

Erythrocytosis in the general population: clinical characteristics and association with clonal hematopoiesis

Hanneke J. C. M. Wouters,^{1,2} René Mulder,³ Isabelle A. van Zeventer,¹ Jan Jacob Schuringa,¹ Melanie M. van der Klauw,² Pim van der Harst,⁴ Arjan Diepstra,⁵ André B. Mulder,³ and Gerwin Huls¹

¹Department of Hematology, ²Department of Endocrinology, ³Department of Laboratory Medicine, ⁴Department of Cardiology, and ⁵Department of Pathology and Medical Biology, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands

Key Points

- Erythrocytosis, only when defined using strict criteria, was associated with cardiovascular morbidity and mortality and all-cause mortality.
- Erythrocytosis was associated with high prevalence of clonal hematopoiesis (38%) (including *JAK2* V617F [5.3%] and *BCOR/BCORL1* [16%]).

Erythrocytosis is a common reason for referral to hematology services and is usually secondary in origin. The aim of this study was to assess clinical characteristics and clonal hematopoiesis (CH) in individuals with erythrocytosis in the population-based Lifelines cohort (n = 147 167). Erythrocytosis was defined using strict (World Health Organization [WHO] 2008/British Committee for Standards in Hematology) and wide (WHO 2016) criteria. Individuals with erythrocytosis (strict criteria) and concurrent leukocytosis and/or thrombocytosis were 1:2 matched with individuals with isolated erythrocytosis and analyzed for somatic mutations indicative of CH ($\geq 5\%$ variant allele frequency). One hundred eighty five males (0.3%) and 223 females (0.3%) met the strict criteria, whereas 4868 males (7.6%) and 309 females (0.4%) met the wide criteria. Erythrocytosis, only when defined using strict criteria, was associated with cardiovascular morbidity (odds ratio [OR], 1.8; 95% confidence interval [CI], 1.2-2.6), cardiovascular mortality (hazard ratio [HR], 2.2; 95% CI, 1.0-4.6), and all-cause mortality (HR, 1.7; 95% CI, 1.2-2.6), independent of conventional risk factors. Mutations were detected in 51 of 133 (38%) evaluable individuals, with comparable frequencies between individuals with and without concurrent cytosis. The *JAK2* V617F mutation was observed in 7 of 133 (5.3%) individuals, all having concurrent cytosis. The prevalence of mutations in *BCOR/BCORL1* (16%) was high, suggesting aberrant epigenetic regulation. Erythrocytosis with CH was associated with cardiovascular morbidity (OR, 9.1; 95% CI, 1.2-68.4) in a multivariable model. Our data indicate that only when defined using strict criteria erythrocytosis is associated with cardiovascular morbidity (especially in the presence of CH), cardiovascular mortality, and all-cause mortality.

Introduction

Erythrocytosis is caused by either a decrease in plasma volume (relative erythrocytosis) or an increase in red blood cell mass (absolute erythrocytosis). Absolute erythrocytosis can be primary or secondary in origin. Primary erythrocytosis results from increased autonomous production of red blood cells, for example, in the presence of myeloproliferative neoplasms (MPN) or rare inherited conditions. Secondary erythrocytosis is used to categorize cases in whom erythropoietin drives the bone marrow to produce more red cells (eg, in case of hypoxia or the presence of an erythropoietin-producing tumor). The most

Submitted 1 September 2020; accepted 12 November 2020; published online 21 December 2020. DOI 10.1182/bloodadvances.2020003323.

The manuscript is based on data from the Lifelines Cohort Study. Lifelines adheres to standards for data availability. The data catalogue of Lifelines is publicly accessible at www.lifelines.nl. All international researchers can obtain data at the Lifelines research office (research@lifelines.nl), for which a fee is required. Data on incident hematologic malignancies were obtained from PALGA, the nationwide network and registry of histology and cytopathology in the Netherlands (for more information: <https://www.palga.nl/>)

[gegevensaanvragen/gegevensaanvragen.html](https://www.palga.nl/)). Results from cause of death analyses are based on calculations by the authors using nonpublic microdata from Statistics Netherlands. Under certain conditions, these microdata are accessible for statistical and scientific research. For further information, please contact microdata@cbs.nl. The full-text version of this article contains a data supplement.

© 2020 by The American Society of Hematology

common cause of acquired primary erythrocytosis is polycythemia vera (PV).¹ In the United States, PV has a prevalence of 44 to 57 per 100 000.²

Multiple studies have assessed morbidity and mortality in individuals with PV. Life expectancy of patients with PV (especially ≤ 50 years) is reduced compared with the general population.³ Arterial and venous thrombosis are important complications among patients with PV and contribute to the increased mortality rate.^{4,5} Besides conventional risk factors, the severity of erythrocytosis, leukocytosis, and thrombocytosis increases the risk of thrombosis in PV.⁵⁻⁸ Although the prevalence of secondary erythrocytosis is believed to be considerably higher than the prevalence of PV, no data on its impact on outcome in the general population are available. For secondary erythrocytosis, there is conflicting evidence regarding the risk of thrombosis.⁹

Clonal hematopoiesis (CH), as defined by the presence of clonal somatic mutations detectable in bone marrow or peripheral blood, is a defining feature of hematologic malignancies.¹⁰ For PV, the presence of a *JAK2* V671F mutation or exon 12 mutations in *JAK2* is 1 of the diagnostic criteria.^{1,11-13} CH is also found to be a consequence of aging. Depending on the sensitivity of the sequencing technique, CH may be detected in 5% to 45% of community-dwelling individuals aged >70 years with normal blood cell counts.^{10,14-19} In this context, the condition is referred to as clonal hematopoiesis of indeterminate potential (CHIP) or age-related clonal hematopoiesis. The most commonly mutated genes in CHIP include *DNMT3A*, *TET2*, and *ASXL1*.^{10,14,15} CHIP confers a 0.5% to 1% annual risk of developing a hematologic malignancy and increases an individual's risk of experiencing a cardiovascular event.^{10,14,20} CH may also be detected in the presence of 1 or more unexplained cytopenias, a condition referred to as clonal cytopenia of unknown significance. Although not (yet) meeting World Health Organization (WHO) criteria for myelodysplastic syndromes, the risk of developing a myeloid malignancy in patients with clonal cytopenia of unknown significance is considerably high (hazard ratio of 14).²¹ Recently, we reported the incidence, mutational spectrum, and evolution of CH in older individuals with anemia.²² To the best of our knowledge, no studies have performed comprehensive mutational screening to detect CH (in genes other than *JAK2*) in individuals with erythrocytosis.

We aimed to assess the mutational spectrum and outcomes associated with erythrocytosis in community-dwelling individuals, using data and biomaterial from the Lifelines cohort, which is an extensive multidimensional population-based cohort study.

Methods

Subjects

This cross-sectional study used data from 147 167 adult subjects participating in the Lifelines cohort study of whom hemoglobin (Hb) and hematocrit (Hct) levels have been measured in a peripheral blood sample at baseline. Lifelines is a multidisciplinary prospective population-based cohort study examining, in a unique 3-generation design, the health and health-related behaviors of 167 729 persons living in the north of The Netherlands. It uses a broad range of investigative procedures in assessing the biomedical, socio-demographic, behavioral, physical, and psychological factors that contribute to the health and disease of the general population, with a special focus on multimorbidity and complex

genetics. All participants provided written informed consent before participating in the study. The study protocol was approved by the medical ethical review committee of the University Medical Center Groningen.^{23,24} Details regarding physical examination, questionnaires, and biochemical measurements are provided in the supplemental Methods.

Definition of erythrocytosis, thrombocytosis, and leukocytosis

Erythrocytosis was defined in 2 ways. The first way was according to the 2008 WHO classification of myeloid neoplasms and the British Committee for Standards in Hematology as Hb concentration >18.5 g/dL or Hct $\geq 52\%$ in males and Hb concentration >16.5 g/dL or Hct $\geq 48\%$ in females (referred to as strict criteria).^{1,11,12} The second way was according to the 2016 WHO classification as a Hb concentration >16.5 g/dL or Hct $>49\%$ in males and Hb concentration >16.0 g/dL or Hct $>48\%$ in females (referred to as wide criteria).¹³ Local laboratory reference intervals were used to define leukocytosis ($\geq 10 \times 10^9/L$) and thrombocytosis ($\geq 400 \times 10^9/L$).

Next-generation sequencing

We hypothesized that individuals meeting the strict criteria for erythrocytosis and having concurrent thrombocytosis and/or leukocytosis had the highest chance of primary erythrocytosis. We performed a subanalysis in which all individuals with erythrocytosis and concurrent leuko- and/or thrombocytosis were matched 1:2 (for age, sex, body mass index [BMI], smoking status, and number of medications used) with individuals having isolated erythrocytosis. Targeted next-generation sequencing (NGS), using the Illumina TruSight Myeloid Sequencing Panel with a filter on 25 genes, was performed in individuals with DNA samples available (supplemental Figure 1). The minimum variant allele frequency (VAF) for variant calling was set to 5.0%. Details of the gene panel, sample processing, library preparation, clustering and sequencing, data analysis, and variant classification are described in the supplemental Methods.

Digital droplet polymerase chain reaction

For additional absolute quantification of the *JAK2* V617F mutation in the sequenced cohort, we used the *JAK2* V617F digital droplet polymerase chain reaction (ddPCR) Mutation Assay (catalog no. 10049550), the QX200 Droplet Digital PCR System, and the QuantaSoft Version 1.7.4 analysis software from Bio-Rad (Bio-Rad Laboratories GmbH), according to manufacturer's instructions. This assay has a detection limit of 0.01%.

Definition of cardiovascular disease

By questionnaire, participants were asked to report their medical history regarding cardiovascular disease (CVD).^{25,26} Thrombosis was defined as self-reported thrombosis or pulmonary embolism. Stroke was defined as self-reported stroke and use of vitamin K antagonists (Anatomical Therapeutic Chemical [ATC] code B01AA) or platelet aggregation inhibitors (ATC code B01AC) or direct thrombin inhibitors (ATC code B01AE). A history of myocardial infarction was defined based on electrocardiogram findings in combination with self-reported medical history of infarction and/or use of antithrombotic medication, as described previously.²⁷

Overall and cause-specific survival

Nearly all participants ($n = 147\,076$) were followed from the moment of their inclusion in the Lifelines cohort until death or up to June 2019 (median follow-up, 100 months; maximum, 163 months). Survival data were obtained from the Municipal Personal Records Database. The primary cause of death was ascertained by linkage to the national death statistics registry (Statistics Netherlands). Based on data availability, cause of death analyses were censored to 31 December 2019. Cause of death was coded according to the International Classification of Diseases, 10th revision (ICD-10). We used ICD-10 codes I00-I99 to classify death from CVD and ICD-10 codes C81-C96 and D45-D47 to classify death from hematologic malignancy.

Diagnosis of hematologic malignancies

Lifelines data were linked to pathology report data from the nationwide network and registry of histo- and cytopathology in the Netherlands Pathologisch Anatomisch Landelijk Geautomatiseerd Archief (PALGA). This databank includes reports from all pathology laboratories in the Netherlands and has nationwide coverage since 1991.²⁸ All Lifelines participants were linked at high certainly level by using pseudonyms based on the first 8 characters of their last name, date of birth, sex, and initials (missing values allowed), and the 4 digits of the postal code. Reports were collected until October 2019. We manually screened all histopathology report conclusions from bone marrow biopsies and reports coded as hematologic malignancy to identify histology reports with diagnosis of hematologic malignant disease including MPN.

Data description and statistical analysis

All statistical analyses were performed using IBM SPSS software, version 23.0 (SPSS Inc., Chicago, IL) and R3.5.2 statistical computing software (supplemental Methods). Between-group differences were evaluated using analysis of variance or, for non-normally distributed variables, using the Kruskal-Wallis test. The χ^2 test was used for statistical comparison of categorical variables. Univariable and multivariable logistical regression analyses were used to study the association between erythrocytosis and cardiovascular events. The composite variable for cardiovascular events was defined as a history of thrombosis and/or stroke and/or myocardial infarction. Survival curves were generated using the Kaplan-Meier method. Univariable and multivariable Cox regression analyses were used to analyze the effect of erythrocytosis on overall survival. For cause-specific survival, competing risk regression according to the method of Fine and Gray was used, with death from other causes considered competing risk. All multivariable analyses concerning cardiovascular events and mortality were performed with adjustment for potential confounders including age, sex, BMI, systolic and diastolic blood pressure, smoking status (nonsmoker, former smoker, and current smoker), medical history of diabetes (based on self-report, elevated fasting glucose [≥ 7.0 mmol/L] or antidiabetic drug use), number of medications used (as a proxy for comorbidity), use of antihypertensive drugs (ATC code C02-4 and C07-9), use of drugs for obstructive airway diseases (ATC code R03), use of androgens (ATC code G03B), family history of CVD (parent, sibling, or child with CVD before age 60 years),²⁹ and low-density lipoprotein (LDL) cholesterol levels. A reduction in plasma volume, which can cause relative erythrocytosis, may be present in individuals with hypovolemia (eg, induced

by the use of diuretics) or in individuals with Gäisbock's syndrome, which is classically described as erythrocytosis in tense/anxious individuals with hypertension and obesity as risk factors.^{30,31} Therefore, as sensitivity analysis, we repeated the analyses in individuals without use of diuretics (ATC code C03) and without metabolic syndrome (defined according to the revised criteria of the National Cholesterol Education Program's Adult Treatment Panel III.^{32,33} Odds ratios (ORs) and hazard ratios (HRs) are reported with 95% confidence interval (CI). All statistical tests were performed 2-sided, and $P < .05$ was considered significant.

Results

Prevalence and characteristics of erythrocytosis in a population-based cohort

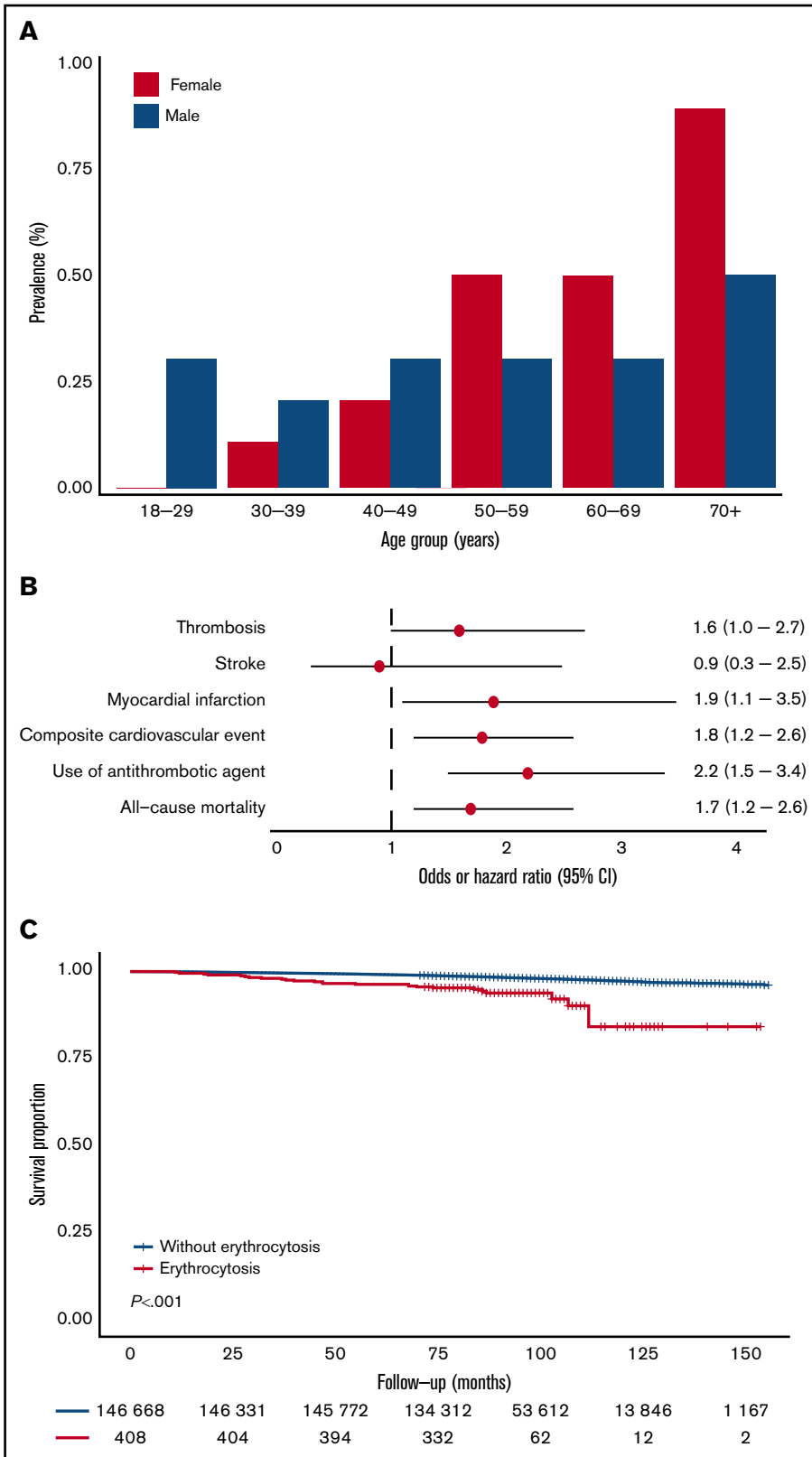
A total of 408 individuals (0.3%), 185 (0.3%) males and 223 (0.3%) females, met the strict criteria for erythrocytosis. The prevalence of erythrocytosis in females gradually increased with age to approximately 1% in the population >70 years. In males, the prevalence was relatively stable across age categories (around 0.3%), with an increase to 0.5% for males >70 years (Figure 1A). When using the wide criteria, 4995 (3.4%) individuals, 4686 (7.6%) males and 309 (0.4%) females, had erythrocytosis. In females, the prevalence of erythrocytosis increased with age, whereas in males, the prevalence decreased (supplemental Figure 2A). Relevant baseline characteristics for individuals with ($n = 408$) and without ($n = 146\,759$) erythrocytosis are shown in Table 1. Individuals with erythrocytosis were older, had a higher BMI, used more medications, including antihypertensive drugs and androgens, were more frequently current smokers, and had higher blood pressure. Total cholesterol, LDL cholesterol, and triglycerides were significantly higher in individuals with erythrocytosis. Furthermore, erythrocytosis was associated with higher mean corpuscular volume (MCV) levels and higher rates of leukocytosis and thrombocytosis. Baseline characteristics for individuals with and without erythrocytosis defined with the wide criteria are shown in supplemental Table 1; in general, they had the same but less pronounced results.

Association between erythrocytosis and cardiovascular events and mortality

Next, we evaluated the association of erythrocytosis with cardiovascular outcomes. In multivariable logistic regression analysis, the presence of erythrocytosis (strict criteria) was associated with myocardial infarction, use of antithrombotic agents, and composite cardiovascular events (Figure 1B; Table 2). During a median follow-up of 100 months, 26 (6.4%) individuals with erythrocytosis and 2826 (1.9%) individuals without erythrocytosis died ($P < .001$). Cause of death could be evaluated in 25 individuals with erythrocytosis. Increased all-cause (HR, 1.7; 95% CI, 1.2-2.6; $P = .003$) and cardiovascular mortality (HR, 2.2; 95% CI, 1.0-4.6; $P = .042$) was observed for individuals with erythrocytosis, using a multivariable model (Figure 1B-C; Table 2; Supplemental Figure 3; Supplemental Table 2). None of the individuals died of hematologic malignancies. Because of low numbers, no cause-specific survival analyses could be performed comparing individuals with erythrocytosis with or without concurrent leuko- and/or thrombocytosis. However, the proportion of individuals with cardiovascular morbidity (12.8% vs 6.4%, $P = .20$) and all-cause mortality (8.5% vs 4.3%, $P = .30$) was twice as high in individuals with

Figure 1. Prevalence of erythrocytosis (strict criteria) as function of sex and age and the association of erythrocytosis with CVD and survival.

(A) Prevalence of erythrocytosis according to sex and age categories. (B) Forest plot for the risk of CVD and all-cause mortality. Logistic regression analyses and Cox proportional hazards regression included age, sex, BMI, systolic and diastolic blood pressure, smoking status, medical history of diabetes, number of medications used, androgen drug use, drugs for obstructive airway disease, antihypertensive drug use, family history of CVD, and LDL cholesterol as covariates. Absence of erythrocytosis was used as reference. Circles indicate the OR/HR, with horizontal lines corresponding to 95% CIs. (C) Kaplan-Meier curve for overall survival, stratified according to the presence of erythrocytosis.



a concurrent leuko- and/or thrombocytosis compared with their matched controls with isolated erythrocytosis, although this was not statistically significant (supplemental Table 3). We additionally

compared individuals with erythrocytosis with and without concurrent cytosis with individuals without erythrocytosis. As shown in supplemental Figure 4, the association between erythrocytosis and

Table 1. Baseline characteristics, CVD history, and all-cause mortality of the study cohort

	General Lifelines population (n = 146 759)	Individuals with erythrocytosis (strict criteria) (n = 408)	P
Male/female, %	41.6/58.4	45.4/54.7	.13
Age, y	44.8 ± 13.1	51.5 ± 13.1	<.001
BMI, kg/m ²	26.1 ± 4.3	27.9 ± 4.9	<.001
Systolic BP, mm Hg	125 ± 15	135 ± 17	<.001
Diastolic BP, mm Hg	74 ± 9	78 ± 10	<.001
Creatinine, μmol/L	73.5 ± 13.7	77.7 ± 15.6	<.001
Number of medications used	1 (0-2)	1 (0-3)	<.001
Use of antihypertensive drugs, %	12.5	24.0	<.001
Use of drugs for obstructive airway diseases, %	7.0	11.0	.001
Use of androgens, %	0.1	0.5	.004
Current smokers, %	20.7	38.2	<.001
Diabetes, %	3.4	7.8	<.001
Metabolic syndrome, %	17.3	35.7	<.001
Family history of CVD, %	8.9	18.2	<.001
Hb, g/dL			
Male	15.1 ± 0.9	17.9 ± 0.6	<.001
Female	13.4 ± 0.9	16.2 ± 0.7	<.001
Hct, %			
Male	44.7 ± 2.5	52.9 ± 1.4	<.001
Female	40.5 ± 2.5	49.0 ± 1.8	<.001
MCV, fL	89.9 ± 4.2	92.2 ± 4.3	<.001
Leukocytes, ×10 ⁹ /L	6.1 ± 1.8	7.4 ± 2.1	<.001
Leukocytosis, %	3.0	10.0	<.001
Platelets, ×10 ⁹ /L	250 ± 57	245 ± 74	.096
Thrombocytosis, %	1.2	2.2	<.001
Fasting glucose, mmol/L	5.0 ± 0.8	5.2 ± 1.0	<.001
HbA1c, %	5.5 ± 0.5	5.6 ± 0.5	<.001
Total cholesterol, mmol/L	5.1 ± 1.0	5.6 ± 1.2	<.001
LDL cholesterol, mmol/L	3.2 ± 0.9	3.7 ± 1.1	<.001
HDL cholesterol, mmol/L			
Male	1.3 ± 0.3	1.3 ± 0.3	.016
Female	1.6 ± 0.4	1.5 ± 0.4	.001
Triglycerides, mmol/L	1.2 ± 0.8	1.5 ± 0.9	<.001
Thrombosis, %	1.7	4.2	<.001
Stroke, %	0.6	0.5	.24
Myocardial infarction, %	1.2	3.4	<.001
Composite cardiovascular events, %	3.3	8.6	<.001
Antithrombotic agent use, %			
Vitamin K antagonists	0.9	3.4	<.001
Platelet aggregation inhibitors	3.2	8.6	<.001
Direct factor Xa inhibitors	0.0	0.0	.83
All-cause mortality, %	1.9	6.4	<.001

Data are given as mean ± standard deviation, median (interquartile range) when not normally distributed, or percentage. Leukocytosis was defined as leukocytes $\geq 10 \times 10^9/L$. Thrombocytosis was defined as a platelet count $\geq 400 \times 10^9/L$.

BP, blood pressure; HDL, high-density lipoprotein; HbA1c, glycated hemoglobin.

cardiovascular outcomes was stronger in those with concurrent cytosis than in those with isolated erythrocytosis. However, isolated erythrocytosis was still significantly associated with composite

cardiovascular events (OR, 1.6; 95% CI, 1.1-2.4; $P = .027$) and all-cause mortality (HR, 1.7; 95% CI; 1.1-2.6; $P = .016$). The analyses were repeated for individuals who met the wide criteria for

Table 2. Results of multivariable regression analysis showing the associations between erythrocytosis (strict criteria) and CVD history and mortality, stratified according to the presence of concurrent leuko- and/or thrombocytosis or clonal hematopoiesis

	Unadjusted OR/HR (95% CI)	Adjusted OR/HR (95% CI)
Erythrocytosis*		
Thrombosis	2.6 (1.6-4.3)†	1.6 (1.0-2.7)
Stroke	1.8 (0.7-4.9)	0.9 (0.3-2.5)
Myocardial infarction	2.9 (1.7-5.0)†	1.9 (1.1-3.5)†
Composite cardiovascular events	2.8 (2.0-4.0)†	1.8 (1.2-2.6)†
Antithrombotic agents	3.0 (2.2-4.1)†	2.2 (1.5-3.4)†
All-cause mortality	3.6 (2.4-5.4)†	1.7 (1.2-2.6)†
Cardiovascular mortality‡	6.3 (3.1-12.6)†	2.2 (1.0-4.6)†
Erythrocytosis with concurrent leuko- and/or thrombocytosis§		
Composite cardiovascular events	1.8 (0.7-4.5)	2.2 (0.8-6.1)
All-cause mortality	1.2 (0.3-3.9)	0.9 (0.2-3.7)
Erythrocytosis with clonal hematopoiesis 		
Composite cardiovascular events	3.4 (1.0-11.9)	9.1 (1.2-68.4)†
All-cause mortality	2.1 (0.5-9.3)	1.1 (0.1-8.5)

Adjusted regression models include sex, age, BMI, systolic and diastolic blood pressure, smoking status, medical history of diabetes, number of medications used, androgen drug use, drugs for obstructive airway disease, antihypertensive drug use, family history of cardiovascular disease and low-density lipoprotein cholesterol as covariates. Data are shown as OR/HR and 95% CI.

*The absence of erythrocytosis was used as reference.

† $P < .05$.

‡Results are based on calculations by the authors using nonpublic microdata from Statistics Netherlands. Under certain conditions, these microdata are accessible for statistical and scientific research. For further information: microdata@cbs.nl.

§Isolated erythrocytosis was used as reference.

||The absence of clonal hematopoiesis in individuals with erythrocytosis was used as reference.

erythrocytosis, without significant associations between composite cardiovascular events and survival (supplemental Figures 2B-C and 3; supplemental Tables 4 and 5). Sensitivity analyses in individuals without use of diuretics and without metabolic syndrome did not materially alter the associations (supplemental Table 6).

Spectrum of clonal hematopoiesis for individuals with erythrocytosis

A total of 47 individuals with erythrocytosis and leuko- or thrombocytosis were matched with 94 individuals with isolated erythrocytosis (strict criteria) (supplemental Figure 1). From this cohort, available DNA samples from 133 individuals were further characterized with NGS to determine the presence and spectrum of CH. We identified 94 clonal somatic mutations in 51 individuals (38%; Figure 2A). The prevalence of CH was comparable for individuals with (18 of 45, 40%) and without (33 of 88, 37.5%) concurrent leuko- and/or thrombocytosis ($P = .78$). Of 51 individuals with CH, 28 (55%) individuals had more than 1 mutation; 18 of 28 had 2 mutations, and 10 of 28 had ≥ 3 mutations (Figure 2B). The maximum detected VAF per individual varied between 5% and 78% (median, 12%).

Our data revealed 3 different mutational spectra (genotypes) associated with erythrocytosis. First, the *JAK2* V617F mutation with

a VAF $\geq 5\%$, which in combination with erythrocytosis is considered as undiagnosed PV, was observed in 7 of 133 (5.3%) individuals. Comutation patterns involving *JAK2* V617F mutation were detected in 2 individuals with both concurrent *TP53* mutations. Second, as expected from previous population-based studies, we identified a high number of individuals with mutations in genes implicated in CHIP (24 of 133, 18%), including *DNMT3A* (19 of 133; 14%), *TET2* (8 of 133; 6%), or *ASXL1* (6 of 133; 5%). Third, 20 of 133 (15%) individuals revealed either a mutation in *BCOR* (13 of 133; 10%) or *BCORL1* (8 of 133; 6%; Figure 2A,D). Of note, in 10 of 20 (50%) cases, the mutation in *BCOR/BCORL1* was truncating (supplemental Figure 5).

All individuals with a *JAK2* V617F mutation had concurrent leuko- and/or thrombocytosis (7 of 45; 15.6%; Figure 2C-D). In addition, individuals with a *JAK2* V617F mutation were older, had a lower MCV, and more frequently had a history of cardiovascular events compared with individuals without mutations (Table 3). The maximum VAF per individual was significantly higher in *JAK2* mutant individuals compared with the other 2 genotypes (median for *JAK2* V617F, 31% [range, 14%-78%]; for mutations in genes implicated in CHIP, 9% [range, 5%-51%]; and for mutated *BCOR/BCORL1*, 13% [range, 5%-45%]; $P = .005$).

Next, we evaluated all 133 individuals for the presence of smaller *JAK2* V617F clones using ddPCR, revealing the presence of *JAK2* V617F mutations with VAF $< 5\%$ in another 7 individuals. However, the clinical characteristics of these individuals were not comparable to individuals with a *JAK2* V617F mutation with VAF $\geq 5\%$ (supplemental Table 7). There were no differences in baseline characteristics for individuals with the other 2 genotypes (ie, genes implicated in CHIP and *BCOR/BCORL1*) compared with those without mutations (Table 3; supplemental Table 8).

Diagnosed hematologic malignancy for individuals with erythrocytosis

Of the 408 individuals with erythrocytosis (strict criteria), 9 (2.2%) were diagnosed with any hematologic malignancy during follow-up compared with 453 (0.3%) of individuals without erythrocytosis. Most malignancies were of myeloid origin (6 of 9). NGS data were available for 5 of 9 individuals. In these, MPNs were exclusively diagnosed in individuals with *JAK2* V617F mutation at VAF $\geq 5\%$ ($n = 4$). Using the wide criteria for erythrocytosis, a total of 26 (0.5%) individuals with erythrocytosis were diagnosed with any hematologic malignancy during follow-up compared with 436 (0.3%) without erythrocytosis. A total of 11 of 26 were of myeloid origin, and all myeloid malignancies were MPNs (supplemental Table 9).

Association between CH and cardiovascular events and mortality for individuals with erythrocytosis

In line with previously observed associations between CH and CVD in the general population,¹⁵ we questioned whether the presence of CH is associated with an additional risk of CVD in individuals with erythrocytosis. Adjusted for potential confounders, individuals with erythrocytosis and CH were at higher risk for composite cardiovascular events (OR, 9.1; 95% CI, 1.2-68.4; $P = .031$) compared with individuals without CH. In contrast, no difference for all-cause mortality (HR, 1.1; 95% CI, 0.1-8.5; $P = .92$) could be observed between individuals with and without CH (Table 2). Hampered by low numbers, we were unable to relate differences in

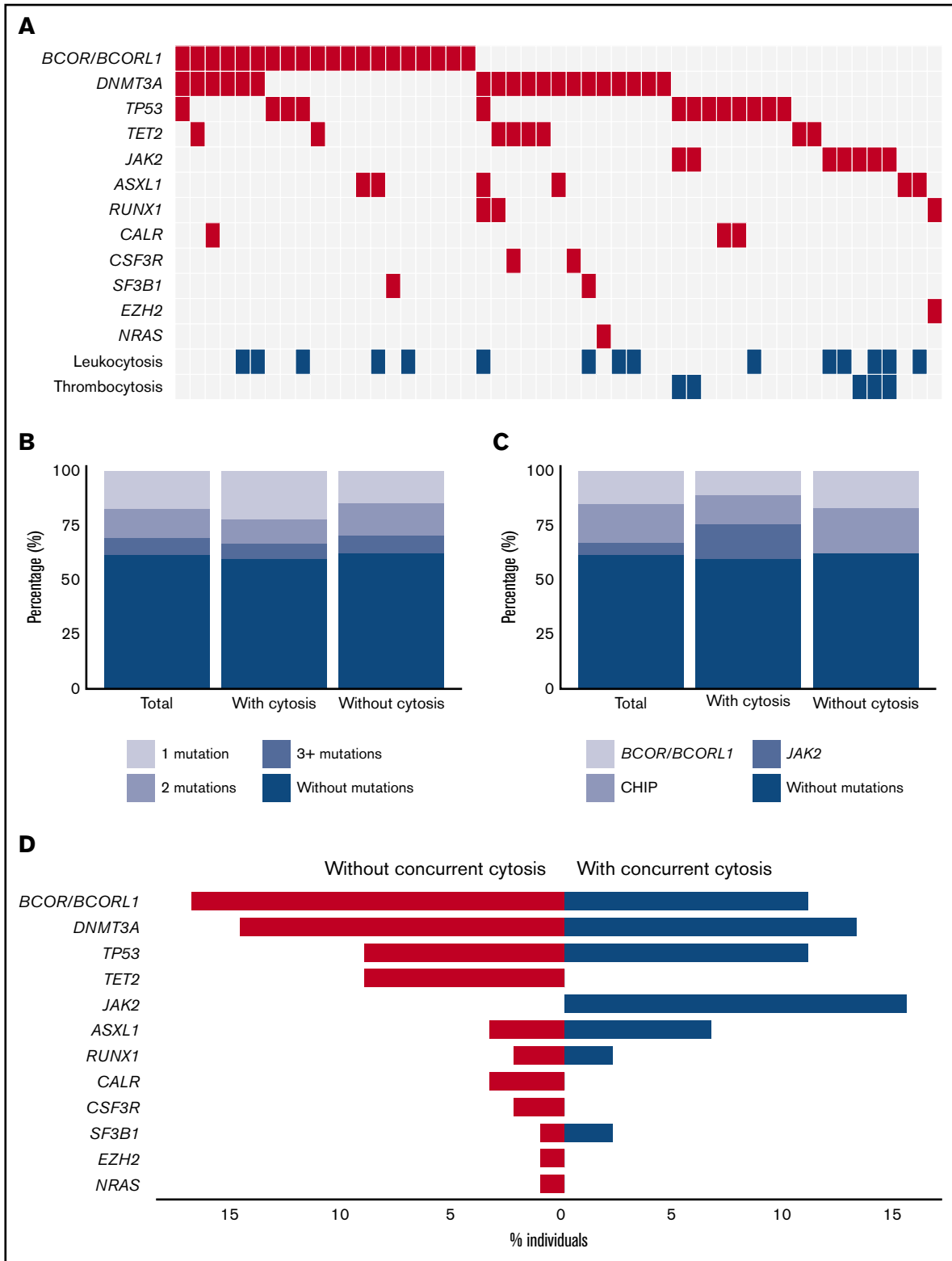


Figure 2. Spectrum of clonal hematopoiesis for individuals with erythrocytosis with and without concurrent leuko- and/or thrombocytosis. (A) Mutational landscape for all somatic mutations detected in sequenced individuals with erythrocytosis ($n = 133$). (B) Proportion of individuals with 1 or more mutations according to the presence of concurrent leuko- and/or thrombocytosis. (C) Proportion of individuals with mutated *JAK2*, *BCOR/BCORL1*, or other mutations (including CHIP genes) according to the presence of concurrent leuko- and/or thrombocytosis. (D) Contribution of individual mutated genes to the spectrum of clonal hematopoiesis detected in individuals with isolated erythrocytosis (red) and individuals with erythrocytosis with concurrent leuko- and/or thrombocytosis (blue). Bars indicate the proportion of individuals with a gene mutation.

Table 3. Baseline characteristics, cardiovascular disease, and all-cause mortality according to genotype

	Without mutations (n = 82)	JAK2 V617F (n = 7)	P	BCOR/BCORL1 (n = 20)	P	Other mutations (n = 24)	P
Male/female, %	30.5/69.5	42.9/57.1	.50	30.0/70.0	.97	16.7/83.3	.18
Age, y	48.4 ± 10.3	61.3 ± 10.4	.002	53.4 ± 12.4	.066	53.3 ± 8.2	.033
Hb, g/dL							
Male	18.5 ± 2.3	17.8 ± 0.7	.18	17.8 ± 0.6	.88	18.4 ± 0.9	.12
Female	16.2 ± 0.4	16.2 ± 0.4	.86	16.3 ± 0.8	.57	16.1 ± 0.5	.35
Hct, %							
Male	52.8 ± 0.7	57.5 ± 9.0	.007	52.8 ± 0.7	.85	53.3 ± 0.1	.20
Female	49.1 ± 2.1	49.6 ± 1.4	.64	48.8 ± 0.7	.62	48.9 ± 1.1	.81
MCV, fL	93.5 ± 4.3	89.3 ± 4.7	.016	93.2 ± 4.0	.79	92.7 ± 4.1	.42
Leukocytes, × 10 ⁹ /L	8.6 ± 2.7	9.8 ± 3.0	.28	7.9 ± 2.4	.29	8.3 ± 2.7	.54
% Leukocytosis	29.3	4/7 (57.1%)	.31	25.0	.81	25.0	.61
Platelets, × 10 ⁹ /L	270 ± 62	524 ± 180	<.001	242 ± 60	.07	254 ± 54	.25
% Thrombocytosis	4.9	5/7 (71.4%)	<.001	0.0	.34	0.0	.46
Composite cardiovascular events, %	4.9	2/7 (28.6%)	.016	15.0	.11	12.5	.19
All-cause mortality, %	4.9	1/7 (14.3%)	.30	10.0	.38	4.2	.89

Data are given as mean ± standard deviation, median (interquartile range) when not normally distributed, or percentage. *P* values were calculated between individuals without mutations and the 3 genotypes. Leukocytosis was defined as leukocytes ≥ 10 × 10⁹/L. Thrombocytosis was defined as a platelet count ≥ 400 × 10⁹/L. The category of other mutations includes mutations in *ASXL1*, *CALR*, *CSF3R*, *DNMT3A*, *EZH2*, *NRAS*, *RUNX1*, *SF3B1*, *TET2*, and *TP53*.

cardiovascular morbidity and mortality to the mutational spectrum of individuals with erythrocytosis.

Discussion

In this comprehensive population-based study, we explored the spectrum of clonal hematopoiesis and cardiovascular outcomes in individuals with erythrocytosis. We demonstrated that erythrocytosis in community-dwelling individuals was associated with cardiovascular morbidity, cardiovascular mortality, and all-cause mortality using strict criteria, but not when using wide criteria, for erythrocytosis. Additionally, the presence of CH in individuals with erythrocytosis was strongly associated with cardiovascular morbidity.

In comparison with the 2008 WHO classification of myeloproliferative neoplasms, in the revision of 2016, the Hb threshold levels were lowered and Hct thresholds were introduced.¹³ The reason for this change was the assumption of underdiagnosis of PV using the 2008 criteria, because it was demonstrated that a considerable proportion of patients with PV had lower Hb/Hct values (masked PV) and were even characterized by inferior overall survival and higher risk of thrombosis compared with overt PV.³⁴ In our study, 5 individuals were diagnosed with a MPN during follow-up who met the wide criteria but not the strict criteria for erythrocytosis. However, the lower thresholds have overlap with the normal values of Hb and Hct, especially in males. Reflecting this, the prevalence of erythrocytosis observed in our study was comparable with previous population-based studies using different criteria.³⁵⁻³⁷ The lowered thresholds have been subject of debate, because it may lead to a large increase of cases suspected of PV and subsequently an increase in investigations. Although for the diagnosis of PV, Hb/Hct is not the only criterium, the lower thresholds are integrated in clinical guidelines for the investigation of erythrocytosis in general.^{37,38} In our study, only erythrocytosis defined using the strict criteria was significantly associated with cardiovascular

morbidity, cardiovascular mortality, and all-cause mortality, after adjustments for conventional cardiovascular risk factors. Our data are in accordance with previous studies, reporting a U-shaped association between Hb and Hct concentrations and mortality.^{39,40} Taken together, these data suggest that the lower Hb and Hct thresholds in the revised WHO classification might be useful to diagnose masked PV but might be too low as screening tool for secondary erythrocytosis given the absence of associations with negative health outcomes. Given the association between erythrocytosis with and without CH and CVD, it could be wondered if these individuals should be considered for preventive intervention (eg, aspirin) of their cardiovascular risk.

Several studies have been performed to assess the prevalence of *JAK2* V617F mutations in the general population and in populations with abnormal hematologic parameters (supplemental Table 10). Prevalence of *JAK2* V617 mutations was 0.1% in a Danish population-based screening (n = 49 488),⁴¹ 0.19% in a cohort of individuals with bipolar disorder or schizophrenia and controls (n = 12 380),¹⁰ and 0.18% in a cohort of mainly patients with diabetes and controls (n = 17 182).¹⁴ A recent Danish population-based cohort (n = 19 958) using ddPCR with an increased sensitivity of 0.009% reported *JAK2* V617F mutations in 3.1% of individuals.⁴² In individuals with abnormal hematologic parameters, the prevalence of *JAK2* V617F ranged from 0% to 4.8% (publications on this topic are summarized in supplemental Table 10). Our data revealed a comparably high prevalence of *JAK2* V617F mutations (VAF ≥ 5%) of 5.3% in individuals with erythrocytosis and 15.6% in individuals having erythrocytosis with concurrent leuko- and/or thrombocytosis. Subsequently, 4 of 7 *JAK2* V617 mutant individuals have been diagnosed with MPN during follow-up. It should be noted that abnormal Hb concentrations at inclusion were communicated with the participant and their general practitioner.

Unexpectedly, we detected a high prevalence of CH (38%) in this relatively young cohort and a remarkably high number of mutations

in *BCOR/BCORL1*. This is in contrast with the spectrum of CH found in previous population-based cohorts, which is mainly characterized by mutations in *DNMT3A*, *ASXL1*, and *TET2*.^{10,14,15} These studies, however, were performed in large cohorts unselected for a hematologic phenotype. To the best of our knowledge, no studies have performed comprehensive mutational profiling (other than *JAK2* V617F) in community-dwelling individuals with erythrocytosis. The high prevalence of *BCOR/BCORL1* mutations seems more comparable to the mutational spectrum of patients with aplastic anemia, with a prevalence of *BCOR/BCORL1* mutations around 10%.⁴³ In aplastic anemia, the presence of *BCOR/BCORL1* mutations was associated with favorable outcome and low rates of progression to clonal myeloid disease.⁴⁴

BCOR/BCORL1 mutations are detected less frequently in other myeloid neoplasms, including MPN.⁴⁵ Moreover, none of the *BCOR/BCORL1* mutant individuals in this cohort developed myeloid neoplasms during follow-up, suggesting that *BCOR/BCORL1* mutations are associated with a low risk of malignant clonal progression also in the context of erythrocytosis. Although data about the role of *BCOR/BCORL1* in hematopoiesis are scarce, it has been shown that functional inactivation of *BCOR* modifies erythroid differentiation,^{46,47} which may possibly be implicated in the context of aplastic anemia (survival advantage) and erythrocytosis (growth advantage). Moreover, (currently unknown) environmental pressures may contribute to selective outgrowth of *BCOR/BCORL1* mutant hematopoietic clones in both conditions.

This study has some strengths and limitations. First, we were able to use information from a large number of participants with a wide range in age, socio-economic status, and comorbidities. All subjects were uniformly characterized. Mutational screening was performed thoroughly and concise; mutations were only reported when validated on 2 bioinformatic platforms, and *JAK2* V617F mutations were confirmed using ddPCR. Additionally, we were able to extend our dataset by linkage to national registries for data on cause of death and diagnosed hematologic malignancies. Several potential limitations should also be acknowledged. Because this is a cross-sectional study, the analyses do not provide information about causality. Additionally, because of the observational nature of the study, it was not possible to exclude the role of unknown or unmeasured confounding variables. Although the agreement between self-report questionnaires and medical record data about cardiovascular disease is substantial,^{25,26} the use of self-reported data may cause both over- and underestimation of the real prevalence. Hb/Hct levels were only measured once, so the erythrocytosis might be transient. Furthermore, no individual examination was performed to exclude the presence of relative erythrocytosis; however, the sensitivity analyses did not materially alter the associations.

References

1. McMullin MF. The classification and diagnosis of erythrocytosis. *Int J Lab Hematol*. 2008;30(6):447-459.
2. Mehta J, Wang H, Iqbal SU, Mesa R. Epidemiology of myeloproliferative neoplasms in the United States. *Leuk Lymphoma*. 2014;55(3):595-600.
3. Passamonti F, Rumi E, Pungolino E, et al. Life expectancy and prognostic factors for survival in patients with polycythemia vera and essential thrombocythemia. *Am J Med*. 2004;117(10):755-761.

In conclusion, erythrocytosis had a prevalence of 0.3% and was associated with cardiovascular morbidity, cardiovascular mortality, and all-cause mortality when defined using strict criteria in this large population-based cross-sectional study. In contrast, no significant associations were present when defining erythrocytosis according to the wide criteria. These data suggest that, although the lower Hb and Hct thresholds in the revised WHO 2016 classification might be useful to diagnose masked PV, they might be too low as a screening tool for secondary erythrocytosis given the absence of associations with cardiovascular morbidity and mortality. The prevalence of CH was relatively high (38%) in this predominantly young cohort of subjects with erythrocytosis and was strongly associated with cardiovascular morbidity. The *JAK2* V617F mutation was exclusively found in individuals with erythrocytosis and concurrent leuko- and/or thrombocytosis, emphasizing the value of these measures in the diagnostic workup. There was an unexpected high prevalence of mutations in *BCOR/BCORL1*, while its role in erythropoiesis and development of MPN is currently unknown.

Acknowledgments

The authors thank Arjan Simpelaar and Bettie Postema-Veenema for performing the NGS and ddPCRs, the services of the Lifelines Cohort Study, the contributing research centers delivering data to Lifelines, PALGA (Dutch Pathology Registry), and all study participants.

This work was supported by a research grant from Novartis, which had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. The Lifelines Biobank initiative has been made possible by subsidy from the Dutch Ministry of Health, Welfare and Sport, the Dutch Ministry of Economic Affairs, the University Medical Center Groningen, University Groningen, and the Northern Provinces of the Netherlands.

Authorship

Contribution: H.J.C.M.W., R.M., A.B.M., and G.H. contributed to the study design; H.J.C.M.W. and I.A.v.Z. performed the statistical analyses; H.J.C.M.W., R.M., I.A.v.Z., J.J.S., M.M.v.d.K., P.v.d.H., A.D., A.B.M., and G.H. contributed to the interpretation of the data and analysis; H.J.C.M.W. wrote the first version of the manuscript; and all authors read and approved the final version of the manuscript.

Conflict-of-interest disclosure: The authors declare no competing financial interests.

ORCID profiles: I.A.v.Z., 0000-0002-0771-1679; A.D., 0000-0001-9239-1050.

Correspondence: Gerwin Huls, Department of Hematology, University Medical Center Groningen, University of Groningen, PO Box 30.001, 9700 RB Groningen, The Netherlands; e-mail: g.huls@umcg.nl.

4. Tefferi A, Rumi E, Finazzi G, et al. Survival and prognosis among 1545 patients with contemporary polycythemia vera: an international study. *Leukemia*. 2013;27(9):1874-1881.
5. Griesshammer M, Kiladjian JJ, Besses C. Thromboembolic events in polycythemia vera. *Ann Hematol*. 2019;98(5):1071-1082.
6. Ohyashiki K, Akahane D, Gotoh A, et al. Uncontrolled thrombocytosis in polycythemia vera is a risk for thrombosis, regardless of JAK2^{V617F} mutational status. *Leukemia*. 2007;21(12):2544-2545.
7. Landolfi R, Di Gennaro L, Barbui T, et al; European Collaboration on Low-Dose Aspirin in Polycythemia Vera (ECLAP). Leukocytosis as a major thrombotic risk factor in patients with polycythemia vera. *Blood*. 2007;109(6):2446-2452.
8. Kroll MH, Michaelis LC, Verstovsek S. Mechanisms of thrombogenesis in polycythemia vera. *Blood Rev*. 2015;29(4):215-221.
9. Bhatt VR. Secondary polycythemia and the risk of venous thromboembolism. *J Clin Med Res*. 2014;6(5):395-397.
10. Genovese G, Kähler AK, Handsaker RE, et al. Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. *N Engl J Med*. 2014;371(26):2477-2487.
11. Tefferi A, Vardiman JW. Classification and diagnosis of myeloproliferative neoplasms: the 2008 World Health Organization criteria and point-of-care diagnostic algorithms. *Leukemia*. 2008;22(1):14-22.
12. McMullin MF, Reilly JT, Campbell P, et al; British Committee for Standards in Haematology. Amendment to the guideline for diagnosis and investigation of polycythaemia/erythrocytosis. *Br J Haematol*. 2007;138(6):821-822.
13. Arber DA, Orazi A, Hasserjian R, et al. The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia [published correction in *Blood*. 2016;128(3):462-463]. *Blood*. 2016;127(20):2391-2405.
14. Jaiswal S, Fontanillas P, Flannick J, et al. Age-related clonal hematopoiesis associated with adverse outcomes. *N Engl J Med*. 2014;371(26):2488-2498.
15. Xie M, Lu C, Wang J, et al. Age-related cancer mutations associated with clonal hematopoietic expansion. *Nat Med*. 2014;20(12):1472-1478.
16. Buscariol M, Provost S, Zada YF, et al. DNMT3A and TET2 dominate clonal hematopoiesis and demonstrate benign phenotypes and different genetic predispositions. *Blood*. 2017;130(6):753-762.
17. McKerrell T, Park N, Moreno T, et al; Understanding Society Scientific Group. Leukemia-associated somatic mutations drive distinct patterns of age-related clonal hemopoiesis. *Cell Rep*. 2015;10(8):1239-1245.
18. van den Akker EB, Pitts SJ, Deelen J, et al; Genome of The Netherlands Consortium. Uncompromised 10-year survival of oldest old carrying somatic mutations in DNMT3A and TET2. *Blood*. 2016;127(11):1512-1515.
19. Cook EK, Izukawa T, Young S, et al. Comorbid and inflammatory characteristics of genetic subtypes of clonal hematopoiesis. *Blood Adv*. 2019;3(16):2482-2486.
20. Jaiswal S, Natarajan P, Silver AJ, et al. Clonal hematopoiesis and risk of atherosclerotic cardiovascular disease. *N Engl J Med*. 2017;377(2):111-121.
21. Malcovati L, Galli A, Travaglino E, et al. Clinical significance of somatic mutation in unexplained blood cytopenia. *Blood*. 2017;129(25):3371-3378.
22. van Zeventer IA, de Graaf AO, Wouters HJCM, et al. Mutational spectrum and dynamics of clonal hematopoiesis in anemia of older individuals. *Blood*. 2020;135(14):1161-1170.
23. Stolk RP, Rosmalen JG, Postma DS, et al. Universal risk factors for multifactorial diseases: LifeLines: a three-generation population-based study. *Eur J Epidemiol*. 2008;23(1):67-74.
24. Scholtens S, Smidt N, Swertz MA, et al. Cohort profile: LifeLines, a three-generation cohort study and biobank. *Int J Epidemiol*. 2015;44(4):1172-1180.
25. Okura Y, Urban LH, Mahoney DW, Jacobsen SJ, Rodeheffer RJ. Agreement between self-report questionnaires and medical record data was substantial for diabetes, hypertension, myocardial infarction and stroke but not for heart failure. *J Clin Epidemiol*. 2004;57(10):1096-1103.
26. Haapanen N, Miilunpalo S, Pasanen M, Oja P, Vuori I. Agreement between questionnaire data and medical records of chronic diseases in middle-aged and elderly Finnish men and women. *Am J Epidemiol*. 1997;145(8):762-769.
27. van der Ende MY, Hartman MHT, Schurer RAJ, et al. Prevalence of electrocardiographic unrecognized myocardial infarction and its association with mortality. *Int J Cardiol*. 2017;243:34-39.
28. Casparie M, Tiebosch AT, Burger G, et al. Pathology databanking and biobanking in The Netherlands, a central role for PALGA, the nationwide histopathology and cytopathology data network and archive. *Cell Oncol*. 2007;29(1):19-24.
29. van der Ende MY, Siland JE, Snieder H, van der Harst P, Rienstra M. Population-based values and abnormalities of the electrocardiogram in the general Dutch population: The LifeLines Cohort Study. *Clin Cardiol*. 2017;40(10):865-872.
30. Smith JK. Gaisböck's syndrome: A case study. *Clin Med Rev Case Rep*. 2017;4(7):175.
31. Stefanini M, Urbas JV, Urbas JE. Gaisböck's syndrome: its hematologic, biochemical and hormonal parameters. *Angiology*. 1978;29(7):520-533.
32. Alberti KG, Eckel RH, Grundy SM, et al; International Association for the Study of Obesity. Harmonizing the metabolic syndrome: a joint interim statement of the International Diabetes Federation Task Force on Epidemiology and Prevention; National Heart, Lung, and Blood Institute; American Heart Association; World Heart Federation; International Atherosclerosis Society; and International Association for the Study of Obesity. *Circulation*. 2009;120(16):1640-1645.
33. Slagter SN, van Vliet-Ostapchouk JV, Vonk JM, et al. Combined effects of smoking and alcohol on metabolic syndrome: the LifeLines cohort study [published correction in *PLoS One*. 2014;9(8):e105157]. *PLoS One*. 2014;9(4):e96406.
34. Barbui T, Thiele J, Gisslinger H, et al. Masked polycythemia vera (mPV): results of an international study. *Am J Hematol*. 2014;89(1):52-54.

35. Ruggeri M, Tosetto A, Frezzato M, Rodeghiero F. The rate of progression to polycythemia vera or essential thrombocythemia in patients with erythrocytosis or thrombocytosis. *Ann Intern Med.* 2003;139(6):470-475.
36. Magnussen K, Hasselbalch HC, Ullum H, Bjerrum OW. Characterization of blood donors with high haemoglobin concentration. *Vox Sang.* 2013;104(2):110-114.
37. Sandes AF, Gonçalves MV, Chauffaille ML. Frequency of polycythemia in individuals with normal complete blood cell counts according to the new 2016 WHO classification of myeloid neoplasms. *Int J Lab Hematol.* 2017;39(5):528-531.
38. Barbui T, Thiele J, Gisslinger H, et al. The 2016 WHO classification and diagnostic criteria for myeloproliferative neoplasms: document summary and in-depth discussion. *Blood Cancer J.* 2018;8(2):15.
39. Lee G, Choi S, Kim K, et al. Association of hemoglobin concentration and its change with cardiovascular and all-cause mortality. *J Am Heart Assoc.* 2018;7(3):e007723.
40. Paul L, Jeemon P, Hewitt J, et al. Hematocrit predicts long-term mortality in a nonlinear and sex-specific manner in hypertensive adults. *Hypertension.* 2012;60(3):631-638.
41. Nielsen C, Birgens HS, Nordestgaard BG, Kjaer L, Bojesen SE. The JAK2 V617F somatic mutation, mortality and cancer risk in the general population. *Haematologica.* 2011;96(3):450-453.
42. Cordua S, Kjaer L, Skov V, Pallisgaard N, Hasselbalch HC, Ellervik C. Prevalence and phenotypes of JAK2 V617F and calreticulin mutations in a Danish general population. *Blood.* 2019;134(5):469-479.
43. Ogawa S. Clonal hematopoiesis in acquired aplastic anemia. *Blood.* 2016;128(3):337-347.
44. Yoshizato T, Dumitriu B, Hosokawa K, et al. Somatic mutations and clonal hematopoiesis in aplastic anemia. *N Engl J Med.* 2015;373(1):35-47.
45. Astolfi A, Fiore M, Melchionda F, Indio V, Bertuccio SN, Pession A. BCOR involvement in cancer. *Epigenomics.* 2019;11(7):835-855.
46. Chesnais V, Arcangeli ML, Delette C, et al. Architectural and functional heterogeneity of hematopoietic stem/progenitor cells in non-del(5q) myelodysplastic syndromes. *Blood.* 2017;129(4):484-496.
47. Kelly MJ, So J, Rogers AJ, et al. Bcor loss perturbs myeloid differentiation and promotes leukaemogenesis. *Nat Commun.* 2019;10(1):1347.