bjh research paper

Blood counts in adult and elderly individuals: defining the norms over eight decades of life

Jakob Zierk,[1,2] Alexander Krebs,[3] Manfred Rauh,[1] Markus Metzler,[1] Astrid Löscher,[4] Erwin Strasser[5] and Stefan W. Krause[6]

[1]Department of Pediatrics and Adolescent Medicine, University Hospital Erlangen, [2]Center of Medical Information and Communication Technology, University Hospital Erlangen, Erlangen, [3]MVZ Labor PD Dr. Volkmann und Kollegen, Karlsruhe, [4]Central Laboratory, University Hospital Erlangen, [5]Department of Transfusion Medicine and Haemostaseology, University Hospital Erlangen and [6]Department of Medicine 5 - Haematology and Oncology, University Hospital Erlangen, Erlangen, Germany

Received 4 September 2019; accepted for publication 3 November 2019

Correspondence: Jakob Zierk, Department of Pediatrics and Adolescent Medicine, University Hospital Erlangen, Loschgestr. 15, 91054 Erlangen, Germany.
E-mail: jakob.zierk@uk-erlangen.de

Summary

The blood count is one of the most common tests used for health assessment. In elderly individuals, selection of a 'healthy' reference population for laboratory assessment is difficult due to the high prevalence of chronic morbidities, leading to uncertainty regarding appropriate reference intervals. In particular, age-specific lower haemoglobin reference limits to define anaemia are controversial. Here, we applied a data mining approach to a large dataset of 3 029 904 clinical routine samples to establish blood count reference intervals. We excluded samples from units/specialists with a high proportion of abnormal blood counts, samples from patients with an unknown or decreased estimated glomerular filtration rate, and samples with abnormal test results in selected other analytes. After sample exclusion, 566 775–572 060 samples from different individuals aged 20–100 years were available for analysis. We then used an established statistical algorithm to determine the distribution of physiological test results and calculated age- and sex-specific reference intervals. Our results show substantial trends with age in haematology analytes' reference intervals. Most notably, haemoglobin and red cell counts decline in men with advanced age, accompanied by increases in red cell volume in both sexes. These findings were confirmed in an independent dataset, and suggest an at least partly physiologic cause.

Keywords: anaemia, haematology, laboratory haematology.

The blood count is one of the most common tests for health assessment with diagnostic and therapeutic implications for a multitude of common to rare and minor to life-threatening conditions. However, despite its frequent clinical use and importance and near-universal availability on the one hand, and an ageing population throughout the world on the other hand, uncertainty regarding appropriate reference intervals for older individuals still exists. In particular, age-specific lower reference limits for haemoglobin, which define anaemia, are controversial, despite substantial consequences for individual patients and considerable public health impact.

A major challenge when establishing reference intervals for individuals >50 years is the increasing proportion of chronic morbidities and medication with age, which leads to exclusion of these subjects from conventional reference interval studies (Adeli *et al.*, 2015; Röhrig *et al.*, 2018). Based on these restrictions, data mining of age-specific reference intervals using laboratory test results collected from routine patient care can be considered a viable complement to conventional reference interval studies (Haeckel *et al.*, 2017; Jones *et al.*, 2018). In contrast to the highly selected population of conventional approaches, a reference interval data mining approach uses exactly the 'real world' population to which the reference intervals are ultimately applied, while still excluding outliers. In addition, the necessity to obtain a patient history and perform a clinical examination restricts the number of individuals who can be recruited for a conventional reference sample, while the number of samples available using data mining is only limited by the amount of test results stored in laboratory information systems. However, great care must be taken to avoid an unwarranted shift or widening of reference intervals in populations with a high prevalence of slightly abnormal test results. In previous work, we have shown that establishment of haematology reference intervals using a sophisticated data mining approach (the reference limit estimator, RLE) is feasible, even in the challenging setting of a tertiary care centre and a high proportion of abnormal test results (Zierk *et al.*, 2018, 2019). In short, this approach assumes a mixture of parametrically distributed samples from healthy individuals and pathologic samples.

 First published online 6 February 2020
British Journal of Haematology, 2020, **189,** 777–789 doi: 10.1111/bjh.16430

From the dominant fraction of physiological test results, the distribution of non-pathological values can be estimated and used to define reference intervals (Arzideh *et al.*, 2008, 2010; Zierk *et al.*, 2018). Only one sample per person is considered to prevent unwarranted over-representation of individuals, and to reduce the impact of patients with disease, who will receive more frequent blood tests than their healthier counterparts.

Here, we report on the establishment of sex- and age-specific reference intervals for individuals of 20–100 years of age for haematology analytes using the RLE on samples from a large laboratory service provider, and validation of the results in an independent dataset from a tertiary care centre. To explore the effect of increasing comorbidity with age on reference intervals, we apply different filtering criteria deduced directly from the laboratory dataset, and measure the age-specific proportion of test results classified as abnormal.

Materials and methods

Study population and selection of samples

We examined blood count measurements performed during routine care of patients and sent to a German laboratory service provider from 01/2012 to 06/2017 (Medizinisches Versorgungszentrum Labor Volkmann, Karlsruhe, Germany). To comply with privacy regulations and ethical considerations, samples were anonymized before data transfer and analysis. Additionally, we examined measurements performed during patient care in the University Hospital Erlangen (Germany) from 01/2013 to 12/2018 to assess the reproducibility of the performed analyses in a tertiary care setting. The influence of renal function and inflammation/acute-phase reactions was assessed by examination of C-reactive protein (CRP) and creatinine concentrations within three days (CRP) and seven days (creatinine) of blood count measurements (if multiple CRP or creatinine measurements were available within the specified time-frame, the maximum concentration was used; see *Sample exclusion* below for details).

Laboratory test results from inpatients and outpatients aged 20–100 years were retrieved from the laboratory databases. Test results from units/specialists with a supposedly high proportion of patients with severe diseases affecting haematology analytes were excluded from the analysis (Table I). The investigated patient population is composed predominantly of Caucasian individuals. Ethnicity of patients was not available in the dataset and no stratification according to ethnicity was performed.

Analytical procedures

We analyzed measurements of haemoglobin, haematocrit, red cell count, red cell indices (mean corpuscular haemoglobin, MCH; mean corpuscular haemoglobin concentration, MCHC; mean corpuscular volume, MCV, platelet count, and white cell count). Haematology analyses were performed on Sysmex (Norderstedt, Germany) XE and XM devices (Labor Volkmann) and Sysmex XE 5000 and XT 2000i devices (University Hospital Erlangen); CRP and creatinine were measured on Roche Diagnostics (Mannheim, Germany) Cobas 8000 devices (Labor Volkmann) and Beckman Coulter Diagnostics (Krefeld, Germany) AU5800 and AU680 devices (University Hospital Erlangen), in accordance with standard operating procedures. Creatinine measurements were performed using a method traceable to reference isotope dilution mass spectroscopy (IDMS). Regular quality control according to German regulations was performed.

Sample selection

Reference intervals were calculated for different age groups (≥20 years to <30 years, ≥30 years to <40 years, …, ≥90 years to <100 years) for men and women. When multiple samples from a single individual in an age group existed, a random sample from that individual was examined. This strategy has proven equivalent in comparison to more complex sample selection strategies, which try to account for the frequency of sampling as a surrogate for 'clinical grounds for retesting' (e.g. selection of the sample 'most isolated in time', i.e. no repeat measurements before or after the sample) in a previous analysis (Zierk *et al.*, 2018).

Calculation of reference intervals

Reference intervals were calculated with an indirect algorithm described and validated previously (Arzideh *et al.*, 2008, 2010; Zierk *et al.*, 2018), ('Reference Limit Estimator' [RLE], developed by the German Society of Clinical Chemistry and Laboratory Medicine's Working Group on Guide Limits, which is freely available as a software package at https://www.dgkl.de/verbandsarbeit/arbeitsgruppen/entscheidungsgrenzen-richtwerte/). In summary, the method estimates reference intervals from an input dataset containing both non-pathologic and pathologic samples. The distribution of non-pathologic samples is modelled with a parametric distribution (so-called Power Normal distribution, a Gaussian distribution after Box–Cox transformation of the data, i.e. a distribution that can accommodate skewed data), whereas no assumptions regarding the distribution of pathologic samples were made. For analyses with a high proportion of abnormally low test results (haemoglobin, haematocrit, red cell count, MCV, MCH, and platelet count) the algorithm settings were adjusted to optimize estimation of the lower boundary of the reference interval ('Pathological values' set to 'Low', indicating that the majority of pathological values are lower than physiological test results, whereas per default pathological values are expected below and above the distribution of normal test results), while for white cell counts the

Table I. Inclusion and exclusion of samples from units/specialists with a supposedly low or high proportion of abnormal test results.

Included units	Excluded units
Laboratory service provider (Labor Volkmann)	
General Practice, Internal Medicine, Obstetrics/Gynaecology, Dermatology, Ophthalmology, Ear/Nose/Throat, Neurology, Radiology, Laboratory medicine, Psychiatry	Surgery, Paediatrics, Orthopaedics, Anaesthesia, Urology, Haematology/Oncology, Intensive care, Dialysis
Validation dataset: Tertiary care centre (University Hospital Erlangen)	
In- and outpatients from neurology, psychiatry, ophthalmology, dermatology, and transfusion medicine; surgical outpatients and preoperative screening (including OB/Gyn, ENT, and urology); internal medicine outpatient departments with a low probability of pathological blood counts	Remaining units, specifically emergency rooms, intensive and intermediate care; surgical wards and intervention departments (including OB/Gyn, ENT, and urology); internal medicine wards and most internal medicine outpatient departments

algorithm settings were adjusted to optimize estimation of the upper boundary.

Sample exclusion

Before the determination of reference intervals with the algorithm outlined above, samples were excluded according to the following criteria (Fig 1):

1. No exclusion (dataset D0).
2. Removal of samples depending on CRP and grossly abnormal creatinine measurements (dataset D1): samples were excluded, if a CRP $\geq$ 5 mg/l was measured within three days before or after the blood count, or if a creatinine $\geq$ 15 mg/l (men) or $\geq$ 14 mg/l (women) was measured within seven days before or after the blood count.
3. All samples with a minimum estimated glomerular filtration rate (eGFR) $\geq$ 60 ml/min $\times$ 1·73 m^2 [according to the simplified Modification of Diet in Renal Disease (MDRD) formula (Levey *et al.*, 2009)] within seven days before/after the blood count were selected (dataset D2) (i.e., samples without corresponding creatinine measurements were implicitly excluded).
4. Removal of samples depending on grossly abnormal test results in other blood count analytes, similar to 'latent abnormal values exclusion' (LAVE) (Ichihara & Boyd, 2010; Ichihara *et al.*, 2017) (dataset D3):
   a Reference intervals were calculated using dataset D0.
   b Test results were discarded if test results in other analytes in the same sample were below the first percentile or above the 99th percentile of the reference limits determined in (a), and reference intervals were calculated again for each analyte. Specifically, haemoglobin, haematocrit, and red cell count test results were discarded if MCV, platelet count, or white cell count results were below the first or above the 99th percentile in D0; for MCH, MCHC, and MCV samples were ignored depending on haemoglobin, platelet count, and white cell count; platelet counts were discarded depending on haemoglobin, MCV, and white cell count; and white cell counts were ignored depending on haemoglobin, MCV, and platelet count. Abnormal values of haemoglobin, haematocrit, and red cell count did not lead to exclusion of the other two closely correlated analytes within this group.
5. Final dataset *D4*: Combination of criteria for D2 and D3: Samples were first excluded according to the *D2* criteria, and then according to the D3 criteria.

Statistical testing of trends of reference intervals with age and differences between sexes

We used the Mann–Kendall test to test for significant changes with age in the percentiles (2.5th, 50th, and 97.5th percentile) and widths of reference intervals. This was performed for all age groups (20–30 years to 90–100 years) and for elderly age groups (50–60 to 90–100 years), to account for continuous changes starting early as well as for changes with a later onset. To estimate the change with age per decade of life, we used Theil–Sen regression. Differences in reference intervals between men and women were tested using the sign test (all age groups' men's 2.5th and 97.5th percentiles versus all women's 2.5th and 97.5th percentiles). All statistical tests were performed using Python (Python Software Foundation, https://www.python.org/) and the scikit-learn (Pedregosa *et al.*, 2011) and pymannkendall (Hussain & Mahmud, 2019) libraries.

Results

We collected 3 029 904 samples from 850 013 different individuals. After exclusion of test results according to requesting unit/specialist (Table I), 2 836 098 samples from 788 112 different individuals remained.

Effect of different sample exclusion strategies on age-specific reference intervals

Exclusion of samples with different filtering strategies, based on CRP, creatinine, estimated glomerular filtration rate (eGFR), and test results in other blood count analytes (Fig 1) substantially reduced the number of samples in an age-

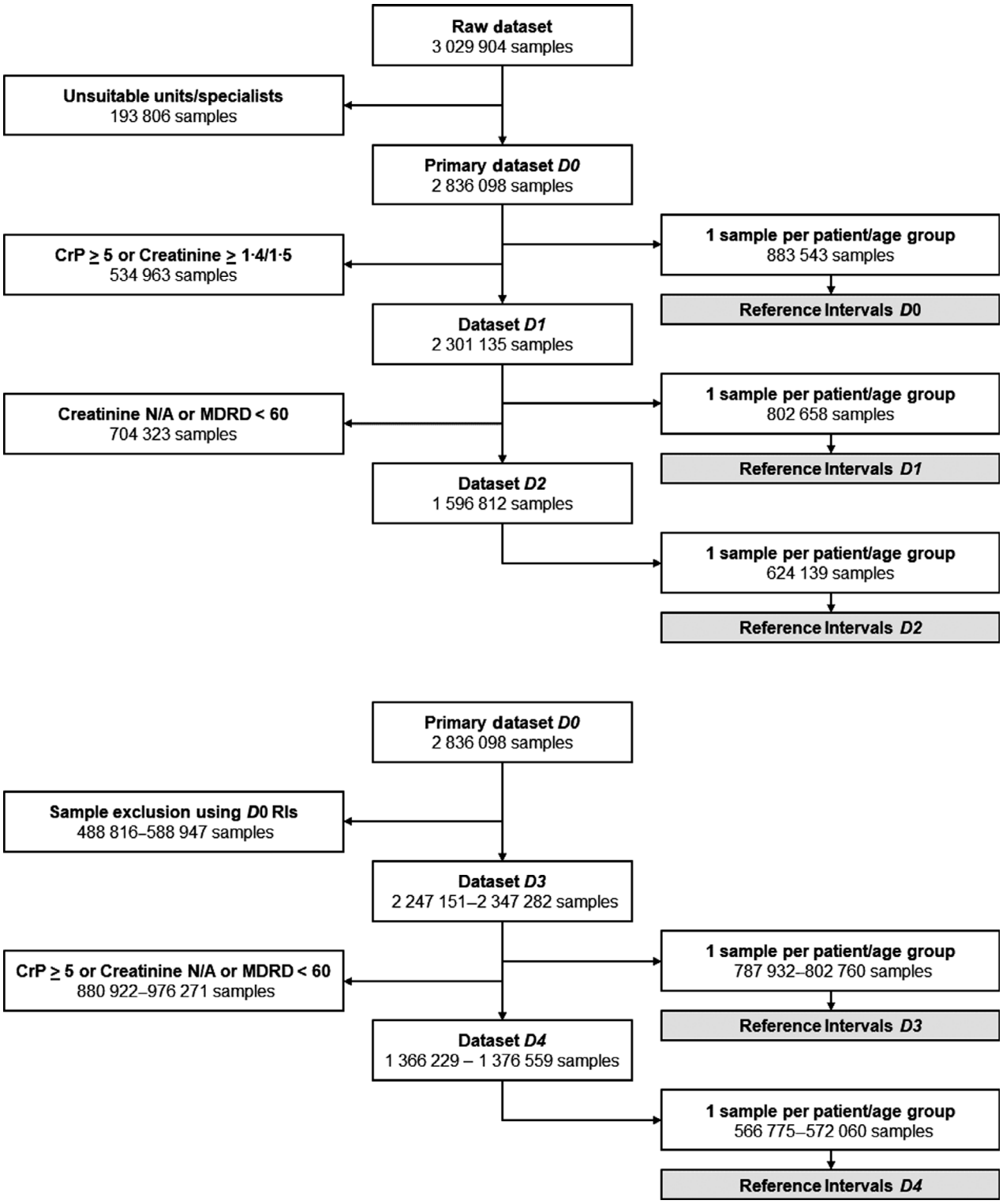

Fig 1. Sample exclusion strategies. MDRD, Modification of Diet in Renal Disease.

specific manner. For example, of 2 836 098 haemoglobin samples, 883 543 were available for analysis after selection of a single sample per age group without further filtering, whereas 570 810 samples remained after application of the strictest exclusion criteria (Table II). Despite extensive exclusion of samples, >10 000 test results were available for analysis for each age and sex group in patients <90 years, while for patients 90–100 years 932–2 461 test results were

Table II. Established age- and sex-specific reference limits. Age- and sex-specific 2.5th percentiles (lower reference limits), 50th percentiles (medians), and 97.5th percentiles (upper reference limits) and the number of samples used for reference interval estimation (*N*) are shown. N_{Raw} denotes the number of samples per age and sex group in the unfiltered raw input dataset.

		Men					Women				
				Percentiles					Percentiles		
	Age	*N*	N_{Raw}	2.5th	50th	97.5th	*N*	N_{Raw}	2.5th	50th	97.5th
Red cell count	20–30	25 112	75 860	4·56	5·16	5·83	31 256	119 204	3·97	4·55	5·19
	30–40	31 701	102 710	4·44	5·10	5·76	38 775	158 977	3·93	4·52	5·14
	40–50	48 078	178 715	4·39	5·03	5·69	56 466	230 595	3·93	4·51	5·15
	50–60	61 190	270 981	4·26	4·94	5·62	69 342	316 915	3·92	4·55	5·18
	60–70	44 403	267 251	4·16	4·84	5·57	52 374	302 922	3·95	4·56	5·22
	70–80	35 735	305 773	3·94	4·70	5·47	43 718	350 625	3·82	4·49	5·19
	80–90	12 095	133 009	3·69	4·54	5·40	17 191	178 506	3·66	4·40	5·15
	90–100	932	10 826	3·50	4·35	5·20	2 442	27 035	3·53	4·32	5·17
Haemoglobin	20–30	25 112	75 860	13·8	15·5	17·4	31 256	119 204	11·7	13·4	15·2
	30–40	31 701	102 710	13·7	15·4	17·3	38 775	158 977	11·7	13·4	15·2
	40–50	48 078	178 715	13·5	15·3	17·1	56 466	230 595	11·8	13·5	15·3
	50–60	61 190	270 981	13·2	15·1	17·0	69 342	316 915	12·1	13·7	15·6
	60–70	44 403	267 251	13·0	14·9	16·9	52 374	302 922	12·0	13·7	15·5
	70–80	35 735	305 773	12·3	14·5	16·7	43 718	350 625	11·7	13·5	15·4
	80–90	12 095	133009	11·5	14·1	16·6	17 191	178 506	11·2	13·3	15·4
	90–100	932	10 826	11·0	13·4	16·4	2 442	27 035	10·7	13·0	15·3
Haematocrit	20–30	25 112	75 860	39·7	44·5	49·9	31 256	119 204	34·8	39·5	44·7
	30–40	31 701	102 710	39·3	44·3	49·6	38 775	158 977	34·8	39·5	44·9
	40–50	48 078	178 715	39·2	44·1	49·5	56 466	230 595	35·0	39·8	45·2
	50–60	61 190	270 981	38·7	43·8	49·4	69 342	316 915	35·6	40·3	45·7
	60–70	44 403	267 251	38·2	43·4	49·2	52 374	302 922	35·5	40·5	45·9
	70–80	35 735	305 773	36·0	42·3	48·7	43 718	350 625	34·4	40·0	45·7
	80–90	12 095	133 009	34·5	41·4	48·4	17 191	178 506	33·1	39·4	45·7
	90–100	932	10 826	32·4	39·7	47·7	2 442	27 035	32·4	38·8	46·6
White cell count	20–30	25 022	75 860	3·95	6·37	10·30	30 948	119 204	3·91	6·54	10·95
	30–40	31 409	102 710	4·01	6·61	10·93	38 399	158 977	3·90	6·56	11·05
	40–50	47 966	178 715	3·97	6·52	10·69	56 052	230 595	3·87	6·57	11·13
	50–60	60 534	270 981	4·00	6·63	10·99	68 855	316 915	3·76	6·26	10·41
	60–70	43 627	267 251	4·09	6·63	10·77	52 022	302 922	3·86	6·24	10·10
	70–80	35 412	305 773	4·09	6·59	10·58	43 475	350 625	3·96	6·35	10·18
	80–90	12 249	133 009	4·15	6·64	10·55	17 395	178 506	3·99	6·47	10·49
	90–100	959	10 826	3·70	6·16	10·05	2 451	27 035	4·12	6·51	9·13
	all	229 876	1 345 125	4·03	6·59	10·78	277 809	1 684 779	3·87	6·41	10·63
MCH	20–30	25 198	75 860	27·8	30·1	32·5	31 723	119 204	27·1	29·6	32·2
	30–40	31 640	102 710	27·7	30·1	32·7	39 252	158 977	27·1	29·8	32·5
	40–50	48 116	178 715	27·9	30·4	33·1	57 425	230 595	27·5	30·1	32·9
	50–60	60 898	270 981	28·0	30·6	33·5	69 613	316 915	27·6	30·1	32·8
	60–70	43 862	267 251	28·1	30·8	33·7	52 347	302 922	27·5	30·1	32·9
	70–80	35 450	305 773	28·2	30·9	33·8	43 518	350 625	27·6	30·1	33·0
	80–90	12 197	133 009	28·0	30·9	33·9	17 402	178 506	27·5	30·1	33·0
	90–100	959	10 826	27·8	30·9	34·3	2 460	27 035	27·1	30·1	33·0
MCHC	20–30	25 198	75 860	32·8	34·8	36·9	31 723	119 204	32·2	34·1	36·1
	30–40	31 640	102 710	32·8	34·8	36·8	39 252	158 977	32·2	34·1	36·1
	40–50	48 116	178 715	32·8	34·7	36·6	57 425	230 595	32·2	34·1	36·0
	50–60	60 898	270 981	32·7	34·6	36·5	69 613	316 915	32·2	34·0	35·9
	60–70	43 862	267 251	32·5	34·4	36·4	52 347	302 922	32·1	33·9	35·7
	70–80	35 450	305 773	32·3	34·3	36·3	43 518	350 625	31·9	33·8	35·7
	80–90	12 197	133 009	31·9	34·0	36·2	17 402	178 506	31·7	33·7	35·7
	90–100	959	10 826	31·8	33·9	36·2	2 460	27 035	31·5	33·6	35·7

Table II. (*Continued*)

		Men					Women				
				Percentiles					Percentiles		
	Age	*N*	N_{Raw}	2.5th	50th	97.5th	*N*	N_{Raw}	2.5th	50th	97.5th
MCV	20–30	25 198	75 860	79·5	86·3	93·7	31 723	119 204	79·7	86·8	94·5
	30–40	31 640	102 710	79·6	86·7	94·6	39 252	158 977	79·7	87·2	95·4
	40–50	48 116	178 715	80·1	87·5	95·6	57 425	230 595	80·3	88·0	96·5
	50–60	60 898	270 981	80·8	88·6	97·0	69 613	316 915	81·0	88·5	96·7
	60–70	43 862	267 251	81·6	89·4	97·9	52 347	302 922	81·3	88·8	97·1
	70–80	35 450	305 773	82·3	90·1	98·7	43 518	350 625	81·5	89·1	97·4
	80–90	12 197	133 009	82·4	90·7	99·8	17 402	178 506	81·1	89·3	98·3
	90–100	959	10 826	82·0	91·0	100·0	2 460	27 035	80·6	89·4	98·6
Platelet count	20–30	25 002	75 860	155	231	339	31 170	119 204	174	262	393
	30–40	31 494	102 710	150	233	346	38 636	158 977	172	261	396
	40–50	48 072	178 715	154	235	348	56 563	230 595	175	266	397
	50–60	61 060	270 981	153	234	356	69 490	316 915	179	264	392
	60–70	43 870	267 251	145	225	349	52 460	302 922	171	258	384
	70–80	35 317	305 773	132	213	339	43 467	350 625	165	251	384
	80–90	12 190	133 009	127	208	340	17 342	178 506	156	246	388
	90–100	949	10 826	133	209	329	2 461	27 035	145	240	398
	all	230 649	1 345 125	150	230	350	279 691	1 684 779	171	260	391

MCH, mean corpuscular haemoglobin; MCHC, mean corpuscular haemoglobin concentration; MCV, mean corpuscular volume.

available. The effect of sex, age, and sample exclusion strategy on the different analytes' reference intervals is shown in Fig 2 and Figure S1. Haemoglobin, haematocrit, and red cell count reference intervals decrease with age in a sex-specific manner (i.e. more pronounced in men than in women). The decrease in age-specific lower reference limits was less pronounced when stricter filtering was applied. Similarly, MCV and MCH reference intervals increase with age and a more pronounced increase is observed in men, with narrower reference intervals in stricter sample exclusion strategies. For platelet counts and white cell counts, age-specific changes in reference intervals are substantially less distinctive, but reference intervals are narrower with stricter filtering. We decided from these results to use the data derived from the strictest sample preselection strategy (i.e., the healthiest population 'D4') for the definition of reference intervals.

Age- and sex-specific reference intervals

Reference intervals established using the strictest sample selection strategy D4 are shown in Fig 3 and Table II. These results show a decrease of sex-specific differences of haemoglobin, haematocrit, and red cell count reference intervals with age. Men's haemoglobin, haematocrit, and red cell count reference intervals decline continuously with age ($P < 0{\cdot}001$, see Table SI), both in lower and upper reference interval limits and in reference interval medians. Changes in women start later and are less pronounced (significant decreases in women's haemoglobin reference intervals after 50 years, $P < 0{\cdot}05$). On the contrary, sex-specific differences in MCH, MCHC and platelet count reference intervals remained relatively stable with age. MCV upper reference limits and medians increase with age in men and women ($P < 0{\cdot}001$). White cell count reference intervals show no substantial sex-specific differences (see Table SII) and a slight decrease of upper reference values with increasing age in women ($P < 0{\cdot}05$). Platelet counts are higher in women than in men ($P < 0{\cdot}001$), and lower reference limits decrease slightly with age in both sexes ($P < 0{\cdot}05$).

Proportion of samples outside the generated reference intervals

The proportion of test results within our primary dataset D0 outside the reference intervals calculated as described above is shown in Fig 4. A marked increase with age in the proportion of test results classified as abnormal despite the shift of reference intervals described above can be observed for haemoglobin, haematocrit, and red cell counts, particularly for men aged 60–90 years. For platelets, a similar, albeit less substantial increase of abnormal test results can be observed (mainly due to thrombocytopenia), while for red cell indices and white cell counts, we observed no clear age-specific trends in the fraction of measurements classified as pathologic.

Reproducibility of reference intervals in a tertiary care centre

To show that the established results are not specific to the primary dataset from a German laboratory service provider, we examined a second dataset from a German tertiary care

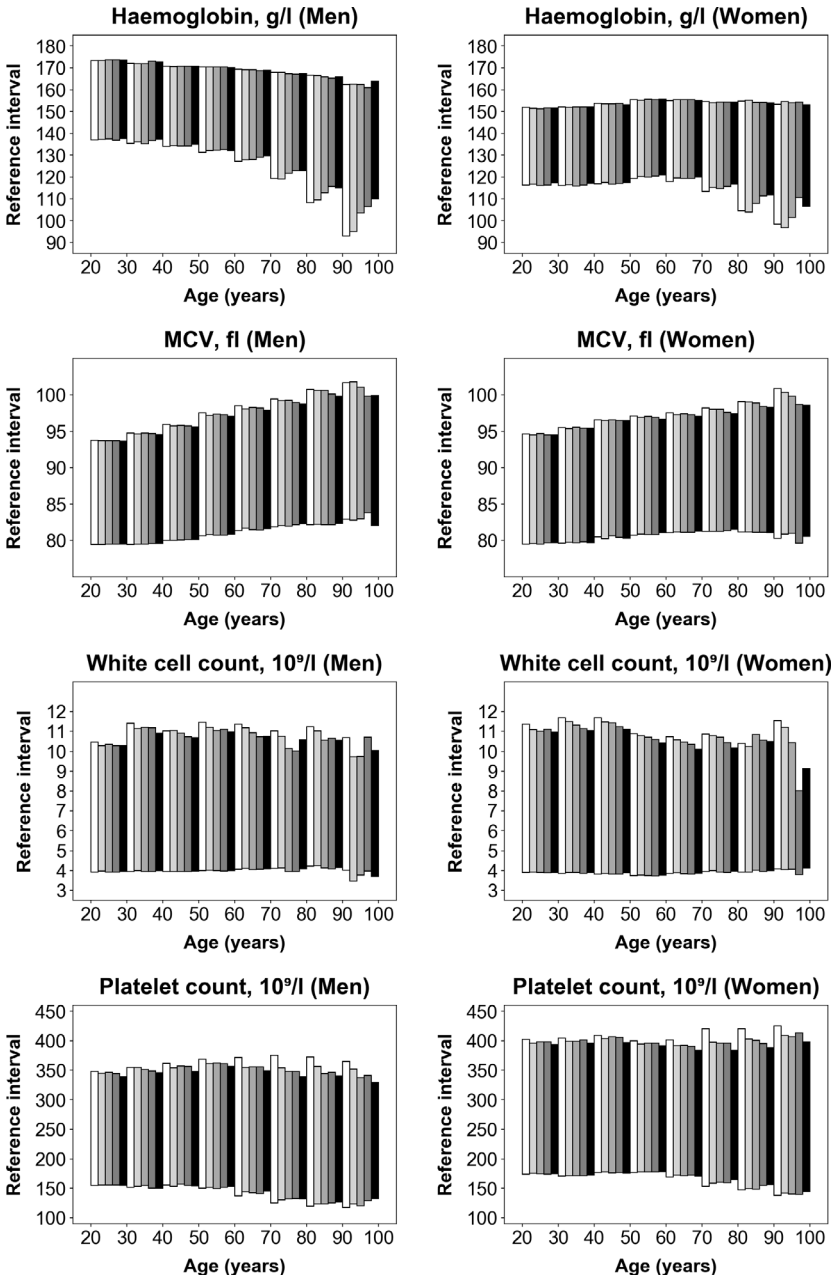

Fig 2. Effect of different filtering strategies on age-specific reference intervals. Bars show the age-specific reference intervals for different filters, with increasing selectivity for each age group [from left to right/white to black: (1., reference intervals D0 in Fig 1) no filter, (2., reference intervals D3) exclusion according to test results below the first or above the 99. percentile in other analytes, (3., reference intervals D1) exclusion according to creatinine and C-reactive protein (CRP), (4., reference intervals D2) exclusion according to CRP, and inclusion only of patients with eGFR (MDRD) ≥60 ml/min × 1·73 m^2, and (5., reference intervals D4) combination of all filters/final reference intervals]. See Figure S1 for results in other analytes. MCV, mean corpuscular volume.

hospital. Starting with 897 576 samples from 198 520 different patients, 269 128 samples from 128 050 patients remained after exclusion by unit [Table I]. Sample exclusion and sample selection were performed as described above and the strictest filtering strategy was used. Thus, 1 537–6 979 samples per age- and sex-specific group were available for the calculation of reference intervals in patients aged 20–90 years, and 110–310 samples were available for patients aged 90–100 years. Results for both centres are shown in Fig 5 and demonstrate comparable results in both datasets, without substantial differences in most analytes and age or sex groups. However, higher upper reference limits for white cell counts are estimated from the tertiary care centre dataset, especially in middle-aged males.

Discussion

Conventional strategies to define reference intervals require the acquisition of a sufficient number of samples from a healthy reference population. If reference intervals change with sex and age, the number of samples necessary is multiplied by the number of cohorts of interest. In elderly patients, recruitment of 'healthy' individuals poses additional challenges due to the high prevalence of chronic diseases and prescription medication: exclusion of individuals according to these criteria leads to the selection of the 'most healthy minority', with questionable appropriateness of resulting reference intervals for the general population, and substantial practical obstacles when recruiting reference individuals. We

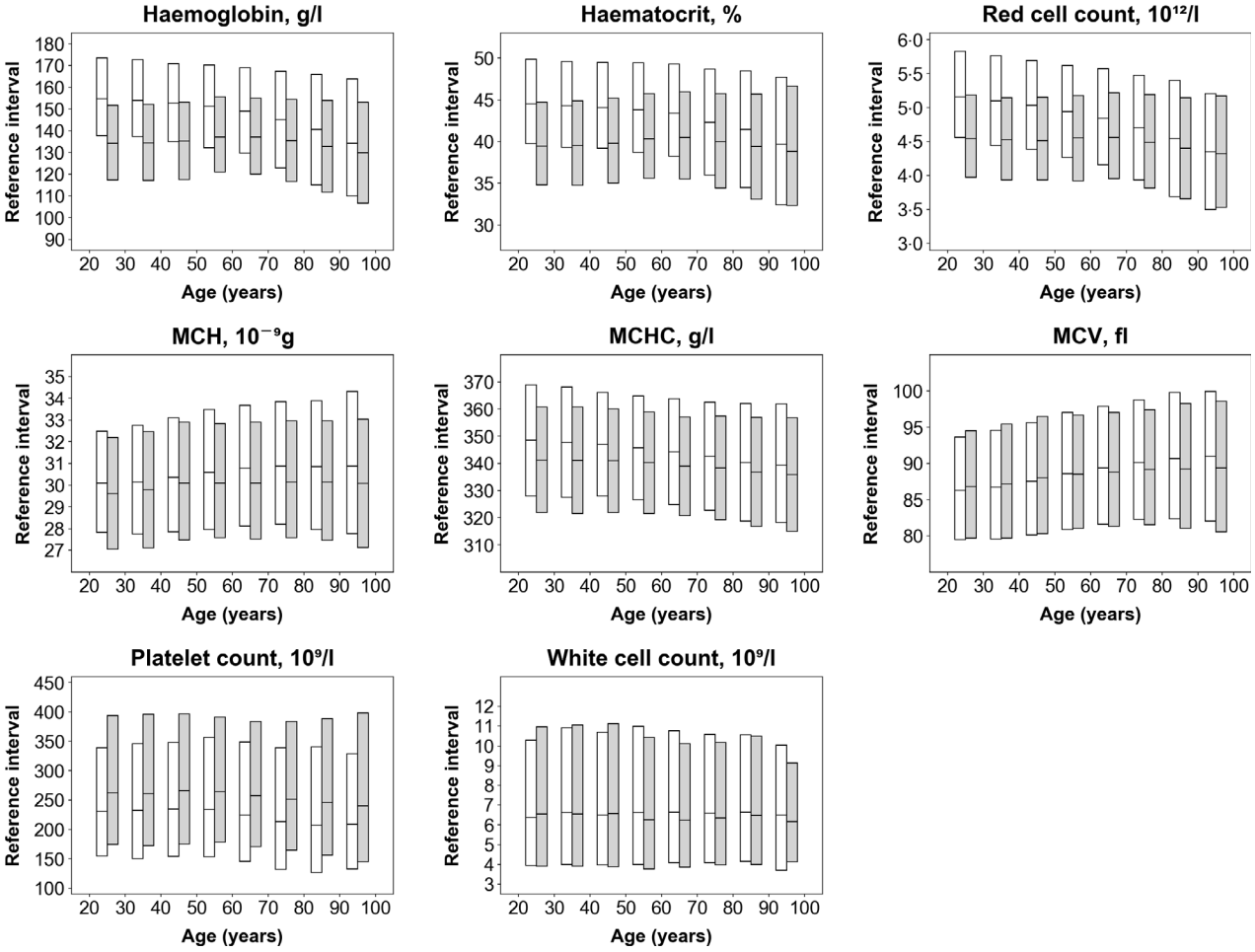

Fig 3. Age- and sex-specific haematology reference intervals for adults aged 20–100 years. Reference intervals (established from dataset D4 after sample exclusion) for men (white bars) and women (grey bars). Median values/50th percentiles are marked within the bars. MCH, mean corpuscular haemoglobin; MCHC, mean corpuscular haemoglobin concentration; MCV, mean corpuscular volume.

circumvented these problems by using a data mining approach and applying an algorithm (Arzideh *et al.*, 2011), that has been successfully used to evaluate paediatric (Zierk *et al.*, 2013, 2015, 2019) and adult (Zierk *et al.*, 2018) reference intervals in haematology, with the aim to define age- and sex-specific reference intervals in the elderly. We did not collect any medical histories, but, starting with two very large datasets of routine clinical samples, we removed pathological samples in a five-step process: (i) we excluded units and specialists caring for patients with a high expected fraction of haematologic abnormalities, (ii) we removed samples from patients with reduced renal function or acute-phase reaction, (iii) we excluded samples with clearly abnormal test results in the same sample in other haematology analytes, (iv) we selected only one sample per patient, thereby reducing over-representation of samples from chronically and critically ill patients, and (v) we calculated the distribution of physiologic measurements using an established algorithm (RLE), separating them from pathologic test results. Using this approach, we can support our findings on sample numbers for each age- and sex-specific cohort that are by magnitudes larger than available with conventional strategies, and that recruit participants using medical histories and clinical examinations (Cheng *et al.*, 2004; Adeli *et al.*, 2015; Röhrig *et al.*, 2018; Fulgoni *et al.*, 2019).

The validity of the reference intervals reported in this manuscript is supported by the largely similar results from two independent cohorts within this work, by the concordance of the results of the middle-aged populations with previous data from others (Pekelharing *et al.*, 2010; Ambayya *et al.*, 2014; Adeli *et al.*, 2015; Koerbin *et al.*, 2017; Ozarda *et al.*, 2017) and from our hospital using a different dataset measured with different haematology analysers in a previous study (Zierk *et al.*, 2018). Sample measurements in this study were performed using Sysmex haematology analysers; however, for the parameters described here, measurements from modern analysers are largely comparable, with subtle differences mainly for platelets, which are of minor clinical relevance from our point of view (Meintker *et al.*, 2013; Bruegel *et al.*, 2015; Zierk *et al.*, 2018). Since our results are derived

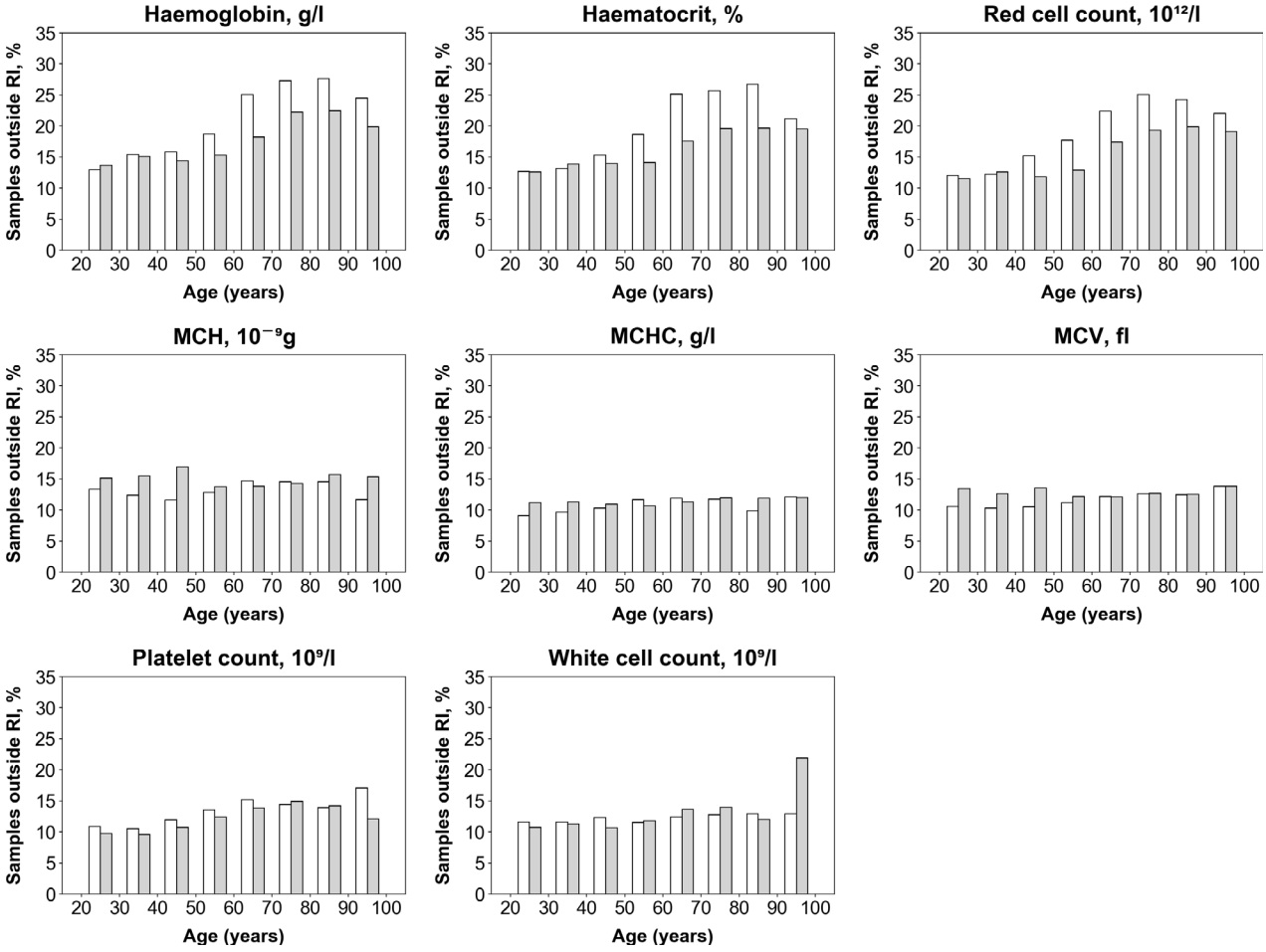

Fig 4. Proportion of samples outside the generated reference intervals. Age- and sex-specific proportion of test results classified as pathological according to the newly established reference intervals (white bars: men, grey bars: women). MCH, mean corpuscular haemoglobin; MCHC, mean corpuscular haemoglobin concentration; MCV, mean corpuscular volume.

from a mainly Caucasian population, it would be interesting to perform similar studies in populations with a different genetic background.

Platelet counts show clear sex-specific differences, with higher measurements in women. Upper reference limits are quite similar across ages, whereas lower limits slightly decrease with age. Due to the minor magnitude of the changes with age in comparison to intra-individual variation, we consider age-specific platelet count reference intervals not necessary from a practical point of view. For white cell counts, no clear sex-specific differences are observed, whereas a slight decline with increasing age in the upper reference limits was noted. This analyte was the only one with a marked difference between the two cohorts in our study: higher upper reference limits were calculated from the hospital cohort despite exclusion of patients with an elevated CRP. The differences between the two datasets can probably be explained by a higher fraction of patients with mild leucocytosis in the hospital dataset, e.g. due to mild immune activation not detected by CRP or due to stress — if this is a frequent but very slight deviation, it is not completely captured by the RLE algorithm that has its strength in the detection of clearly abnormal measurements (Zierk *et al.*, 2018). From a practical point of view, we suggest the results for the outpatient cohort are applied without sex- and age-specific separation.

Age- and sex-related variation of erythrocyte parameters are of special interest. For the definition of anaemia, most authors use the World Health Organization (WHO) definitions of 120 g/l (women) and 130 g/l (men), without stratification by age in adults (WHO Scientific Group on Nutritional Anaemias & World Health Organization, 1968). However, the validity of these cut-off values (which have been developed more than 40 years ago by a WHO expert group) for older subjects is controversial (Beutler & Waalen, 2006). Data from the National Health and Nutrition Examination Survey (NHANES) 1999 to 2012 derived from 20 497 adults aged 20 to 79 years (41% exclusion rate from the primary population-based sample) shows decreasing haemoglobin reference intervals with age in men aged 20–79 years, while female reference intervals are essentially stable with a

slight increase after menopause (Fulgoni *et al.*, 2019). Similar trends had been reported before from the earlier NHANES III data based on 6 472 samples from healthy adults (exclusion rate 62%) (Cheng *et al.*, 2004). The Canadian Health Measures Survey reports changes in reference intervals mainly in children. While 4 070 samples were available from individuals aged 15–79 years, 79 % of screened individuals aged 60–79 years old were excluded (Adeli *et al.*, 2015; Horowitz, 2015). This makes the judgement of specific changes in the elderly difficult, although, a slight increase of MCV and a decrease of haemoglobin in men can be observed with increasing age (Adeli *et al.*, 2015).

Overall, the reference intervals from the publications cited above are in accordance with our data regarding age- and sex-specific differences as well as absolute magnitudes. However, due to our different approach starting with routine data, we can calculate reference intervals from cohorts that are much larger and thus give a much more exact picture. In men, medians and upper and lower reference limits of haemoglobin, haematocrit and red cell count decrease with age. This change is most pronounced for red cell counts, because we observe a parallel increase in MCH and MCV. All of these changes occur in a continuous fashion between the ages of 30 and 90 years. In women, we see a slight rise of haemoglobin and haematocrit after menopause and a decline in the lower reference limits thereafter. In both sexes, we observe widening reference intervals in the elderly.

In contrast to our results and to the studies cited above, a 2018 report by Röhrig *et al.* shows reference intervals in the three age decades of individuals between 60 and 90 years old to be essentially identical and also similar to younger individuals in a study specifically devoted to analytes of

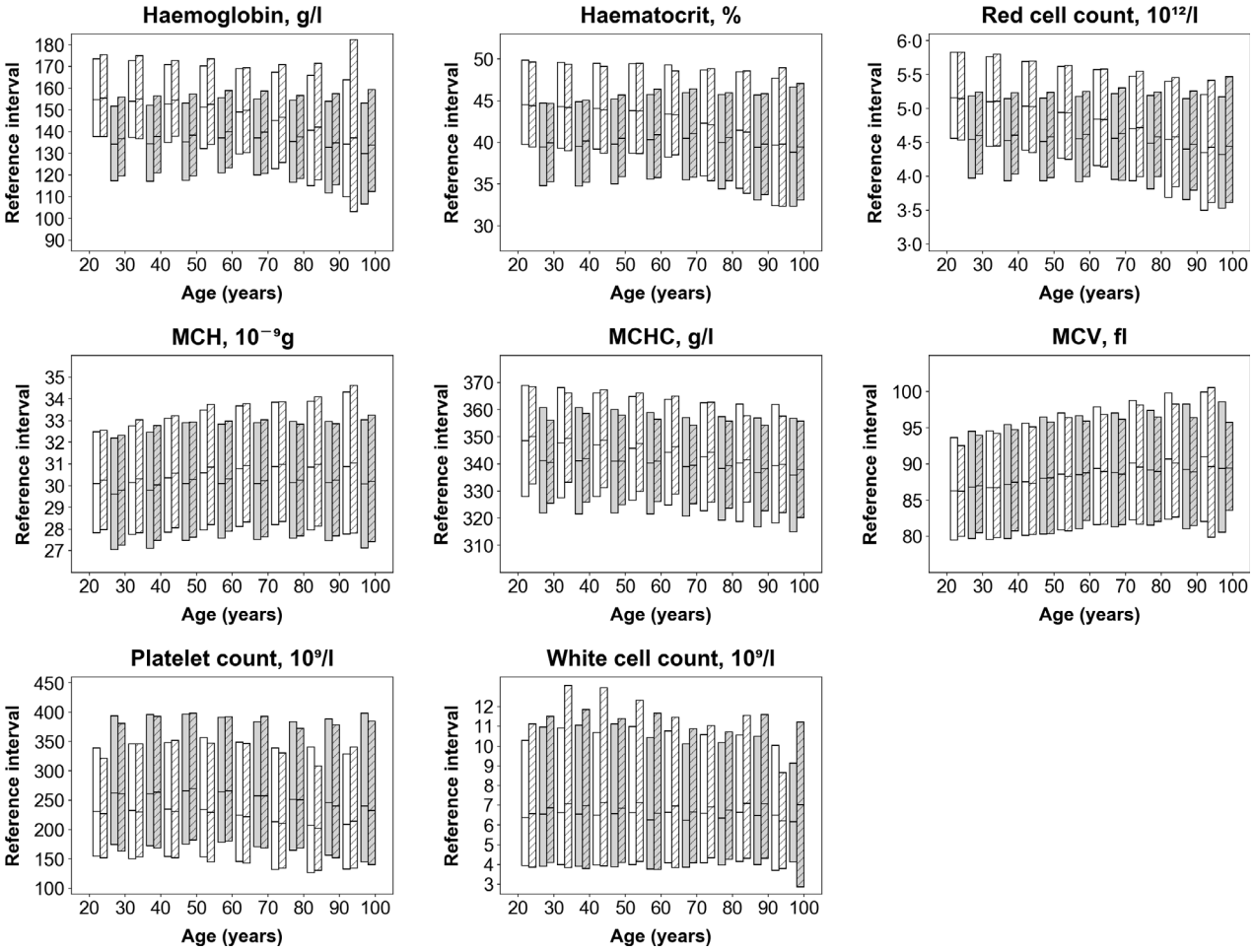

Fig 5. Reproducibility of reference intervals in a tertiary care centre. Age-specific reference intervals for haematology analytes for men (white bars, left pair) and women (grey bars, right pair) in different datasets [no bar pattern/left bar: final dataset (D4) used for reference interval estimation in this report (laboratory service provider, mainly outpatients), bar with pattern background/right bar: validation dataset from a tertiary care centre (mainly inpatients)]. Reference intervals for men and women aged 90–100 years in the tertiary care centre have to be considered cautiously due to the reduced number of samples available in this group. MCH, mean corpuscular haemoglobin; MCHC, mean corpuscular haemoglobin concentration; MCV, mean corpuscular volume.

[Correction added on 6 June 2020, after online publication: In Figures 4 and 5, the figure legends were previously swapped and have been corrected in this current version].

erythropoiesis in the elderly (Röhrig *et al.*, 2018). These results have to be considered critically due to methodological aspects. As in other studies, patients were excluded based on history and laboratory indicators of disease. However, in this study samples were also excluded if any blood count test result was outside the normal range defined in a German guideline (https://www.onkopedia.com/de/onkopedia/guidelines/eisenmangel-und-eisenmangelanaemie), i.e. haemoglobin from a sample was excluded from further analysis if the haematocrit from the same sample was not in the predefined normal range. Since haemoglobin, haematocrit and red cell count are tightly correlated, we consider this strategy misleading and biased towards the predefined reference intervals, pushing the remaining cohort artificially towards 'super-healthy' normal values.

Based on our results and in accordance with published data, we conclude that with increasing age, 'physiologic' values of erythropoiesis decline, as we observe in men, whereas in women this shift is counteracted by blood loss during menstruation and possibly additional hormonal effects, leading to lower values in younger age, followed by a slight rise after menopause with a final drop of lower haemoglobin reference intervals in women >80 years.

These findings are in line with the current understanding of the physiology and pathophysiology of haematopoietic stem cell (HSC) ageing (Groarke & Young, 2019): Although the number of HSCs increases in the elderly and their differentiation is skewed towards myeloid progenitors, these stem cells show reduced regenerative capacity and loss in function (Geiger *et al.*, 2013). Increasing levels of erythropoietin are believed to partly compensate for decreased erythropoietin sensitivity (Ershler *et al.*, 2005), and overall erythroid cell output of aged HSCs is substantially reduced (Geiger *et al.*, 2013; Chung & Park, 2017). The age-specific reference intervals reported by us reflect these established physiologic effects of ageing. Importantly, when using the reported reference intervals — despite their shift with age — a higher prevalence of anaemia is still detected in the elderly, as exemplified by increasing percentages of abnormal measurements in the input dataset (Fig 4).

Having established a physiological drop of haemoglobin with increasing age, this leads to a conflict with results from several studies showing an association between anaemia and morbidity and mortality in the elderly (Culleton *et al.*, 2006; Penninx *et al.*, 2006; Lee *et al.*, 2018b; Wouters *et al.*, 2019). According to these studies, anaemic elderly individuals have an increased risk of death in the following years even after correction of major confounding factors. This holds true if anaemia is defined with age-independent lower limits as suggested by the WHO, but some increase in risk is already observed with haemoglobin concentrations in the low-normal range at least in women (i.e., HB 120–130 g/l). Interestingly, a similar phenomenon was not found in individuals below the age of 60 in one population-based study (Wouters *et al.*, 2019), but was also observed in younger individuals in another study (Lee *et al.*, 2018b, 2018a). In 2019, Wouters *et al.* grouped individuals by the type of anaemia according to additional serum measurements into anaemia of nutritional deficiency (including iron deficiency), anaemia of chronic inflammation, and unexplained anaemia, and found that anaemia of chronic inflammation is strongly correlated with inferior quality of life and increased mortality, whereas the influence of anaemia of other causes is much less pronounced (Wouters *et al.*, 2019).

Interpreting our findings in the context of the aforementioned reports, we conclude that lower red blood cell-related reference intervals in the elderly are a physiological phenomenon. At the same time, measurements in the sub-normal and low-normal range are linked to decreased quality of life and increased risk of death, probably by correlation rather than by causation. Therefore, it would be prudent to check for a treatable cause of anaemia in elderly individuals with haemoglobin values below 130 g/l (i.e. anaemic or in the low-normal range by current definitions). Intra-individual trends of red blood cell-related analytes in sequential tests may give further information; however, this topic is beyond the scope of this report. If no disease is found, an individual may just have his individual 'set point' in the lower range of the physiologic age-dependent distribution.

Acknowledgements

This work was conducted within the MIRACUM consortium. MIRACUM is funded by the German Ministry for Education and Research (BMBF) [funding number FKZ01ZZ1801A].

Conflicts of interest

The authors declare no other relevant competing financial interests.

Author contributions

JZ designed and performed research, performed statistical analyses, analyzed and interpreted data, and wrote the manuscript. SWK designed and performed research, analyzed and interpreted data, and wrote the manuscript. AK, ES, and AL collected data. MR and MM analyzed and interpreted data.

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Fig S1. Effect of different filtering strategies on age-specific reference intervals. Bars show the age-specific reference intervals for different filters, with increasing selectivity for each age group [from left to right/white to black: (1., reference intervals D0 in Fig 1) no filter, (2., reference intervals D3) exclusion according to test results below the first or above the 99. percentile in other analytes, (3., reference intervals

D1) exclusion according to creatinine and C-reactive protein (CRP), (4., reference intervals D2) exclusion according to CRP, and inclusion only of patients with eGFR (MDRD) ≥ 60 ml/min × 1·73 m^2, and (5., reference intervals D4) combination of all filters/final reference intervals]. See Fig 2 for results in other analytes.

Table SI. Significance of trends with age and extent of change (per decade of life) in reference intervals' 2.5th, 50th, 95.5th percentiles and reference intervals' widths.

Table SII. Significance of differences between men and women and extent of differences in reference intervals' 2.5th, 50th, 95.5th percentiles.

References

Adeli, K., Raizman, J.E., Chen, Y., Higgins, V., Nieuwesteeg, M., Abdelhaleem, M., Wong, S.L. & Blais, D. (2015) Complex biological profile of hematologic markers across pediatric, adult, and geriatric ages: establishment of robust pediatric and adult reference intervals on the basis of the Canadian health measures survey. *Clinical Chemistry*, **61**, 1075–1086.

Ambayya, A., Su, A.T., Osman, N.H., Nik-Samsudin, N.R., Khalid, K., Chang, K.M., Sathar, J., Rajasuriar, J.S. & Yegappan, S. (2014) Haematological reference intervals in a multiethnic population. *PLoS ONE*, **9**, e91968.

Arzideh, F. (2008) Estimation of Medical Reference Limits by Truncated Gaussian and Truncated Power Normal Distributions. http://elib.suub.uni-bremen.de/diss/docs/00011259.pdf

Arzideh, F., Wosniok, W. & Haeckel, R. (2010) Reference limits of plasma and serum creatinine concentrations from intra-laboratory data bases of several German and Italian medical centres: comparison between direct and indirect procedures. *Clinica Chimica Acta*, **411**, 215–221.

Arzideh, F., Wosniok, W. & Haeckel, R. (2011) Indirect reference intervals of plasma and serum thyrotropin (TSH) concentrations from intra-laboratory data bases from several German and Italian medical centres. *Clinical Chemistry and Laboratory Medicine*, **49**, 659–664.

Beutler, E. & Waalen, J. (2006) The definition of anemia: what is the lower limit of normal of the blood hemoglobin concentration? *Blood*, **107**, 1747–1750.

Bruegel, M., Nagel, D., Funk, M., Fuhrmann, P., Zander, J. & Teupser, D. (2015) Comparison of five automated hematology analyzers in a university hospital setting: Abbott Cell-Dyn Sapphire, Beckman Coulter DxH 800, Siemens Advia 2120i, Sysmex XE-5000, and Sysmex XN-2000. *Clinical Chemistry and Laboratory Medicine*, **53**, 1057–1071.

Cheng, C.K.-W., Chan, J., Cembrowski, G.S. & van Assendelft, O.W. (2004) Complete blood count reference interval diagrams derived from NHANES III: stratification by age, sex, and race. *Laboratory Hematology*, **10**, 42–53.

Chung, S.S. & Park, C.Y. (2017) Aging, hematopoiesis, and the myelodysplastic syndromes. *Blood Advances*, **1**, 2572–2578.

Culleton, B.F., Manns, B.J., Zhang, J., Tonelli, M., Klarenbach, S. & Hemmelgarn, B.R. (2006) Impact of anemia on hospitalization and mortality in older adults. *Blood*, **107**, 3841–3846.

Ershler, W.B., Sheng, S., McKelvey, J., Artz, A.S., Denduluri, N., Tecson, J., Taub, D.D., Brant, L.J., Ferrucci, L. & Longo, D.L. (2005) Serum erythropoietin and aging: a longitudinal analysis. *Journal of the American Geriatrics Society*, **53**, 1360–1365.

Fulgoni, V.L., Agarwal, S., Kellogg, M.D. & Lieberman, H.R. (2019) Establishing pediatric and adult RBC reference intervals with NHANES data using piecewise regression. *American Journal of Clinical Pathology*, **151**, 128–142.

Geiger, H., de Haan, G. & Florian, M.C. (2013) The ageing haematopoietic stem cell compartment. *Nature Reviews. Immunology*, **13**, 376–389.

Groarke, E.M. & Young, N.S. (2019) Aging and hematopoiesis. *Clinics in Geriatric Medicine*, **35**, 285–293.

Haeckel, R., Wosniok, W., Arzideh, F., Zierk, J., Gurr, E. & Streichert, T. (2017) Critical comments to a recent EFLM recommendation for the review of reference intervals. *Clinical Chemistry and Laboratory Medicine (CCLM)*, **55**. Available at: https://www.degruyter.com/view/j/cclm.ahead-of-print/cclm-2016-1112/cclm-2016-1112.xml?format=INT [Accessed February 3, 2017].

Horowitz, G.L. (2015) The power of asterisks. *Clinical Chemistry*, **61**, 1009–1011.

Hussain, Md & Mahmud, I. (2019) pyMannKendall: a python package for non parametric Mann Kendall family of trend tests. *Journal of Open Source Software*, **4**, 1556.

Ichihara, K. & Boyd, J.C. (2010) An appraisal of statistical procedures used in derivation of reference intervals. *Clinical Chemistry and Laboratory Medicine*, **48**, 1537–1551.

Ichihara, K., Ozarda, Y., Barth, J.H., Klee, G., Qiu, L., Erasmus, R., Borai, A., Evgina, S., Ashavaid, T., Khan, D., Schreier, L., Rolle, R., Shimizu, Y., Kimura, S., Kawano, R., Armbruster, D., Mori, K. & Yadav, B.K.; Committee on Reference Intervals and Decision Limits; International Federation of Clinical Chemistry and Laboratory Medicine. (2017) A global multicenter study on reference values: 1. Assessment of methods for derivation and comparison of reference intervals. *Clinica Chimica Acta*, **467**, 70–82.

Jones, G.R.D., Haeckel, R., Loh, T.P., Sikaris, K., Streichert, T., Katayev, A., Barth, J.H. & Ozarda, Y. (2018) Indirect methods for reference interval determination – review and recommendations.. *Clinical Chemistry and Laboratory Medicine (CCLM)*, **57**, 20–29. https://doi.org/10.1515/cclm-2018-0073

Koerbin, G., Potter, J.M., Andriolo, K., West, N.P., Glasgow, N., Hawkins, C., Cavanaugh, J.A. & Hickman, P.E. (2017) 'Aussie Normals': an a priori study to develop reference intervals in a healthy Australian population using the Beckman Coulter LH 750 Haematology Analyser as candidates for harmonised values. *Pathology*, **49**, 518–525.

Lee, G., Choi, S., Kim, K., Yun, J.-M., Son, J.S., Jeong, S.-M., Kim, S.M., Kim, Y.-Y., Park, S.Y., Koh, Y., Hwang, S.-S. & Park, S.M. (2018a) Association between changes in hemoglobin concentration and cardiovascular risks and all-cause mortality among young women. *Journal of the American Heart Association*, **7**, e008147.

Lee, G., Choi, S., Kim, K., Yun, J.-M., Son, J.S., Jeong, S.-M., Kim, S.M. & Park, S.M. (2018b) Association of hemoglobin concentration and its change with cardiovascular and all-cause mortality. *Journal of the American Heart Association*, **7**, e007723.

Levey, A.S., Stevens, L.A., Schmid, C.H., Zhang, Y.L., Castro, A.F., Feldman, H.I., Kusek, J.W., Eggers, P., Van Lente, F., Greene, T., Coresh, J.; CKD-EPI (Chronic Kidney Disease Epidemiology Collaboration). (2009) A new equation to estimate glomerular filtration rate. *Annals of Internal Medicine*, **150**, 604–612.

Meintker, L., Ringwald, J., Rauh, M. & Krause, S.W. (2013) Comparison of automated differential blood cell counts from Abbott Sapphire, Siemens Advia 120, Beckman Coulter DxH 800, and Sysmex XE-2100 in normal and pathologic samples. *American Journal of Clinical Pathology*, **139**, 641–650.

Ozarda, Y., Ichihara, K., Bakan, E., Polat, H., Ozturk, N., Baygutalp, N.K., Taneli, F., Guvenc, Y., Ormen, M., Erbayraktar, Z., Aksoy, N., Sezen, H., Demir, M., Eskandari, G., Polat, G., Mete, N., Yuksel, H., Vatansev, H., Gun, F., Akin, O., Ceylan, O., Noyan, T., Gozlukaya, O., Aliyazicioglu, Y., Kahraman, S., Dirican, M., Tuncer, G.O., Kimura, S. & Eker, P. (2017) A nationwide multicentre study in Turkey for establishing reference intervals of haematological parameters with novel use of a panel of whole blood. *Biochemia Medica*, **27**, 350–377.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M. & Duchesnay, É. (2011) Scikit-learn: machine learning in python. *Journal of Machine Learning Research*, **12**, 2825–2830.

Pekelharing, J.M., Hauss, O., De Jonge, R., Lokhoff, J., Sodikromo, J., Spaans, M., Brouwer, R., De Lathouder, S. & Hinzmann, R. (2010) Haematology reference intervals for established and

novel parameters in healthy adults. *Sysmex Journal International*, **20**, 1–9.

Penninx, B.W.J.H., Pahor, M., Woodman, R.C. & Guralnik, J.M. (2006) Anemia in old age is associated with increased mortality and hospitalization. *The Journals of Gerontology. Series A, Biological Sciences and Medical Sciences*, **61**, 474–479.

Röhrig, G., Becker, I., Gutensohn, K. & Nebe, T. (2018) Red blood cell counts and indices in the elderly German population. *Journal of Laboratory Medicine*, **42**, 131–139.

WHO Scientific Group on Nutritional Anaemias & World Health Organization. (1968) Nutritional anaemias : report of a WHO scientific group [meeting held in Geneva from 13 to 17 March 1967]. Geneva: World Health Organization. Available at: https://apps.who.int/iris/handle/10665/40707 [Accessed March 25, 2019].

Wouters, H.J.C.M., van der Klauw, M.M., de Witte, T., Stauder, R., Swinkels, D.W., Wolffenbuttel, B.H.R. & Huls, G. (2019) Association of anemia with health-related quality of life and survival: a large population-based cohort study. *Haematologica*, **104**, 468–476.

Zierk, J., Arzideh, F., Haeckel, R., Rascher, W., Rauh, M. & Metzler, M. (2013) Indirect determination of pediatric blood count reference intervals. *Clinical Chemistry and Laboratory Medicine*, **51**, 863–872.

Zierk, J., Arzideh, F., Rechenauer, T., Haeckel, R., Rascher, W., Metzler, M. & Rauh, M. (2015) Age- and sex-specific dynamics in 22 hematologic and biochemical analytes from birth to adolescence. *Clinical Chemistry*, **61**, 964–973.

Zierk, J., Arzideh, F., Haeckel, R., Rauh, M., Metzler, M., Ganslandt, T. & Krause, S.W. (2018) Indirect determination of hematology reference intervals in adult patients on Beckman Coulter UniCell DxH 800 and Abbott CELL-DYN Sapphire devices. *Clinical Chemistry and Laboratory Medicine (CCLM)*, **57**, 730–739.

Zierk, J., Hirschmann, J., Toddenroth, D., Arzideh, F., Haeckel, R., Bertram, A., Cario, H., Frühwald, M.C., Groß, H.-J., Groening, A., Grützner, S., Gscheidmeier, T., Hoff, T., Hoffmann, R., Klauke, R., Krebs, A., Lichtinghagen, R., Mühlenbrock-Lenter, S., Neumann, M., Nöllke, P., Niemeyer, C.M., Razum, O., Ruf, H.-G., Steigerwald, U., Streichert, T., Torge, A., Rascher, W., Prokosch, H.-U., Rauh, M. & Metzler, M. (2019) Next-generation reference intervals for pediatric hematology. *Clinical Chemistry and Laboratory Medicine (CCLM)*, **57**, 1595–1607.

