - effect allele frequency; Phet - p value for between-study heterogeneity in association between SNP and telomere length; †from a meta-analysis of Mangino[18] and Gu[20] performed in the present study.

Table 2. Study characteristics for primary non-communicable diseases

	No. cases	No. controls	No. SNPs	Statistical power	Pop.	Study / First author
Cancer						
Bladder cancer	1601	1819	10	0.62	EUR	NBCS[76]
Breast cancer	48155	43612	13	1.00	EUR	BCAC[17,77]
Estrogen receptor –ve	7465	42175	13	1.00	EUR	BCAC[17,77]
Estrogen receptor +ve	27074	41749	13	1.00	EUR	BCAC[17,77]
Colorectal cancer	14537	16922	9	1.00	EUR	CORECT/GECCO[64,78]
Endometrial cancer	6608	37925	12	1.00	EUR	ECAC[79,80]
Esophageal SCC	1942	2111	11	0.64	EA	Abnet[81]
Glioma	1130	6300	12	0.72	EUR	Wrensch[82] & Walsh[66]
Head & neck cancer	2082	3477	12	1.00	EUR	McKay et al[83]
Kidney cancer	2461	5081	12	0.99	EUR	KIDRISK[84]
Lung cancer	11348	15861	13	1.00	EUR	ILCCO[85]
Adenocarcinoma	3442	14894	13	1.00	EUR	ILCCO[85]
Squamous cell carcinoma	3275	15038	13	1.00	EUR	ILCCO[85]
Skin cancer						
Melanoma	12814	23203	13	1.00	EUR	MC[86]
Basal cell carcinoma	3361	11518	13	1.00	EUR	NHS/HPFS[87]
Neuroblastoma	2101	4202	12	0.87	EUR	Diskin[88]
Ovarian cancer	15397	30816	13	1.00	EUR	OCAC[17,89]
Clear cell	1016	30816	13	0.76	EUR	OCAC[17,89]
Endometriod	2154	30816	13	0.98	EUR	OCAC[17,89]
Mucinous	1643	30816	13	0.94	EUR	OCAC[17,89]
Serous invasive	9608	30816	13	1.00	EUR	OCAC[17,89]
Serous LMP	972	30816	13	0.73	EUR	OCAC[17,89]
Pancreatic cancer	5105	8739	12	1.00	EUR	PanScan (incl. EPIC)[90]
Prostate cancer	22297	22323	11	1.00	EUR	PRACTICAL[91,92]
Testicular germ cell cancer	986	4946	11	0.52	EUR	Turnbull[93] & Rapley[94]
Autoimmune/inflammatory diseases						
Alopecia areata	2332	5233	7	0.60	EUR	Betz[95]
Atopic dermatitis	10788	30047	13	1.00	EUR	EAGLE[96]
Celiac disease	4533	10750	3	0.82	EUR	Dubois[97]
Inflammatory bowel disease						
Crohn's disease	5956	14927	11	1.00	EUR	IIBDGC[98]
Ulcerative colitis	6968	20464	12	1.00	EUR	IIBDGC[98]
Juvenile idiopathic arthritis	1866	14786	11	0.87	EUR	Thompson[99]†
Multiple sclerosis	14498	24091	3	1.00	EUR	IMSGC[100]
Aggressive periodontitis	888	6789	13	0.63	EUR	Schaefer[101]
Rheumatoid arthritis	5538	20163	11	1.00	EUR	Stahl[102]
Cardiovascular diseases						
Abdominal aortic aneurysm	4972	99858	13	1.00	EUR	AC[103–108]
Coronary heart disease	22233	64762	13	1.00	EUR	CARDIoGRAM[109]
Heart failure	2526	20926	13	0.99	EUR	CHARGE-HF[110]
Hemorrhagic stroke	2963	5503	12	0.96	EUR	METASTROKE/ISGC[111]
Ischemic stroke	12389	62004	13	1.00	EUR	METASTROKE/ISGC[112,113]
large vessel disease	2167	62004	13	0.99	EUR	METASTROKE/ISGC[112,113]
small vessel disease	1894	62004	13	0.97	EUR	METASTROKE/ISGC[112]
cardioembolic	2365	62004	13	0.99	EUR	METASTROKE/ISGC[112]
Sudden cardiac arrest	3954	21200	13	1.00	EUR	Unpublished
Diabetes						
Type 1 diabetes	7514	9045	6	0.95	EUR	T1DBase[114115]
Type 2 diabetes	10415	53655	11	1.00	EUR	DIAGRAM[116]
Eye disease						

AMD	7473	51177	13	1.00	EUR	AMD Gene[117]
Retinopathy	1122	18289	12	0.75	EUR	Jensen[118]
Lung diseases						
Asthma	13034	20638	4	1.00	EUR	Ferreira/GABRIEL[119,120]
COPD	2812	2534	12	0.85	EUR	COPDGene[121]
Interstitial lung disease	1616	4683	9	0.60	EUR	Fingerlin[122]
Neurological / psychiatric diseases						
ALS	6100	7125	12	1.00	EUR	SLAGEN/ALSGEN[123]
Alzheimer's disease	17008	37154	12	1.00	EUR	IGAP[124]
Anorexia nervosa	2907	14860	9	0.93	EUR	GCAN[125]
Autism	4949	5314	7	0.82	EUR	PGC[126]
Bipolar disorder	7481	9250	9	1.00	EUR	PGC[127]
Major depressive disorder	9240	9519	8	0.99	EUR	PGC[128]
Schizophrenia	35476	46839	12	1.00	EUR	PGC[129]
Tourette syndrome	1177	4955	13	0.74	EUR	TICG/TSAICG[130]
Other						
Chronic kidney disease	5807	56430	13	1.00	EUR	CKDGen[131]
Endometriosis	4604	9393	11	1.00	Mix	Nyholt[132]

Study acronyms: AC, the aneurysm consortium; **ALSGEN,** the International Consortium on Amyotrophic Lateral Sclerosis Genetics**; AMD Gene**, Age-related Macular Degeneration Gene Consortium; **BCAC**, Breast Cancer Association Consortium; **CARDIoGRAM**, Coronary ARtery DIsease Genome wide Replication and Meta-analysis; **CHARGE-HF**, Cohorts for Heart and Aging Research in Genomic Epidemiology Consortium – Heart Failure Working Group; **COPDGene**, The Genetic Epidemiology of Chronic Obstructive Pulmonary Disease; **CKDGen**, Chronic Kidney Disease Genetics consortium; **CORECT**, ColoRectal Transdisciplinary Study; **DIAGRAM**, DIAbetes Genetics Replication And Meta-analysis; **EAGLE**, EArly Genetics & Lifecourse Epidemiology Eczema Consortium (excluding 23andMe); **ECAC**, Endometrial Cancer Association Consortium; **EPIC**, European Prospective Investigation into Cancer and Nutrition study; **GABRIEL**, Multidisciplinary Study to Identify the Genetic and Environmental Causes of Asthma in the European Community; **GCAN**, Genetic Consortium for Anorexia Nervosa; **GECCO**, Genetics and Epidemiology of Colorectal Cancer Consortium; **IGAP**, International Genomics of Alzheimer's Project; **HPFS**, Health Professionals Follow-Up Study; **ILCCO**, International Lung Cancer Consortium; **IMSGC**, International Multiple Sclerosis Genetic Consortium; **IIBDGC**, International Inflammatory Bowel Disease Genetics Consortium; **KIDRISK**, Kidney cancer consortium; **MC**, the melanoma meta-analysis consortium; **METASTROKE/ISGC**, METASTROKE project of the International Stroke Genetics Consortium; **NBCS**, Nijmegen Bladder Cancer Study; **NHS**, Nurses' Health Study; **OCAC**, Ovarian Cancer Association Consortium; **PanScan**, Pancreatic Cancer Cohort Consortium; **PGC**, Psychiatric Genomics Consortium; **PRACTICAL**, Prostate Cancer Association Group to Investigate Cancer Associated Alterations in the Genome; **SLAGEN**, Italian Consortium for the Genetics of Ayotrophic Lateral Sclerosis; **T1DBase**, type 1 diabetes database; **TICG** (Tourette International Collaborative-Genetics); **TSAICG** (Tourette Syndrome Association International Consortium for Genetics);. **Abbreviations**: **ALS**, amyotrophic lateral sclerosis; **AMD**, age-related macular degeneration; **COPD**, chronic obstructive pulmonary disease; **EUR**, European; **EA**, East Asian; **LMP**, low malignant potential; **No.**, number; **Pop**., population; **SCC**, squamous cell carcinoma; **SNP**, single nucleotide polymorphism; **-ve**, negative; **+ve**, positive; †plus previously unpublished data.

Figure 1. The association between genetically increased telomere length and odds of primary non-communicable diseases

Legend to Figure 1

*P value for association between genetically increased telomere length and disease from maximum likelihood; the effect estimate for heart failure is a hazard ratio (all others are odds ratios); P_{het}, P-value for heterogeneity amongst SNPs within the instrument; COPD, chronic obstructive pulmonary disease; SNP, single nucleotide polymorphism; CI, confidence interval; LMP, low malignancy potential; ER, estrogen receptor; -VE, negative; +VE, positive.

Figure 2. The association between genetically increased telomere length and odds of cancer as a function of selected characteristics

Legend to Figure 2

The plotted data show how the strength of the relationship between genetically increased telomere length and cancer varies by the selected characteristic. The R^2 statistic indicates how much of the variation between cancers can be explained by the selected characteristic. P-values are from meta-regression models. Circle sizes are proportional to the inverse of the variance of the log odds ratio. The hashed line indicates the null of no association between telomere length and cancer (i.e. an odds ratio of 1). Data for percentage survival 5 years after diagnosis, cancer incidence and median age-at-diagnosis was downloaded from the Surveillance, Epidemiology, and End Results Program.[33] Data for average lifetime number of stem cell divisions was downloaded from Tomasetti and Vogelstein.[34] Not all cancers had information available for the selected characteristics (hence the number of cancers varies across the subplots). Information was available for 9 cancers for tissue-specific rates of stem cell division, 13 cancers for percentage surviving 5 years post-diagnosis, 17 cancers for cancer incidence and 13 cancers for median age-at-diagnosis. SD, standard deviation; OR, Odds ratio.

Figure 3. Comparison of genetic and prospective observational studies[†] of the association between telomere length and disease

Legend to Figure 3

*from fixed-effects meta-analysis of independent observational studies described in Table S3; [†]search strategy and characteristics for observational studies are described in Tables S3 and S4; ‡CCHS and CGPS; +PLCO, ATBC & SWHS (acronyms explained in Table S3); CI, confidence interval

REFERENCES

1. Blackburn EH, Epel ES, Lin J. Human telomere biology: A contributory and interactive factor in aging, disease risks, and protection. Science (80-) 2015;350(6265):1193–8.

2. López-Otín C, Blasco MA, Partridge L, Serrano M, Kroemer G. The hallmarks of aging. Cell 2013;153(6):1194–217.

3. Samani NJ, van der Harst P. Biological ageing and cardiovascular disease. Heart 2008;94(5):537–9.

4. Weischer M, Bojesen SE, Nordestgaard BG. Telomere shortening unrelated to smoking, body weight, physical activity, and alcohol intake: 4,576 general population individuals with repeat measurements 10 years apart. PLoS Genet 2014;10(3):e1004191.

5. Houben JMJ, Moonen HJJ, van Schooten FJ, Hageman GJ. Telomere length assessment: biomarker of chronic oxidative stress? Free Radic Biol Med 2008;44(3):235–46.

6. Marchesi V. Risk factors: Short telomeres: association with cancer survival and risk. Nat Rev Clin Oncol 2013;10(5):247.

7. Haycock PC, Heydon EE, Kaptoge S, Butterworth a. S, Thompson A, Willeit P. Leucocyte telomere length and risk of cardiovascular disease: systematic review and meta-analysis. BMJ 2014;349(jul08 3):g4227–g4227.

8. Rode L, Nordestgaard BG, Bojesen SE. Peripheral blood leukocyte telomere length and mortality among 64,637 individuals from the general population. J Natl Cancer Inst 2015;107(6):djv074.

9. Zhao J, Miao K, Wang H, Ding H, Wang DW. Association between telomere length and type 2 diabetes mellitus: a meta-analysis. PLoS One 2013;8(11):e79993.

10. Weischer M, Nordestgaard BG, Cawthon RM, Freiberg JJ, Tybjærg-Hansen A, Bojesen SE. Short telomere length, cancer survival, and cancer risk in 47102 individuals. J Natl Cancer Inst 2013;105(7):459–68.

11. Ma H, Zhou Z, Wei S, et al. Shortened telomere length is associated with increased risk of cancer: a meta-analysis. PLoS One 2011;6(6):e20466.

12. Wentzensen IM, Mirabello L, Pfeiffer RM, Savage SA. The association of telomere length and cancer: a meta-analysis. Cancer Epidemiol Biomarkers Prev 2011;20(6):1238–50.

13. Pooley KA, Sandhu MS, Tyrer J, et al. Telomere length in prospective and retrospective cancer case-control studies. Cancer Res 2010;70(8):3170–6.

14. Hou L, Joyce BT, Gao T, et al. Blood Telomere Length Attrition and Cancer Development in the Normative Aging Study Cohort. EBioMedicine 2015;2(6):591–6.

15. Welter D, MacArthur J, Morales J, et al. The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. Nucleic Acids Res 2014;42(Database issue):D1001-6.

16. Burdett T, Hall P, Hastings E, et al. The NHGRI-EBI Catalog of published genome-wide association studies [Internet]. [cited 2015 Jan 15];Available from: www.ebi.ac.uk/gwas

17. Pooley KA, Bojesen SE, Weischer M, et al. A genome-wide association scan (GWAS) for mean telomere length within the COGS project: identified loci show little association with hormone-related cancer risk. Hum Mol Genet 2013;22(24):5056–64.

18. Mangino M, Hwang S-J, Spector TD, et al. Genome-wide meta-analysis points to CTC1 and ZNF676 as genes regulating telomere homeostasis in humans. Hum Mol Genet 2012;21(24):5385–94.

19. Prescott J, Kraft P, Chasman DI, et al. Genome-wide association study of relative telomere length. PLoS One 2011;6(5):e19635.

20. Gu J, Chen M, Shete S, et al. A genome-wide association study identifies a locus on chromosome 14q21 as a predictor of leukocyte telomere length and as a marker of susceptibility for bladder cancer. Cancer Prev Res (Phila) 2011;4(4):514–21.

21. Codd V, Nelson CP, Albrecht E, et al. Identification of seven loci affecting mean telomere length and their association with disease. Nat Genet 2013;45(4):422–7.

22. Codd V, Mangino M, van der Harst P, et al. Common variants near TERC are associated with mean telomere length. Nat Genet 2010;42(3):197–9.

23. Liu Y, Cao L, Li Z, et al. A genome-wide association study identifies a locus on TERT for mean telomere length in Han Chinese. PLoS One 2014;9(1):e85043.

24. Saxena R, Bjonnes A, Prescott J, et al. Genome-wide association study identifies variants in casein kinase II (CSNK2A2) to be associated with leukocyte telomere length in a Punjabi Sikh diabetic cohort. Circ Cardiovasc Genet 2014;7(3):287–95.

25. Levy D, Neuhausen SL, Hunt SC, et al. Genome-wide association identifies OBFC1 as a locus involved in human leukocyte telomere biology. Proc Natl Acad Sci U S A 2010;107(20):9293–8.

26. Hindorff LA LA, MacArthur J, Morales J, et al. A catalog of published genome-wide association studies [Internet]. [cited 2015 Jan 15];Available from: www.genome.gov/gwastudies

27. Chene G, Thompson SG. Methods for Summarizing the Risk Associations of Quantitative Variables in Epidemiologic Studies in a Consistent Form. Am J Epidemiol 1996;144(6):610–21.

28. Burgess S, Scott RA, Timpson NJ, Davey Smith G, Thompson SG, EPIC-InterAct Consortium. Using published data in Mendelian randomization : a blueprint for efficient identification of causal risk factors. Eur J Epidemiol 2015;30(7):543–52.

29. Thomas DC, Lawlor DA, Thompson JR. Re: Estimation of bias in nongenetic observational studies using “Mendelian triangulation” by Bautista et al. Ann Epidemiol 2007;17(7):511–3.

30. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. Int J Epidemiol 2015;44(2):512–25.

31. VanderWeele TJ, Tchetgen Tchetgen EJ, Cornelis M, Kraft P. Methodological challenges in mendelian randomization. Epidemiology 2014;25(3):427–35.

32. Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent estimation in Mendelian randomization with some invalid instruments using a weighted median estimator. Genet Epidemiol

33. National Cancer Institute. Surveillance, Epidemiology, and End Results Program [Internet]. [cited 2015 Aug 1];Available from: www.seer.cancer.gov

34. Tomasetti C, Vogelstein B. Variation in cancer risk among tissues can be explained by the number of stem cell divisions. Science 2015;347(6217):78–81.

35. U.S. Census Bureau [Internet]. [cited 2016 Jul 11];Available from: U.S. Census Bureau

36. R Core Team. A language and environment for statistical computing. 2013;

37. Burgess S, Thompson SG. Bias in causal estimates from Mendelian randomization studies with weak instruments. Stat Med 2011;30(11):1312–23.

38. Davey Smith G, Ebrahim S. What can mendelian randomisation tell us about modifiable behavioural and environmental exposures? BMJ Br Med J 2005;330(7499):1076–9.

39. Anic GM, Sondak VK, Messina JL, et al. Telomere length and risk of melanoma, squamous cell carcinoma, and basal cell carcinoma. Cancer Epidemiol 2013;37(4):434–9.

40. Pellatt AJ, Wolff RK, Torres-Mejia G, et al. Telomere length, telomere-related genes, and breast cancer risk: the breast cancer health disparities study. Genes Chromosomes Cancer 2013;52(7):595–609.

41. Caini S, Raimondi S, Johansson H, et al. Telomere length and the risk of cutaneous melanoma and non-melanoma skin cancer: a review of the literature and meta-analysis. J Dermatol Sci 2015;80(3):168–74.

42. Sanchez-Espiridion B, Chen M, Chang JY, et al. Telomere length in peripheral blood leukocytes and lung cancer risk: a large case-control study in Caucasians. Cancer Res 2014;74(9):2476–86.

43. Campa D, Mergarten B, De Vivo I, et al. Leukocyte telomere length in relation to pancreatic cancer risk: a prospective study. Cancer Epidemiol Biomarkers Prev 2014;23(11):2447–54.

44. Cui Y, Cai Q, Qu S, et al. Association of leukocyte telomere length with colorectal cancer risk: nested case-control findings from the Shanghai Women's Health Study. Cancer Epidemiol Biomarkers Prev 2012;21(10):1807–13.

45. De Vivo I, Prescott J, Wong JYY, Kraft P, Hankinson SE, Hunter DJ. A prospective study of relative telomere length and postmenopausal breast cancer risk. Cancer Epidemiol Biomarkers Prev 2009;18(4):1152–6.

46. Han J, Qureshi AA, Prescott J, et al. A prospective study of telomere length and the risk of skin cancer. J Invest Dermatol 2009;129(2):415–21.

47. Hofmann JN, Lan Q, Cawthon R, et al. A prospective study of leukocyte telomere

length and risk of renal cell carcinoma. Cancer Epidemiol Biomarkers Prev 2013;22(5):997–1000.

48. Julin B, Shui I, Heaphy CM, et al. Circulating leukocyte telomere length and risk of overall and aggressive prostate cancer. Br J Cancer 2015;112(4):769–76.

49. Kim S, Sandler DP, Carswell G, et al. Telomere length in peripheral blood and breast cancer risk in a prospective case-cohort analysis: results from the Sister Study. Cancer Causes Control 2011;22(7):1061–6.

50. Lan Q, Cawthon R, Shen M, et al. A prospective study of telomere length measured by monochrome multiplex quantitative PCR and risk of non-Hodgkin lymphoma. Clin Cancer Res 2009;15(23):7429–33.

51. Lee I-M, Lin J, Castonguay AJ, Barton NS, Buring JE, Zee RYL. Mean leukocyte telomere length and risk of incident colorectal carcinoma in women: a prospective, nested case-control study. Clin Chem Lab Med 2010;48(2):259–62.

52. Liang G, Qureshi AA, Guo Q, De Vivo I, Han J. No association between telomere length in peripheral blood leukocytes and the risk of nonmelanoma skin cancer. Cancer Epidemiol Biomarkers Prev 2011;20(5):1043–5.

53. Lynch SM, Major JM, Cawthon R, et al. A prospective analysis of telomere length and pancreatic cancer in the alpha-tocopherol beta-carotene cancer (ATBC) prevention study. Int J Cancer 2013;133(11):2672–80.

54. McGrath M, Wong JYY, Michaud D, Hunter DJ, De Vivo I. Telomere length, cigarette smoking, and bladder cancer risk in men and women. Cancer Epidemiol Biomarkers Prev 2007;16(4):815–9.

55. Nan H, Du M, De Vivo I, et al. Shorter telomeres associate with a reduced risk of melanoma development. Cancer Res 2011;71(21):6758–63.

56. Prescott J, McGrath M, Lee I-M, Buring JE, De Vivo I. Telomere length and genetic analyses in population-based studies of endometrial cancer risk. Cancer 2010;116(18):4275–82.

57. Qu S, Wen W, Shu X-O, et al. Association of leukocyte telomere length with breast cancer risk: nested case-control findings from the Shanghai Women's Health Study. Am J Epidemiol 2013;177(7):617–24.

58. Risques RA, Vaughan TL, Li X, et al. Leukocyte telomere length predicts cancer risk in Barrett's esophagus. Cancer Epidemiol Biomarkers Prev 2007;16(12):2649–55.

59. Seow WJ, Cawthon RM, Purdue MP, et al. Telomere length in white blood cell DNA and lung cancer: a pooled analysis of three prospective cohorts. Cancer Res 2014;74(15):4090–8.

60. Shen M, Cawthon R, Rothman N, et al. A prospective study of telomere length measured by monochrome multiplex quantitative PCR and risk of lung cancer. Lung Cancer 2011;73(2):133–7.

61. Walcott F, Rajaraman P, Gadalla SM, et al. Telomere length and risk of glioma. Cancer Epidemiol 2013;37(6):935–8.

62. Zee RYL, Castonguay AJ, Barton NS, Buring JE. Mean telomere length and risk of

incident colorectal carcinoma: a prospective, nested case-control approach. Cancer Epidemiol Biomarkers Prev 2009;18(8):2280–2.

63. Willeit P, Willeit J, Mayr A, et al. Telomere length and risk of incident cancer and cancer mortality. Jama 2010;304(1538–3598 (Electronic)):69–75.

64. Zhang C, Doherty J a., Burgess S, et al. Genetic determinants of telomere length and risk of common cancers: a Mendelian randomization study. Hum Mol Genet 2015;24(18):5356–66.

65. Iles MM, Bishop DT, Taylor JC, et al. The effect on melanoma risk of genes previously associated with telomere length. J Natl Cancer Inst 2014;106(10).

66. Walsh KM, Codd V, Rice T, et al. Longer genotypically-estimated leukocyte telomere length is associated with increased adult glioma risk. Oncotarget 2015;6(40):42468–77.

67. Ojha J, Codd V, Nelson CP, et al. Genetic Variation Associated with Longer Telomere Length Increases Risk of Chronic Lymphocytic Leukemia. Cancer Epidemiol Biomarkers Prev 2016;25(7):1043–9.

68. Armanios M, Blackburn EH. The telomere syndromes. Nat Rev Genet 2012;13(10):693–704.

69. Armanios M. Syndromes of telomere shortening. Annu Rev Genomics Hum Genet 2009;10(46):45–61.

70. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell 2011;144(5):646–74.

71. Blasco MA. Telomere length, stem cells and aging. Nat Chem Biol 2007;3(10):640–9.

72. Stone RC, Horvath K, Kark JD, Susser E, Tishkoff SA, Aviv A. Telomere Length and the Cancer-Atherosclerosis Trade-Off. PLoS Genet 2016;12(7):e1006144.

73. Pierce BL, Burgess S. Efficient design for Mendelian randomization studies: subsample and 2-sample instrumental variable estimators. Am J Epidemiol 2013;178(7):1177–84.

74. Burgess S, Butterworth AS, Thompson JR. Beyond Mendelian randomization: how to interpret evidence of shared genetic predictors. J Clin Epidemiol 2015;1–9.

75. Centers for Disease Control and Prevention [Internet]. [cited 2016 Jul 14];Available from: http://www.cdc.gov/nchs/fastats/deaths.htm

76. Rafnar T, Sulem P, Thorleifsson G, et al. Genome-wide association study yields variants at 20p12.2 that associate with urinary bladder cancer. Hum Mol Genet 2014;23(20):5545–57.

77. Michailidou K, Hall P, Gonzalez-Neira A, et al. Large-scale genotyping identifies 41 new loci associated with breast cancer risk. Nat Genet 2013;45(4):353–61, 361-2.

78. Schumacher FR, Schmit SL, Jiao S, et al. Genome-wide association study of colorectal cancer identifies six new susceptibility loci. Nat Commun 2015;6:7138.

79. Spurdle AB, Thompson DJ, Ahmed S, et al. Genome-wide association study identifies a common variant associated with risk of endometrial cancer. Nat Genet

2011;43(5):451–4.

80. Painter JN, O'Mara TA, Batra J, et al. Fine-mapping of the HNF1B multicancer locus identifies candidate variants that mediate endometrial cancer risk. Hum Mol Genet 2015;24(5):1478–92.

81. Abnet CC, Freedman ND, Hu N, et al. A shared susceptibility locus in PLCE1 at 10q23 for gastric adenocarcinoma and esophageal squamous cell carcinoma. Nat Genet 2010;42(9):764–7.

82. Wrensch M, Jenkins RB, Chang JS, et al. Variants in the CDKN2B and RTEL1 regions are associated with high-grade glioma susceptibility. Nat Genet 2009;41(8):905–8.

83. McKay JD, Truong T, Gaborieau V, et al. A genome-wide association study of upper aerodigestive tract cancers conducted within the INHANCE consortium. PLoS Genet 2011;7(3):e1001333.

84. Purdue MP, Johansson M, Zelenika D, et al. Genome-wide association study of renal cell carcinoma identifies two susceptibility loci on 2p21 and 11q13.3. Nat Genet 2010;43(1):60–5.

85. Wang Y, McKay JD, Rafnar T, et al. Rare variants of large effect in BRCA2 and CHEK2 affect risk of lung cancer. Nat Genet 2014;46(7).

86. Law MH, Bishop DT, Lee JE, et al. Genome-wide meta-analysis identifies five new susceptibility loci for cutaneous malignant melanoma. Nat Genet 2015;47(9):987–95.

87. Zhang M, Song F, Liang L, et al. Genome-wide association studies identify several new loci associated with pigmentation traits and skin cancer risk in European Americans. Hum Mol Genet 2013;22(14):2948–59.

88. Diskin SJ, Capasso M, Schnepp RW, et al. Common variation at 6q16 within HACE1 and LIN28B influences susceptibility to neuroblastoma. Nat Genet 2012;44(10):1126–30.

89. Pharoah PDP, Tsai Y-Y, Ramus SJ, et al. GWAS meta-analysis and replication identifies three new susceptibility loci for ovarian cancer. Nat Genet 2013;45(4):362–70, 370-2.

90. Wolpin BM, Rizzato C, Kraft P, et al. Genome-wide association study identifies multiple susceptibility loci for pancreatic cancer. Nat Genet 2014;46(9):994–1000.

91. Eeles RA, Olama AA Al, Benlloch S, et al. Identification of 23 new prostate cancer susceptibility loci using the iCOGS custom genotyping array. Nat Genet 2013;45(4):385–91, 391-2.

92. Al Olama AA, Kote-Jarai Z, Berndt SI, et al. A meta-analysis of 87,040 individuals identifies 23 new susceptibility loci for prostate cancer. Nat Genet 2014;46(10):1103–9.

93. Turnbull C, Rapley E a, Seal S, et al. Variants near DMRT1, TERT and ATF7IP are associated with testicular germ cell cancer. Nat Genet 2010;42(7):604–7.

94. Rapley EA, Turnbull C, Al Olama AA, et al. A genome-wide association study of testicular germ cell tumor. Nat Genet 2009;41(7):807–10.

95. Betz RC, Petukhova L, Ripke S, et al. Genome-wide meta-analysis in alopecia areata resolves HLA associations and reveals two new susceptibility loci. Nat Commun 2015;6:5966.

96. EArly Genetics and Lifecourse Epidemiology (EAGLE) Eczema Consortium, Australian Asthma Genetics Consortium(AAGC), Australian Asthma Genetics Consortium AAGC. Multi-ancestry genome-wide association study of 21,000 cases and 95,000 controls identifies new risk loci for atopic dermatitis. Nat Genet 2015;47(12):1449–56.

97. Dubois PC a, Trynka G, Franke L, et al. Multiple common variants for celiac disease influencing immune gene expression. Nat Genet 2010;42(4):295–302.

98. Liu JZ, van Sommeren S, Huang H, et al. Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. Nat Genet 2015;47(9):979–86.

99. Thompson SD, Marion MC, Sudman M, et al. Genome-wide association analysis of juvenile idiopathic arthritis identifies a new susceptibility locus at chromosomal region 3q13. Arthritis Rheum 2012;64(8):2781–91.

100. Beecham AH, Patsopoulos NA, Xifara DK, et al. Analysis of immune-related loci identifies 48 new susceptibility variants for multiple sclerosis. Nat Genet 2013;45(11):1353–60.

101. Schaefer AS, Richter GM, Nothnagel M, et al. A genome-wide association study identifies GLT6D1 as a susceptibility locus for periodontitis. Hum Mol Genet 2010;19(3):553–62.

102. Stahl EA, Raychaudhuri S, Remmers EF, et al. Genome-wide association study meta-analysis identifies seven new rheumatoid arthritis risk loci. Nat Genet 2010;42(6):508–14.

103. Bown MJ, Jones GT, Harrison SC, et al. Abdominal aortic aneurysm is associated with a variant in low-density lipoprotein receptor-related protein 1. Am J Hum Genet 2011;89(5):619–27.

104. Gretarsdottir S, Baas AF, Thorleifsson G, et al. Genome-wide association study identifies a sequence variant within the DAB2IP gene conferring susceptibility to abdominal aortic aneurysm. Nat Genet 2010;42(8):692–7.

105. Jones GT, Bown MJ, Gretarsdottir S, et al. A sequence variant associated with sortilin-1 (SORT1) on 1p13.3 is independently associated with abdominal aortic aneurysm. Hum Mol Genet 2013;22(14):2941–7.

106. Harrison SC, Smith AJP, Jones GT, et al. Interleukin-6 receptor pathways in abdominal aortic aneurysm. Eur Heart J 2013;34(48):3707–16.

107. Elmore JR, Obmann MA, Kuivaniemi H, et al. Identification of a genetic variant associated with abdominal aortic aneurysms on chromosome 3p12.3 by genome wide association. J Vasc Surg 2009;49(6):1525–31.

108. Borthwick K, Smelser D, Bock J, et al. Ephenotyping for Abdominal Aortic Aneurysm in the Electronic Medical Records and Genomics (eMERGE) Network: Algorithm Development and Konstanz Information Miner Workflow. Int J Biomed Data Min 2015;4(1).

741 109. Schunkert H, König IR, Kathiresan S, et al. Large-scale association analysis identifies
742 13 new susceptibility loci for coronary artery disease. Nat Genet 2011;43(4):333–8.

743 110. Smith NL, Felix JF, Morrison AC, et al. Association of genome-wide variation with
744 the risk of incident heart failure in adults of European and African ancestry: a
745 prospective meta-analysis from the cohorts for heart and aging research in genomic
746 epidemiology (CHARGE) consortium. Circ Cardiovasc Genet 2010;3(3):256–66.

747 111. Woo D, Falcone GJ, Devan WJ, et al. Meta-analysis of genome-wide association
748 studies identifies 1q22 as a susceptibility locus for intracerebral hemorrhage. Am J
749 Hum Genet 2014;94(4):511–21.

750 112. Malik R, Freilinger T, Winsvold BS, et al. Shared genetic basis for migraine and
751 ischemic stroke: A genome-wide analysis of common variants. Neurology
752 2015;84(21):2132–45.

753 113. Traylor M, Farrall M, Holliday EG, et al. Genetic risk factors for ischaemic stroke and
754 its subtypes (the METASTROKE collaboration): a meta-analysis of genome-wide
755 association studies. Lancet Neurol 2012;11(11):951–62.

756 114. Barrett JC, Clayton DG, Concannon P, et al. Genome-wide association study and
757 meta-analysis find that over 40 loci affect risk of type 1 diabetes. Nat Genet
758 2009;41(6):703–7.

759 115. Burren OS, Adlem EC, Achuthan P, Christensen M, Coulson RMR, Todd JA.
760 T1DBase: update 2011, organization and presentation of large-scale data sets for type
761 1 diabetes research. Nucleic Acids Res 2011;39(Database):D997–1001.

762 116. Morris ADPDPA, Voight BFB, Teslovich TMT, et al. Large-scale association analysis
763 provides insights into the genetic architecture and pathophysiology of type 2 diabetes.
764 Nat Genet 2012;44(9):981–90.

765 117. Fritsche LG, Chen W, Schu M, et al. Seven new loci associated with age-related
766 macular degeneration. Nat Genet 2013;45(4):433–9, 439-2.

767 118. Jensen RA, Sim X, Li X, et al. Genome-wide association study of retinopathy in
768 individuals without diabetes. PLoS One 2013;8(2):e54232.

769 119. Moffatt MF, Kabesch M, Liang L, et al. Genetic variants regulating ORMDL3
770 expression contribute to the risk of childhood asthma. Nature 2007;448(7152):470–3.

771 120. Ferreira MAR, Matheson MC, Duffy DL, et al. Identification of IL6R and
772 chromosome 11q13.5 as risk loci for asthma. Lancet (London, England)
773 2011;378(9795):1006–14.

774 121. Cho MH, McDonald M-LN, Zhou X, et al. Risk loci for chronic obstructive pulmonary
775 disease: a genome-wide association study and meta-analysis. Lancet Respir Med
776 2014;2(3):214–25.

777 122. Fingerlin TE, Murphy E, Zhang W, et al. Genome-wide association study identifies
778 multiple susceptibility loci for pulmonary fibrosis. Nat Genet 2013;45(6):613–20.

779 123. Fogh I, Ratti A, Gellera C, et al. A genome-wide association meta-analysis identifies a
780 novel locus at 17q11.2 associated with sporadic amyotrophic lateral sclerosis. Hum
781 Mol Genet 2014;23(8):2220–31.

124. Lambert JC, Ibrahim-Verbaas CA, Harold D, et al. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. Nat Genet 2013;45(12):1452–8.

125. Boraska V, Franklin CS, Floyd JAB, et al. A genome-wide association study of anorexia nervosa. Mol Psychiatry 2014;19(10):1085–94.

126. Smoller JW, Hospital MG. Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. Lancet 2013;381(9875):1371–9.

127. Sklar P, Ripke S, Scott LJ, et al. Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near ODZ4. Nat Genet 2011;43(10):977–83.

128. Ripke S, Wray NR, Lewis CM, et al. A mega-analysis of genome-wide association studies for major depressive disorder. Mol Psychiatry 2013;18(4):497–511.

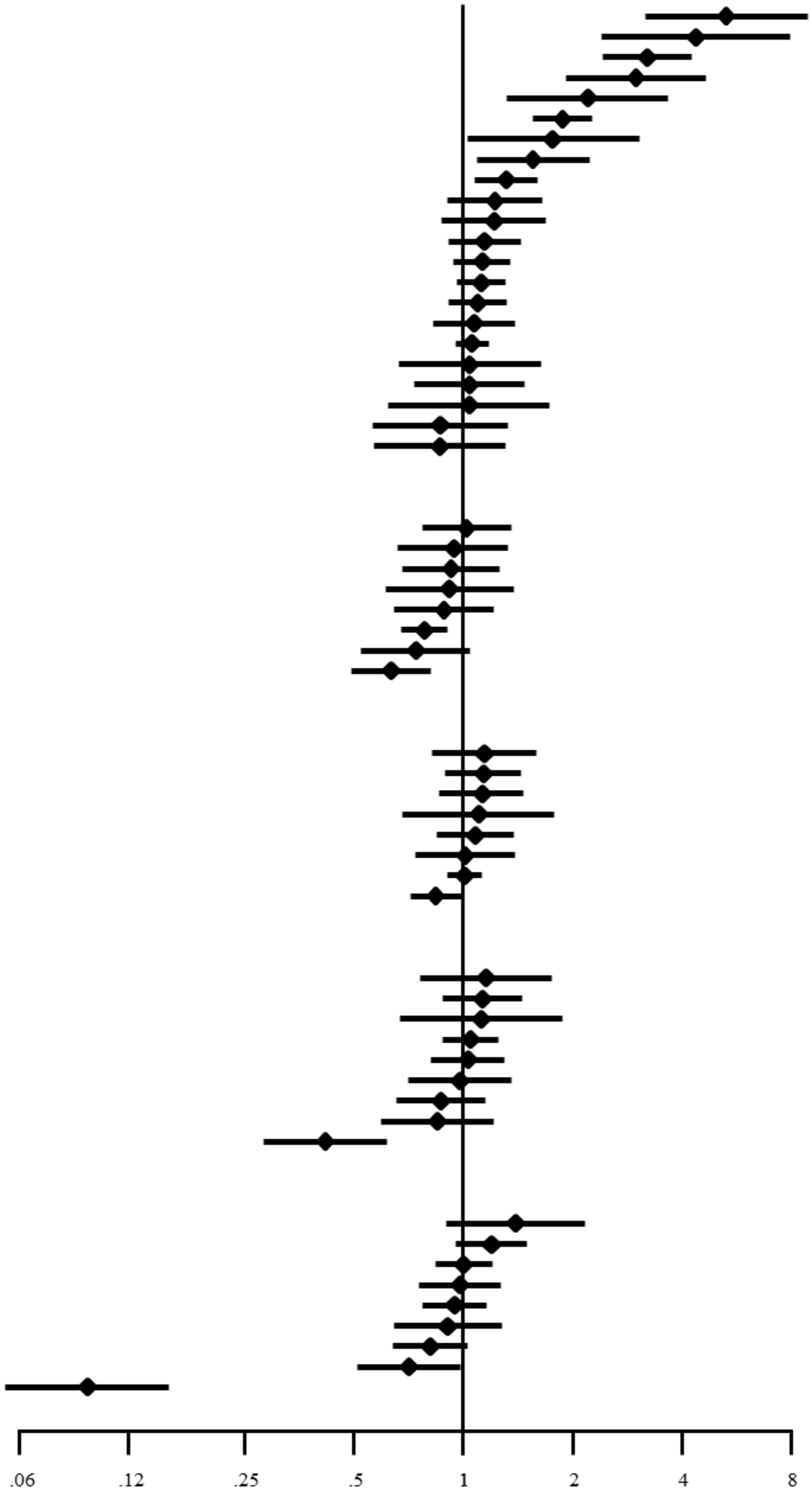
129. Ripke S, Neale BM, Corvin A, et al. Biological insights from 108 schizophrenia-associated genetic loci. Nature 2014;511(7510):421–7.

130. Scharf JM, Yu D, Mathews CA, et al. Genome-wide association study of Tourette's syndrome. Mol Psychiatry 2013;18(6):721–8.

131. Köttgen A, Pattaro C, Böger C a, et al. New loci associated with kidney function and chronic kidney disease. Nat Genet 2010;42(5):376–84.

132. Nyholt DR, Low S-K, Anderson CA, et al. Genome-wide association meta-analysis identifies new endometriosis risk loci. Nat Genet 2012;44(12):1355–9.

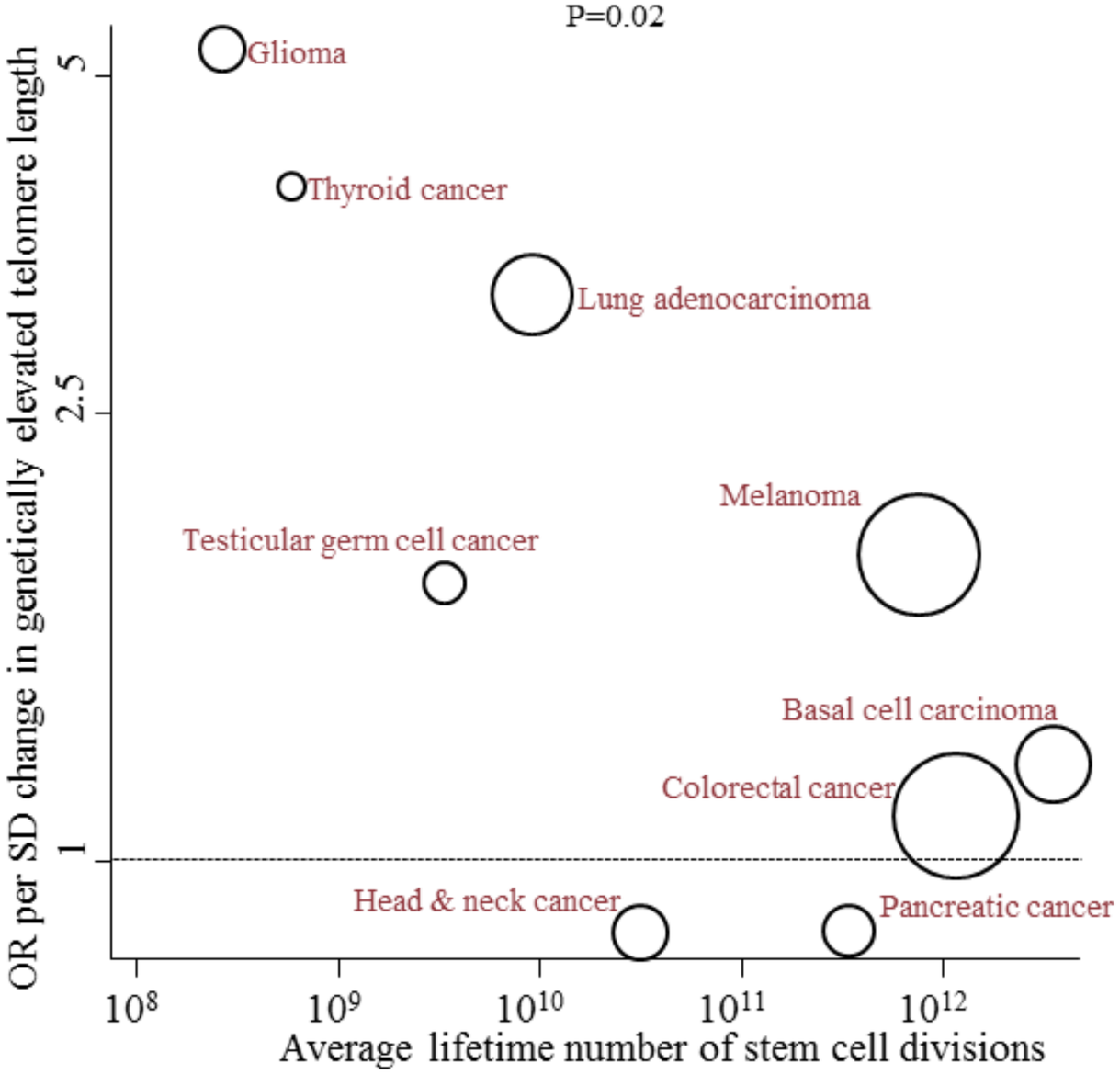
	Tumor/disease subtype	No. of cases	No. of SNPs	P*	P_{het}
Cancer					
Glioma	-	1130	12	$2.45x10^{-10}$	.0135
Ovarian cancer	Serous LMP	972	13	$1.62x10^{-6}$	.082
Lung cancer	Adenocarcinoma	3442	13	$6.66x10^{-16}$	.0001
Neuroblastoma	-	2101	12	$1.11x10^{-6}$	.3799
Bladder cancer	-	1601	10	.0026	.2541
Skin cancer	Melanoma	12814	13	$5.79x10^{-11}$	.086
Testicular germ cell cancer	-	985	11	.0423	.3002
Kidney cancer	-	2461	12	.0164	.251
Endometrial cancer	-	6607	12	.0091	.1436
Skin cancer	Basal cell carcinoma	3361	13	.203	.0398
Ovarian cancer	Endometriod	2154	13	.2499	.3319
Breast cancer	ER-ve	7465	13	.2542	<0.0001
Ovarian cancer	Serous invasive	9608	13	.2088	.0407
Prostate cancer	-	22297	11	.1533	<0.0001
Colorectal cancer	-	14537	9	.3436	.0156
Lung cancer	Squamous cell carcinoma	3275	13	.6212	.9042
Breast cancer	ER+	27074	13	.2912	.0228
Ovarian cancer	Clear cell	1016	13	.8676	.3226
Ovarian cancer	Mucinous	1643	13	.8396	.3147
Esophageal cancer	Squamous cell carcinoma	1942	8	.9055	.4513
Pancreatic cancer	Adenocarcinoma	5105	12	.5009	.0016
Head & neck cancer	-	2082	12	.4767	.0763
Cardiovascular diseases					
Heart failure	-	2526	13	.8931	.5348
Ischemic stroke	Small vessel disease	1894	13	.7141	.2358
Sudden cardiac arrest	-	3954	13	.62	.5400
Intracerebral stroke	-	5503	12	.6719	.0002
Ischemic stroke	Cardioembolic	2365	13	.4464	.0374
Coronary heart disease	-	22233	13	.0009	.2441
Ischemic stroke	Large vessel disease	2167	13	.0801	.1578
Abdominal aortic aneurysm	-	4972	13	.0003	.0617
Neurological / psychiatric diseases					
Anorexia Nervosa	-	2907	9	.4406	.1797
Bipolar disorder	-	7481	9	.315	.1427
Amyotrophic lateral sclerosis	-	6100	12	.4049	.6331
Tourette syndrome	-	1177	13	.6936	.4978
Major depressive disorder	-	9240	8	.5646	.254
Autism	-	4949	7	.9514	.8232
Schizophrenia	-	9379	12	.9245	.0053
Alzheimer's disease	-	17008	12	.0296	.1451
Autoimmune/inflammatory diseases					
Alopecia areata	-	2332	7	.4994	.9505
Inflammatory bowel disease	Crohn's disease	5956	11	.3586	.0585
Periodontitis	-	888	13	.6702	.0066
Atopic dermatitis	-	10788	13	.6094	.1194
Inflammatory bowel disease	Ulcerative colitis	6968	12	.8303	.0984
Multiple sclerosis	-	14498	3	.8885	.1455
Rheumatoid arthritis	-	5538	11	.3184	.1828
Juvenile idiopathic arthritis	-	1866	11	.3669	.922
Celiac disease	-	4533	3	$1.00x10^{-5}$	.0028
Other diseases					
Retinopathy	-	1126	12	.1453	.3831
Age-related macular degeneration		7473	13	.1206	.209
Type 2 diabetes	-	10415	11	.9837	.6811
Endometriosis	-	4604	11	.8606	.0576
Chronic kidney disease	-	5807	13	.5859	.6972
COPD	-	2812	12	.5598	.3807
Asthma	-	13034	4	.0779	.3213
Type 1 diabetes	-	7514	6	.0378	.0586
Interstitial lung disease	-	1616	9	$2.02x10^{-19}$	.0014



Average lifetime number of stem cell divisions

Adjusted R^2=52.63%

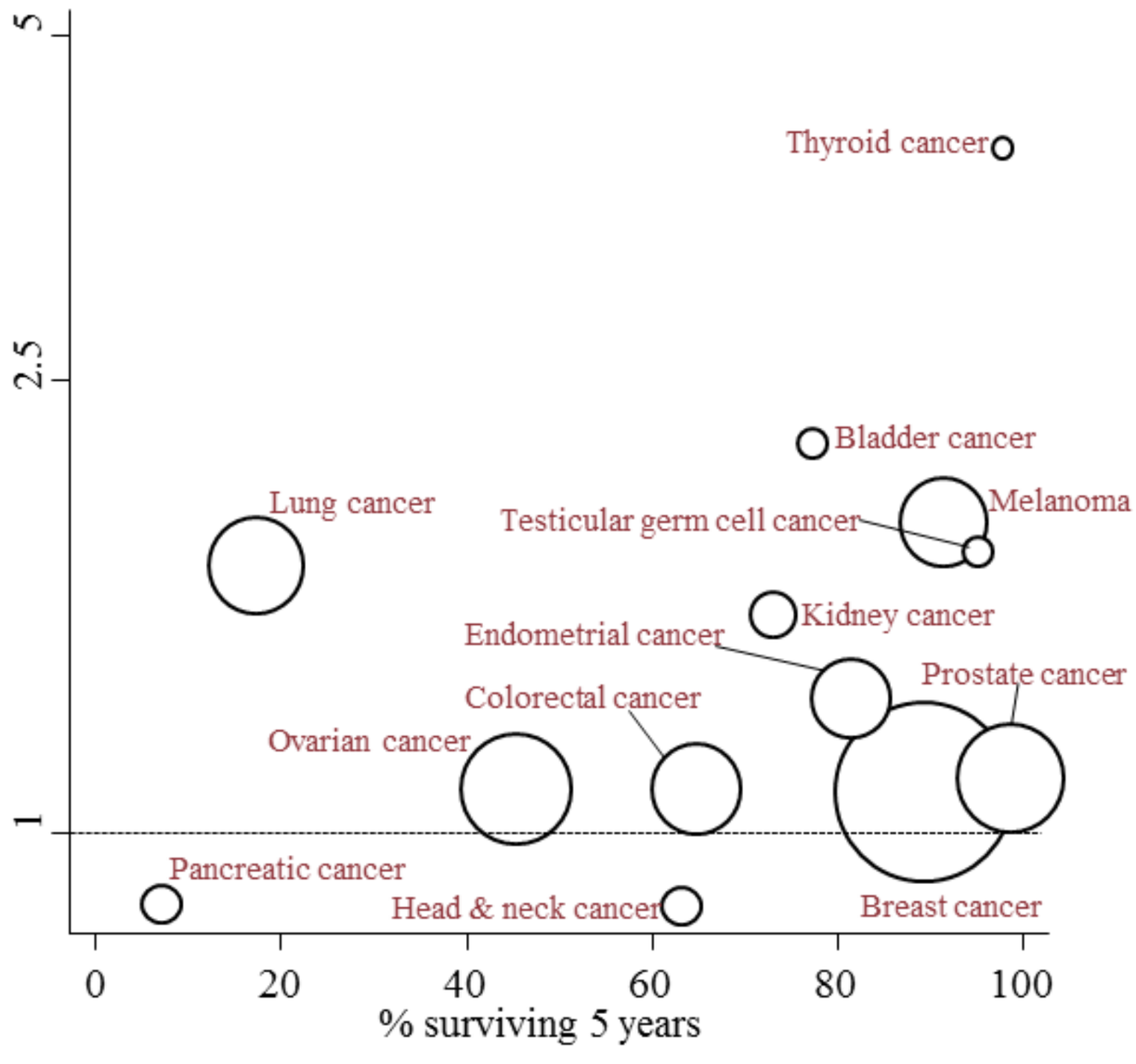
P=0.02



% surviving 5 years

Adjusted R^2=-19.49%

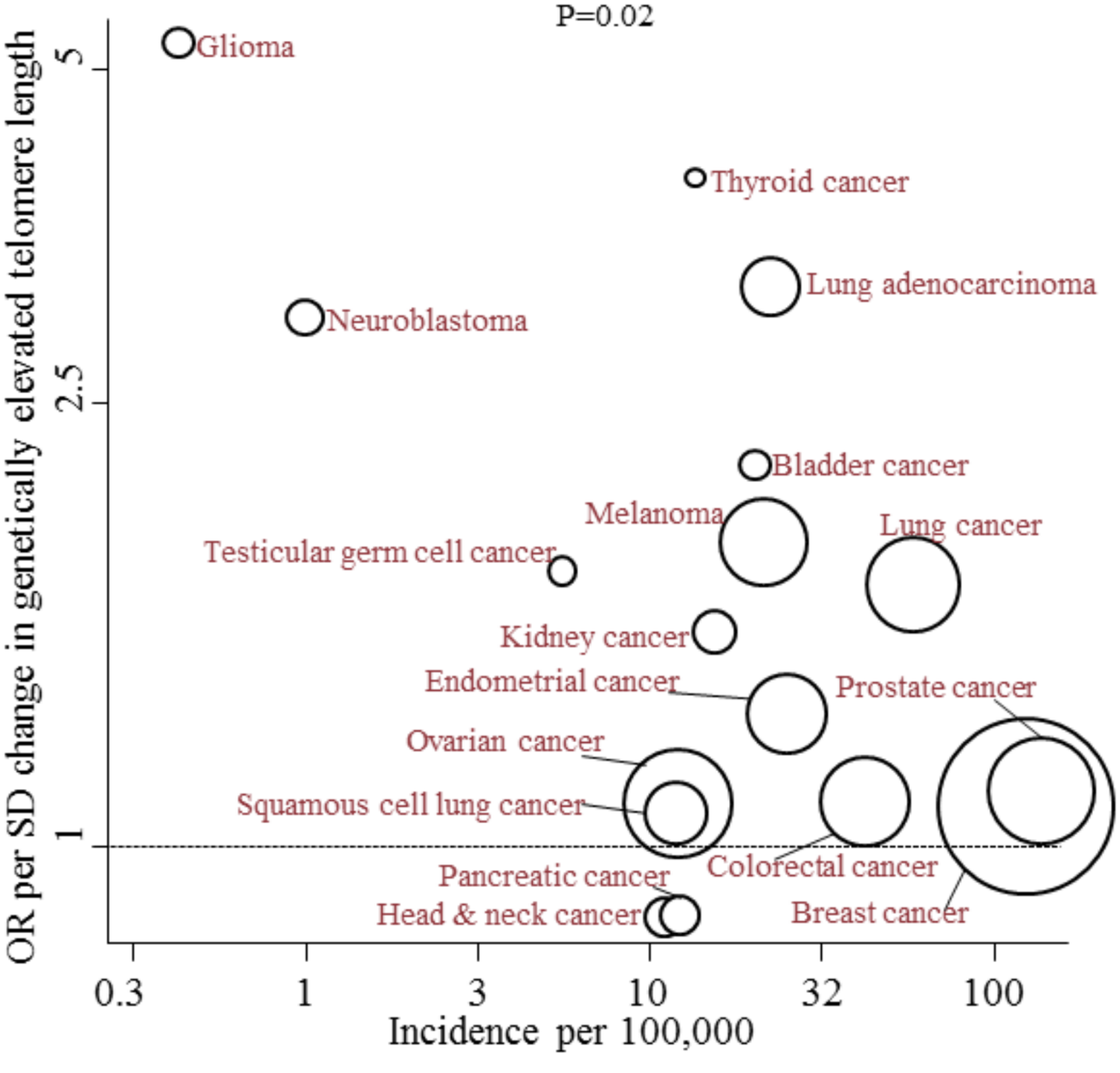
P=0.37



Cancer incidence

Adjusted R^2=32.12%

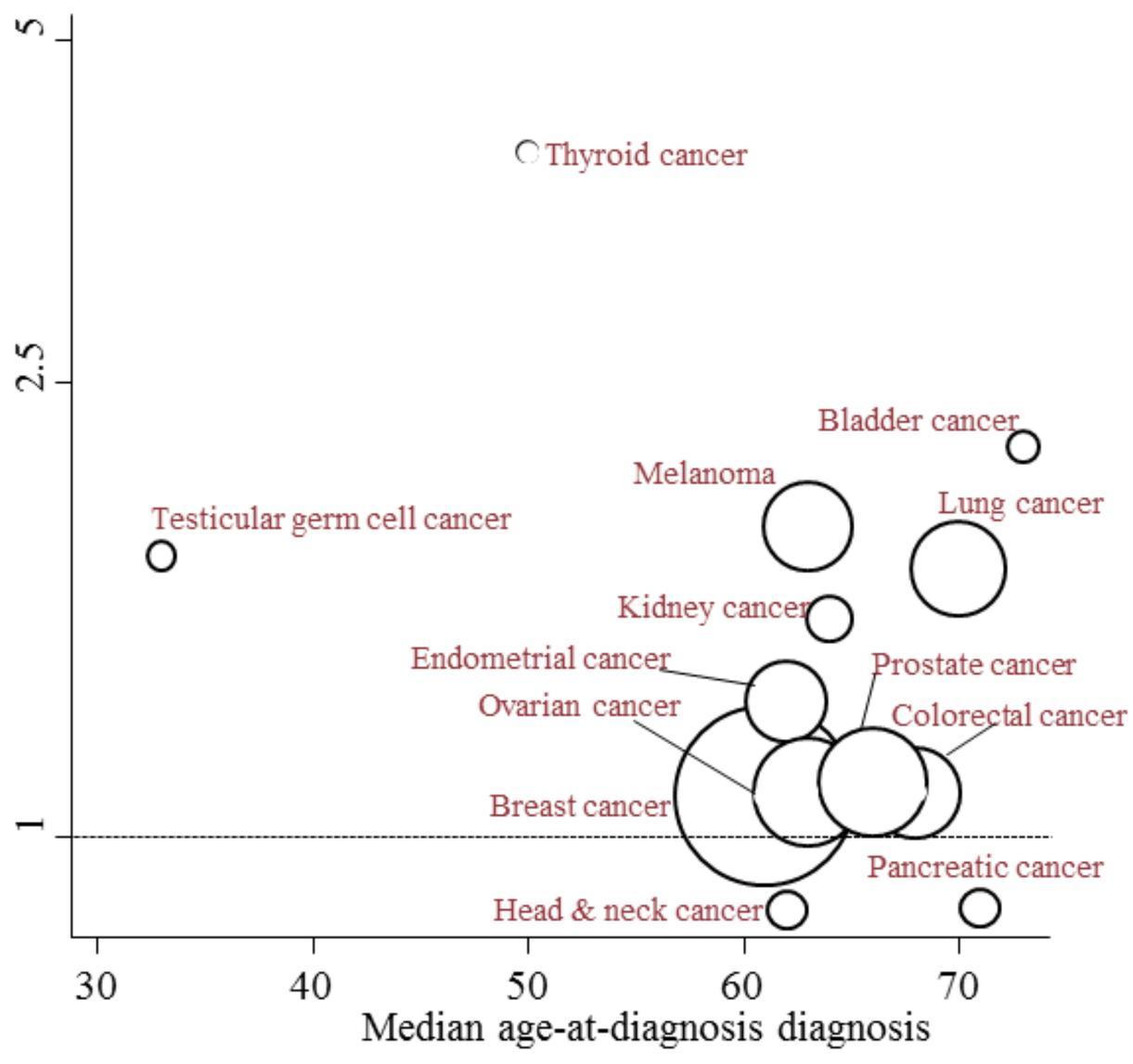
P=0.02



Median age-at-diagnosis diagnosis

Adjusted R^2=-8.86%

P=0.41



	No. of cases	Odds ratio (95% CI) per standard deviation increase in telomere length
Cancer		
Breast cancer		
Genetic study	48155	1.08 (0.99, 1.19)
Observational study*	1716	1.02 (0.99, 1.05)
Prostate cancer		
Genetic study	22297	1.12 (0.96, 1.30)
Observational study*	1340	1.07 (1.01, 1.14)
Ovarian cancer		
Genetic study	15397	1.09 (0.94, 1.27)
Observational study	96	1.13 (0.98, 1.32)
Colorectal cancer		
Genetic study	14537	1.09 (0.91, 1.31)
Observational study*	1447	1.04 (0.97, 1.11)
Lung cancer		
Genetic study	11348	1.71 (1.44, 2.04)
Observational study‡	522	0.94 (0.87, 1.02)
Observational study+	847	1.28 (1.12, 1.46)
Endometrial cancer		
Genetic study	6608	1.31 (1.07, 1.61)
Observational study*	382	1.06 (0.95, 1.19)
Pancreatic cancer		
Genetic study	5105	0.86 (0.56, 1.32)
Observational study*	648	1.05 (0.95, 1.17)
Lung adenocarcinoma		
Genetic study	3442	3.19 (2.40, 4.22)
Observational study	288	1.44 (1.14, 1.82)
Skin basal cell carcinoma		
Genetic study	3361	1.22 (0.90, 1.65)
Observational study	363	0.96 (0.85, 1.09)
Lung squamous cell carcinoma		
Genetic study	3275	1.07 (0.82, 1.39)
Observational study	163	1.05 (0.78, 1.42)
Kidney cancer		
Genetic study	2461	1.55 (1.08, 2.23)
Observational study*	268	0.94 (0.81, 1.10)
Head & neck cancer		
Genetic study	2082	0.86 (0.57, 1.30)
Observational study	76	0.89 (0.72, 1.09)
Melanoma		
Genetic study	12814	1.88 (1.54, 2.28)
Observational study*	734	1.17 (1.06, 1.29)
Bladder cancer		
Genetic study	1601	2.19 (1.32, 3.66)
Observational study	184	1.28 (1.02, 1.61)
Testicular cancer		
Genetic study	986	1.76 (1.02, 3.04)
Observational study	10	0.94 (0.56, 1.55)
Glioma		
Genetic study	1130	5.27 (3.15, 8.81)
Observational study	101	0.90 (0.68, 1.18)
Non-neoplastic diseases		
Coronary heart disease		
Genetic study	22233	0.78 (0.67, 0.90)
Observational study	2272	0.86 (0.78, 0.94)
Ischemic stroke		
Genetic study	12389	0.85 (0.73, 1.00)
Observational study	824	0.94 (0.82, 1.08)
Type 2 diabetes		
Genetic study	10415	1.00 (0.84, 1.20)
Observational study	2011	0.90 (0.83, 0.97)