

Supplementary Appendix

Supplement to: van Zeventer IA, de Graaf AO, van der Klauw MM, Vellenga E, van der Reijden BA, Schuringa JJ, Diepstra A, Malcovati L, Jansen JH and Huls G. *Peripheral blood cytopenias in the ageing general population and risk of incident hematological disease and mortality.*

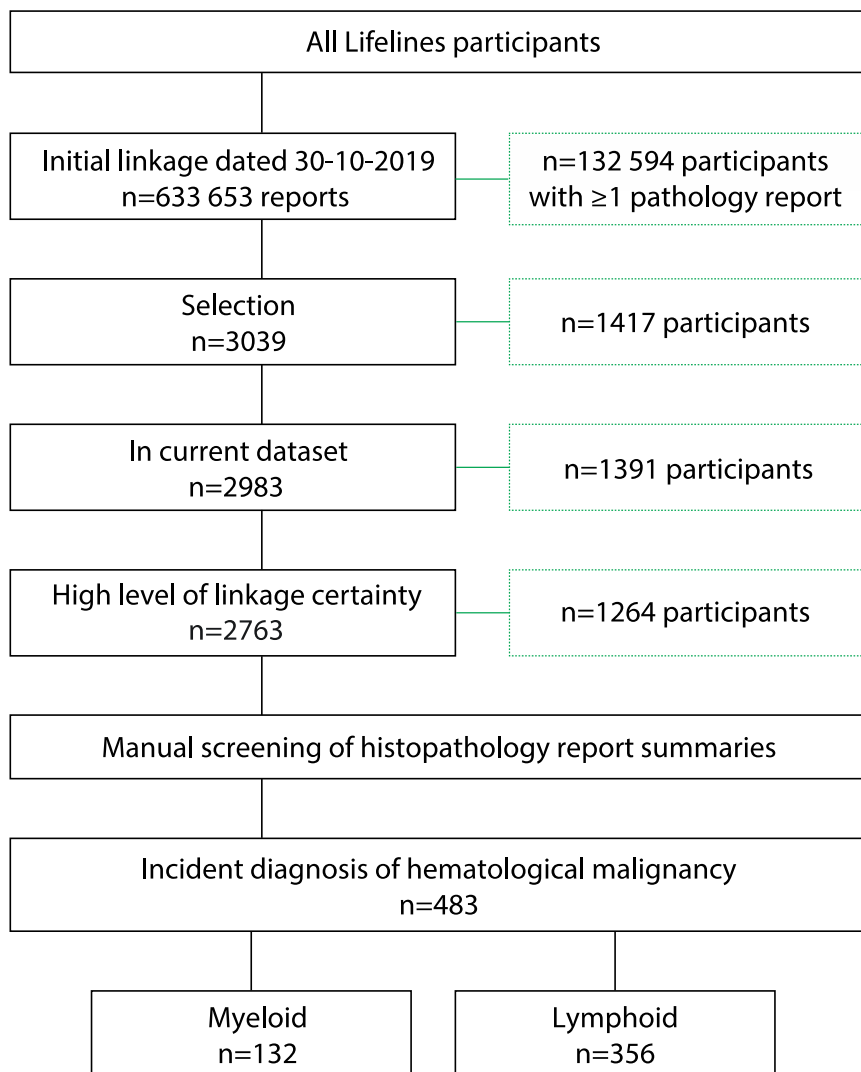
Recruitment and representativeness of the Lifelines cohort study.....	1
Retrieval of pathology reports from the Dutch Nationwide Network and Registry of Histo- and Cytopathology (PALGA).....	2
Linkage to the national death statistics registry (Statistics Netherlands) to retrieve causes of death.....	5
Supplementary Table 1.....	7
Supplementary Table 2.....	8
Supplementary Table 3.....	9
Supplementary Table 4.....	10
Supplementary Table 5.....	11
Supplementary Table 6.....	12
Supplementary Table 7.....	13
Supplementary Table 8.....	14
Supplementary Figure 1.....	15
Supplementary Figure 2.....	16
Supplementary Figure 3.....	17
Supplementary Figure 4.....	18
References.....	19

Recruitment and representativeness of the Lifelines cohort study

Different recruitment strategies were used for the Lifelines cohort study, aimed at a three-generation structure. First, an index population (25-49 years of age) was recruited via all general practitioners in the northern provinces of the Netherlands. Importantly, in the Netherlands, all inhabitants are registered at a general practitioner. Thus, the recruitment was independent of a recent visit at the general practitioner for health complaints. Individuals with limited life expectancy, severe psychiatric or physical illness or insufficient knowledge of the Dutch language were not invited for participation. Subsequently, the index population was asked to invite their family members to participate in the study. In addition, individuals could self-register as a participant, and were also asked to invite their family members.^{1,2}

The representativeness of the Lifelines population has previously been compared to the total adult background population in the north of the Netherlands. As a result of the recruitment strategy, middle-aged individuals were overrepresented. The Lifelines cohort comprised a lower proportion of elderly persons compared to the general population, which may be due to a higher rate of disabling diseases or less access to the internet for self-registration. Further, immigrants were relatively underrepresented. However, apart from demographic differences that were probably introduced by the recruitment strategy, Lifelines participants were found broadly representative of the background population in the north of the Netherlands with respect to socioeconomic characteristics, lifestyle factors, chronic diseases and overall health.² This indicates that the risk of selection bias is low and risk estimates are likely representing the associations in the general population.

Retrieval of pathology reports from the Dutch Nationwide Network and Registry of Histo- and Cytopathology (PALGA)

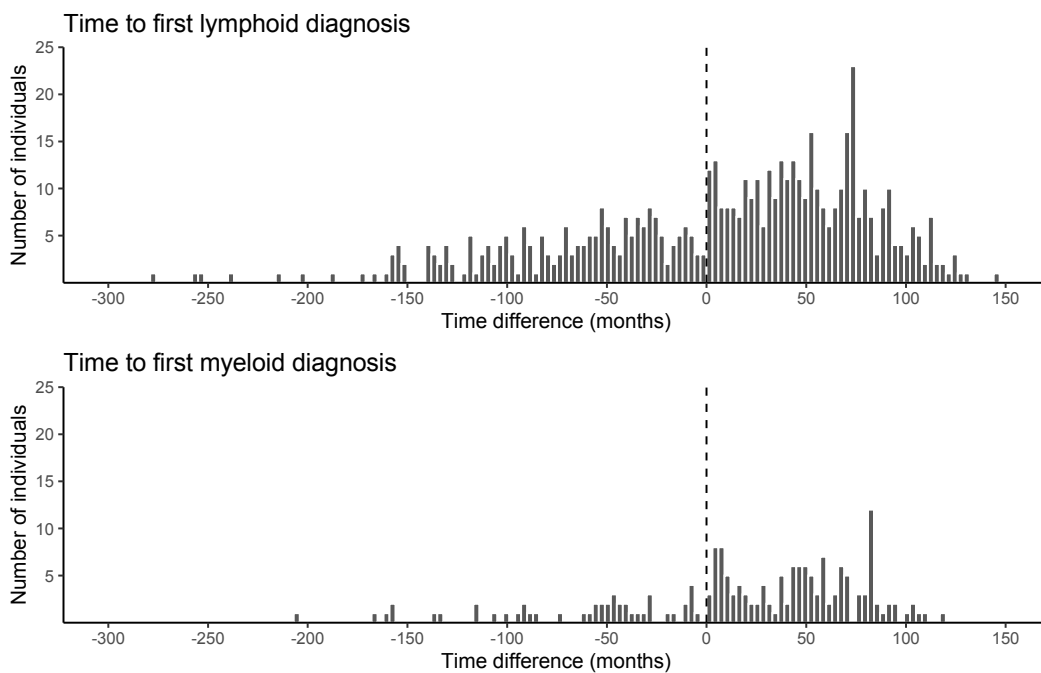


The Lifelines cohort was linked to the Dutch Nationwide Network and Registry of Histo- and Cytopathology (PALGA). This databank includes all reports from histopathology and cytopathology generated in pathology laboratories in the Netherlands and has a complete academic and non-academic national coverage since 1991.³ This study was approved by the Scientific Board of PALGA complying with the regulations for pseudonymized (epidemiological) studies.

The linkage strategy is outlined above. Individuals were linked by using coded personal information by which reports in the registry could be tracked. Reports were retrieved until 10-2019. To identify PALGA reports of interest, selection criteria were applied to include all available reports from bone marrow material, as well as reports with a diagnosis code (based on SNOMED nomenclature) in the following categories: 49 (malignant lymphomas), 50 (non-Hodgkin lymphoma), 51 (Hodgkin lymphoma), 52 (leukemia) and 60 (myeloid). After selection of relevant reports, the linkage procedure yielded 2983 reports for 1391 Lifelines participants in the current dataset. We further restricted to reports meeting criteria for high certainty linkage, based on pseudonyms of the first 8

characters of the last name, date of birth, sex, initials (missing values allowed) and the first 4 digits of the postal code. In total, 2763 reports for 1264 individuals were included for subsequent analyses.

All reports were manually screened and reviewed, and subsequently classified for the presence of a malignant hematological disorder. The diagnosis date was set at the first retrieved histopathology report for myeloid or lymphoid malignancy. Time to first myeloid diagnosis, relative to inclusion in the study, ranged from -205 to 118 months. For first lymphoid diagnosis, the timeframe ranged from -276 to 145 months.



Distribution in time differences between first diagnosis of lymphoid (upper panel) or myeloid (lower panel) malignancy and inclusion in the Lifelines cohort.

Prevalent malignancies, with a first diagnostic report before inclusion in the Lifelines study, were identified for 254 participants, n=47 myeloid and n=207 lymphoid.

We excluded incident reports which were considered non-diagnostic:

- Absence of malignancy n=198
- Minor or major criterium for mastocytosis (bone marrow or skin biopsy) n=15
- Solid malignancy in bone marrow n=7
- Irrelevant report (erroneously linked by miscoding) n=6
- Inadequate material n=9
- No classifying diagnosis (inconclusive description by pathologist) n=25

In total, 483 individuals experienced any first myeloid and/or lymphoid malignancy after the inclusion in the study, n=132 myeloid and n=356 lymphoid, with 5 individuals being diagnosed with both a first lymphoid and first myeloid malignant disorder.

To validate the linkage procedure, the number of linked records was compared to data from the PALGA databank across the Netherlands. We restricted the search to individuals with available pseudonyms for the first 4 characters of the last name, date of birth and sex. This was corresponding to the initial search strategy for Lifelines participants. The number of nationwide records from 1991 to 2019 was assessed relative to the total adult population in the Netherlands, as retrieved from Statistics Netherlands.⁴ The total number of PALGA records was comparable to the Dutch population. However, a higher percentage of records meeting the selection criteria was identified in the total Dutch population, which might be explained by the relative overrepresentation of middle-aged individuals in the Lifelines cohort and a lower proportion of elderly persons², and possibly also by regional health differences.

	Lifelines participants[#]	Total Dutch population*	Total Dutch population ≥18 years*
Total number of individuals	165106	17407585	14070340
Reference date	1-11-2019	1-1-2020	1-1-2020
Number of PALGA records	633653	73771598	72302154
Number of individuals	132594 (80.3%)	12055207 (69.3%)	11436920 (81.3%)
Number of PALGA records meeting selection criteria	3039	789606	766671
Number of individuals	1417 (0.9%)	341369 (2.0%)	330616 (2.3%)

[#]Linkage of records was performed in 10/2019.

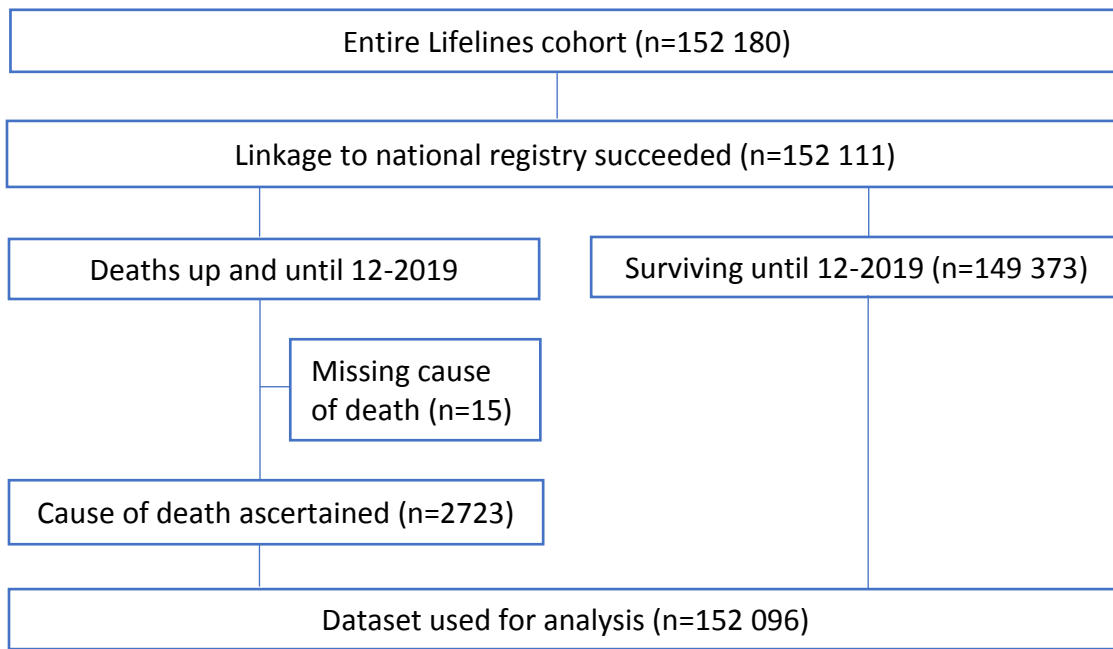
*Records were counted until 12/2019.

Linkage to the national death statistics registry (Statistics Netherlands) to retrieve causes of death*

For all individuals that die in the Netherlands, a death certificate is completed by a coroner or physician. Subsequently, all deaths are coded by Statistics Netherlands according to the International Classification of Diseases, tenth revision (ICD-10).⁵ The reliability of manual coding by Statistics Netherlands for major causes of death, including cancers, is high.⁶ In addition, automatic coding was implemented in 2013, which is thought to further improve reliability. Automatic coding was also found to be highly concordant with manual coding with respect to cancer-related deaths.⁷ To identify death from hematological neoplasms, ICD-10 codes C81x-C96x and D45x-D47x were used, as outlined below:

C81-C96		
Malignant neoplasms, stated or presumed to be primary, of lymphoid, haematopoietic and related tissue		
	C81	Hodgkin lymphoma
	C82	Follicular lymphoma
	C83	Non-follicular lymphoma
	C84	Mature T/NK-cell lymphomas
	C85	Other and unspecified types of non-Hodgkin lymphoma
	C86	Other specified types of T/NK-cell lymphoma
	C88	Malignant immunoproliferative diseases
	C90	Multiple myeloma and malignant plasma cell neoplasms
	C91	Lymphoid leukaemia
	C92	Myeloid leukaemia
	C93	Monocytic leukaemia
	C94	Other leukaemias of specified cell type
	C95	Leukaemia of unspecified cell type
	C96	Other and unspecified malignant neoplasms of lymphoid, haematopoietic and related tissue
D37-48		
Neoplasms of uncertain or unknown behaviour		
	D45	Polycythaemia vera
	D46	Myelodysplastic syndromes
	D47	Other neoplasms of uncertain or unknown behaviour of lymphoid, haematopoietic, and related tissue

Linkage to the national registry succeeded for 152 111 Lifelines participants. Data on primary cause of death were available for the entire inclusion period up and until 12-2019. Therefore, analyses on cause of death were censored to this date. We were able to retrieve the cause for 2723 out of 2738 deaths in this period. Individuals with missing cause of death were excluded, leaving a cohort of n=152 096 for further analyses.



Cause of death analyses were performed in the subset of individuals for which data on the respective blood cell lineage were available:

- Anemia: n=147 093
- Thrombocytopenia: n=146 997
- Neutropenia: n=144 603
- >1 cytopenia: n=144 512

*The results are based on calculations by the authors using non-public microdata from Statistics Netherlands. Under certain conditions, these microdata are accessible for statistical and scientific research. For further information: microdata@cbs.nl.

Supplementary Table 1. Proposed ICUS cut-off criteria

There is no general agreement with regard to cut-off criteria to define idiopathic cytopenias of undetermined significance (ICUS). A summary of proposed criteria is given below.

	1	2	3	4
Hemoglobin concentration	<12 g/dL for women, <13 g/dL for men	<11 g/dL	<10 g/dL	<10 g/dL
Platelet count	<150 x 10 ⁹ /L	<100 x 10 ⁹ /L	<100 x 10 ⁹ /L	<100 x 10 ⁹ /L
Absolute neutrophil count	<1.8 x 10 ⁹ /L	<1.5 x 10 ⁹ /L	<1.8 x 10 ⁹ /L	<0.8 x 10 ⁹ /L

1. Generally accepted/institutional reference values. Argued for by Greenberg et al. to define cytopenias in myelodysplastic syndromes (MDS).⁸ Used to define ICUS in 2017.⁹
2. Used to define ICUS in 2007.¹⁰
3. World Health Organization criteria to define cytopenias in MDS.¹¹
4. Alternative criteria to define cytopenias relevant for prognosis in MDS (IPSS-R).¹²

Supplementary Table 2. Prevalence of severe cytopenias according to NCI criteria.

		N	Total cohort	< 60 years	≥ 60 years
Anemia					
Grade 3	< 8.0 g/dL	147170	45 (0.03%)	43 (0.03%)	2 (0.01%)
Thrombocytopenia					
Grade 2	< 75 x 10 ⁹ /L	147074	65 (0.04%)	53 (0.04%)	12 (0.05%)
Grade 3	< 50 x 10 ⁹ /L	147074	27 (0.02%)	20 (0.02%)	7 (0.03%)
Grade 4	< 25 x 10 ⁹ /L	147074	9 (0.01%)	5 (0.00%)	4 (0.02%)
Neutropenia					
Grade 3	< 1.0 x 10 ⁹ /L	144676	107 (0.07%)	87 (0.07%)	20 (0.09%)
Grade 4	< 0.5 x 10 ⁹ /L	144676	4 (0.00%)	4 (0.00%)	0 (0.00%)

Alternative cutoffs were based on the National Cancer Institute toxicity criteria to define cytopenias¹³ and displayed in this table when not included as a cutoff in the main manuscript. Prevalence of cytopenias is displayed as number (%). N, number of evaluable individuals.

Supplementary Table 3. Risk of all-cause mortality according to the presence of cytopenias, stratified by age category (<60 and ≥60 years).

	N	Unadjusted		Adjusted*		Adjusted**		p-value for interaction [#]
		Hazard ratio (95% CI)	p-value	Hazard ratio (95% CI)	p-value	Hazard ratio (95% CI)	p-value	
Anemia	147 170							<0.001
<60 years	5403	1.21 (0.93-1.57)	0.161	1.43 (1.10-1.86)	0.009	1.35 (1.04-1.76)	0.027	
≥60 years	756	2.72 (2.30-3.22)	<0.001	1.87 (1.58-2.22)	<0.001	1.63 (1.37-1.94)	<0.001	
Thrombocytopenia	147 074							0.298
<60 years	1777	2.31 (1.65-3.23)	<0.001	2.13 (1.52-2.99)	<0.001	2.10 (1.50-2.95)	<0.001	
≥60 years	631	1.94 (1.57-2.40)	<0.001	1.37 (1.11-1.69)	0.004	1.33 (1.08-1.65)	0.008	
Neutropenia	144 676							0.935
<60 years	5643	0.84 (0.62-1.15)	0.276	0.78 (0.57-1.07)	0.118	0.84 (0.61-1.14)	0.260	
≥60 years	1282	0.76 (0.61-0.95)	0.017	0.99 (0.79-1.24)	0.935	1.06 (0.84-1.33)	0.624	

Generally accepted reference values were used to define cytopenias: anemia, hemoglobin concentration <12.0 g/L in women or <13.0 g/L in men; thrombocytopenia, platelet counts <150 x 10⁹/L; neutropenia, neutrophil counts <1.8x10⁹/L. Individuals with absence of cytopenia for the respective lineage were used as a reference.

*The multivariable model included age and sex as covariates. **The multivariable model included age, sex and the number of medications used as covariates. [#]To evaluate the effect of age on the association between cytopenias and OS, we used a Wald χ^2 test for the interaction term for the respective cytopenia and age category (<60 and ≥60 years). Interaction models included sex as a covariable. CI, confidence interval; N, number of individuals.

Supplementary Table 4. Risk of all-cause mortality according to the presence of multilineage cytopenias.

	N	Unadjusted		Adjusted*		Adjusted**	
		Hazard ratio (95% CI)	p-value	Hazard ratio (95% CI)	p-value	Hazard ratio (95% CI)	p-value
Anemia and thrombocytopenia	147 073						
<i>None</i>	138 623	Ref		Ref		Ref	
<i>Anemia</i>	6043	1.66 (1.43-1.92)	<0.001	1.76 (1.52-2.04)	<0.001	1.53 (1.32-1.78)	<0.001
<i>Thrombocytopenia</i>	2296	2.63 (2.17-3.19)	<0.001	1.47 (1.21-1.78)	<0.001	1.42 (1.17-1.73)	<0.001
<i>Anemia and thrombocytopenia</i>	111	9.54 (6.00-15.16)	<0.001	4.75 (2.98-7.55)	<0.001	4.08 (2.56-6.49)	<0.001
Anemia and neutropenia	144 668						
<i>None</i>	132 213	Ref		Ref		Ref	
<i>Anemia</i>	5530	1.70 (1.46-1.98)	<0.001	1.80 (1.55-2.09)	<0.001	1.56 (1.34-1.82)	<0.001
<i>Neutropenia</i>	6400	0.92 (0.76-1.12)	0.402	0.88 (0.73-1.07)	0.198	0.95 (0.78-1.15)	0.564
<i>Anemia and neutropenia</i>	525	1.08 (0.60-1.96)	0.790	1.29 (0.71-2.33)	0.401	1.25 (0.69-2.26)	0.461
Thrombocytopenia and neutropenia	144 578						
<i>None</i>	135 604	Ref		Ref		Ref	
<i>Thrombocytopenia</i>	2053	2.79 (2.30-3.40)	<0.001	1.51 (1.24-1.84)	<0.001	1.44 (1.18-1.76)	<0.001
<i>Neutropenia</i>	6644	0.84 (0.69-1.02)	0.085	0.82 (0.67-0.99)	0.042	0.88 (0.72-1.07)	0.183
<i>Thrombocytopenia and neutropenia</i>	277	3.15 (1.93-5.15)	<0.001	2.16 (1.32-3.54)	0.002	2.34 (1.43-3.82)	<0.001
Pancytopenia							
<i>None</i>	130 156	Ref		Ref		Ref	
<i>1 cytopenia</i>	13 568	1.45 (1.29-1.62)	<0.001	1.30 (1.16-1.45)	<0.001	1.25 (1.12-1.40)	<0.001
<i>>1 cytopenia</i>	828	1.90 (1.31-2.73)	<0.001	1.62 (1.12-2.34)	0.010	1.58 (1.10-2.28)	0.014
<i>Pancytopenia</i>	26	11.57 (4.81-27.83)	<0.001	11.92 (4.96-28.67)	<0.001	11.87 (4.93-28.54)	<0.001

Generally accepted reference values were used to define cytopenias: anemia, hemoglobin concentration <12.0 g/L in women or <13.0 g/L in men; thrombocytopenia, platelet counts <150 x 10⁹/L; neutropenia, neutrophil counts <1.8x10⁹/L. Individuals with absence of cytopenia in the respective lineages were used as the reference. *The multivariable model included age and sex as covariates. **The multivariable model included age, sex and the number of medications used as covariates. CI, confidence interval; N, number of individuals; ref, reference group.

Supplementary Table 5. Number of individuals with incident diagnosis of hematological malignancies, according to age category (<60 or ≥60 years) and the presence of cytopenias.

		N	Incident malignancy – n (%)		
			All	Lymphoid	Myeloid
All individuals					
	<60 years	129 728	267 (0.21%)	195 (0.15%)	75 (0.06%)
	≥60 years	22 452	216 (0.96%)	161 (0.72%)	57 (0.25%)
Anemia	<12.0 g/dL (women) or <13.0 g/dL (men)	6159	44 (0.71%)	28	18
Moderate	< 11.0 g/dL	1272	6 (0.47%)	4	2
Severe	< 10.0 g/dL	442	1 (0.23%)	1	0
Thrombocytopenia	< 150 × 10 ⁹ /L	2408	33 (1.37%)	25	9
Moderate	< 100 × 10 ⁹ /L	153	6 (3.92%)	4	2
Neutropenia	< 1.8 × 10 ⁹ /L	6925	39 (0.56%)	27	14
Moderate	< 1.5 × 10 ⁹ /L	2050	20 (0.98%)	12	9
Severe	< 0.8 × 10 ⁹ /L	30	1 (3.33%)	1	0
>1 cytopenia		861	14 (1.63%)	9	6

N, number of individuals

Supplementary Table 6. Subdistribution hazard ratios for incident diagnosis of hematological malignancies according to the presence of cytopenias.

		Unadjusted		Adjusted*	
		Hazard ratio (95%CI)	p-value	Hazard ratio (95%CI)	p-value
Anemia	<12.0 g/dL (women) or <13.0 g/dL (men)	2.32 (1.70-3.17)	<0.001	2.47 (1.80-3.39)	<0.001
Thrombocytopenia	< 150 x 10 ⁹ /L	4.67 (3.28-6.66)	<0.001	3.35 (2.32-4.83)	<0.001
Neutropenia	< 1.8 x 10 ⁹ /L	2.06 (1.49-2.87)	<0.001	1.97 (1.41-2.74)	<0.001
>1 cytopenia		5.69 (3.33-9.72)	<0.001	5.13 (3.00-8.79)	<0.001

Generally accepted reference values were used to define cytopenias: anemia, hemoglobin concentration <12.0 g/L in women or <13.0 g/L in men; thrombocytopenia, platelet counts <150 x 10⁹/L; neutropenia, neutrophil counts <1.8x10⁹/L. Individuals with absence of cytopenia in the respective lineage were used as a reference.

*The multivariable Fine Gray regression model included age and sex as covariates. CI, confidence interval.

Supplementary Table 7. Subdistribution hazard ratios for incident diagnosis of hematological malignancies according to the presence of cytopenias and age category (<60 or ≥60 years).

	Unadjusted		Adjusted*	
	Hazard ratio (95%CI)	p-value	Hazard ratio (95%CI)	p-value
Anemia				
<60 years	2.02 (1.30-3.13)	0.002	2.35 (1.48-3.72)	<0.001
≥60 years	3.27 (2.10-5.10)	<0.001	3.02 (1.93-4.72)	<0.001
Thrombocytopenia				
<60 years	4.10 (2.39-7.02)	<0.001	3.86 (2.24-6.66)	<0.001
≥60 years	3.41 (2.12-5.47)	<0.001	3.13 (1.91-5.15)	<0.001
Neutropenia				
<60 years	2.40 (1.57-3.66)	<0.001	2.31 (1.52-3.52)	<0.001
≥60 years	1.41 (0.83-2.39)	0.200	1.52 (0.89-2.61)	0.130
>1 cytopenia				
<60 years	6.07 (2.99-12.30)	<0.001	6.26 (3.08-12.72)	<0.001
≥60 years	4.32 (1.90-9.82)	<0.001	4.23 (1.86-9.59)	<0.001

Generally accepted reference values were used to define cytopenias: anemia, hemoglobin concentration <12.0 g/L in women or <13.0 g/L in men; thrombocytopenia, platelet counts <150 x 10⁹/L; neutropenia, neutrophil counts <1.8x10⁹/L. Individuals with absence of cytopenia in the respective lineage were used as a reference.

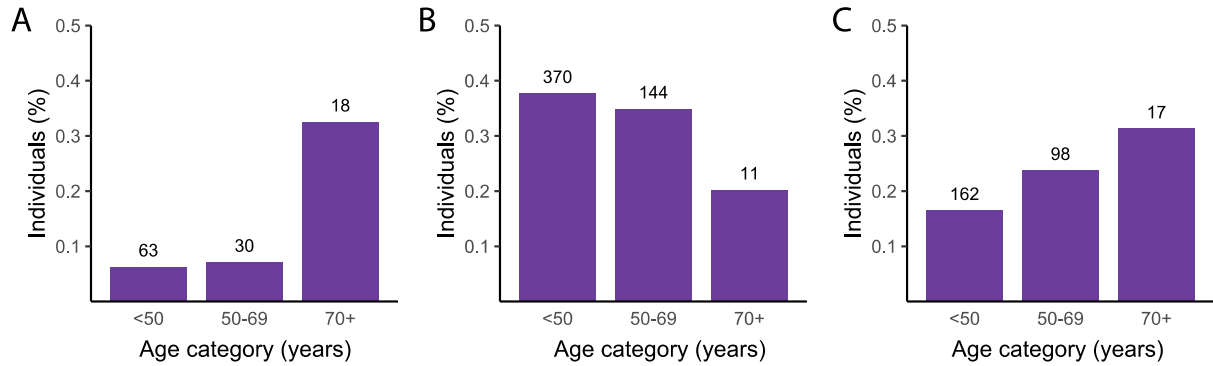
*The multivariable Fine Gray regression model included age and sex as covariates. CI, confidence interval; N, number of individuals.

Supplementary Table 8. Hematological causes of death (n=110) for all Lifelines participants with successful linkage to the national registry and cause of death data being available (n=152 096).

	Entire Lifelines cohort (n=152 096)	<60 years (n=129 662)	≥60 years (n=22 434)
Lymphoma (C81-C85)	31 (0.02%)	9 (0.007%)	22 (0.10%)
Plasma cell dyscrasia (C90)	22 (0.01%)	8 (0.006%)	14 (0.06%)
Lymphoid leukemia (C91)	5 (0.003%)		
Myeloid malignancies - AML, MDS, MPN (C92, C93, D46, D47)	35 (0.02%)	14 (0.01%)	21 (0.09%)
Leukemia, unspecified (C95)	17 (0.01%)		

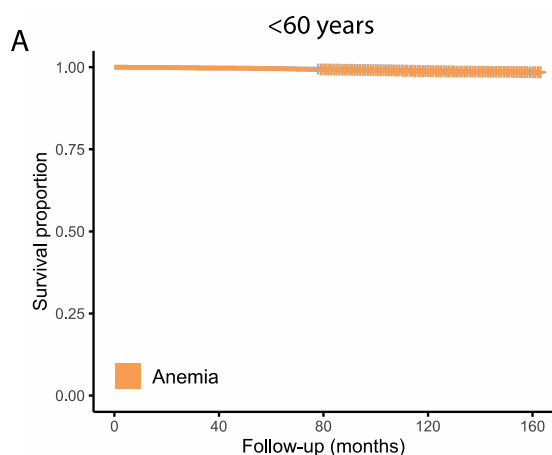
Supplementary Figure 1. Presence of multilineage cytopenias stratified by age

The graphs indicate the percentage of individuals with concurrent anemia and thrombocytopenia (A), anemia and neutropenia (B) and thrombocytopenia and neutropenia (C), stratified by age category. The absolute number of individuals with the respective multilineage cytopenia for each age category is given. Anemia, hemoglobin concentration <12.0 g/L in women or <13.0 g/L in men; thrombocytopenia, platelet counts <150 x 10⁹/L; neutropenia, neutrophil counts <1.8x10⁹/L.



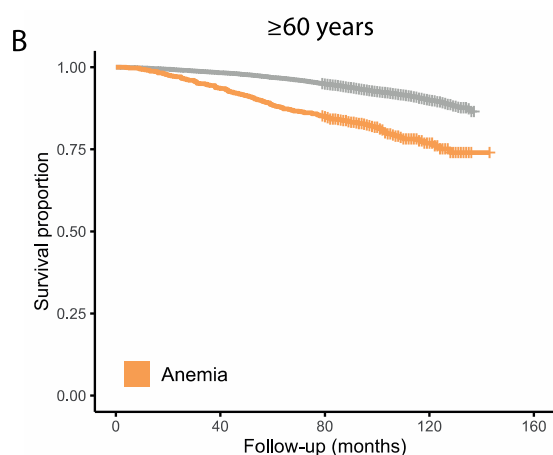
Supplementary Figure 2. Kaplan-Meier graphs for overall survival stratified by age category

Kaplan-Meier graphs are shown for individuals with anemia (<13.0 g/dL in men and <12.0 g/dL in women, A-B), thrombocytopenia (<150 x 10⁹/L, C-D) and neutropenia (<1.8 x 10⁹/L, E-F). Individuals with absence of the respective cytopenia are shown in grey. Survival curves are shown separately for individuals <60 (left panel) and ≥60 years (right panel).



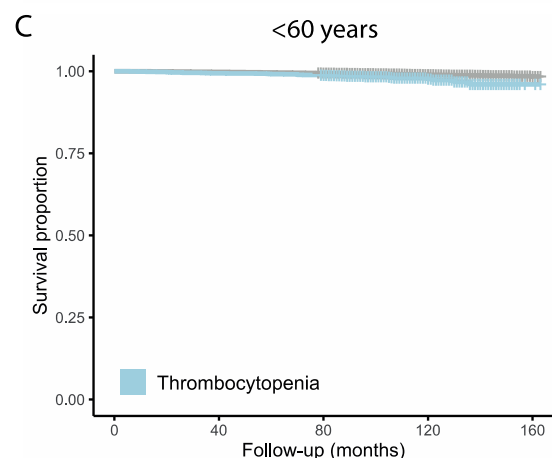
Number at risk

—	119659	119397	116383	23537	841
—	5403	5387	5268	1356	28



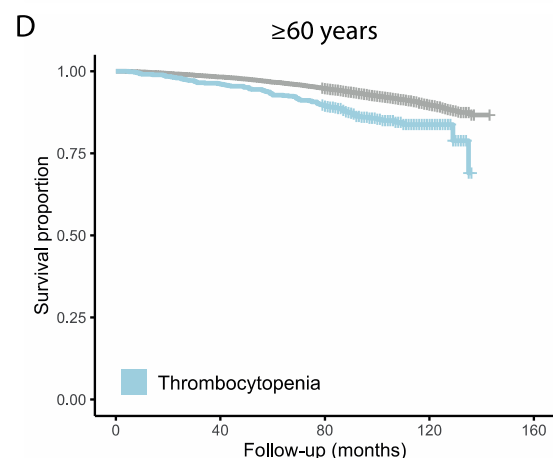
Number at risk

—	21352	21000	19718	2651	0
—	756	708	630	102	0



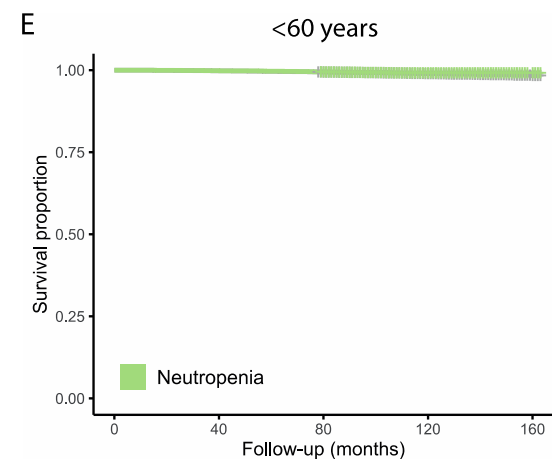
Number at risk

—	123209	122942	119846	24521	863
—	1777	1766	1729	344	6



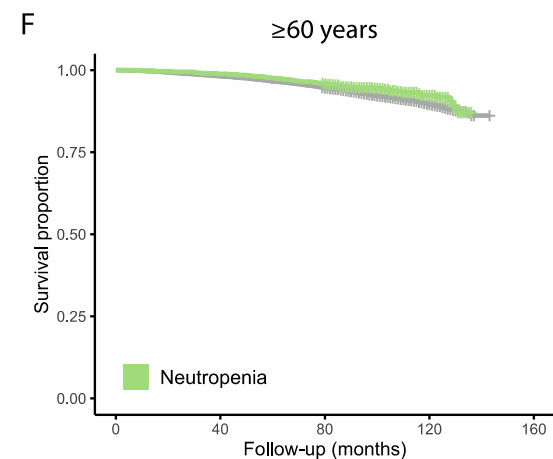
Number at risk

—	21457	21081	19774	2678	0
—	631	607	554	72	0



Number at risk

—	117304	117047	114047	23339	831
—	5643	5632	5536	1060	21

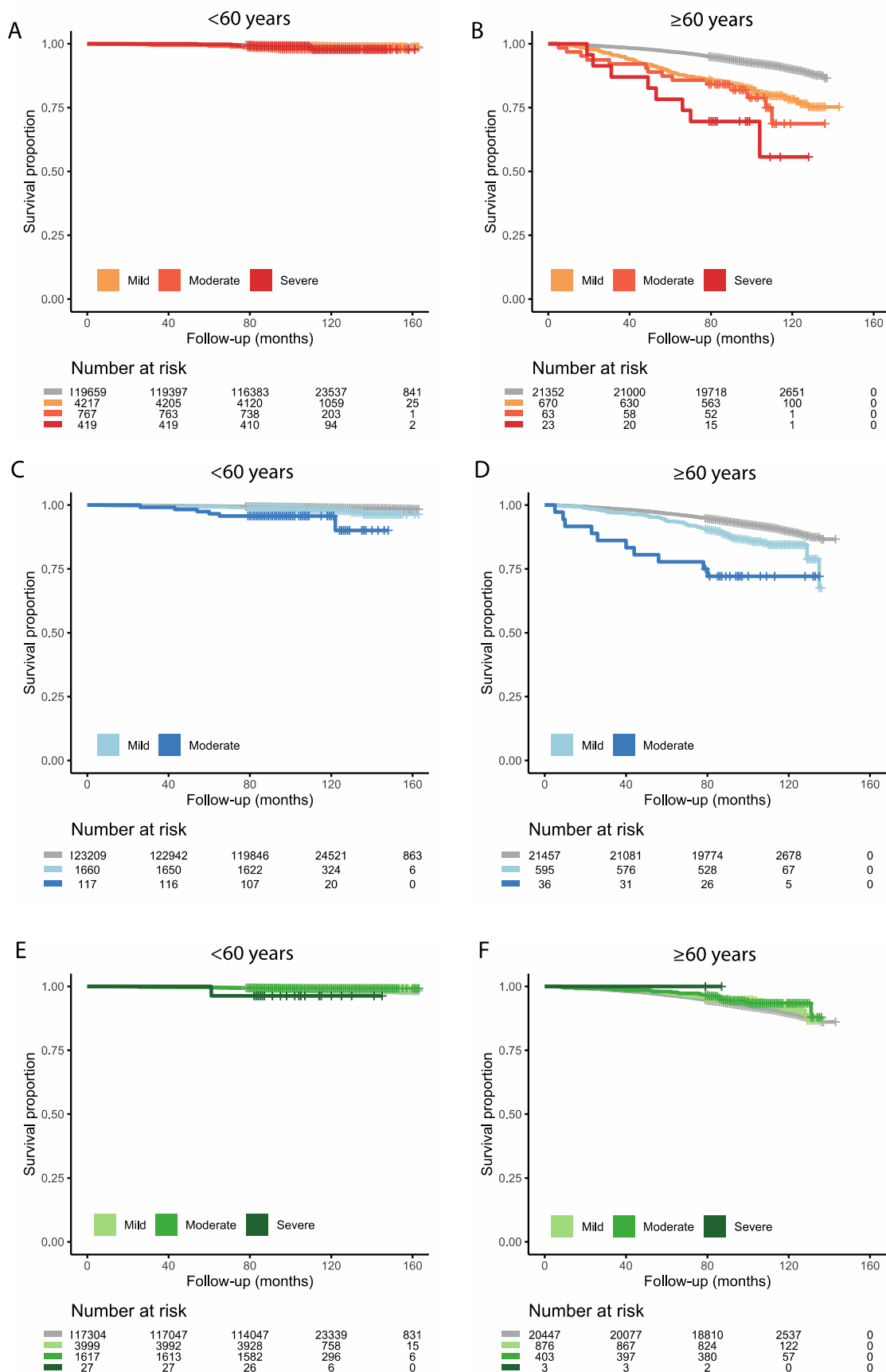


Number at risk

—	20447	20077	18810	2537	0
—	1282	1267	1206	179	0

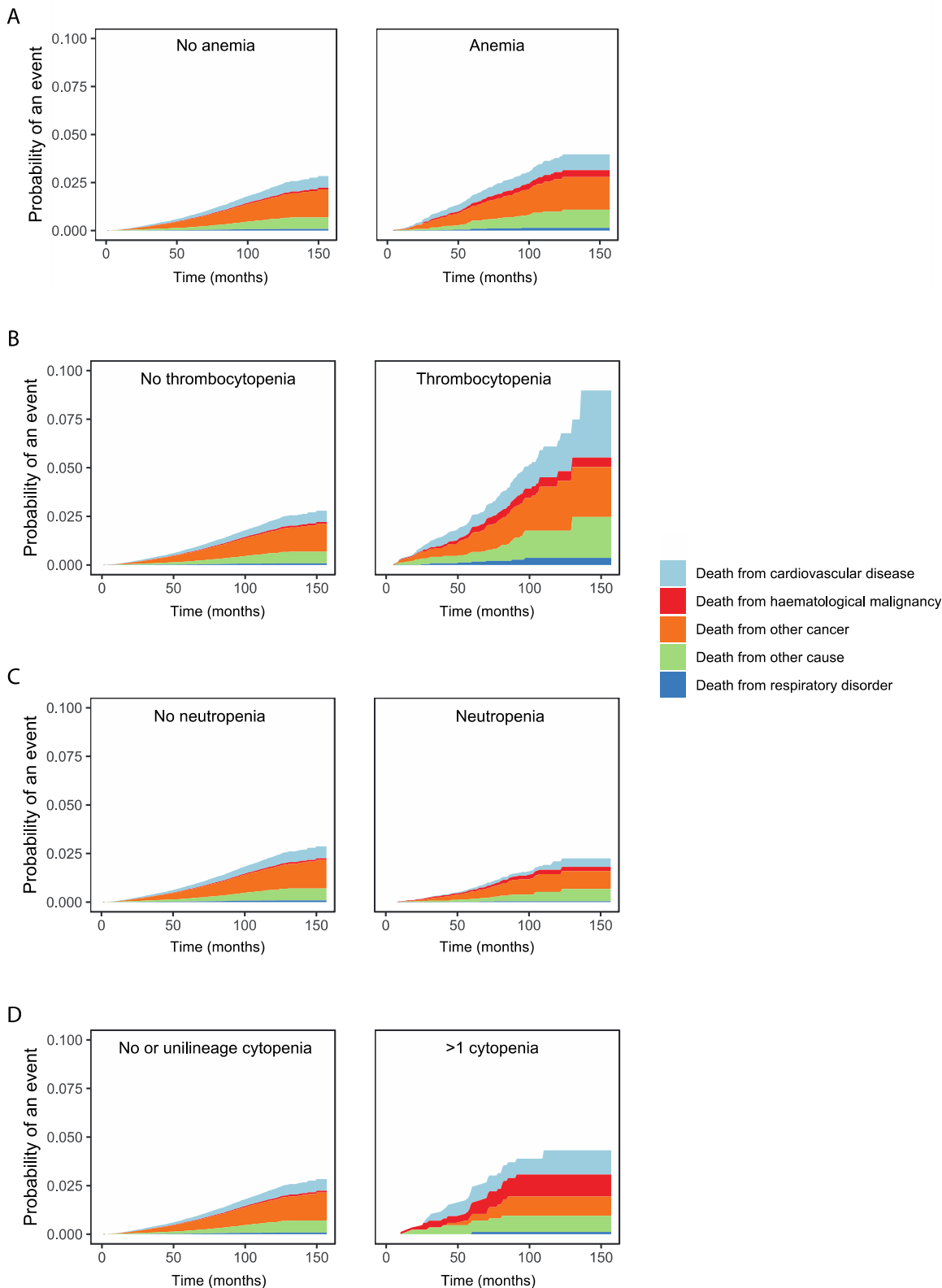
Supplementary Figure 3. Kaplan-Meier graphs for overall survival according to cytopenia severity and stratified by age category

Anemia (A-B) was classified as mild (<13.0 g/dL in men and <12.0 g/dL in women), moderate (<11.0 g/dL) or severe (<10.0 g/dL). Thrombocytopenia (C-D) was classified as mild (<150 x 10⁹/L) or moderate (<100 x 10⁹/L). Neutropenia (E-F) was classified as mild (<1.8 x 10⁹/L), moderate (<1.5 x 10⁹/L) or severe (<0.8 x 10⁹/L). Individuals with absence of the respective cytopenia are shown in grey. Survival curves are shown separately for individuals <60 and ≥60 years.



Supplementary Figure 4. Major other causes of death for individuals with peripheral cytopenias

Cumulative incidence graphs show probabilities for cause-specific mortality according to the presence of anemia (A), thrombocytopenia (B), neutropenia (C) or >1 cytopenia (D). Death due to hematological malignancies was classified by ICD-10 codes C81x-C96x and D45x-D47x. Other causes of death were subclassified as death from cardiovascular disease (ICD-10 codes Ix), respiratory disorders (ICD-10 codes Jx), solid cancers (ICD-10 codes Cx except for C81x-C96x) and other causes. Anemia, hemoglobin concentration <12.0 g/L in women or <13.0 g/L in men; thrombocytopenia, platelet counts <150 x 10⁹/L; neutropenia, neutrophil counts <1.8x10⁹/L. Results based on calculations by the authors using non-public microdata from Statistics Netherlands.



References

1. Scholtens S, Smidt N, Swertz MA, et al. Cohort Profile: LifeLines, a three-generation cohort study and biobank. *Int J Epidemiol*. 2015;44:1172-1180.
2. Klijs B, Scholtens S, Mandemakers JJ, Snieder H, Stolk RP, Smidt N. Representativeness of the Lifelines Cohort Study. *PLoS One*. 2015;10(9):e0137203.
3. 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;29:19-24.
4. Centraal Bureau voor de Statistiek. Statline: Bevolking; geslacht, leeftijd en burgerlijke staat, 1 januari. Available from <https://opendata.cbs.nl/statline/#/CBS/nl/dataset/7461BEV/table>. Last accessed 22-03-2021.
5. World Health Organization (WHO). International Statistical Classification of Diseases and Related Health Problems, Tenth Revision; World Health Organization: Geneva, Switzerland, 1992.
6. Harteloh P, de Bruin K, Kardaun J. The reliability of cause-of-death coding in The Netherlands. *Eur. J. Epidemiol*. 2010;25:531-538.
7. Harteloh P. The implementation of an automated coding system for cause-of-death statistics. *Inform Health Soc Care*. 2020;45(1):1-14.
8. Greenberg PL, Tuechler H, Schanz J, et al. Cytopenia levels for aiding establishment of the diagnosis of myelodysplastic syndromes. *Blood*. 2016;128:2096–2097.
9. Valent P, Orazi A, Steensma DP, et al. Proposed minimal diagnostic criteria for myelodysplastic syndromes (MDS) and potential pre-MDS conditions. *Oncotarget* 2017;8:73483-73400.
10. Valent P, Horny H-P, Bennett JM, et al. Definitions and standards in the diagnosis and treatment of the myelodysplastic syndromes: consensus statements and report from a working conference. *Leuk Res* 2007;31:727-36.
11. Arber DA, Orazi A, Hasserjian R, et al. The 2016 revision to the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia. *Blood*. 2016;127:2391-405.
12. Greenberg PL, Tuechler H, Schanz J, et al. Revised International Prognostic Scoring System for Myelodysplastic Syndromes. *Blood* 2012;120:2454-2465.
13. National Institutes of Health, National Cancer Institute. Common Terminology Criteria for Adverse Events (CTCAE) v5.0. Available from https://ctep.cancer.gov/protocoldevelopment/electronic_applications/ctc.htm. Last accessed 23-03-2021.