Smoking, blood cells and myeloproliferative neoplasms
meta-analysis and Mendelian randomization of 2.3 million people

Jayasuriya, Nimesh A.; Kjaergaard, Alisa D.; Pedersen, Kasper M.; Sørensen, Anders L.; Bak, Marie; Larsen, Morten K.; Nordestgaard, Børge G.; Bojesen, Stig E.; Çolak, Yunus; Skov, Vibe; Kjær, Lasse; Tolstrup, Janne S.; Hasselbalch, Hans C.; Ellervik, Christina

Published in:
British Journal of Haematology

DOI:
10.1111/bjh.16321

Publication date:
2020

Document version
Accepted manuscript

Citation for published version (APA):
Smoking, blood cells, and myeloproliferative neoplasms: meta-analysis and mendelian randomization of 2.3 million people

Short title: Smoking, blood cells, and myeloproliferative neoplasms

Nimesh Amrith Jayasuriya1,2, Alisa D. Kjaergaard1,3, Kasper M. Pedersen4,5,6, Anders Lindholm Sørensen4,5, Marie Bak5, Morten Kranker Larsen5,7, Børge G. Nordestgaard4,8, Stig E. Bojesen4,8, Yunus Çolak4,8, Vibe Skov5, Lasse Kjær5, Janne Schurmann Tolstrup8,10, Hans Carl Hasselbalch4,5, Christina Ellervik1,4,11,12

1Department of Laboratory Medicine, Boston Children’s Hospital, Boston, USA
2School of Medicine, University of Glasgow, Glasgow, United Kingdom
3Department of Clinical Epidemiology, Aarhus University Hospital and Aarhus University, Aarhus, Denmark
4Faculty of Health and Medical Sciences, University of Copenhagen, Denmark
5Department of Hematology, Zealand University Hospital, Roskilde, Denmark

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/BJH.16321

This article is protected by copyright. All rights reserved
Meta-analyses and mendelian randomization (MR) may clarify the associations of smoking, blood cells, and myeloproliferative neoplasms (MPN). We investigated the association of smoking with blood cells in the Danish General Suburban Population Study (GESUS, N=11,083), meta-analyses (including GESUS) of 92 studies (N=531,741), and MR of smoking variants CHRNA3(rs1051730[A]) and BDNF(rs6484320[A]) in UK Biobank, and with MPN in a meta-analysis of 6 studies (N(total/cases):1,425,529/2187), totaling 2,307,745 participants.
In the meta-analysis the random effects SMD in current smokers vs. non-smokers was 0.82(0.75-0.89,p=2.0*10^{-108}) for leukocytes, 0.09(-0.02-0.21,p=0.12) for erythrocytes, 0.53(0.42-0.64,p=8.0*10^{-22}) for haematocrit, 0.42(0.34-0.51,p=7.1*10^{-21}) for haemoglobin, 0.19(0.08-0.31,p=1.2*10^{-3}) for MCH, 0.29(0.19-0.39,p=1.6*10^{-8}) for MCV, and 0.04(-0.04-0.13,p=0.34) for platelets with trends for ever/ex-/current smokers, light/heavy smokers, and female/male smokers. Analyses presented high heterogeneity but low publication bias. Per allele in BDNF, ever-smoking vs. never smoking was associated with increased blood cell counts (neutrophils, erythrocytes, platelets), haematocrit, haemoglobin, and PCT. Per allele in CHRNA3, cigarettes per day in current smokers were associated with increased blood cell counts (leukocytes, neutrophils), MCH, RDW, and MCV. Pooled fixed effects OR for MPN was 1.44(95%CI:1.33-1.56;p=1.8*10^{-19};I^2=0%) in current smokers, 1.29(1.15-1.44;p=8.0*10^{-6};I^2=0%) in ex-smokers, 1.49(1.26-1.77;p=4.4*10^{-6};I-square=0%) in light-smokers, and 2.04(1.74-2.39,p=2.3*10^{-18};I^2=51%) in heavy-smokers compared with non-smokers. Smoking is observationally and genetically associated with blood cell counts in current and ex-smokers and observationally with risk of MPN.

**Keywords:**

Blood cells, tobacco smoking, myeloproliferative neoplasms, essential thrombocytopenia, polycythaemia vera, myelofibrosis, meta-analysis, haematocrit, haemoglobin, leukocytes, erythrocytes, thrombocytes

**Abbreviations (in alphabetic order):**

CBC: Complete blood count

ET: Essential thrombocytopenia

GESUS: The Danish General Suburban Population Study

This article is protected by copyright. All rights reserved
HCT: Haematocrit
Hgb: Haemoglobin
IRF: Immature reticulocyte fraction
MCH: Mean corpuscular haemoglobin
MCHC: Mean corpuscular haemoglobin concentration
MCV: Mean corpuscular volume
MF: Myelofibrosis
MPN: Myeloproliferative neoplasms
MPV: Mean platelet volume
MRV: Mean reticulocyte volume
NLR: Neutrophil lymphocyte ratio
PCT: Platelet crit
PDW: Platelet distribution width
PLT: Platelet
PV: Polycythaemia vera
RBC: Red blood cell
RDW: Red blood cell distribution width (CV: Coefficient of variation (%))
WBC: White blood cell

Introduction

The classic Philadelphia-negative chronic myeloproliferative neoplasms (MPN) are acquired stem cell diseases that include essential thrombocythaemia (ET), polycythaemia vera (PV), primary myelofibrosis (MF) and unclassifiable MPN (MPN-U)(Spivak 2017). Patients with MPN have an increased comorbidity burden and exhibit shorter survival compared to the general population(Hultcrantz, et al 2015). Chronic inflammation may be a driving force for clonal

This article is protected by copyright. All rights reserved


As observational studies may be prone to reverse causation, the genome-wide association studies (GWAS) CHRNA3 (The Cholinergic Receptor Nicotinic Alpha 3 Subunit, rs1051730 [A]), which associates with number of cigarettes smoked per day (Tobacco and Genetics 2010), and BDNF (brain-derived neurotrophic factor rs6484320[A]) (Pedersen, et al 2019), which associates with smoking initiation, may provide an opportunity to assess associations to CBC through an instrumental variable approach termed Mendelian Randomization (MR).

We investigated the association of cigarette smoking with CBC in a Danish observational study (N=11,083), meta-analyses of 92 studies (N=531,741), and MR analysis of CHRNA3 and BDNF in UKbiobank (N=350,475), and the association of smoking with MPN in a meta-analysis of 6 studies (N=1,425,529), totaling 2,307,745 participants.

Methods

This article is protected by copyright. All rights reserved
The Danish General Suburban Population Study (GESUS) is a cross-sectional general population study of 21,205 adults (participation rate 43%) from 2010-2013 (Bergholdt, et al 2013, Cordua, et al 2019). We included 11,083 healthy participants (mean age was 52 years (SD: 12), 53.5% women) with no cardiovascular disease or cancer. CBC were measured on fresh EDTA whole blood on Sysmex XE-5000 within 24h of blood draw. We calculated adjusted means for CBC by smoking status (never, current, ex-, and ever smoker) from multivariable linear regression models adjusted for potential confounders that a priori could be associated with tobacco smoking and/or blood cell values (Supplementary methods).

The meta-analysis protocol was registered at the International Prospective Register of Systematic Reviews (Prospero), https://www.crd.york.ac.uk/PROSPERO/, CRD42018102571 (Page, et al 2018). A systemic literature search with assistance from Countway Library at Harvard Medical School was conducted on the databases Pubmed, EMBASE, Web of Science and World Health Organisation Global Health Library for all relevant papers published on smoking, complete blood cell count indices, and myeloproliferative neoplasms until June 11th, 2018 (Supplementary methods). An article was relevant if it originated from cross-sectional, case control, retrospective, population, prospective, or trial studies at baseline and reported original data on CBC or MPN. For CBC, we calculated the pooled effect size using the unitless standardized mean difference (SMD, 95% CI) and the pooled non-standardized (weighted) mean differences (WMD, 95% CI). For CBC, we used the random-effects model (DerSimonian and Laird) due to heterogeneity. For MPN, we calculated the fixed effects pooled odds ratio (95% CI) as there was no study heterogeneity. Heterogeneity, sensitivity analyses, and publication bias are described in Supplementary Methods. We assessed double-sided p-values with Bonferroni correction for multiple comparisons by dividing the p-values with 24 cell type measurements to obtain an overall global significance of double-sided p<0.002.

We used instrumental variable analysis in a two-sample Mendelian randomization approach with summary-level exposure of smoking initiation (ever vs. never smoker) using brain-derived neurotrophic factor BDNF (rs6484320[A]) from The Tobacco and Genetics Consortium (Tobacco and Genetics 2010) and cigarettes per day using Cholinergic Receptor Nicotinic Alpha 3 Subunit CHRNA3 (rs1051730 [A]) (Liu, et al 2010). We extracted genetic
summary statistics for CBC from UK Biobank (Bycroft, et al 2018). CBC were measured on fresh EDTA whole blood on Beckman Coulter LH750 within 24h of blood draw.

CBC was reported using SI units(Brereton, et al 2016). MPN was defined as any of the following: ET, PV, pre-MF, MF, and MPN-U, by ICD coding from registries, WHO coding by haematologists, or by bone marrow examination by pathologists(Barbui, et al 2018).

Results

Complete Blood Count

*The Danish General Suburban Population Study (GESUS)*

Characteristics of the study participants based on smoking status are shown in Supplementary Table 1. In current smokers, ex-smokers, or ever smokers compared to never smokers, leukocyte counts were increased (Figure 1, Supplementary Figures 1-2); results were similar for all differential subpopulations, except for NLR and basophils. In current smokers, ex-smokers, or ever smokers compared to never smokers, erythrocyte counts were decreased (Figure 1, Supplementary Figures 1-2). In current smokers compared to never smokers, haematocrit, haemoglobin, MCH, MCV, and RDW were increased (Supplementary Figure 1), but in ex-smokers, haematocrit and haemoglobin were decreased (Supplementary Figure 2). In current smokers, ex-smokers, or ever smokers compared to never smokers, platelet counts and PCT were increased (Figure 1, Supplementary Figures 1-2), but MPV and PDW were not different.

Meta-analysis

For the meta-analysis, we included 92 study populations (including GESUS) on CBC published from 1974-2018 comprising 531,741 individuals with 214,703 current smokers, 90,174 ex-smokers, and 222,600 never or non-smokers between the ages of 15 and 99 years from 33
countries (Figure 2, Table 1, Supplementary Table 2). Males were overrepresented in the meta-analyses with a male:female ratio of 6:1.

Random effects non-standardized weighted mean differences (WMD) in current smokers vs. non-smokers per cell type are shown in Supplementary Figures 3-26. The random effects SMD of leukocytes in current smokers was 0.82 (0.75-0.89, p=2.0*10^{-10}) overall (Figure 3) equivalent to a mean difference (WMD) of 1.25*10^9/L (Supplementary Figure 3). Compared with non-smokers, the SMD of leukocytes was 1.56 (1.34-1.79, p=7.2*10^{-41}) in current heavy smokers, 0.87 (0.69-1.05, p=2.2*10^{-21}) in current light smokers, 0.84 (0.75-0.93, p=1.1*10^{-72}) in current male smokers, and 0.65 (0.57-0.73, p=3.5*10^{-60}) in current female smokers (Supplementary Figures 27-30). Compared with never smokers, the SMD was 0.16 (0.11-0.21, p=8.0*10^{-12}) in ex-smokers, and 0.60 (0.52-0.68, p=2.2*10^{-53}) in ever-smokers (Supplementary Figures 31-32). SMD of absolute counts, but not percentages, for neutrophils, eosinophils, basophils, monocytes, and lymphocytes were similar (Figure 3, Supplementary Figures 27-32).

Random effects SMD of haematocrit in current smokers was 0.53 (0.41-0.65, p=6.7*10^{-19}) overall (Figure 3) equivalent to a WMD of 0.02 (L/L) (Supplementary Figure 17). Compared with non-smokers, the SMD of haematocrit was 1.08 (0.77-1.40, p=1.4*10^{-11}) in current heavy smokers, 0.54 (0.21-0.87, p=1.4*10^{-3}) in current light smokers, 0.29 (0.20-0.37, p=3.2*10^{-11}) in current male smokers, and 0.33 (0.11-0.55, p=3.3*10^{-3}) in current female smokers. Compared with never smokers, the SMD was 0.47 (0.12-0.82, p=7.8*10^{-3}) in ex-smokers, and 0.19 (0.10-0.28, p=7.3*10^{-5}) in ever-smokers (Supplementary Figures 27-32). Results were similar and in the same direction for erythrocytes, haemoglobin, MCH, MCV, and RDW (Figure 3, Supplementary Figures 27-32). Results for MCHC and reticulocytes were not consistent (Figure 3, Supplementary Figures 27-32).

Compared with non-smokers, random effects SMD of platelets was 0.04 (-0.04-0.13, p=0.34) overall in current smokers (Figure 3), 0.22 (0.09-0.34, p=5.2*10^{-4}) in current heavy smokers, 0.09 (-0.05-0.23, p=0.21) in current light smokers, 0.05 (-0.06-0.16, p=0.40) in current male smokers, and 0.01 (-0.14-0.16, p=0.88) in current female smokers. Compared with never smokers, the SMD was 0.06 (0.00-0.11, p=0.04) in ex-smokers, and 0.15 (0.10-0.19, p=4.6*10^{-10}) in ever-smokers (Supplementary Figures 27-32). Random effects SMD of MPV and PDW were not significant in any analyses (Figure 3, Supplementary Figures 27-32).
All analyses presented very high heterogeneity, reflected by the $I^2$ and Q-statistic. Egger’s tests for publication bias revealed no publication bias except in analyses of leukocytes and platelets in current smokers. Analyses by ethnicity revealed that erythrocyte counts were higher in current smokers compared to non-smokers in Middle Eastern countries [SMD(95%CI): 0.49 (0.12,0.85), p=0.009] whereas erythrocyte counts were lower in current smokers compared to non-smokers in Western European countries [SMD(95%CI): -0.23(-0.39,-0.07), p=0.006] (p-difference=0.02) (Supplementary Table 3); the SMD by ethnicity was not different for leukocytes, haematocrit, haemoglobin, or platelets. Analyses by laboratory method did not reveal any differences and did not remove heterogeneity (Supplementary Table 4). Meta-regressions did not reveal impact of moderator variables (study size, publication year, longitude or latitude of study location, or mean ages) (Supplementary Table 5). Leave-one-study-out analyses for leukocytes, erythrocytes, haemoglobin, haematocrit, and platelets did not reveal any exaggerated influences of studies (Supplementary Tables 6-10). A sub-group analysis of all cell types in current smokers vs. never smokers revealed similar but accentuated results compared to analyses in current smokers vs. non-smokers (Supplementary Figure 33).

Mendelian randomization study in UK Biobank

The mean(SD) of cigarette smoking per day was 15.5(8.4) in current smokers and 18.9(10.9) in ex-smokers. The $\beta$(SE) and instrument strength F was [0.0853(0.00439), 376] per SD of cigarettes per day for CHRNA3 (rs1051730[A]) and [0.0614(0.011),31] for smoking initiation for BDNF (rs6484320[A]). BDNF (rs6484320[A]) associated with increased BMI and CHRNA (rs1051730[A]) associated with chronic obstructive lung disease and lung cancer in UK Biobank (Supplementary Table 11).

For each rs6484320[A] in BDNF, ever-smoking compared to never smoking was associated with standard deviation (SD) increases in neutrophil counts of 0.11, erythrocytes of 0.13, haematocrit of 0.13, in haemoglobin of 0.12, platelets of 0.11, and PCT of 0.12, but decreased values of eosinophil %, monocyte %, reticulocytes and indices, and PDW (Supplementary Figure 34, Supplementary Table 12). For each rs1051730[A] in CHRNA3 in current smokers, one SD increase in cigarettes per day (~10 cigarettes per day) were associated

This article is protected by copyright. All rights reserved
with SD increases in leukocytes of 0.07, neutrophils of 0.09, MCH of 0.15, MCV of 0.16, RDW of 0.8, and reticulocytes of 0.07 (Figure 3), with results for male higher than in female, and with results persisting but attenuated in ex-smokers (Supplementary Figures 35-37).

Meta-analysis of Myeloproliferative Neoplasms

We included 6 studies, 1,425,529 individuals with 2187 MPN cases (0.2%) (Figure 2, Table 1, Supplementary Table 13) published between 1997-2018. The pooled fixed effects odds ratios for MPN were 1.44(95%CI: 1.33-1.56; p=1.8*10^{-19}; I^2=0%) in current smokers, 1.49(1.26-1.77; p=4.4*10^{-6}; I-square=0%) in light-smokers, 2.04(1.74-2.39, p=2.3*10^{-18}; I^2=51%) in heavy-smokers vs. non-smokers, and 1.29(1.15-1.44; p=8.0*10^{-6}; I^2=0%) in ex-smokers and 1.39(1.30-1.48, p=8.3*10^{-23}; I^2=35%) in ever smokers vs. never-smokers (Figure 5 and 6). There was no evidence of publication bias. Subgroup analyses for PV and ET showed similar results for current smokers, which had overlapping confidence intervals with ex-smokers and ever-smokers; there were only two studies including patients with primary MF (Supplementary Table 14).

Discussion

We examined the association between smoking, blood counts, and risk of MPN in 2.3 million participants. In the Danish population study and in the meta-analysis, current smoking was associated with increased leukocyte counts (including, neutrophils, eosinophils, monocytes, lymphocytes), haematocrit, haemoglobin, MCH, and MCV, and increased risk of MPN compared with non-smokers. Leukocyte counts were impacted the most relative to other cell types. Compared to non-smokers, there was a dose-trend for cell counts, indices, and MPN for ever/ex-/current smokers and light/heavy smokers, and male/female smokers. All analyses for CBC presented high between-study heterogeneity but low publication bias, whereas analyses for MPN presented low to moderate between-study heterogeneity with no publication bias. In the MR study, ever smoking compared with never smoking was associated with increased neutrophils, erythrocytes, haematocrit, haemoglobin, MCV, platelets, and PCT, while higher cigarette consumption in current smokers was associated with increased leucocytes, neutrophils, MCH, MCV, RDW, and reticulocytes.
The chemical compounds found in tobacco, such as nicotine, tar, carbon monoxide, and hydrogen cyanide, are associated with genotoxic, cytotoxic and cancerogenic effects (Lohler and Wollenberg 2019). Thus, the observed associations between smoking and increased CBCs could merely reflect production of cells in reaction and secondary to smoking in an otherwise healthy cell lineage. It could also reflect acquired mutations in hematopoietic stem cells in the bone marrow with clonal hematopoiesis of indeterminate potential or MPN, as the associations persisted despite smoking cessation (Baxter, et al 2005, Nielsen, et al 2013, Weinberg, et al 2012).

Smoking and MPNs share several clinical, biochemical and molecular characteristics, which may all be mediated by chronic inflammation and oxidative stress (Hasselbalch 2015). Smoking is associated with genotoxic effects, namely the acquired somatic driver mutation JAK2 V617F for MPN (Baxter, et al 2005, Nielsen, et al 2013, Weinberg, et al 2012), and oxidative damage to DNA leading to increased reactive oxygen species (Prieme, et al 1998). The JAK-STAT and NF-κB signaling pathways, associated with carcinogenesis, are activated and perpetuated by reactive oxygen species in both smokers and in patients with MPNs (Bjorn and Hasselbalch 2015, Hasselbalch 2015). Smokers, as well as patients with MPN, have a high risk of thrombosis, which is likely explained by an elevated haematocrit and leukocytosis and in vivo leukocyte, platelet and endothelial activation (Hasselbalch 2015, Spivak 2017). Interestingly, patients with MPNs have increased risk of secondary cancers, particularly lung cancer and urinary tract cancers, which are also related to smoking (Frederiksen, et al 2011, Landtblom, et al 2018).

Some CBC results were discordant in the general population study, the meta-analysis, and the MR study. In the MR study, ever smoking was associated with increased erythrocyte counts, but in the general population study current smoking was inversely associated with erythrocyte counts. There can be several explanations for these findings. First, smoking is associated with hypoxia as carbon monoxide binds to haemoglobin and thereby reduces the oxygen-binding capacity which theoretically would lead to increased erythrocytes as a compensatory mechanism (Goldstein 2008). Second, the lower counts of erythrocytes in current smokers compared to non-smokers observed in the general population study along with an increased heterogeneity in red cell size (RDW) could be caused by a decreased erythropoiesis, increased cell destruction, bone marrow depression, and chronic inflammation (Asgary, et al
2005, Li, et al 2017); thus, in the MR analysis, smoking intensity in current smokers was associated with decreased erythrocyte counts. Third, the inverse relationship between smoking and erythrocyte counts could also be explained by chronically ill smokers being less likely to participate. Fourth, in a recent MR study, smoking was associated with increased leukocytes, haematocrit, haemoglobin, and MCV, but not erythrocytes or platelets(Pedersen, et al 2019).

In GESUS and the meta-analysis, current smoking was associated with increased MCV (i.e. macrocytosis) and RDW, and likewise in the MR study, the smoking intensity was associated with increased MCV and RDW in current smokers. The findings for MCV can be explained by smoking and cotinine, a nicotine metabolite, also being associated with decreased levels of folate and cobalamin and increased levels of homocysteine(Haj Mouhamed, et al 2011). An increased RDW is associated with poor overall survival and poor cancer-specific survival in patients with cancer(Hu, et al 2017), particularly in patients with CML(Iriyama, et al 2015), but no studies have investigated RDW as a prognostic factor for MPN.

For CBC, the SMD allowed us to compare all cell types on the same scale, and the random effects statistic accounted for the high heterogeneity in study designs across studies. We investigated study characteristics such as smoking intensity (heavy vs. light), smoking status (current, ex, ever), sex, lab testing methodology, study size, year published, age distribution, or geographical region as culprits for heterogeneity. In the meta-analysis, we were not able to address pack-years, smoking duration, years since smoking cessation, or differences between filter vs. non-filter cigarettes, as reporting in studies were too diverse or lacking on these factors. A recent study showed that within ex-smokers, years since smoking cessation was inversely associated with white blood cells and the differentials, but values were still elevated compared to never smokers(Pedersen, et al 2019). Lastly, many unreported pre-analytical factors in the studies in the meta-analyses may have affected CBC such as seasonal and diurnal variations, fasting vs. non-fasting sampling, time since last cigarette smoked, position at blood draw (supine or sitting), altitude, physical activity, possible dormant infections, allergies, undiagnosed diseases, nutritional status, menstrual cycle and menopausal status.

For MPN, there was not enough information to address sex differences. Also, the description of MPN was not completely identical in the studies. Leal used international classification of diseases 9 (ICD-9)(Leal, et al 2014) and Pedersen ICD-8 and ICD-10(Pedersen,
et al 2018). Sørensen recruited MPN patients from a single-institution based on WHO-5 MPN criteria (Lindholm Sorensen and Hasselbalch 2016). Kroll used ICD-10 and other classification systems and included myelodysplastic syndrome, CML, and other myeloproliferative disorders (Kroll, et al 2012), but it was not possible to filter out these diagnoses. Pasqualetti did not specify the MPN diagnosis or coding any further (Pasqualetti, et al 1997). None of the studies explicitly mentioned bone marrow pathology as part of the MPN diagnosis. Incidence rates (IR per 100,000 person-years) of MPN varied across the follow-up studies with an IR of 98 in Leal (Leal, et al 2014) with a median of 11 years of follow-up ( ~ 8.9 annually per 100,000), an IR of 13 in Pedersen with a mean of 6.8 years of follow-up ( ~ 1.9 annually per 100,000) (Pedersen, et al 2018), and an IR of 16 in Kroll with a mean of 10 years of follow-up ( ~1.6 annually per 100,000) (Kroll, et al 2012). In comparison, a recent meta-analysis reported an incidence rate of 2.6 annually per 100,000 for MPN (Titmarsh, et al 2014).

In conclusion, the association of smoking with hyperproliferation of blood cells and MPN were supported by the strength of the associations, the consistency across the large number of observational studies, the MR analysis, and the dose-response relationship. Furthermore, previous complementary studies support the mechanistic specificity of smoking with genotoxicity forgoing and being the driver of MPN. Future studies should elucidate when increased cell counts are merely secondary to smoking or reflect an underlying cancer. Our findings of an association between smoking and MPNs should alert clinicians that elevated cell counts in smokers may also reflect an underlying myeloproliferative cancer.

Acknowledgements

Search terms and creation of Endnote library with 18,516 references for the meta-analyses: Paul Bain, Countway Library of Medicine, 10 Shattuck street, Boston (paul_bain@hms.harvard.edu).

UK-biobank: extraction of data: Mathias N. Kjaergaard, Technology and Innovation, Terma A/S, Lystrup, Denmark (mnk@mnk.dk).

Authorship contributions

This article is protected by copyright. All rights reserved
Conception of study design: NAJ, MB, MKL, BGN, SEB, YC, VHS, LK, HCH, CE. Study PIs for GESUS (SJ-113, SJ114, SJ452): CE, HCH, VHS, LK. Collection of participants for GESUS: CE, MKL. Meta-analysis: identification of articles, extraction of data, statistical coding, analyses, figures, tables: NAJ, CE. UK-biobank: statistical coding, analyses, figures, tables: ADK. Study PIs providing data for MPN meta-analyses from previously published studies: ALS, KMP, JST, HCH. Manuscript draft: NAJ and CE. Critical revision of manuscript and final approval: All authors.

Conflicts of interest disclosures

The authors do not have any conflicts of interest.

Funding

The Danish General Suburban Population Study was funded by the Region Zealand Foundation, Naestved Hospital Foundation, Johan and Lise Boserup Foundation, TrygFonden, Johannes Fog’s Foundation, Region Zealand, Naestved Hospital, The National Board of Health, and the Local Government Denmark Foundation. Kasper M. Pedersen, Christina Ellervik, and Børge G. Nordestgaard was funded by Karen Elise Jensen’s foundation.
Table 1. Characteristics of study designs on chronic cigarette smoking, complete blood count, and MPN

<table>
<thead>
<tr>
<th>Study populations</th>
<th>Complete blood counts</th>
<th>MPN</th>
<th>All</th>
<th>B+C+D</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>General population</td>
<td>Meta-analysis**</td>
<td>Mendelian Randomization (C)</td>
<td>Meta-analysis</td>
</tr>
<tr>
<td>Countries</td>
<td>GESUS</td>
<td>Denmark</td>
<td>92</td>
<td>33</td>
</tr>
<tr>
<td>Publication years (yyyy-yyyy)</td>
<td>-</td>
<td>1974-2018</td>
<td>2018</td>
<td>1997-2018</td>
</tr>
<tr>
<td>Age range, years</td>
<td>20-99</td>
<td>15-99</td>
<td>40-69</td>
<td>16-99</td>
</tr>
<tr>
<td>Individuals, N*</td>
<td>11083</td>
<td>531741</td>
<td>350475</td>
<td>1425529</td>
</tr>
<tr>
<td></td>
<td>Men, N</td>
<td>Women, N</td>
<td>Mendelson smokers, N</td>
<td>Current smokers, N</td>
</tr>
<tr>
<td></td>
<td>5155</td>
<td>5928</td>
<td>2217</td>
<td>690</td>
</tr>
<tr>
<td></td>
<td>454003</td>
<td>77252</td>
<td>214703</td>
<td>11881</td>
</tr>
<tr>
<td></td>
<td>162399</td>
<td>188076</td>
<td>25348</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>31377</td>
<td>1391966</td>
<td>267942</td>
<td>17610</td>
</tr>
<tr>
<td></td>
<td>1425529</td>
<td>1391966</td>
<td>267942</td>
<td>17610</td>
</tr>
</tbody>
</table>

*Total individuals included were the sum of current smokers, ex-smokers, non-smokers, and never-smokers. If articles did not provide N for exposed/unexposed, total individuals were the number listed as total in the article.

**The meta-analysis of blood counts included GESUS. Never-smokers were life-long never-smokers. Non-smokers were non-current smokers. Men+women does not add up to total individuals, as some studies did not provide information on sex.

#UK biobank genotypes from Haplotype Reference Consortium (HRC) plus UK10K & 1000 Genomes reference panels as released by UK Biobank in March 2018. See text for details. Never smokers is approximately 49% of total, but the exact number for those with genotype+hematology was not publically available.

##Heavy smokers: >=20 cigarettes per day. Light-smokers <20 cigarettes per day.


This article is protected by copyright. All rights reserved
Figure Legend

Figure 1. Standardized adjusted mean differences in current smokers vs. never smokers – The Danish General Suburban Population Study

Figure 2. PRISMA flow chart. EMBASE: Excerpta Medica database. GESUS: The Danish General Suburban Population Study. WHO-GHL: World Health Organization Global Health Library.

Figure 3. Meta-analyses of complete blood cell counts - Current smokers vs non-smokers

Figure 4. Mendelian Randomization of complete blood cell counts using CHRNA3 genotype – SMD per genetically one additional cigarette per day in current smokers in UKBiobank. CHRNA: Cholinergic Receptor Nicotinic Alpha 3 Subunit (rs1051730 [A]). SMD: standardized mean difference (adjusted for sex, age, principal components).

Figure 5. Meta-analysis of Philadelphia-negative myeloproliferative neoplasms – Current smokers vs. non-smokers.

Figure 6. Meta-analyses of Philadelphia-negative chronic myeloproliferative neoplasms – By smoking status. P-Egger: publication bias p-value according to Eggers test. p-het: heterogeneity p-value. SMD: standardized mean difference.

References


This article is protected by copyright. All rights reserved


This article is protected by copyright. All rights reserved


The Danish General Suburban Population Study

Standardized adjusted mean differences in complete blood cell counts
in current smokers and never smokers

<table>
<thead>
<tr>
<th>Type</th>
<th>Mean Adjusted Difference</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Red Blood</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White Blood</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Platelets</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Standardized mean difference 95% CI

bjh_16321_f1.tif

This article is protected by copyright. All rights reserved
### Meta-analysis of Philadelphia-negative chronic myeloproliferative neoplasms

#### Current smokers vs. non-smokers

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Odds Ratio (95% CI)</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ito</td>
<td>2012</td>
<td>1.42 (1.31, 1.55)</td>
<td>0.00</td>
</tr>
<tr>
<td>Ueda</td>
<td>2013</td>
<td>1.61 (1.35, 2.00)</td>
<td>0.01</td>
</tr>
<tr>
<td>Landau</td>
<td>2014</td>
<td>1.03 (0.57, 2.27)</td>
<td>0.00</td>
</tr>
<tr>
<td>Goorissen</td>
<td>2015</td>
<td>0.00 (0.00, 0.00)</td>
<td>0.00</td>
</tr>
<tr>
<td>Goorissen</td>
<td>2016</td>
<td>1.68 (1.12, 2.53)</td>
<td>0.03</td>
</tr>
<tr>
<td>Goorissen</td>
<td>2017</td>
<td>1.68 (1.12, 2.53)</td>
<td>0.03</td>
</tr>
</tbody>
</table>

**Overall (fixed-effects model)**

Odds Ratio (95% CI): 1.44 (1.33, 1.55)

Weight: 100.00
<table>
<thead>
<tr>
<th>Source</th>
<th>Relative Risk</th>
<th>95% CI</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test A</td>
<td>1.2</td>
<td>0.9-1.5</td>
<td>0.1</td>
</tr>
<tr>
<td>Test B</td>
<td>1.5</td>
<td>1.2-2.0</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Test C</td>
<td>1.0</td>
<td>0.8-1.2</td>
<td>0.5</td>
</tr>
<tr>
<td>Test D</td>
<td>0.8</td>
<td>0.6-1.1</td>
<td>0.1</td>
</tr>
</tbody>
</table>