The inverse association of body mass index with lung cancer: exploring residual confounding, metabolic aberrations and within-person variability in smoking in the Me-Can 2.0 study

Angela M Wood[1], Håkan Jonsson[2], Gabriele Nagel[3,4], Christel Häggström[5,6], Jonas Manjer[7], Hanno Ulmer[8], Anders Engeland[9,10], Emanuel Zitt[4,11], Sylvia Jochems[12], Sara Ghaderi[10], Pär Stattin[6], Tone Bjørge[10,13], Tanja Stocks[12]

[1]Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK

[2]Department of Radiation Sciences, Umeå University, Umeå, Sweden

[3]Institute of Epidemiology and Medical Biometry, Ulm University, Ulm, Germany

[4]Agency for Preventive and Social Medicine, Bregenz (aks), Austria

[5]Department of Public Health and Clinical Medicine, Umeå University, Umeå, Sweden

[6]Department of Surgical Sciences, Uppsala University, Uppsala, Sweden

[7]Department of Surgery, Skåne University Hospital, Lund University, Malmö, Sweden

[8]Department of Medical Statistics, Informatics and Health Economics, Innsbruck Medical University, Innsbruck, Austria

[9]Department of Chronic Diseases and Ageing, Norwegian Institute of Public Health, Bergen, Norway

[10]Department of Global Public Health and Primary Care, University of Bergen, Bergen, Norway

[11]Department of Internal Medicine, Academic Teaching Hospital Feldkirch, Feldkirch, Austria

[12]Department of Clinical Sciences Lund, Lund University, Lund, Sweden

[13]Cancer Registry of Norway, Oslo, Norway

Running title: Exploring the inverse association of BMI with lung cancer

Keywords: body mass index, metabolic syndrome, smoking, regression dilution bias, cancer risk

Financial support: This work was supported by the Crafoord Foundation (no. 20180660 to TS) and the Swedish Cancer Society (no. 2017/1019 to TS).

Correspondence to: Angela Wood, University of Cambridge, Department of Public Health and Primary Care, Strangeways Research Laboratory, Cambridge, CB1 8RN, UK, email: amw79@medschl.cam.ac.uk, or to Tanja Stocks, Lund University, Department of Clinical Sciences Lund, Barngatan 4, 221 85 Lund, Sweden, email: tanja.stocks@med.lu.se

Conflicts of interest: The authors declare no potential conflicts of interest.

Word count: 3,374

Numbers of Tables and Figures: 1 Table, 4 Figures.

Abstract

Background: The inverse observational association between body mass index (BMI) and lung cancer risk remains unclear. We assessed whether the association is explained by metabolic aberrations, residual confounding and within-person variability in smoking, and compared against other smoking-related cancers.

Methods: We investigated the association between BMI, and its combination with a metabolic score (MS) of mid-blood pressure, glucose and triglycerides, with lung cancer and other smoking-related cancers in 778,828 individuals. We used Cox regression, adjusted and corrected for within-person variability in smoking (status/pack-years), calculated from 600,201 measurements in 221,958 participants.

Results: Over a median follow-up of 20 years, 20,242 smoking-related cancers (6,735 lung cancers) were recorded. Despite adjustment and correction for substantial within-person variability in smoking, BMI remained inversely associated with lung cancer (hazard ratio per standard deviation increase, 0.87 [95% confidence interval 0.85-0.89]). Individuals with BMI<25 kg/m^2 and high MS had the highest risk (hazard ratio 1.52 [1.44-1.60] *vs* BMI≥25 with low MS). These associations were weaker and non-significant amongst non-smokers. Similar associations were observed for head and neck cancers and oesophageal squamous cell carcinoma, whereas for other smoking-related cancers, we generally observed positive associations with BMI.

Conclusions: The increased lung cancer risk with low BMI and high MS is unlikely due to residual confounding and within-person variability in smoking. However, similar results for other cancers strongly related to smoking suggests a remaining, unknown, effect of smoking.

Impact: Extensive smoking-adjustments may not capture all the effects of smoking on the relationship between obesity-related factors and risk of smoking-related cancers.

Introduction

Body mass index (BMI), a surrogate measure of obesity, has been related to higher risks of many cancer forms with some of the strongest associations found for smoking-related cancers, especially oesophageal (adenocarcinoma), renal cell and liver cancer (1-3). In contrast, a consistent inverse association has been reported for lung cancer (1-5). As smoking is strongly related to both lung cancer and lower body weight (6, 7), the inverse association has been proposed to be caused by residual confounding by insufficient adjustment for smoking (3, 8). Mendelian randomization studies, which under certain assumptions estimate causal associations (9), have not supported the inverse association with lung cancer (10-12), and several large observational studies showed no association among non-smokers (3, 5, 13). However, some studies did show an inverse association among non-smokers (14), or among smokers even after accounting for detailed smoking information (4, 5). Therefore, as alternative explanation for the inverse relationship between BMI and lung cancer, Renehan *et al.* suggested larger measurement error of smoking than of BMI (8). To our knowledge, this has not yet been investigated.

In contrast to BMI, waist circumference has been positively related to lung cancer risk (4, 15). A large prospective study by Yu *et al.* showed the highest lung cancer risk for low BMI and high waist circumference combined (4). This phenotype was associated with heavy smoking, but it was suggested that central obesity with lower muscle mass, and therefore retained BMI, and more metabolic aberrations potentially associated with lung cancer, could underlie these findings (4, 15). Investigating BMI and metabolic aberrations jointly, with extensive control for smoking habits, could clarify whether the association between BMI and lung cancer is dependent on the presence of metabolic aberrations, and whether the increased lung cancer risk

with low BMI could potentially be reflective of a sarcopenic phenotype with metabolic aberrations.

In this study, we investigated the association between BMI and its combination with a metabolic score comprising mid-blood pressure and circulating triglycerides and glucose, with lung cancer and other smoking-related cancers. By leveraging serial measurements on BMI and smoking information, we corrected for within-person variability (including measurement-errors and short- and long-term within-person variability) (16). Our aim was to determine how metabolic aberrations, residual confounding and within-person variability in smoking influence associations between BMI and lung cancer and other smoking-related cancers.

Materials and methods

Population

We focused our study on participants from Me-Can 2.0, constituting six population-based cohorts (Oslo study I, Norwegian Counties Study, Age 40-Programme, Västerbotten Intervention Project, Malmö Preventive Project, and Vorarlberg Health Monitoring and Prevention Programme [VHM&PP]), from Norway, Sweden and Austria. Data were available from 843,531 individuals with 1,557,855 serial health examinations, collected during 1972-2014. Me-Can 2.0 is a continuation of Me-Can 1.0 that has been previously described in detail (17). Compared to Me-Can 1.0 (publications in 2009-2015), Me-Can 2.0 does not include Cohort of Norway, but it includes the full Oslo I, Norwegian Counties Study and 40-year programme, and additional individuals and observations in the Västerbotten Intervention Project in 2006-2014 and in the VHM&PP in 2003-2005. Measurement methods of anthropometrics and metabolic factors have been previously described (17). However, in the Västerbotten Intervention Project, the measurements for blood pressure and triglycerides

changed on Sept 1, 2009. Before that date, blood pressure was measured in supine position, and was thereafter measured in sitting position. Serum triglycerides were measured on a Reflotron bench-top analyzer at the examining health care centre before Sept 1, 2009, and were thereafter measured at the Clinical Chemistry department at the nearest hospital using standard enzymatic methods. Formula for transformation of blood pressure and triglyceride levels measured before versus after Sept 1, 2009 have been calculated (n individuals=838 for triglycerides, and 648 for blood pressure), and these were applied in the present study. Triglyceride levels measured on Sept 1, 2009 onwards were converted to old measurement levels by: 0.177 + (0.932 × triglyceride level). Formula for blood pressure were age- and sex-specific (18).

Data on smoking status were included in Me-Can, and further information on tobacco smoking, including frequency and number of years smoking, were added to the Me-Can 2.0 database. In the Norwegian and Swedish cohorts, smoking information was retrieved from questionnaires. In the VHM&PP, smoking habits were requested orally by the physician performing the health examination, but the recording did not distinguish between missing information and non-smoker status. This likely led to a weaker association between baseline smoking and lung cancer risk in the VHM&PP than in other cohorts (**Supplementary Figure 1**). Consequently, participants in the VHM&PP cohort were excluded from subgroup analyses of non-smokers.

The study was conducted in accordance with the Declaration of Helsinki, and was approved by ethical committees in each country.

Follow-up

Linkages of participants were performed with the respective national or regional (VHM&PP) cancer register, the national cause of death register, and the national total population register

(Norway and Sweden) for information on cancer diagnoses, deaths, and emigration, including the year 2012 in Norway and 2014 in Sweden and Austria.

Categorisation of smoking-related cancers

Smoking-related cancers were defined as those listed by the International Agency for Research on Cancer (monograph 83 in 2004 (19) and additions in 2012 (20)) as probably caused by smoking and with a relative risk of smokers *vs* non-smokers of at least 1.2. *In situ* cancers were excluded except urothelial carcinoma *in situ*, which were included because they frequently progress to higher stage (21). We used International Classification of Diseases (ICD) codes to categorise cancers by topography (**Supplementary Table 1**), and morphology or histology codes were used if necessary for further subgrouping. In our population, the hazard ratio (HR) of cancer in smokers *vs* non-smokers was 7.4 or higher for cancers of the lung, larynx, and oesophageal squamous cell carcinoma (SCC) and 3.4 or lower for all other cancers.

Selection criteria

To be eligible for the study, participants had to have information on BMI, smoking status and pack-years, and no baseline history of a malignant cancer (excluding basalioma and *in situ* tumours, but including blood cancers [ICD-7 200-209] of uncertain or unknown behavior), resulting in 778,828 individuals with 1,264,393 serial measurements. For the combined analysis of BMI and metabolic score, we further analysed a subset of individuals with additional information on blood pressure, triglycerides and glucose (488,659 individuals with 876,618 serial measurements).

Our aim was to determine how metabolic aberrations, residual confounding and within-person variability in smoking influence associations between BMI and lung cancer and other smoking-related cancers.

Statistical analysis

The specific analyses performed to investigate how 1) metabolic aberrations, 2) residual confounding and 3) within-person variability in smoking influence associations between BMI and lung cancer and other smoking-related cancers are outlined in **Supplementary Table 2**, and are described in further detail below.

HR of cancer by BMI level was calculated using Cox regression with follow-up from baseline until the first cancer diagnosis, emigration, death, or end of follow-up, whichever came first. We used attained age as time-scale, stratified by cohort, sex and birth decade, with adjustment for baseline age, fasting time, smoking status (non-smoker, smoker and ex-smoker) and pack-years. There was no evidence of violation of the proportional hazards assumption, which was tested by including time interactions with the exposures. To assess possible reverse causation, we repeated the analyses excluding the first 5-20 years of follow-up. We assessed between study heterogeneity by performing analyses separately by cohort and with the I^2 statistic (22). We investigated effect modification by sex and baseline smoking status with Wald tests for interaction. To further visualise BMI-smoking interactions, we calculated associations between categories of BMI (<20, 20-22.49, 22.5-24.99, 25-27.49, 27.5-29.99, ≥30 kg/m^2) and smoking (status/pack-years) jointly in relation to cancer risk.

To correct for within-person variability in BMI and smoking status and pack-years, we estimated long-term average levels ("usual levels") using multi-level regression calibration and

information from up to 600,201 serial assessments in up to 230,454 individuals. This was achieved using linear mixed effects models, by regressing repeat measurements on baseline measures, adjusted for duration of follow-up and baseline levels of age, sex, fasting time (as appropriate), and for baseline levels of BMI, smoking status and pack-years (when not the independent variable) with random effects for cohort and repeat measurement (16, 23). The regression dilution ratio (RDR), i.e. the calibration slope, which measures the overall extent of within-person variability, was extracted from the calibration model. Further analyses allowing for age-dependent RDR corrections produced similar results and are not reported. Estimated usual levels of BMI, smoking status and pack-years were used directly in the Cox regression models to calculate corrected HRs.

We assessed the shape of the association with lung cancer by calculating HRs within tenths of BMI and plotted them against mean usual (and baseline) levels within each group. We estimated 95% confidence intervals (CIs) for each group (including the reference group) that corresponded to the amount of information underlying each group (24, 25). For each cancer outcome we determined the best fitting 1^{st} or 2^{nd} order fractional polynomial (FP) to describe the relationship with BMI at baseline (using a 1% significance level as evidence for a 2^{nd} order FP over a 1^{st} order FP) using Cox regression models stratified by sex, cohort and birth decade (26). We also analysed BMI assuming a linear relationship with cancer, expressing results per standard deviation (SD) in BMI levels.

We further investigated the associations with cancer risk for BMI combined with a metabolic score comprising mid-blood pressure ([systolic+diastolic blood pressure]/2), triglycerides and glucose. Each metabolic factor was standardised by cohort, sex, and fasting time (except mid-blood pressure), before being summed up into a metabolic score. Four categories were analysed

with the division of BMI at 25 kg/m^2 and the metabolic score at the median value, using high BMI with low MS as reference to mimic the analysis of BMI and waist circumference by Yu *et al*[4]. Because statistical power was reduced in this analysis due to the smaller dataset with complete information on metabolic factors and the division of the exposure into four categories, we combined cancers as relevant based on our findings for BMI.

Analyses using Stata (version 14) involved 2-sided p-values and 95% CIs.

Results

Among the 778,828 participants in the study, the mean age at baseline was 44 years (SD=9) and 382,853 (49%) were men (**Table 1**). At baseline, 357,146 (46%) participants were categorised as non-smokers, 182,858 (23%) as ex-smokers, 181,244 (23%) as light-moderate smokers (defined as <20 pack-years) and 57,579 (7%) as heavy smokers (defined as ≥20 pack-years). On average, ex-smokers had higher BMI levels whereas heavy smokers had higher levels of metabolic factors. A larger proportion of smokers than non-smokers and ex-smokers had a BMI below 25 kg/m^2 with high MS.

Participants with repeated measurements were younger and were more likely to be non-smokers than those without repeated measurements (**Supplementary Table 3**). The RDR was 0.95 (95% CI, 0.94-0.95) for BMI and 0.43 (0.37-0.49), 0.23 (0.32-0.44) and 0.51 (0.46-0.57) for smoking, ex-smoking and pack-years, respectively (**Supplementary Figure 2**).

During a median follow-up of 20 years, 20,242 individuals were diagnosed with a smoking-related cancer, of which 6,735 (33%) were lung cancer. In relation to lung cancer risk, there was an inverse and curvilinear association with BMI, with the lowest risk for those with a BMI

around 30 kg/m^2 (**Figure 1**). These associations attenuated but persisted after adjustment for smoking, and did not further attenuate after correction for within-person variability in smoking. Exclusion of the first 5-20 years of follow-up did not change the results. The HR of lung cancer per SD higher BMI was 0.87 (95% CI, 0.85-0.89) in fully adjusted models.

Associations between BMI and other smoking-related cancers than lung cancer generally attenuated after adjustment for smoking and correction for within-person variability, but to a lesser degree than seen for lung cancer (not shown). In fully-adjusted FP analyses of BMI and specific smoking-related cancers, BMI was inversely associated with cancers of the lung, larynx, oesophageal SCC, pharynx and oral cavity, i.e. cancers that are most strongly related to smoking and head and neck cancers, and was positively associated with cancer of the liver, pancreas, oesophageal adenocarcinoma, stomach-cardia, cervix uteri SCC, renal cell carcinoma and myeloid leukemia (**Supplementary Figure 3**).

Baseline smoking status modified the associations between BMI and several smoking-related cancers (**Figure 2**). For all smoking-related cancers combined, there was a positive association with BMI among non-smokers (HR per SD higher BMI, 1.14 [95% CI, 1.10-1.18]), but an inverse association among smokers (HR 0.94 [0.92-0.96], p for interaction<0.001). Inverse associations were observed for lung, head and neck cancers and oesophageal SCC and were restricted to smokers. Interaction analyses of BMI categories and smoking in greater detail showed interaction for head and neck cancers and oesophageal SCC combined (p for interaction<0.001), but not for lung cancer (p for interaction=0.09) (**Figure 3**). To explore whether the interaction was general for body size, we repeated the analysis for height, but found no interaction with smoking (**Supplementary Figure 4**).

The relationships between BMI and risk of cancers were generally similar for men and women, except for cancers of the stomach non-cardia, large cell lung cancer and urinary bladder (**Supplementary Figure 5**).

A BMI below 25 kg/m^2 combined with a high metabolic score was associated with an increased lung cancer risk compared to a BMI above 25 kg/m^2 and a low metabolic score (HR 1.52 [95% CI, 1.44-1.60] (**Figure 4**). Cancers of the head and neck and oesophageal SCC combined, but not other cancers combined, showed a similar pattern of association for the metabolic score by BMI level to that observed for lung cancer. The pattern of association for lung cancer was similar for men and women and did not change after exclusion of the first 5-20 years of follow-up (**Supplementary Figure 6**). Among non-smokers, however, the association was weak and non-significant after excluding the VHM&PP (HR 1.21 [95% CI, 0.91-1.63].

Discussion

The main finding of this pooled cohort study confirms an inverse and curvilinear association between BMI and lung cancer risk, which persisted after extensive adjustment for smoking information and correction for within-person variability in smoking, and after excluding initial years of follow-up. Further, we found diverse associations for smokers and non-smokers such that higher BMI among non-smokers and lower BMI among smokers were associated with higher risks of combined smoking-related cancers. This inverse association among smokers was found for cancers of the lung, head and neck and oesophageal SCC. The highest risk of these cancers, but not of other cancers combined, was observed for low BMI combined with a high metabolic score.

BMI has been consistently shown to be inversely associated with lung cancer risk (1-5), but the lack of a biological explanation and a weak or no association among non-smokers have led some to suspect bias from residual confounding in smoking. However, in two large prospective studies (4, 5), the inverse association among smokers persisted after adjustment for smoking in detail, which questions the possibility of confounding bias due to insufficient adjustment for smoking. In 2012, Renehan *et al.* showed with data simulation that greater measurement error of smoking than of BMI might explain the association (8). However, despite correction for the substantial within-person variability in smoking information in our study, the curvilinear and inverse association between BMI and lung cancer risk persisted, probably due to a weak relationship between BMI and smoking. Additionally, consistent with other studies (4, 5), excluding up to 20 years of follow-up showed no evidence of reverse causation in the observed associations.

In contrast to BMI, waist circumference has been positively related to the risk of lung cancer (4, 15). Yu *et al.* observed the highest risk for low BMI combined with high waist circumference (4), which in our study was observed for the combination of low BMI and a high metabolic score. A sarcopenic phenotype of central obesity with lower muscle mass, but with more metabolic aberrations, could explain this BMI-waist controversy (4, 15). However, low BMI combined with high waist circumference in Yu *et al*, or with high metabolic score in our study, were associated with smoking. Furthermore, the strikingly similar findings in our study for BMI, with and without the metabolic score, in relation to lung cancer and other cancers strongly related to smoking and of organs directly smoke-exposed at inhalation, i.e. cancers of the head and neck and oesophageal SCC, suggest that smoking may be at play even after extensive control for it. Moreover, in joint analysis of BMI and smoking, cancers of the head and neck and oesophageal SCC combined were more strongly related to smoking for each lower BMI

category, which is supported by a study of head and neck cancers, larger than ours, which additionally adjusted for alcohol intake (27). We speculated this to reflect higher per-cigarette concentration of tobacco carcinogens in persons of small than large body and organ size, which could have a larger impact on organs directly exposed to smoke at inhalation than on more distant organs. As these organs typically grow lengthwise, such theory would be supported by a similar interaction between height and smoking, which we did not observe, but was found in the aforementioned study (27) and in a pooled analysis of case-control studies of head and neck cancers (28). We conclude that it is difficult to find an explanation that accommodates both a differential association with lung cancer risk between waist circumference and BMI, and the uniform results in our study for the cancers strongly related to smoking and head and neck cancers.

Our study's access to individual-participant data from large cohort studies with linkage to high-quality cancer registers (29-31), and extensive follow-up of study participants enabled us to investigate rarer cancers, cancer subtypes, and non-smokers separately. The simultaneous investigation of lung cancer with other cancers further clarified the understanding of this common cancer. To correct for misclassification and within-person variability in smoking information and BMI, we also used extensive information on serial assessments, from which we could refute the hypothesis of large measurement error of smoking influencing the results of BMI and lung cancer (8). To limit and assess reverse causality, we focused on individuals without baseline cancer and investigated omittance of the initial periods of follow-up.

Nevertheless, our study has some limitations. We lacked data on potentially important confounders such as socioeconomic factors, physical activity, diet, alcohol intake, medications, and specific infections, such as helicobacter pylori in relation to stomach cancer and human

papillomavirus in cervix cancer. Self-reported smoking data are prone to bias and are challenging to harmonise across studies that used varying methods to record such data. We made substantial efforts to validate the smoking data, which led to exclusion of one cohort with ambiguous data on non-smokers. We have not accounted for competing risks and although it is possible that our results are influenced by censoring high-risk individuals with premature deaths from non-cancer causes, this does not explain the differential associations across the smoking-related cancers. Because some individuals who reduced their BMI due to health complications are included in our analysis, we cannot fully exclude the effects of reverse causation. Therefore, additional study designs such as non-linear Mendelian randomization studies will be important to establish causal associations with particular emphasis on effect modification by smoking (9). So far, Mendelian randomization studies have refuted an inverse linear association between BMI and lung cancer risk (10-12) and have supported a positive association between BMI and renal cell (32) and pancreatic cancer (33, 34).

In conclusion, our study shows inverse and curvilinear associations between BMI and lung cancer risk despite extensive adjustment and correction for within-person variability in smoking habits. The highest lung cancer risk was observed for low BMI and high metabolic score combined. However, these associations were more evident among smokers, and similar associations were observed for cancers of the head and neck and oesophageal SCC, but not for cancers less strongly related to smoking. From our investigation, we conclude that i) completely controlling for the effect of smoking on cancers of the lung, head and neck, and oesophageal SCC in an observational analysis appears difficult even with detailed and repeat smoking information, ii) the remaining influence of smoking on the association between obesity, metabolic factors, and these cancer forms remains unclear, and iii) with the available

methodology to date, these associations may be best investigated by Mendelian randomization, ideally incorporating non-linear associations.

Acknowledgements

We thank all participants of the cohorts. In Norway, we thank the screening team of the former National Health Screening Service of Norway, now the Norwegian Institute of Public Health. We also thank all scientists and organisations behind the VIP cohort and Åsa Ågren and her team at the department of biobank research for coordinating the VIP data. In the MPP, we thank Anders Dahlin, database manager of the cohort. In the VHM&PP, we thank Elmar Stimpfl and Karin Parschalk for excellent technical support as well as Markus Wallner, Christian Bernhard, and Gabriela Dür from the Vorarlberg State Government.

References

1. Kyrgiou M, Kalliala I, Markozannes G, Gunter MJ, Paraskevaidis E, Gabra H, et al. Adiposity and cancer at major anatomical sites: umbrella review of the literature. BMJ. 2017;356:j477.

2. Fang X, Wei J, He X, Lian J, Han D, An P, et al. Quantitative association between body mass index and the risk of cancer: A global Meta-analysis of prospective cohort studies. Int J Cancer. 2018;143(7):1595-1603.

3. Bhaskaran K, Douglas I, Forbes H, dos-Santos-Silva I, Leon DA, Smeeth L. Body-mass index and risk of 22 specific cancers: a population-based cohort study of 5.24 million UK adults. Lancet. 2014;384(9945):755-65.

4. Yu D, Zheng W, Johansson M, Lan Q, Park Y, White E, et al. Overall and Central Obesity and Risk of Lung Cancer: A Pooled Analysis. J Natl Cancer Inst. 2018;110(8):831-42.

5. Smith L, Brinton LA, Spitz MR, Lam TK, Park Y, Hollenbeck AR, et al. Body mass index and risk of lung cancer among never, former, and current smokers. J Natl Cancer Inst. 2012;104(10):778-89.

6. Chiolero A, Faeh D, Paccaud F, Cornuz J. Consequences of smoking for body weight, body fat distribution, and insulin resistance. Am J Clin Nutr. 2008;87(4):801-9.

7. Carreras-Torres R, Johansson M, Haycock PC, Relton CL, Davey Smith G, Brennan P, et al. Role of obesity in smoking behaviour: Mendelian randomisation study in UK Biobank. BMJ. 2018;361:k1767.

8. Renehan AG, Leitzmann MF, Zwahlen M. Re: body mass index and risk of lung cancer among never, former, and current smokers. J Natl Cancer Inst. 2012;104(21):1680-1; author reply 1.

9. Lawlor DA, Harbord RM, Sterne JA, Timpson N, Davey Smith G. Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. Stat Med. 2008;27(8):1133-63.

10. Carreras-Torres R, Johansson M, Haycock PC, Wade KH, Relton CL, Martin RM, et al. Obesity, metabolic factors and risk of different histological types of lung cancer: A Mendelian randomization study. PLoS One. 2017;12(6):e0177875.

11. Carreras-Torres R, Haycock PC, Relton CL, Martin RM, Smith GD, Kraft P, et al. The causal relevance of body mass index in different histological types of lung cancer: A Mendelian randomization study. Scientific reports. 2016;6:31121.

12. Gao C, Patel CJ, Michailidou K, Peters U, Gong J, Schildkraut J, et al. Mendelian randomization study of adiposity-related traits and risk of breast, ovarian, prostate, lung and colorectal cancer. Int J Epidemiol. 2016;45(3):896-908.

13. Reeves GK, Pirie K, Beral V, Green J, Spencer E, Bull D. Cancer incidence and mortality in relation to body mass index in the Million Women Study: cohort study. Bmj. 2007;335(7630):1134.

14. Duan P, Hu C, Quan C, Yi X, Zhou W, Yuan M, et al. Body mass index and risk of lung cancer: Systematic review and dose-response meta-analysis. Scientific reports. 2015;5:16938.

15. Hidayat K, Du X, Chen G, Shi M, Shi B. Abdominal Obesity and Lung Cancer Risk: Systematic Review and Meta-Analysis of Prospective Studies. Nutrients. 2016;8(12).

16. Fibrinogen Studies Collaboration. Correcting for multivariate measurement error by regression calibration in meta-analyses of epidemiological studies. Stat Med. 2009;28(7):1067-92.

17. Stocks T, Borena W, Strohmaier S, Bjorge T, Manjer J, Engeland A, et al. Cohort Profile: The Metabolic syndrome and Cancer project (Me-Can). Int J Epidemiol. 2010;39(3):660-7.

18. Häggstrom C, Jonsson H, Bjorge T, Nagel G, Manjer J, Ulmer H, et al. Linear age-course effects on the associations between body mass index, triglycerides, and female breast and male liver cancer risk: An internal replication study of 800,000 individuals. Int J Cancer. 2020;146(1):58-67.

19. World Health Organization/International Agency for Research on Cancer. Tobacco smoke and involuntary smoking. IARC Monogr Eval Carcinog Risks Hum. 2004;83:1-1438.

20. World Health Organization/International Agency for Research on Cancer. Personal habits and indoor combustions. Volume 100 E. A review of human carcinogens. IARC Monogr Eval Carcinog Risks Hum. 2012;100(Pt E):1-538.

21. Kirkali Z, Chan T, Manoharan M, Algaba F, Busch C, Cheng L, et al. Bladder cancer: epidemiology, staging and grading, and diagnosis. Urology. 2005;66(6 Suppl 1):4-34.

22. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med. 2002;21(11):1539-58.

23. Wood AM, White I, Thompson SG, Lewington S, Danesh J. Regression dilution methods for meta-analysis: assessing long-term variability in plasma fibrinogen among 27,247 adults in 15 prospective studies. Int J Epidemiol. 2006;35(6):1570-8.

24. Easton DF, Peto J, Babiker AG. Floating absolute risk: an alternative to relative risk in survival and case-control analysis avoiding an arbitrary reference group. Stat Med. 1991;10(7):1025-35.

25. Plummer M. Improved estimates of floating absolute risk. Stat Med. 2004;23(1):93-104.

26. Royston P, Ambler G, Sauerbrei W. The use of fractional polynomials to model continuous risk variables in epidemiology. Int J Epidemiol. 1999;28(5):964-74.

27. Gaudet MM, Kitahara CM, Newton CC, Bernstein L, Reynolds P, Weiderpass E, et al. Anthropometry and head and neck cancer:a pooled analysis of cohort data. Int J Epidemiol. 2015;44(2):673-81.

28. Leoncini E, Ricciardi W, Cadoni G, Arzani D, Petrelli L, Paludetti G, et al. Adult height and head and neck cancer: a pooled analysis within the INHANCE Consortium. Eur J Epidemiol. 2014;29(1):35-48.

29. Barlow L, Westergren K, Holmberg L, Talback M. The completeness of the Swedish Cancer Register - a sample survey for year 1998. Acta Oncol. 2009;48(1):27-33.

30. Larsen IK, Smastuen M, Johannesen TB, Langmark F, Parkin DM, Bray F, et al. Data quality at the Cancer Registry of Norway: an overview of comparability, completeness, validity and timeliness. Eur J Cancer. 2009;45(7):1218-31.

31. Oberaigner W, Stuhlinger W. Record linkage in the Cancer Registry of Tyrol, Austria. Methods Inf Med. 2005;44(5):626-30.

32. Johansson M, Carreras-Torres R, Scelo G, Purdue MP, Mariosa D, Muller DC, et al. The influence of obesity-related factors in the etiology of renal cell carcinoma-A mendelian randomization study. PLoS Med. 2019;16(1):e1002724.

33. Carreras-Torres R, Johansson M, Gaborieau V, Haycock PC, Wade KH, Relton CL, et al. The Role of Obesity, Type 2 Diabetes, and Metabolic Factors in Pancreatic Cancer: A Mendelian Randomization Study. J Natl Cancer Inst. 2017;109(9).

34. Lu Y, Gentiluomo M, Lorenzo-Bermejo J, Morelli L, Obazee O, Campa D, et al. Mendelian randomisation study of the effects of known and putative risk factors on pancreatic cancer. J Med Genet. 2020;57(12):820-828.

Table 1. Baseline characteristics of participants in the Metabolic syndrome and Cancer project 2.0, overall and by categories of smoking

Characteristic	Overall (n=778 828)	Non-smoker (n=357 147)	Ex-smoker (n=182 858)	Smoker, <20 pack-years (n=181 244)	Smoker, ≥20 pack-years (n=57 579)
Cohort, n (%)					
VHM&PP	161 232 (21)	129 949 (36)	0 (0)	23 246 (13)	7947 (14)
VIP	94 205 (12)	60 328 (17)	19 366 (12)	10 062 (6)	4449 (8)
MPP	20 971 (3)	11 634 (3)	0 (0)	4005 (2)	5332 (9)
40-y	394 722 (51)	119 435 (34)	142 219 (78)	102 067 (56)	31 001 (54)
Oslo	17 039 (2)	3576 (1)	4227 (2)	6724 (4)	2512 (4)
NCS	90 659 (12)	32 225 (9)	16 956 (9)	35 140 (19)	6338 (11)
Age, mean (SD) years	44 (9)	44 (11)	46 (6)	42 (6)	47 (7)
Male, n (%)	382 853 (49)	159 024 (44)	96 453 (53)	87 485 (48)	39 891 (69)
Decade of birth, n (%)					
1890-1930	62 034 (8)	32 623 (9)	9520 (5)	12 074 (7)	7 817 (14)
1930-1950	299 957 (39)	138 251 (39)	52 400 (29)	80 256 (44)	29 050 (50)
1950-1970	383 530 (49)	159 450 (44)	119 783 (66)	83 807 (46)	20 490 (36)
1970-1990	33 307 (4)	26 823 (8)	1155 (1)	5107 (3)	222 (<1)
BMI, mean (SD) kg/m^2	25.0 (3.9)	25.0 (4.1)	25.6 (3.7)	24.4 (3.7)	25.2 (3.8)
BMI category, kg/m^2, n (%)					
<20	47 069 (6)	22 971 (6)	5782 (3)	15 029 (8)	3288 (6)
20-22.49	200 814 (26)	91 384 (26)	40 082 (22)	55 301 (31)	14 047 (24)
22.5-24.99	181 630 (23)	80 799 (23)	44 327 (24)	43 099 (24)	13 405 (23)
25-27.49	178 351 (23)	79 265 (22)	47 839 (26)	37 553 (21)	13 694 (24)
27.5-29.99	92 765 (12)	42 945 (12)	25 237 (14)	17 198 (9)	7385 (13)
≥30	78 199 (10)	39 783 (11)	19 591 (11)	13 064 (7)	5761 (10)
MS, mean (SD)	-0.05 (1.99)	-0.09 (1.99)	-0.01 (1.92)	-0.16 (1.93)	0.36 (2.23)
BMI-MS category[a], n (%)					
BMI<25 kg/m^2-high MS	93 941 (19)	45 061 (18)	21 957 (19)	20 054 (23)	6 869 (24)
BMI<25 kg/m^2-low MS	165 776 (34)	89 631 (35)	34 251 (30)	34 014 (38)	7 880 (27)
BMI≥25 kg/m^2-high MS	147 523 (31)	77 108 (31)	38 158 (33)	22 249 (25)	10 008 (35)
BMI≥25 kg/m^2-low MS	77 977 (16)	40 104 (16)	21 116 (18)	12 703 (14)	4 054 (14)
Mid-blood pressure, mean (SD) mmHg[b]	105 (13)	106 (13)	107 (12)	104 (12)	107 (13)
Serum triglycerides, mean (SD) mmol/l[b,c]	1.4 (1.0)	1.3 (1.0)	1.5 (0.9)	1.5 (1.1)	1.7 (1.3)
Serum glucose, mean (SD) mmol/l[b,d]	5.4 (1.2)	5.4 (1.2)	5.4 (1.0)	5.3 (1.0)	5.6 (1.3)
Diabetes[e], n (%)	6007 (1)	2262 (1)	1921 (1)	1205 (1)	619 (1)

Oslo, Oslo study I; NCS, Norwegian Counties Study; 40-y, Age 40-Programme; VHM&PP, Vorarlberg Health Monitoring and Prevention Programme; VIP, Västerbotten Intervention Project; MPP, Malmö Preventive Project; SD, standard deviation; BMI, body mass index; MS, metabolic score.

[a]The metabolic score (MS) is a sum of z scores for mid-blood pressure, triglycerides and glucose, each standardised by cohort, sex, and fasting time (except mid-blood pressure), calculated for 485 217 (62%) individuals. The median was used as cut-point for low/high MS.

[b]Mean and SD calculated separately by cohort and then combined using random effects meta-analysis.

[c]Based on 286 249 (37%) individuals with eight hours or more of fasting.

[d]Based on 260 038 (33%) individuals with eight hours or more of fasting and with glucose measured in serum or plasma (excludes the MPP where glucose was measured in whole blood).

[e]Self-reported diabetes. Excludes the VHM&PP that lacked the information.

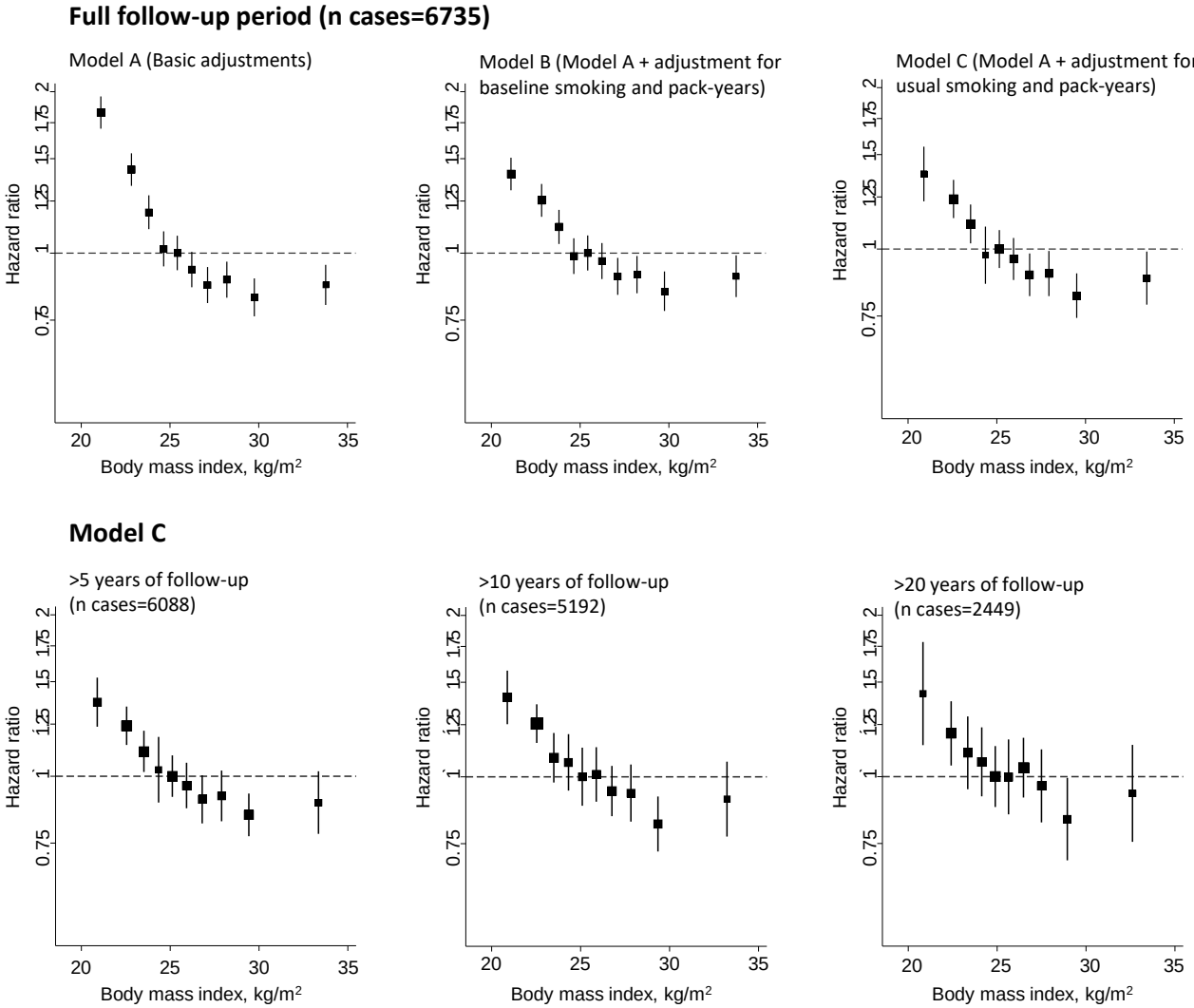

Figure 1. Hazard ratio (95% confidence interval) of lung cancer by usual BMI levels in deciles (n participants=778 828) by use of different models and follow-up periods. Model A shows hazard ratios from Cox regression with attained age as time-scale, stratified by cohort, sex, and birth decade, and adjusted for baseline age and fasting time. The fifth decile is the referent group.

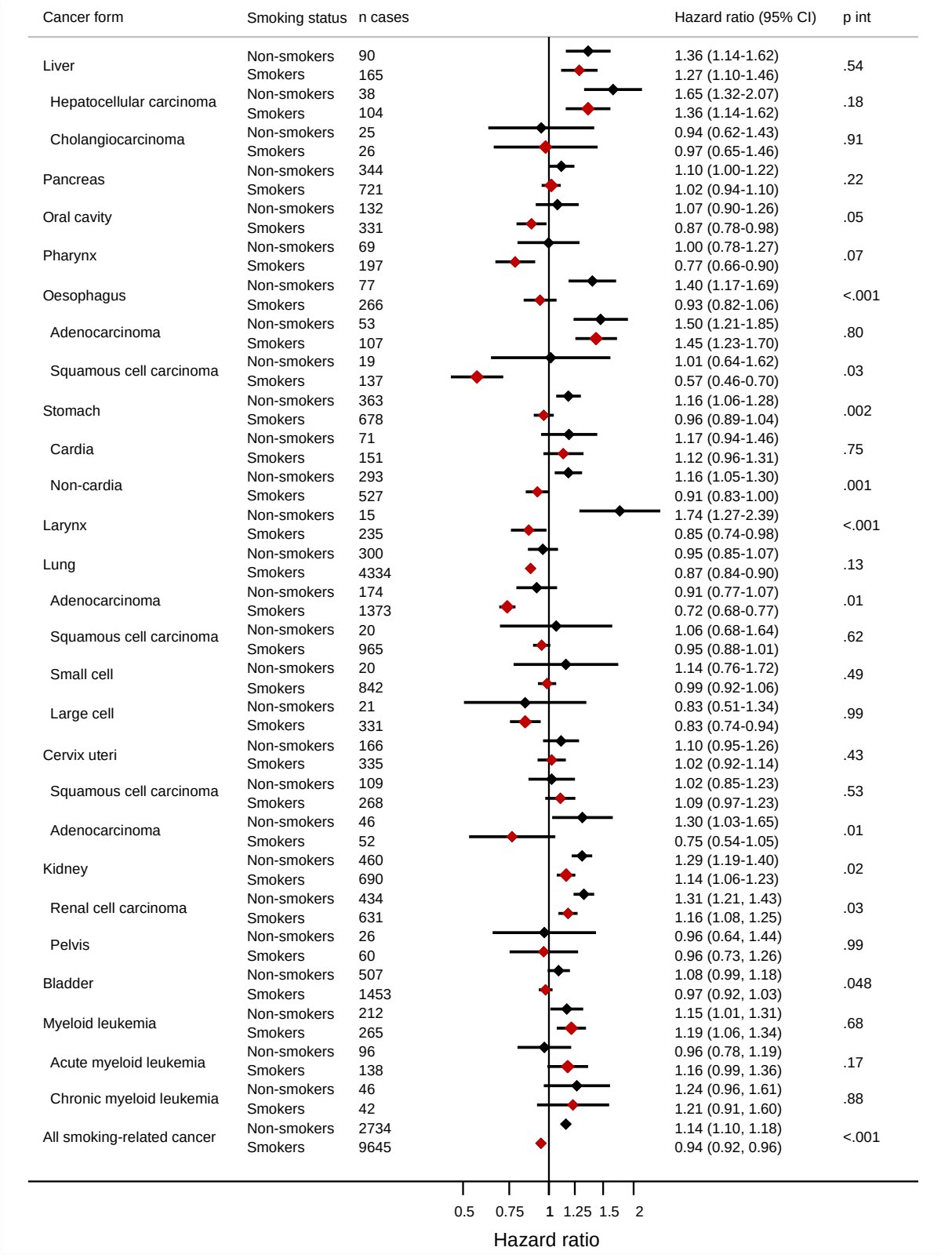

Figure 2. Hazard ratio (95% confidence interval) of cancer per standard deviation of usual body mass index (3.9 kg/m^2) among non-smokers (n=227 198) and smokers (n=207 630), respectively. The analysis of non-smokers excludes the Vorarlberg Health Monitoring and Prevention Programme. Hazard ratios were calculated using Cox regression with attained age as time-scale, stratified by cohort, sex, and birth decade, and adjusted for baseline age, fasting time, usual smoking status, and usual pack-years. P int denotes the p-value for interaction between body mass index and smoking status (smoking *vs* non-smoking).

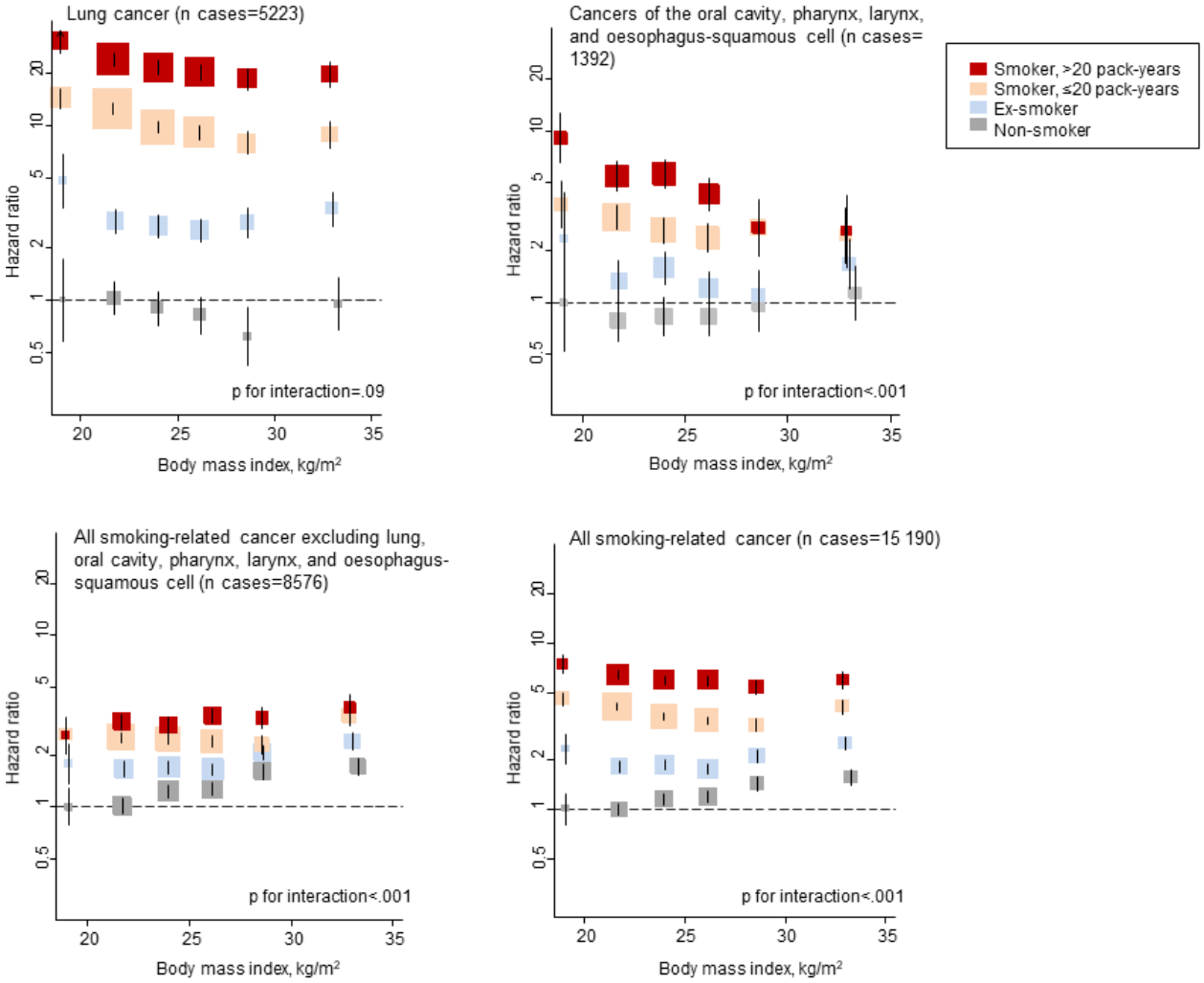

Figure 3. Hazard ratio (95% confidence interval) of cancer across categories of baseline body mass index in subgroups of smoking (n participants=617 596), excluding the Vorarlberg Health Monitoring and Prevention Programme. Hazard ratios were calculated using Cox regression with attained age as time-scale, stratified by cohort, sex, and birth decade, and adjusted for baseline age, fasting status, usual smoking status, and usual pack-years. Non-smokers with BMI below 20 kg/m^2 is the referent group. P for interaction denotes the p-value for the difference in linear trends of hazard ratios across body mass index level between smoking groups.

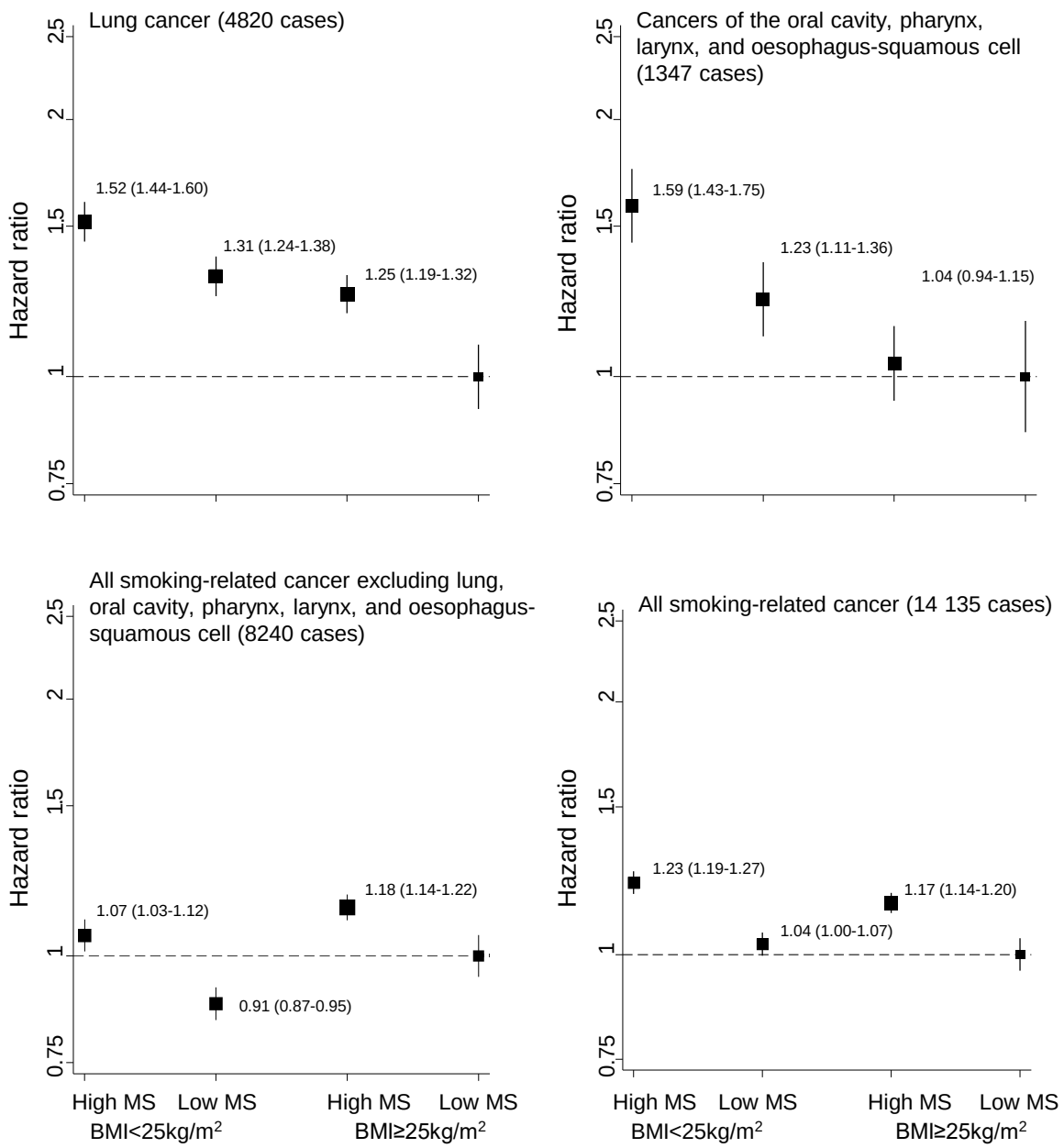

Figure 4. Hazard ratio (95% confidence interval) of cancer by level of body mass index and metabolic score combined (n participants=503 167). The metabolic score comprises mid-blood pressure, serum triglycerides and glucose, each standardised by cohort, sex, and fasting time (except mid-blood pressure) before being summed up and divided at the median (high/low MS). Hazard ratios were calculated using Cox regression with attained age as time-scale, stratified by cohort, sex, and birth decade, and adjusted for baseline age, usual smoking status, and usual pack-years. BMI ≥25 kg/m^2 with low MS is the referent group.