Clinica Chimica Acta 499 (2019) 108–114

Contents lists available at ScienceDirect

Clinica Chimica Acta

journal homepage: www.elsevier.com/locate/cca

Data mining of reference intervals for coagulation screening tests in adult patients

Jakob Zierk[a,b,*], Thomas Ganslandt[c], Manfred Rauh[a], Markus Metzler[a], Erwin Strasser[d]

[a] Department of Pediatrics and Adolescent Medicine, University Hospital Erlangen, Erlangen, Germany
[b] Center of Medical Information and Communication Technology, University Hospital Erlangen, Erlangen, Germany
[c] Heinrich-Lanz-Center, Ruprecht-Karls-University Heidelberg, Mannheim University Medicine, Mannheim, Germany
[d] Department of Transfusion Medicine and Haemostaseology, University Hospital of Erlangen, Erlangen, Germany

ARTICLE INFO

Keywords:
Reference intervals
Coagulation screening tests
Data mining

ABSTRACT

Background: Appropriate reference intervals are essential when evaluating laboratory test results. However, establishment of reference intervals is challenging, especially for coagulation screening tests, and uncertainty exists regarding age- and sex-dependency of test results. Data mining of laboratory information systems is an emerging approach to reference interval determination, and we evaluated its applicability to coagulation tests.
Methods: We analyzed measurements of activated partial thromboplastin time (aPTT), prothrombin time (PT), international normalized ratio (INR), thrombin time (TT), and fibrinogen performed during clinical care in the University Hospital Erlangen, Germany (1,778,738 samples from 116,754 adult patients, 45,577-509,859 samples per analyte). We identified the proportion of samples from healthy individuals using an established statistical approach (Reference Limit Estimator), in which the distribution of physiological test results is approximated using a parametrical function, and used for the calculation of reference intervals.
Results: We established age- and sex specific reference intervals for aPTT, PT, INR, TT, and fibrinogen, and created batch- and reagent-specific aPTT-reference intervals. Additionally, we evaluated the sensitivity of the established aPTT reference intervals for the detection of factor VIII, IX, XI, XII deficiencies.
Conclusion: Data mining of laboratory test results allows the creation of age- and sex-reference intervals for coagulation tests that are specific to the examined population, analytical framework, and reagent. This approach can complement conventional methods when establishing reference intervals and improve clinical decision-making based on coagulation tests. The reference intervals established in this study show only minor variation with sex and age, supporting the practice of providing a common reference interval for adult women and men.

1. Introduction

Activated partial thromboplastin time (aPTT), prothrombin time (PT), international normalized ratio (INR), thrombin time (TT), and fibrinogen are frequently performed coagulation tests measured using one-stage clotting assays. For diagnostic and therapeutic decisions, individual patients' coagulation test results are interpreted in comparison to reference intervals provided by the laboratory [1–3]. Appropriate assessment of test results therefore requires reference intervals specific to the examined population, analytical framework, and reagent [1,4]. For coagulation screening tests, intra-laboratory establishment or validation of reference intervals is particularly important, due to increased reagent- and batch-dependency of test results in comparison to e.g. biochemical analyses. Additionally, uncertainty exists regarding age- and sex-dependency of test results, although an ageing population throughout the world and a high frequency of elderly patients undergoing laboratory testing rely on appropriate interpretation of test results [2,5–7].

Established direct approaches for the calculation of reference intervals, as e.g. defined by the Clinical & Laboratory Standards Institute (CLSI) [8], require measurement of a large number of samples (in most cases > 120) in a carefully selected homogenous population of healthy reference individuals [3,8,9], a procedure which is time-consuming and costly. Recruitment of a reference population of healthy volunteers often poses considerable challenges, and in many cases samples from healthy blood donors are used [9,10]. As a consequence, the reference population frequently varies considerably from the population to which the test results are ultimately applied. Often, patients are substantially older, receiving medication, and have a high prevalence of co-morbidities [11]. This leads to uncertainty regarding the appropriateness of

* Corresponding author at: Department of Pediatrics and Adolescent Medicine, University Hospital Erlangen, Loschgestr. 15, Erlangen 91054, Germany.
E-mail address: jakob.zierk@uk-erlangen.de (J. Zierk).

https://doi.org/10.1016/j.cca.2019.09.006
Received 14 June 2019; Received in revised form 31 August 2019; Accepted 7 September 2019
Available online 09 September 2019
0009-8981/

reference intervals created using a population of young and healthy adults and has prevented the establishment of reference intervals for older adults. Additionally, the width of reference intervals' confidence intervals in conventional approaches is often unacceptable, especially when the distribution of test results is non-normal [12].

Indirect approaches have been established as an alternative to reference interval determination [12–14]. These methods are based on the assumption that the majority of samples measured during patient care are not pathological and can be used to create reference intervals [14]. To this end, advanced statistical methods are applied, which identify the proportion of physiological samples in a mixed dataset, containing both pathological and physiological samples. The availability of samples from laboratory information systems allows the creation of reference intervals which are specific to the examined population, age-group, analytical framework, batches and reagents. Indirect methods have been extensively used in the pediatric population, which poses extraordinary challenges when establishing reference intervals [13,15], and specifically, pediatric reference intervals for coagulation tests have been established using indirect methods [16]. However, indirect methods have not been studied in the context of adult coagulation reference intervals so far. The aim of this report is the creation of reference intervals for coagulation screening tests (aPTT, PT, INR, TT and fibrinogen) for different age groups using an established indirect method, and to evaluate the potential of this approach to identify batch- and reagent-specific differences.

2. Methods

2.1. Study population and selection of samples

We analyzed measurements of aPTT, prothrombin time, INR, thrombin time, and fibrinogen performed during clinical care of patients in the University Hospital Erlangen from July 2014 to June 2017. Samples from inpatients and outpatients aged 20–90 years, including patients from intensive care units and specialty units, were retrieved from the laboratory database. This resulted in 421,903, 496,889, 200,544, 44,814, and 495,290 test results from 97,510, 112,309, 67,278, 17,830, and 112,086 different patients for aPTT, PT, INR, TT, and fibrinogen, respectively. To explore the age-dependency of reference intervals, samples were split into fine-grained age groups of 10 years (20–30 years, 30–40 years, …, 80–90 years) and a combined age group (20–90 years). We selected 10-year partitions, as statistical approaches to age partitioning, e.g. the Harris & Boyd test, are not applicable to mixed populations of pathological and physiological test results in non-Gaussian distributions, and can potentially be easily integrated into laboratory information systems. The investigated patient population is composed predominantly of Caucasian individuals and no stratification according to ethnicity was performed.

2.2. Analytical procedures

Measurements were performed on an automated coagulation analyzer STA-R (STAGO, France), in accordance with manufacturer instructions and laboratory standard operating procedures. INR was calculated based on the PT using the batch-specific International Sensitivity Index (ISI) values provided by the manufacturer. APTT was measured using STA Neoplastine (STAGO, France, July 2014 to December 2014) and different batches of Cephascreen (STAGO, France, December 2014 to June 2017) reagents, differences between reagents and batches are analyzed in a separate analysis. For the final aPTT reference intervals, only aPTT measurements using Cephascreen reagents were used. The Clauss fibrinogen assay (STA-Liquid FIB, STAGO, France) is based on the thrombin clotting time [17]. External quality control was performed at least four times a year. Analyte stability over time during the study period was analyzed by Teil-Senn regression of quarterly median values (Supplemental Table 1). While Teil-Senn regression shows slight shifts during the study period in aPTT, the magnitude of these shifts is not clinically relevant in relation to guidelines for laboratory quality control [18,19] and the range of established reference intervals, and we thus included all time intervals for further analysis.

2.3. Calculation of reference intervals

Reference intervals were calculated with an indirect algorithm described and validated previously [14] and freely available as a software package (“Reference Limit Estimator (RLE)”, developed by the German Society of Clinical Chemistry and Laboratory Medicine's Working Group on Guide Limits, available at http://www.dgkl.de/). 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 assumed to follow a parametric distribution (a Gaussian distribution after Box-Cox transformation of the data, i.e. a distribution that can accommodate skewed data), whereas pathologic samples are assumed to be scattered randomly. Using an elaborate statistical process, the distribution of physiological samples is isolated from the mixed dataset and used to calculate reference intervals. We analyzed datasets containing only one sample per patient; if more than one sample per individual was available we selected one randomly (in accordance to previous studies, which show no benefit of more systematic sample selection strategies [20]). Reference intervals for aPTT, prothrombin time, INR, thrombin time were estimated using RLE setting “Pathological values” set to “High”, fibrinogen reference intervals using “Pathological values” set to “Both”. (“High” indicates that the majority of pathological values are higher than physiological test results, whereas “Both” indicates that pathological values are expected below and above the distribution of normal test results.) Samples were anonymized before calculation of reference intervals

2.4. Validation of reference intervals in a filtered dataset

To validate the established reference intervals, we calculated reference intervals in a second dataset, in which we included only test results from patients with only one test result per calendar year (resulting in elimination of patients in which retesting was deemed appropriate by the treating physicians, i.e. most patients on anticoagulants and patients from intensive care units.)

2.5. Evaluation of aPTT sensitivity

To evaluate the sensitivity of the calculated aPTT upper reference limits with respect to the detection of coagulation factor deficiencies, we investigated blood samples where aPTT and coagulation factor VIII, IX, XI, or XII were measured simultaneously (n = 5725, 3857, 3711, and 2727). The sensitivity of the calculated aPTT upper reference limits as a cut-off level to detect factor deficiencies was calculated for decreasing residual factor activities (measured activity 100%–0%).

2.6. Influence of age- and sex-specific reference intervals on test result classification

To examine the hypothetical effect of the newly established reference intervals on classification of samples as healthy or pathologic, we applied the generated reference intervals to the input dataset and calculated the fraction of samples below and above the 2.5th and 97.5th percentiles. We compared classification of age- and sex-specific reference intervals, sex-specific (i.e. not age-specific) reference intervals, and an overall reference interval.

3. Results

Age- and sex-specific reference intervals and combined reference

Table 1
Age- and sex-specific reference intervals for activated partial thromboplastin time (aPTT), international normalized ratio (INR), thrombin time, fibrinogen, and prothrombin time.

Age	Men				Women				Men & Women			
	RI	50th	λ	n	RI	50th	λ	n	RI	50th	λ	n
Activated partial thromboplastin time (aPTT), s												
20–30	25.4–34.4	29.9	1.0	5303	25.4–33.7	29.6	1.0	6613	25.4–34.1	29.7	1.0	11,916
30–40	25.2–34.6	29.6	0.12	4876	25.1–33.9	29.5	1.0	6860	25.1–34.0	29.5	0.92	11,736
40–50	25.0–32.9	29.0	1.0	6216	25.0–33.7	29.1	0.16	6677	24.9–33.1	29.0	0.97	12,893
50–60	25.0–33.8	29.1	0.0	9753	24.6–33.8	28.8	0.0	8550	24.8–33.9	29.0	0.0	18,303
60–70	24.7–33.3	29.0	1.0	9183	24.5–33.0	28.4	0.0	7820	24.7–33.8	28.9	0.0	17,003
70–80	25.1–34.9	29.6	0.0	9501	24.3–33.6	28.6	0.0	8600	24.6–34.5	29.1	0.0	18,101
80–90	25.4–35.2	29.9	0.0	4339	24.0–34.4	29.2	1.0	5213	24.7–35.0	29.6	0.44	9552
20–90	25.1–34.3	29.4	0.13	48,122	24.8–34.3	29.1	0.0	49,386	24.9–34.4	29.3	0.0	97,508
International normalized ratio (INR), %												
20–30	0.88–1.17	1.02	0.69	6197	0.86–1.14	1.0	1.0	7803	0.88–1.15	1.01	0.69	14,000
30–40	0.86–1.14	0.99	0.0	5680	0.86–1.12	0.99	0.57	7969	0.86–1.13	0.99	0.19	13,649
40–50	0.86–1.11	0.98	0.0	7297	0.85–1.12	0.98	0.0	7838	0.85–1.12	0.98	0.0	15,135
50–60	0.85–1.13	0.98	0.0	11,165	0.85–1.09	0.96	0.0	9892	0.85–1.11	0.97	0.0	21,057
60–70	0.85–1.13	0.99	1.0	10,509	0.84–1.1	0.96	0.0	8980	0.85–1.13	0.98	0.0	19,489
70–80	0.87–1.19	1.02	0.0	10,928	0.84–1.12	0.98	0.86	9939	0.86–1.14	1.0	0.43	20,867
80–90	0.89–1.25	1.06	0.0	4890	0.88–1.19	1.02	0.0	5994	0.88–1.21	1.03	0.0	10,884
20–90	0.86–1.16	1.0	0.0	55,199	0.85–1.14	0.98	0.0	57,110	0.86–1.15	0.99	0.0	112,309
Thrombin time, s												
20–30	14.7–17.8	16.3	1.0	2746	14.6–17.2	15.9	1.0	4239	14.6–17.5	16.0	1.0	6985
30–40	14.7–17.8	16.2	1.0	2918	14.5–17.3	15.9	0.0	5101	14.6–17.6	16.0	0.0	8019
40–50	14.8–17.8	16.3	1.0	4000	14.7–17.2	16.0	0.89	4980	14.7–17.5	16.1	0.89	8980
50–60	14.7–17.8	16.2	1.0	6598	14.9–17.6	16.2	0.62	6282	14.8–17.7	16.2	1.0	12,880
60–70	14.7–17.9	16.3	0.1	6492	14.9–17.8	16.4	1.0	5653	14.8–17.9	16.3	0.79	12,145
70–80	14.7–18.2	16.3	0.0	6743	14.9–18.1	16.5	0.0	6092	14.8–18.1	16.4	0.0	12,835
80–90	14.6–18.3*	16.4	0.0	2942	14.8–18.1*	16.5	1.0	3508	14.7–18.4	16.5	0.0	6450
20–90	14.7–17.9	16.3	0.83	31,932	14.7–17.8	16.2	0.0	35,346	14.7–17.8	16.2	0.33	67,278
Fibrinogen*, mg/dl												
20–30	166–326	246	1	414	195–473	304	0	1690	187–472	297	0	2104
30–40	185–407	275	0	471	176–566	316	0	2325	176–555	313	0	2796
40–50	185–411	298	1	873	198–518	321	0	1362	197–508	316	0	2235
50–60	180–632	337	0	1782	226–521	344	04	1465	200–571	340	04	3247
60–70	193–636	351	0	1877	202–678	378	0.12	1344	197–684	367	0	3221
70–80	186–706	363	0	1823	204–565	351	0.25	1398	202–636	358	0	3221
80–90	185–784	382	01	625	189–730	372	0	626	184–751	373	01	1251
20–90	184–602	332	0	7767	193–579	334	0	10,064	189–590	334	0	17,831
Prothrombin time, s												
20–30	11.8–14.8	13.3	1.0	6178	11.7–14.6	13.1	1.0	7784	11.8–14.7	13.2	1.0	13,962
30–40	11.7–14.4	13.0	0.3	5671	11.6–14.3	13.0	1.0	7946	11.6–14.3	13.0	1.0	13,617
40–50	11.7–14.2	12.9	0.0	7286	11.6–14.3	12.9	0.0	7818	11.6–14.2	12.9	0.0	15,104
50–60	11.6–14.4	12.9	0.0	11,138	11.5–14.1	12.7	0.0	9871	11.5–14.3	12.8	0.0	21,009
60–70	11.8–14.5	13.0	0.0	10,492	11.5–14.1	12.7	0.0	8968	11.5–14.4	12.9	0.0	19,460
70–80	11.8–14.9	13.3	0.0	10,907	11.7–14.3	12.9	0.0	9923	11.7–14.6	13.1	0.0	20,830
80–90	11.9–15.2	13.5	0.72	4888	11.8–15.0	13.3	0.0	5981	11.8–15.1	13.4	0.32	10,869
20–90	11.7–14.6	13.1	0.0	55,099	11.6–14.4	12.9	0.0	56,990	11.6–14.6	13.0	0.0	112,089

50th denotes the 50th percentile/median, λ denotes the Box-Cox transformation parameter ("skewness"). Reference intervals marked with an asterisk (fibrinogen, and thrombin time in individuals 80–90 years) are established using < 4000 samples and should therefore be considered cautiously. The proportion of test results classified as pathological is available in Table 3.

intervals for aPTT, PT, INR, TT, and fibrinogen are shown in Table 1 and Fig. 1. aPTT and thrombin time reference intervals exhibit only minor variation with sex and age. INR and prothrombin time reference intervals show marginally lower upper reference intervals in women than in men and a slight increase in upper reference intervals in the age group 80–90 years. Fibrinogen reference intervals show an increasing trend in upper reference intervals with age, however, the number of samples available is reduced in comparison to the other analytes. Comparison of reference intervals generated using the complete input dataset with a filtered dataset, in which test results from patients with multiple samples per calendar year were removed, shows a trend to narrower reference intervals in the filtered population, although the differences are minor (Supplemental Table 2).

Reference intervals for different aPTT batches using Cephascreen and STA Neoplastine reagents are shown in Table 2. Upper aPTT reference intervals differ considerably between STA Neoplastine and Cephascreen reagents, specifically, the difference between sex-specific Neoplastine and Cephascreen upper aPTT reference intervals exceeds the "Permissible relative deviation of a single result" of 10.5% according to German Guidelines ("Rili-BAEK") [18,19] in all batches. In contrast, sex- and batch-specific upper Cephascreen aPTT reference interval differences are within the specifications mentioned above in 29/30 cases. (Upper female reference intervals for batch #1 [31.9 s] exceed the allowed deviation to batch #3 [35.3 s])

A comparison of the magnitude of the changes with age and sex on the one hand (upper aPTT reference interval range 32.7–35.2 s), and the batch-specific variation in Cephascreen aPTT reference intervals on the other hand (upper aPTT reference interval range 33.4–36.1 s, Table 2), shows that the effects of age and gender are minor in comparison to batch effects.

The sensitivity of the calculated aPTT reference intervals to detect deficiencies in factor VIII, IX, XI, and XII is displayed in Fig. 2. These

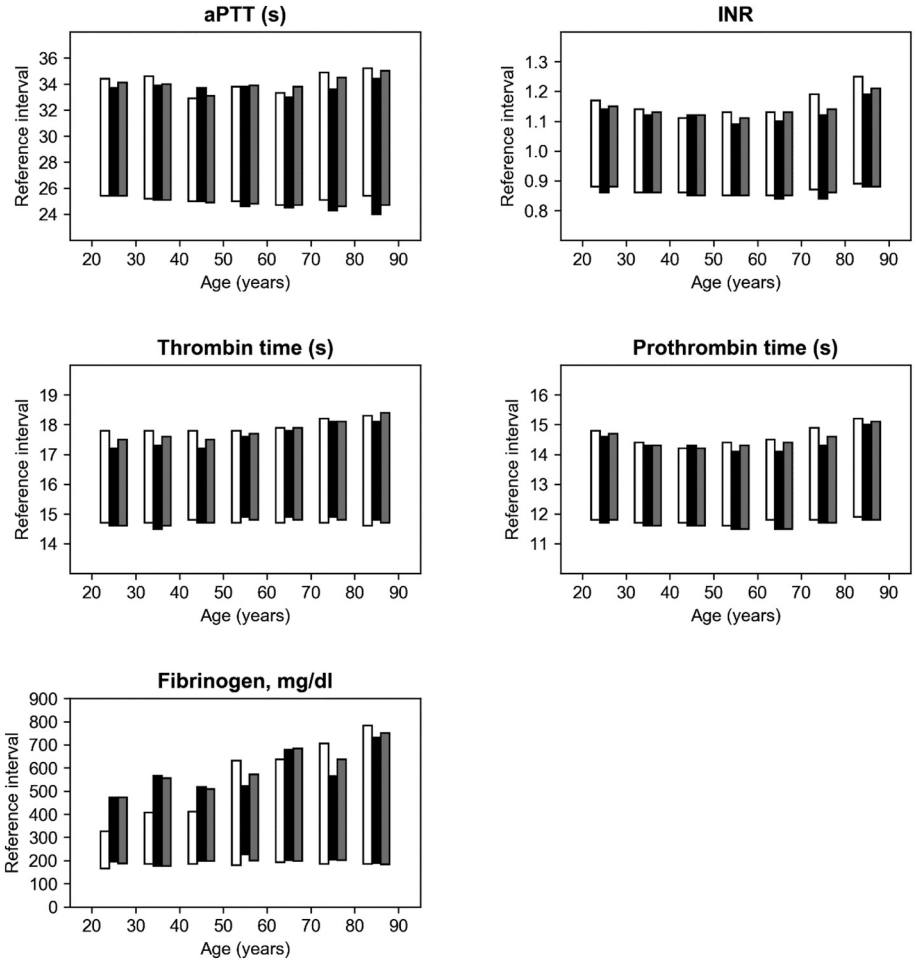

Fig. 1. Age- and sex-specific reference intervals for activated partial thromboplastin time (aPTT), international normalized ratio (INR), thrombin time, prothrombin time, and fibrinogen. Reference intervals for men (white boxes), women (black boxes), and men and women combined (gray boxes) for specific age groups.

results show a high sensitivity of the established reference intervals, specifically, 100% sensitivity to detect deficiencies < 30% in factors VIII, IX, and XI when using age- and sex-specific reference intervals. For 35% factor levels, aPTT sensitivity is 100.0%, 98.8%, 100.0%, and 98.7% for factors VIII, IX, XI, and XII when using age- and sex-specific reference intervals, while sensitivity is 99.1%, 97.6%, 100.0%, and 97.4% when using combined reference intervals for men and women across all age groups. While these results show the highest sensitivity to detect factor deficiencies when using age- and sex-specific aPTT reference limits (in contrast to a combined upper aPTT reference limit for all age groups), the used methodology does not allow for significance testing between reference intervals' sensitivities. Regarding the coagulation cascade, sensitivity is highest for factor XI, followed by factors VIII and IX, while sensitivity to detect factor XII deficiencies < 40% is comparatively low. The drop in factor VIII sensitivity marked with an asterisk is due to a factor VIII activity measurement of 3.0% with an aPTT of 34.3 s in a 63-year-old male, which is only identified as abnormal using the established age- and sex-specific reference intervals. However, this finding should not be overrated, as sex-specific male reference intervals and overall reference limits are only marginally wider than the age- and sex-specific reference interval (25.1–34.3 s and 24.9–34.4 s vs. 24.7–33.3 s).

Table 2
Reference intervals (RI) for activated partial thromboplastin time (aPTT) for different Cephascreen reagent batches and a STA Neoplastine reagent.

Batch	Men		Women		Men & Women	
	RI	n	RI	n	RI	n
STA Neoplastine	28.3–41.0	11,844	27.8–40.1	11,708	28.0–40.6	23,552
Cephascreen batch #1	25.1–35.0	2411	24.4–31.9	2041	24.6–35.1	4452
Cephascreen batch #2	26.1–34.9	2282	25.2–33.4	2166	25.4–33.9	4448
Cephascreen batch #3	25.6–36.1	6026	25.3–35.3	6094	25.4–35.7	12,120
Cephascreen batch #4	25.0–33.9	11,439	24.6–33.7	11,086	24.7–33.9	22,525
Cephascreen batch #5	25.6–34.2	7722	24.7–33.5	7614	25.0–33.6	15,336
Cephascreen batch #6	25.3–34.0	10,275	25.2–34.5	10,139	25.1–33.9	20,414
Cephascreen batch #7	25.1–34.1	6682	24.7–34.7	6537	24.8–34.5	13,219
Cephascreen batch #8	25.0–36.0	1908	24.5–33.4	1711	24.8–35.1	3619
Cephascreen batch #9	25.4–34.4	8265	24.9–33.9	8099	25.1–34.1	16,364
Cephascreen batch #10	24.9–33.6	12,632	24.3–33.4	12,490	24.6–33.7	25,122

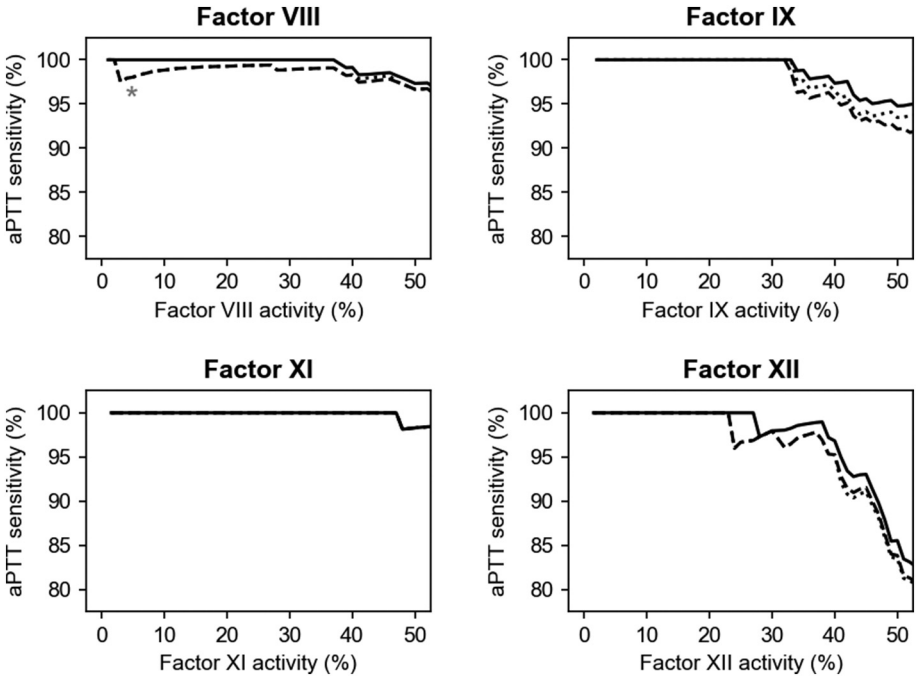

Fig. 2. Sensitivity of activated partial thromboplastin time (aPTT) to detect deficiencies in factors VIII, IX, XI, and XII. Sensitivity is defined as the ratio of aPTT measurements above the reference interval (i.e. pathologically high) to all aPTT measurements as a function of factor VIII, IX, XI, or XII activity measured in the same sample (n = 5725, 3857, 3711, and 2727). Solid lines show sensitivity when age- and sex-specfic aPTT reference ranges are used, dashed lines show sensitivity when sex-specific (i.e. not age-specific) aPTT reference ranges are used, and dotted lines show sensitivity when a single overall (i.e. neither sex- nor age-specific) aPTT reference range is used (Table 1).
* indicates a drop in factor VIII sensitivity due to a factor VIII activity measurement of 3.0% with an aPTT of 34.3 s in a 63-year-old male, which is only flagged as pathologically high using the established age-specific reference intervals.

Table 3 shows the influence of test result classification using age- and sex-specific reference intervals, sex-specific reference intervals, or a single overall reference interval. The differences between the different classification approaches are minor in most age groups and analytes, however, use of age- and sex-specific reference intervals tends to increase the proportion of test results classified as abnormal in patients aged 20–70 years, while decreasing the proportion of pathological test results in patients > 70–80 years. This analysis includes all test results from all patients, even if individual patients had multiple samples taken (unlike the reference intervals, which were established using one sample per patient). In comparison to conventional reference interval approaches (i.e. one sample per patient), this results in an over-representation of pathological test results (e.g. > 40% pathologically high test results for aPTT and INR in patients > 50 years) and an underrepresentation of other results (e.g. ≤ 1% of prothrombin test results and INR test results are abnormally low).

4. Discussion

Interpretation of coagulation test results requires availability of appropriate reference intervals for diagnostic and therapeutic decision making. We established reference intervals for coagulation screening tests in adult patients with an indirect approach ("Reference Limit Estimator", RLE). The dataset underlying the created reference intervals is composed of samples from in- and outpatients in a tertiary care university hospital. This allowed analysis of a large number of patients (n = 17,830-112,309, depending on the analyte) under real-world conditions and the analysis of changes with age. In contrast to conventional approaches to establish reference intervals, no selection and recruitment of a healthy reference population had to be performed, which is a challenging and expensive task, often resulting in limitations regarding the number of available samples and the ability to stratify for age and gender. Additionally, the variability of coagulation test results is larger than that of biochemical analytes due to intra-individual and analytical factors. Importantly, direct approaches also lead to substantial differences between the reference population and the patients to which the reference intervals are ultimately applied – both in- and outpatients are in most cases older, on medication, and have medical conditions not existing in the reference population. The reference intervals reported here account for these factors, making them possibly more appropriate for this complex mixed population of in- and outpatients to identify patients with relevant coagulation disorders than classical "health-associated" reference intervals. Conversely, the established reference intervals are probably not appropriate for a more selective setting, e.g. outpatients treated exclusively in private practice.

Use of data mining methods to establish reference intervals is controversial, as the analyzed sample set contains an unknown yet substantial proportion of pathological test results [21,22]. Establishment of reference intervals for coagulation tests using hospital patient data has to be performed with particular caution, as a substantial proportion of in- and out-patients are receiving medication (i.e. anticoagulants) or undergoing biological processes (i.e. acute-phase reactions and bleeding) that influence the coagulation cascade. However, we have shown in previous studies that the employed algorithm is able to correctly identify the distribution of physiological test results even in challenging settings. Einwallner et al. have used the same statistical approach to establish reference intervals for coagulation screening tests in pediatric patients [16]. The relative stability of reference intervals with age –despite an assumed increase in the proportion of pathological test results due to age-related diseases and e.g. anticoagulant medication– is a further indicator of the method's stability (Fig. 1). Additionally, evaluation of a "filtered" dataset, in which we removed all test results from patients with multiple samples per calendar year (i.e. those patients where repeat testing was considered necessary from a clinical perspective), shows only marginally narrower reference intervals (Supplemental Table 2). The applicability of the employed data mining approach is also highlighted in Fig. 2, which shows a high sensitivity of the established reference intervals to detect deficiencies in coagulation factors VIII, IX, XI, and XII. Importantly, the sensitivity of the established aPTT reference intervals is > 95% for the clinically highly relevant factor deficiencies < 30% in factors VIII, IX and XI, while sensitivity is considerable lower for factor XII, which is not associated with a bleeding phenotype. The established reference intervals for aPTT, PT, INR, TT show only minor variation with age, except for patients aged 80–90 years (Table 1 and Fig. 1). Importantly, the effects of age and sex on reference intervals are less pronounced than the inevitable effects of inter-batch variation (Tables 1 and 2). Similarly, the change in test results classified as abnormal when using age- and sex-specific reference intervals in comparison to a single overall reference interval is neglectable in most age groups (Table 3). From a clinical

Table 3
Proportion of abnormal test results depending on the specifity of the used reference intervals.

Age	n	Age- and sex-specific RIs		Sex-specific RIs		Overall RI	
		%low	%high	%low	%high	%low	%high
Activated partial thromboplastin time (aPTT), s							
20–30	25,935	3.2	24.9	2.2	24.0	2.1	24.1
30–40	26,847	2.7	25.2	2.7	23.8	2.6	23.9
40–50	42,243	2.6	39.6	2.5	35.2	2.4	35.4
50–60	77,916	2.3	43.4	2.7	41.0	2.6	41.3
60–70	95,136	1.7	53.1	2.1	49.6	2.1	49.9
70–80	101,184	1.5	54.5	1.7	53.2	1.6	53.4
80–90	52,642	0.9	53.3	1.2	56.7	1.2	56.8
International normalized ratio (INR), %							
20–30	30,762	1.0	20.9	0.4	22.8	0.5	23.5
30–40	31,859	0.8	25.5	0.6	22.4	0.8	22.9
40–50	50,523	0.6	35.2	0.6	30.3	0.8	31.5
50–60	92,100	0.5	40.8	0.8	35.5	1.0	37.1
60–70	109,991	0.3	47.3	0.4	42.1	0.6	44.1
70–80	120,207	0.3	49.4	0.1	49.8	0.2	51.3
80–90	61,447	0.3	44.0	0.1	58.0	0.1	59.3
Thrombin time, s							
20–30	12,567	4.3	11.8	5.1	9.3	5.1	9.3
30–40	14,983	4.0	12.1	4.2	10.4	4.2	10.4
40–50	21,287	4.3	21.0	3.9	18.6	3.9	18.6
50–60	37,327	4.1	23.8	3.3	22.7	3.3	22.7
60–70	43,386	4.1	28.9	3.0	29.9	3.0	29.9
70–80	47,133	2.6	30.5	2.1	33.6	2.1	33.6
80–90	23,861	2.5	31.6	2.0	34.8	2.0	34.8
Fibrinogen, mg/dl							
20–30	3720	5.7	19.9	6.2	7.1	6.5	6.8
30–40	4732	6.8	10.4	6.9	6.5	7.0	6.0
40–50	5715	15.0	17.8	13.5	10.8	14.2	11.1
50–60	9940	12.8	13.8	13.1	13.4	14.0	13.5
60–70	10,224	10.3	12.5	9.4	15.2	9.8	15.4
70–80	7863	6.0	16.0	5.9	16.6	6.1	16.8
80–90	2620	5.3	5.6	6.3	16.0	6.3	16.1
Prothrombin time, s							
20–30	30,655	1.0	21.5	0.6	23.7	0.5	23.7
30–40	31,751	0.8	24.8	0.8	22.9	0.7	22.9
40–50	50,368	0.9	35.3	0.9	31.2	0.7	31.6
50–60	91,800	0.8	40.9	1.0	36.7	0.9	37.2
60–70	109,602	0.7	47.6	0.6	43.3	0.5	44.0
70–80	119,870	0.3	47.0	0.2	50.9	0.2	51.3
80–90	61,244	0.2	44.5	0.1	59.2	0.1	59.3

Note that this analysis includes all test results from all patients, even if patients had multiple samples, resulting in an overrepresentation of pathological test results.

point-of-view, this finding is reassuring, as most laboratories do not provide age-specific reference intervals for coagulation tests in adults. Our findings support this practice.

Comparison of the established reference intervals to reference intervals reported by others using direct approaches and comparable analytical pipelines shows generally comparable results. INR reference ranges established by us are narrower than those reported (0.80–1.20) by Monagle et al. [23] in 22/24 age- and sex-groups, while the established prothrombin reference intervals closely match. The adult aPTT reference intervals reported by Summerhayes et al. (29–41 s) [24] exceed ours; despite the use of an identical assay (STAGO Cephascreen). On the one hand, these findings confirm the validity of our approach - despite the analysis of a mixed hospital/outpatient population; on the other hand, this underlines the need for intra-laboratory establishment of reference intervals, especially for aPTT. Upper reference intervals for fibrinogen from our study exceed those from Monagle et al. and others [5,25]. This is possibly due to an increased proportion of patients with acute-phase-reactions, which are a substantial proportion of patients in a tertiary care hospital. Comparison of our results to reference intervals established using the “filtered” dataset (only samples from patients with ≤1 test result per year) shows less pronounced change with age in the filtered dataset, confirming this hypothesis. To remove this confounding factor, fibrinogen samples could additionally be filtered based on the test results of other acute-phase proteins or inflammation markers (e.g. C-reactive protein, procalcitonin, or white cell count). However, the reduced number of samples available for fibrinogen did not allow us to further exclude samples based on other analytes' test results, the established upper fibrinogen reference intervals can therefore not be used to differentiate between patients with acute-phase reaction or inflammation and without. Nevertheless, our results show an increase in fibrinogen with age, which is in line with clinical experience and literature [25]. Importantly, lower reference limits for fibrinogen - which are important from a clinical point-of-view when evaluating bleeding tendencies - show only minor change with age and between unfiltered and filtered datasets.

4.1. Limitations

In our study, we analyzed patients' coagulation test result without associated medical information, i.e. no information regarding medical conditions, medications, recent surgery, etc. was available, and reference intervals were calculated using a purely statistical algorithm. We used this approach as accuracy and completeness of the mentioned information is unknown in most hospital information systems, and the statistical method employed has been validated in previous studies in populations with a similar proportion of pathological test results [14,20]. Additionally, efforts to make the mentioned information available in a standardized and readily useable manner are underway, e.g. the Germany medical informatics initiative [26], and will allow more precise selection of patients and stratification for in future studies.

We have established age-specific reference intervals and can show trends with age in coagulation tests, however, the used statistical approach does not allow testing for statistically significant differences between age- and sex-groups. Specifically, the calculation of confidence intervals is not possible with the current implementation of the used indirect algorithm – however, a revised algorithm implementation is currently being prepared and will allow calculation of confidence intervals in future studies. For the time being, the differences between age- and sex-groups have to be evaluated in the context of absolute differences between reference limits and clinical knowledge and experience.

Our results are the first reference intervals for coagulation screening tests established using the indirect method developed by Arzideh et al. While our results show the validity of the used approach in the examined patient setting, more widespread use of this method requires validation of our findings in other cohorts, especially before integration into clinical decision-making. Additionally, coagulation test results are subject to analyzer- and batch-effects, especially aPTT results; application of the reported reference intervals in other settings therefore requires transference of the published reference intervals according to guidelines [8]. On the other hand, the magnitude of changes in reference intervals with age is a physiological phenomenon rather than an effect of methodological or analytical specifics, these findings can therefore be generalized to other settings.

5. Conclusions

Data mining of laboratory test results enables the creation of reference intervals for coagulation tests that are specific to the examined population, age group, analytical framework, and reagent. This approach can complement conventional methods when establishing reference intervals and improve clinical decision-making based on coagulation tests. The age- and sex-specific reference intervals established in this study show only minor variation with sex and age, and support the current practice of providing a common reference interval for adult women and men.

Acknowledgements

Supported by the Interdisciplinary Center for Clinical Research (IZKF) at the University Hospital of the University of Erlangen-Nuremberg by a "Laboratory Rotation programme" for JZ. JZ received a Microsoft Azure for Research Grant to use Microsoft Azure computing resources, which was used for the calculation of reference intervals.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cca.2019.09.006.

References

[1] D.D. Castellone, Establishing reference intervals in the coagulation laboratory, Int. J. Lab. Hematol. 39 Suppl 1 (2017) 121–127, https://doi.org/10.1111/ijlh.12661.
[2] I. Mackie, P. Cooper, A. Lawrie, S. Kitchen, E. Gray, M. Laffan, British Committee for Standards in Haematology, Guidelines on the laboratory aspects of assays used in haemostasis and thrombosis, Int. J. Lab. Hematol. 35 (2013) 1–13, https://doi.org/10.1111/ijlh.12004.
[3] Y. Ozarda, K. Sikaris, T. Streichert, J. Macri, IFCC Committee on Reference intervals and Decision Limits (C-RIDL, Distinguishing reference intervals and clinical decision limits - a review by the IFCC committee on reference intervals and decision limits, Crit. Rev. Clin. Lab. Sci. 55 (2018) 420–431, https://doi.org/10.1080/10408363.2018.1482256.
[4] M. Milos, D.C. Herak, R. Zadro, Discrepancies between APTT results determined with different evaluation modes on automated coagulation analyzers, Int. J. Lab. Hematol. 32 (2010) 33–39, https://doi.org/10.1111/j.1751-553X.2008.01111.x.
[5] U. Nowak-Göttl, V. Limperger, G. Kenet, F. Degenhardt, R. Arlt, J. Domschikowski, H. Clausnizer, J. Liebsch, R. Junker, D. Steppat, Developmental hemostasis: a lifespan from neonates and pregnancy to the young and elderly adult in a European white population, Blood Cells Mol. Dis. (2016), https://doi.org/10.1016/j.bcmd.2016.11.012.
[6] G. Siest, J. Henny, R. Gräsbeck, P. Wilding, C. Petitclerc, J.M. Queraltó, P. Hyltoft Petersen, The theory of reference values: an unfinished symphony, Clin. Chem. Lab. Med. 51 (2013) 47–64, https://doi.org/10.1515/cclm-2012-0682.
[7] J. Henny, A. Vassault, G. Boursier, I. Vukasovic, P. Mesko Brguljan, M. Lohmander, I. Ghita, F.A.B. Andreu, C. Kroupis, L. Sprongl, M.H.M. Thelen, F.J.L.A. Vanstapel, T. Vodnik, W. Huisman, M. Vaubourdolle, Working Group Accreditation and ISO/CEN standards (WG-A/ISO) of the EFLM, Recommendation for the review of biological reference intervals in medical laboratories, Clin. Chem. Lab. Med. 54 (2016) 1893–1900, https://doi.org/10.1515/cclm-2016-0793.
[8] CLSI, Defining, Establishing, and Verifying Reference Intervals in the Clinical Laboratory; Approved Guideline - Third Edition, Clinical and Laboratory Standards Institute, Wayne, PA, 2008.
[9] G.R.D. Jones, R. Haeckel, T.P. Loh, K. Sikaris, T. Streichert, A. Katayev, J.H. Barth, Y. Ozarda, Indirect methods for reference interval determination – review and recommendations, Clin. Chem. Lab. Med. CCLM. 0 (2018), https://doi.org/10.1515/cclm-2018-0073.
[10] R.C. Friedberg, R. Souers, E.A. Wagar, A.K. Stankovic, P.N. Valenstein, College of American Pathologists, The origin of reference intervals, Arch. Pathol. Lab. Med. 131 (2007) 348–357, https://doi.org/10.1043/1543-2165(2007)131[348:TOORI]2.0.CO;2.
[11] G.L. Horowitz, The power of asterisks, Clin. Chem. 61 (2015) 1009–1011, https://doi.org/10.1373/clinchem.2015.243048.
[12] R. Haeckel, W. Wosniok, F. Arzideh, J. Zierk, E. Gurr, T. Streichert, Critical comments to a recent EFLM recommendation for the review of reference intervals, Clin. Chem. Lab. Med. 0 (2017), https://doi.org/10.1515/cclm-2016-1112.
[13] J. Zierk, F. Arzideh, T. Rechenauer, R. Haeckel, W. Rascher, M. Metzler, M. Rauh, Age- and sex-specific dynamics in 22 hematologic and biochemical analytes from birth to adolescence, Clin. Chem. 61 (2015) 964–973, https://doi.org/10.1373/clinchem.2015.239731.
[14] F. Arzideh, W. Wosniok, R. Haeckel, Indirect reference intervals of plasma and serum thyrotropin (TSH) concentrations from intra-laboratory data bases from several German and Italian medical centres, Clin. Chem. Lab. Med. 49 (2011) 659–664, https://doi.org/10.1515/CCLM.2011.114.
[15] J. Zierk, F. Arzideh, R. Haeckel, W. Rascher, M. Rauh, M. Metzler, Indirect determination of pediatric blood count reference intervals, Clin. Chem. Lab. Med. 51 (2013) 863–872, https://doi.org/10.1515/cclm-2012-0684.
[16] C. Weidhofer, E. Meyer, R. Ristl, H. Wiedemann, J. Cadamuro, U. Kipman, J. Zierk, C. Male, P. Quehenberger, E. Haschke-Becher, E. Einwallner, Dynamic reference intervals for coagulation parameters from infancy to adolescence, Clin. Chim. Acta Int. J. Clin. Chem. 482 (2018) 124–135, https://doi.org/10.1016/j.cca.2018.04.003.
[17] A. Clauss, Rapid physiological coagulation method in determination of fibrinogen, Acta Haematol. 17 (1957) 237–246, https://doi.org/10.1159/000205234.
[18] Neufassung der, Richtlinie der Bundesärztekammer zur Qualitätssicherung laboratoriumsmedizinischer Untersuchungen – Rili-BÄK – Richtlinie der Bundesärztekammer zur Qualitätssicherung laboratoriumsmedizinischer Untersuchungen, Dtsch Arztebl Int 111 (2014) (A-1583).
[19] Revision of the "Guideline of the German Medical Association on Quality Assurance in Medical Laboratory Examinations – Rili-BAEK" (unauthorized translation), LaboratoriumsMedizin. 39 (2015) 26, https://doi.org/10.1515/labmed-2014-0046.
[20] J. Zierk, F. Arzideh, R. Haeckel, M. Rauh, M. Metzler, T. Ganslandt, S.W. Krause, Indirect determination of hematology reference intervals in adult patients on Beckman Coulter UniCell DxH 800 and Abbott CELL-DYN Sapphire devices, Clin. Chem. Lab. Med. (2018), https://doi.org/10.1515/cclm-2018-0771.
[21] Y. Zhang, W. Ma, G. Wang, Y. Lv, Y. Peng, X. Peng, Limitations of the Hoffmann method for establishing reference intervals using clinical laboratory data, Clin. Biochem. 63 (2019) 79–84, https://doi.org/10.1016/j.clinbiochem.2018.11.005.
[22] J.L.V. Shaw, A. Cohen, D. Konforte, T. Binesh-Marvasti, D.A. Colantonio, K. Adeli, Validity of establishing pediatric reference intervals based on hospital patient data: a comparison of the modified Hoffmann approach to CALIPER reference intervals obtained in healthy children, Clin. Biochem.). doi:https://doi.org/10.1016/j.clinbiochem.2013.11.008.
[23] P. Monagle, C. Barnes, V. Ignjatovic, J. Furmedge, F. Newall, A. Chan, L.D. Rosa, S. Hamilton, P. Ragg, S. Robinson, A. Auldist, C. Crock, N. Roy, S. Rowlands, Developmental haemostasis, Thromb. Haemost. 95 (2006) 362–372, https://doi.org/10.1160/TH05-01-0047.
[24] R. Summerhayes, V. Ignjatovic, M. Hall, P. Monagle, Age-related reference ranges for neoplastine R (PT) and cephascreen (APTT) in healthy children, J. Thromb. Haemost. 5 (2007), https://doi.org/10.1111/j.1538-7836.2007.tb00018.x P-M-010-P-M-327.
[25] K. Adeli, J.E. Raizman, Y. Chen, V. Higgins, M. Nieuwesteeg, M. Abdelhaleem, S.L. Wong, D. Blais, 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, Clin. Chem. 61 (2015) 1075–1086, https://doi.org/10.1373/clinchem.2015.240531.
[26] S.C. Semler, F. Wissing, R. Heyder, German medical informatics initiative, Methods Inf. Med. 57 (2018) e50–e56, https://doi.org/10.3414/ME18-03-0003.