

J. Clin. Chem. Clin. Biochem.  
Vol. 23, 1985, pp. 157–166

## Determination of Reference Intervals for 10 Serum Proteins Measured by Rate Nephelometry, Taking into Consideration Different Sample Groups and Different Distribution Functions

By *W. Behr, G. Schlimok*

*Institut für Laboratoriumsmedizin des Zentralklinikums Augsburg*

*V. Firchau and H. A. Paul*

*Lehrstuhl für Statistik der Universität Augsburg*

(Received June 13/October 10, 1984)

**Summary:** Reference intervals were established for 10 serum proteins (IgA, IgG, IgM, transferrin, haptoglobin, complement C3, complement C4,  $\alpha_1$ -acid-glycoprotein,  $\alpha_1$ -antitrypsin,  $\alpha_2$ -macroglobulin) measured by rate nephelometry. The reference individuals – 200 blood donors – were divided into 5 subgroups: men aged 19–39 and 40–60 years, women aged 19–39 and 40–60 years and women aged 19–48 years using oral contraceptives. Where possible, two or more subgroups were combined to give reference sample groups. Criteria for this procedure are given.

The reference limits of the sample groups were estimated by parametric methods. Assuming that for a specified serum protein the type of distribution is the same in each subgroup, the data were standardized with estimated group specific parameter values and combined into one big sample. This permitted an improved determination of the underlying type of distribution. As a possible form of distribution we also considered the normal distribution truncated on the left side at  $c \geq 0$ . In some cases, after determination of an optimal  $c$ , this unusual distribution fitted the data significantly better than the generally used normal or log-normal distribution.

*Bestimmung von Referenzbereichen für 10 kinetisch nephelometrisch gemessene Serumproteine unter Berücksichtigung verschiedener Stichprobengruppen und Verteilungsfunktionen*

**Zusammenfassung:** Referenzbereiche für 10 kinetisch nephelometrisch gemessene Serumproteine (IgA, IgG, IgM, Transferrin, Haptoglobin, Komplement C3, Komplement C4, saures  $\alpha_1$ -Glykoprotein,  $\alpha_1$ -Antitrypsin,  $\alpha_2$ -Makroglobulin) wurden festgelegt. Die Referenzpersonen – 200 Blutspender – wurden in 5 Untergruppen eingeteilt: Männer im Alter von 19–39 und 40–60 Jahren, Frauen im Alter von 19–39 und 40–60 Jahren und Frauen im Alter von 19–48 Jahren mit Einnahme oraler Kontrazeptiva. Wenn möglich, wurden mehrere Untergruppen zu einer Referenzgruppe vereinigt. Kriterien für dieses Verfahren werden angegeben.

Die Referenzgrenzen der Gruppen wurden mit Hilfe parametrischer Methoden geschätzt. Unter der Annahme, daß der Verteilungstyp für ein Serumprotein in jeder Untergruppe gleich ist, wurden die Meßwerte mit Hilfe geschätzter gruppenspezifischer Parameter standardisiert und zu einer großen Stichprobe zusammengefaßt. Damit konnte der zugrundeliegende Verteilungstyp besser festgelegt werden. Als mögliche Verteilung wurde auch die auf der linken Seite bei  $c \geq 0$  abgeschnittene Normal-Verteilung in Betracht gezogen. Nachdem ein optimaler Wert  $c$  bestimmt war, paßte sich diese unübliche Verteilung in einigen Fällen den Meßdaten signifikant besser an als die allgemein gebräuchliche Normal- oder log-Normal-Verteilung.

## Introduction

With the introduction of rate nephelometry for the quantitation of serum proteins, discrepancies between values measured by radial immunodiffusion and rate nephelometry for immunoglobulins and some other proteins in the same specimen have been noted (1). Therefore reference intervals for serum proteins determined by rate nephelometry are necessary.

Only a few investigations have been reported on this subject (2–5). Some of the reported reference limits differ from each other, possibly as a result of the different statistical methods used, or the high variation of the analytical data attributable to the dependence of many serum proteins upon age, sex and use of oral contraceptives (6–9).

The purpose of our work is to establish reference intervals for 10 serum proteins (i. e. IgA, IgG, IgM, transferrin, haptoglobin, complement C3, complement C4,  $\alpha_1$ -acid-glycoprotein,  $\alpha_1$ -antitrypsin,  $\alpha_2$ -macroglobulin) measured by rate nephelometry. If, for one of the above reasons, significant differences exist in the concentration of a serum protein, the reference intervals should be estimated for separate subclasses, based on age, sex and use of oral contraceptives.

## Materials and Methods

### Reference individuals

The reference individuals were blood donors living in Augsburg and its near environs. Blood was collected on several occasions between October 1982 and February 1983, the collections were organized by the Red Cross. The blood donors were not in a fasting state, because the Red Cross fixtures always took place in the evening. The selection of the donors was based on a detailed questionnaire, a short examination by the physicians of Red Cross and a measurement of haemoglobin and body weight according to published recommendations (10).

### Sample groups

From the collective of blood donors 200 reference individuals were selected and divided into 5 subgroups (tab. 1). The criteria for this selection and distribution were age, sex and the use of oral contraceptives.

### Specimen collection

After venipuncture and donation of 500 ml, about 7 ml blood were allowed to drop into a plastic tube without the use of a tourniquet. During the specimen collection the reference individuals were in a supine posture. The blood samples were allowed to clot and stored overnight at 4°C. In the next morning the serum was separated after centrifugation at 2000 g for 10 min and stored at 4°C. The analyses were carried out within 48 h after specimen collection. Subsequently the sera were stored at -18°C.

### Analytical method

The concentrations of the 10 serum proteins were measured by rate nephelometry using an automated immunochemistry system (Auto-ICS, Beckman Instruments, Munich). The determinations were performed as specified by the manufacturer, using antisera, calibrators, buffer solutions and diluents provided by the manufacturer.

Accuracy and precision were together checked by a human blood-based control serum with known concentrations of the investigated serum proteins, predetermined by rate nephelometry (Electrophoresis Control, AHS/Deutschland GmbH, Munich, Lot No. EC-120). An accuracy control serum with assigned pathological concentrations was not commercially available. The control specimen for accuracy and precision was analysed at the start of every run and also after every ten donor specimens.

In order to exclude extreme values caused by analytical errors, the two lowest and two highest values of each subgroup were repeated after thawing the sera. Since the results of repeated tests in all cases agreed within the  $\pm 5\%$  range, all measured values could be accepted.

For the elimination of outliers we used the criterion described by Dixon (11). We only had to exclude 3 blood-donors, all because of very low  $\alpha_1$ -antitrypsin values, possibly caused by a heterozygote hereditary deficiency. Instead of these donors 3 other reference individuals of equivalent age and sex were selected.

### Statistical methods

#### Combining of subgroups to form reference sample groups

Estimation of reference limits becomes increasingly imprecise as the sample size decreases. Therefore we tried to combine subgroups in order to form reference sample groups. This is only permissible, if there are no significant differences either in location or in dispersion of the data of the subgroups. A different combination may be appropriate for each serum protein.

Tab. 1. Classification of the reference individuals into 5 subgroups.

	Subgroup				
	1	2	3	4	5
Number of reference individuals	40	40	40	40	40
Age (years)	19–39	40–60	19–39	19–48	40–60
Average age	29	49	29	26	49
Sex	male	male	female	female	female
Use of oral contraceptives	–	–	no	yes	no

To decide which subgroups may be combined, the two samples *Kolmogorov-Smirnov* test (12) can be used; this is suitable because it is a distribution free test. The use of this test and criteria for the selection of suitable subgroups are shown in more detail in appendix A.

#### Determination of reference intervals

##### General considerations

As usual the reference limits were estimated as the 0.025 and 0.975 fractiles of the reference distribution accompanied by the 0.90 confidence interval (13). Fractiles may be estimated by parametric or by non parametric methods. The non parametric estimates are free of erroneous assumptions about the underlying distribution type, but have a greater variance (14) and require sample sizes of  $N \geq 120$  (15). Therefore we decided to use parametric methods, providing the level of significance of the goodness-of-fit test with the proposed distributions is reasonable high, e.g. greater than 0.25. Thus the selection of a suitable distribution is of great importance.

The confidence intervals of the fractiles  $f$  were calculated by the formula.

$$f \pm 2.81 \frac{S_x}{\sqrt{N}}$$

given in (13).

$S_x$  is the standard deviation of the  $N$  reference values.

##### Considered distributions

As a first assumption, clinical data are regarded to have normal or log-normal distributions. As is well known, log-normal data  $x_i$  may be transformed to normal distributed data  $y_i$  by the function  $y_i = \log(x_i - b)$ , in which  $b$  has to be suitably determined. Two different cases should be considered:  $b = 0$  and  $-b_0 < b < b_0$  with  $b_0 = \min\{x_i\}$ . The data for some serum proteins could not be fitted properly to one of the above mentioned distributions. The histograms indicated that the data follow a normal distribution on the right side, whereas they seem to be cut off on the left. The normal distribution truncated on the left side at  $c \geq 0$  has these properties (fig. 1). As this is an unusual distribution, the formulas of its density function, expectation value and variance are presented in appendix B. Thus four types of distribution as shown in table 2 were taken into consideration.

##### Selection of the best fitted distribution

The *Kolmogorov-Smirnov* test (16) as modified by *Lilliefors* (17) was applied separately to every reference sample group to compute the set of parameter values providing the best fit to the four tested distributions. For the distribution 1 and 3 the usual estimators  $\hat{\mu}$  and  $\hat{\sigma}$  were used. In order to determine  $c$  and  $b$  a small increment  $h$  was used to evaluate a discrete set of  $c$  and  $b$  in the interval  $0 \leq c < \min\{x_i\}$  or  $b < \min\{x_i\}$ , e.g.:

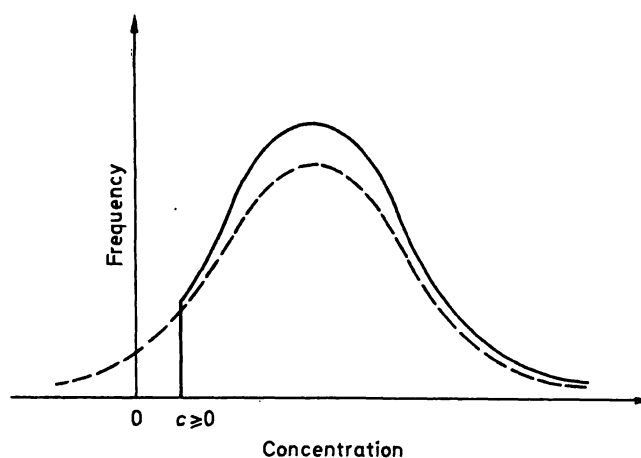
$$c_0 = 0; c_1 = h; \dots c_{i+1} = c_i + h.$$


Fig. 1. Normal distribution truncated on the left side at  $c \geq 0$  in comparison with the normal distribution (interrupted line).

For all these values of  $c$  and  $b$ , the corresponding estimators  $\hat{\mu}_c$ ,  $\hat{\sigma}_c$  (see appendix B) or  $\hat{\mu}_b$ ,  $\hat{\sigma}_b$  and the goodness-of-fit of the data to these distributions were calculated. The sets  $(\hat{c}, \hat{\mu}_c, \hat{\sigma}_c)$  and  $(\hat{b}, \hat{\mu}_b, \hat{\sigma}_b)$ , which provide the best fit, were regarded as an estimation of the unknown parameters for type 2 or type 4 distribution.

The same calculations were performed in every reference sample group of a serum protein.

It is assumed that for a specified serum protein the type of distribution is the same in every reference sample group, only differing in parameter values. The data for each group were therefore standardized with the above estimated specific parameter values, combined into a big sample, and a goodness-of-fit test over all the reference values of a serum protein was made. The distribution type with the highest level of significance  $\alpha_F$  ( $F = \text{fit}$ ) at the *Kolmogorov-Smirnov* test, including the *Lilliefors* correction, were regarded as underlying distribution. According to the *Lilliefors* correction an additional correction<sup>1)</sup>  $\bar{D}_2 = 1.06 \cdot D_2$ , or  $\bar{D}_4 = 1.075 \cdot D_4$  had to be taken into account for distribution type 2 or type 4, respectively.

Thus, having selected the presumed distribution for a specified serum protein, the reference limits and their confidence intervals were calculated separately in every reference sample group. All calculations were done on a Siemens 7007 at the University of Augsburg.

<sup>1)</sup> The factors 1.06 and 1.075 result from special investigations (18) made separately on the influence of the additional parameters  $a$  or  $b$ .

Tab. 2. Considered distributions.

Number	1	2	3	4
Type of distribution	Normal distribution	Truncated normal distribution	log-normal distribution ( $b = 0$ )	log-normal distribution $-b_0 < b < b_0$ with $b_0 = \min\{x_i\}$
Parameter	$\mu, \sigma$	$c, \mu_c, \sigma_c$	$\mu, \sigma$	$b, \mu_b, \sigma_b$

## Results

### Accuracy and precision

Table 3 shows the results of inaccuracy and between-day imprecision. The coefficients of variation (CV) of within-day imprecision are not presented. They were all below the CV of between-day imprecision. Analytical trends could not be observed by visual inspection of the control chart.

### Measured values

Table 4 presents the mean of the reference values, the standard deviation and the minimal and maximal reference value for each investigated serum protein and each subgroup.

### Combining of subgroups to form reference sample groups

The combining of subgroups to form reference sample groups showed congruent results only in some of the 10 serum proteins (IgA +  $\alpha_2$ -macroglobulin and C3 + transferrin). Many subgroups could not be combined with others. They were therefore handled as reference sample groups (tab. 5). Particularly the subgroups 4 and 3 – young women with and without use of oral contraceptives – have an exceptional position.

In order to demonstrate the effect of combining subgroups with reference sample groups, the histograms of the haptoglobin values and the levels of significance  $\alpha_F$  of the *Kolmogorov-Smirnov* test are shown in figure 2. This protein was selected, because its distribution has been discussed in detail in the literature (15, 19). The histograms of the subgroups generally did not allow reliable estimation of the underlying distribution type. Because of the small sample size, the sensitivity to the randomness of the sample is too great. After combining subgroups by means of a non parametric test, however, it was possible to display the distribution type graphically with greater certainty. This effect could be observed in all investigated serum proteins.

### Reference limits and their confidence intervals

After standardization and combination of the reference values for each serum protein, the levels of significance of the *Kolmogorov-Smirnov* test including the Lilliefors and the b-c-correction were found as shown in table 6. The underlined values show the distribution type selected for further calculation. As all levels of significance were greater than 0.25, all reference limits and their confidence intervals could be calculated using parametric methods. The results are given in table 7.

Tab. 3. Inaccuracy and between-day imprecision.

Constituent	Control serum			Measured values*) n = 12			
	Assigned value mg/l	Tolerance limits mg/l	Mean mg/l	Inaccuracy %	Imprecision CV %	$x_{\min}$ mg/l	$x_{\max}$ mg/l
IgA	2270	1910–2630	2350	3.52	4.59	2180	2510
IgG	10750	9100–12400	11370	5.77	4.21	10700	12000
IgM	1320	1120–1520	1338	1.36	4.47	1220	1410
Transferrin	3180	2540–3820	3283	3.24	3.08	3100	3470
Haptoglobin	1250	1000–1500	1162	7.04	7.13	1040	1290
C3	1300	1040–1560	1353	4.08	7.28	1200	1460
C4	280	230–330	261.7	6.54	5.51	236	290
$\alpha_1$ -Acid glycoprotein	600	470–730	606.5	1.08	4.77	569	660
$\alpha_1$ -Antitrypsin	1890	1510–2270	1868	1.16	5.19	1690	2050
$\alpha_2$ -Macroglobulin	1900	1510–2290	1843	3.00	5.53	1680	2010

CV = coefficient of variation

$x_{\min}$  = lowest value

$x_{\max}$  = highest value

\*) The data of accuracy and between-day precision were taken from the control samples positioned immediately after the calibrator samples.

Tab. 4. Mean value and standard deviation (s. d.) of the reference values

Constituent	Subgroup*)	Mean $\pm$ s. d. mg/l	$x_{\min}$ mg/l	$x_{\max}$ mg/l
IgA	1	2034 $\pm$ 689	940	3940
	2	2733 $\pm$ 818	990	3870
	3	1767 $\pm$ 605	740	3470
	4	1518 $\pm$ 525	570	3040
	5	2131 $\pm$ 689	680	3520
IgG	1	11020 $\pm$ 1680	8030	14500
	2	10890 $\pm$ 2390	7200	19000
	3	11510 $\pm$ 2110	7160	16100
	4	10560 $\pm$ 2060	5910	15300
	5	10890 $\pm$ 1710	6220	15700
IgM	1	1313 $\pm$ 486	610	2480
	2	1157 $\pm$ 501	410	2560
	3	1730 $\pm$ 579	850	3070
	4	1430 $\pm$ 563	600	2770
	5	1439 $\pm$ 600	550	3100
Transferrin	1	3287 $\pm$ 494	2430	4420
	2	3372 $\pm$ 458	2490	4460
	3	3398 $\pm$ 524	2350	4560
	4	3796 $\pm$ 599	2770	5180
	5	3423 $\pm$ 522	2530	4590
Haptoglobin	1	591 $\pm$ 312	100	1320
	2	712 $\pm$ 325	178	1370
	3	579 $\pm$ 253	85	1180
	4	703 $\pm$ 408	110	1790
	5	728 $\pm$ 275	205	1470
C3	1	1079 $\pm$ 200	750	1490
	2	1081 $\pm$ 254	680	1770
	3	1009 $\pm$ 199	760	1600
	4	1120 $\pm$ 186	710	1470
	5	1052 $\pm$ 234	750	1600
C4	1	221 $\pm$ 61	103	322
	2	232 $\pm$ 65	124	409
	3	183 $\pm$ 46	98	285
	4	173 $\pm$ 58	93	325
	5	201 $\pm$ 56	117	328
$\alpha_1$ -Acid glycoprotein	1	556 $\pm$ 140	302	881
	2	576 $\pm$ 153	388	1130
	3	467 $\pm$ 104	329	860
	4	386 $\pm$ 93	214	595
	5	533 $\pm$ 149	317	898
$\alpha_1$ -Antitrypsin	1	1527 $\pm$ 226	1090	2080
	2	1402 $\pm$ 197	1080	1780
	3	1640 $\pm$ 296	1160	2370
	4	2229 $\pm$ 382	1450	3010
	5	1644 $\pm$ 323	1070	2360
$\alpha_2$ -Macroglobulin	1	1353 $\pm$ 310	790	1950
	2	1130 $\pm$ 239	760	1950
	3	1417 $\pm$ 268	710	1960
	4	1570 $\pm$ 413	910	2430
	5	1304 $\pm$ 278	790	2150

\*) Cf. table 1.

Tab. 5. Combining of subgroups to form reference sample groups according to the criteria given in appendix A.

	IgA	IgG	IgM	Trans- ferrin	Hapto- globin	C3	C4	$\alpha_1$ -Acid glycoprotein	$\alpha_1$ -Anti- trypsin	$\alpha_2$ -Macro- globulin
Reference sample groups consisting of sets of 1-4 subgroups	1 + 5 2 3 4	1 + 2 + 4 + 5 3	1 + 4 + 5 2 3	1 + 2 + 3 + 5 4	1 + 3 2 + 4 5	1 + 2 + 3 + 5 4	1 + 2 3 4 5	1 + 2 + 5 3 4	1 2 3 + 5 4	1 + 5 2 3 4
$\alpha_G$	0.41	0.27	0.27	0.41	0.41	0.27	0.77	0.27	0.92	0.41

$\alpha_G$  (G = Group) is the minimum of the computed  $\alpha_G$  values of all subsets, which belong to the selected set of subgroups (see appendix A).

Tab. 6. Levels of significance  $\alpha_F$  of the Kolmogorov-Smirnov test after standardization and combination of the reference values.

Serum protein	Type of distribution			
	Normal distribution	Truncated normal distribution	log-normal distribution (b = 0)	log-normal distribution (-b <sub>0</sub> < b < b <sub>0</sub> )
IgA	0.01	<u>0.35</u>	0.02	0.02
IgG	0.41	<u>0.45</u>	0.07	0.11
IgM	0.00	<u>0.02</u>	0.31	<u>0.43</u>
Transferrin	0.08	<u>0.54</u>	0.28	<u>0.22</u>
Haptoglobin	0.04	<u>0.37</u>	0.01	0.01
C3	0.00	<u>0.00</u>	0.04	<u>0.31</u>
C4	0.00	0.00	0.83	<u>0.35</u>
$\alpha_1$ -Acid glycoprotein	0.00	0.02	<u>0.38</u>	0.32
$\alpha_1$ -Antitrypsin	0.02	0.03	<u>0.40</u>	<u>0.47</u>
$\alpha_2$ -Macroglobulin	0.01	0.04	0.50	<u>0.66</u>

The selected type of distribution is underlined.

Tab. 7. Reference limits and their 0.90 confidence intervals (in parentheses) of 10 serum proteins in mg/l (rounded values).

Serum protein	Subgroup				
	1 men 19-39 years	2 men 40-60 years	3 women 19-39 years	4 women 19-48 years with use of oral contraceptives	5 women 40-60 years
IgA	760 (540-980) 3450 (3230-3670)	1150 (740-1560) 4250 (3840-4660)	820 (490-1140) 3080 (2760-3400)	650 (380-930) 2630 (2360-2900)	like subgroup 1
IgG	7020 (6580-7460) 14750 (14310-15190)	like subgroup 1	7730 (6720-8840) 15840 (14830-16850)	like subgroup 1	like subgroup 1
IgM	620 (570-980) 2810 (2530-3120)	520 (470-580) 2880 (2200-3770)	840 (720-980) 3140 (2720-3630)	like subgroup 1	like subgroup 1
Transferrin	2480 (2360-2600) 4390 (4270-4510)	like subgroup 1	like subgroup 1	2620 (2360-2880) 4970 (4700-5240)	like subgroup 1
Haptoglobin	110 (10-210) 1180 (1080-1280)	100 (0-200) 1480 (1350-1610)	like subgroup 1	like subgroup 2	260 (120-400) 1300 (1160-1440)
C3	750 (730-770) 1700 (1580-1830)	like subgroup 1	like subgroup 1	780 (710-860) 1520 (1420-1630)	like subgroup 1
C4	120 (110-130) 380 (350-420)	like subgroup 1	110 (100-120) 290 (260-320)	90 (80-100) 310 (270-350)	110 (100-120) 330 (290-370)
$\alpha_1$ -Acid glycoprotein	320 (300-340) 890 (840-950)	like subgroup 1	310 (290-340) 670 (620-740)	230 (210-260) 600 (540-670)	like subgroup 1
$\alpha_1$ -Antitrypsin	1120 (1040-1210) 2000 (1890-2120)	1040 (970-1110) 1810 (1720-1910)	1170 (1120-1220) 2400 (2230-2580)	1580 (1470-1700) 3110 (2860-3380)	like subgroup 3
$\alpha_2$ -Macroglobulin	890 (850-930) 2100 (1910-2310)	780 (730-830) 1700 (1520-1900)	930 (840-1030) 2010 (1860-2170)	900 (800-1010) 2580 (2280-2920)	like subgroup 1

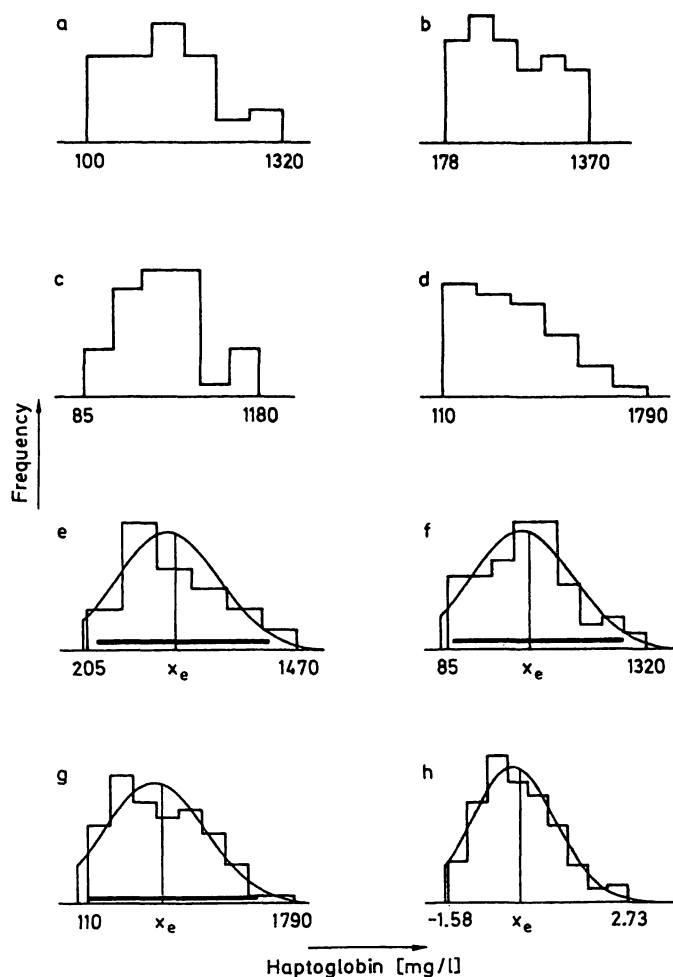


Fig. 2. Histograms of the data of haptoglobin and levels of significance  $\alpha_F$  of the *Kolmogorov-Smirnov* test. Only  $\alpha_F$  of the best fitting distribution, i. e. the truncated normal one, and in comparison the alternative log-normal distribution with  $-b_0 < b < b_0$  are presented.

- a) Subgroup 1 of the reference individuals: men, 19–39 years  
truncated normal: 0.83;  $\log_b$ -normal: 0.02
- b) Subgroup 2 of the reference individuals: men, 40–60 years  
truncated normal: 0.43;  $\log_b$ -normal: 0.28
- c) Subgroup 3 of the reference individuals: women, 19–39 years  
truncated normal: 0.67;  $\log_b$ -normal: 0.01
- d) Subgroup 4 of the reference individuals: women, 19–48 years, with use of oral contraceptives  
truncated normal: 0.30;  $\log_b$ -normal: 0.09
- e) Subgroup 5 of the reference individuals: women, 40–60 years  
truncated normal 0.01;  $\log_b$ -normal: 0.06
- f) Reference sample groups 1 + 3 (fig. 2a + 2c)  
truncated normal: 0.67;  $\log_b$ -normal: 0.01
- g) Reference sample groups 2 + 4 (fig. 2b + 2d)  
truncated normal: 0.29;  $\log_b$ -normal: 0.03
- h) After standardization and union of all reference values  
truncated normal: 0.37;  $\log_b$ -normal: 0.01

—: density function of the assumed distribution, i. e. the truncated normal distribution  
 —: calculated reference interval  
 $x_e$ : expected value

## Discussion

All the investigated serum proteins showed variations in concentrations associated with age, sex or use of oral contraceptives. In accordance with other investigators (2, 4, 7, 8) an increase with age was found for IgA and haptoglobin, whereas sex associated differences were seen for IgM, IgG,  $\alpha_1$ -antitrypsin (higher values for women) and IgA, C3, C4,  $\alpha_1$ -acid-glycoprotein (higher values for men), although not in all age groups. As the sex associated differences first of all are based on the exceptional position of subgroup 3 (women aged 19–39 years), they may be conditioned by the hormonal state of women capable of bearing children.

The subgroup 4, women using oral contraceptives, showed lower values in the concentrations of IgA, IgG, IgM and  $\alpha_1$ -acid-glycoprotein and higher values in the concentrations of transferrin, haptoglobin, C3,  $\alpha_1$ -antitrypsin and  $\alpha_2$ -macroglobulin compared with the corresponding female group not using hormonal contraceptives. These results correspond well with other investigations with the exception of IgG, IgM and haptoglobin, where no or different effects of oral contraceptives were reported (2, 4, 6, 9).

The standard deviations of the serum proteins given in table 4 show a great variety. To achieve just sufficient accuracy for the reference interval of each serum protein the sample sizes should be different in every subgroup e. g. proportional to the variance in every subgroup. In some cases the confidence interval of the lower reference limit is too large compared with the value of the lower reference limit, e. g. haptoglobin. In this case the sample size ought to be raised.

Neglect of the possible differences in reference sample groups caused by age, sex etc. leads for some serum proteins to a different choice of the underlying distribution type. Alternative calculations, the result of which is not presented here, showed that, with the exception of haptoglobin, the log-normal distribution would be preferred. The level of significance  $\alpha_F$  of the *Kolmogorov-Smirnov* test diminishes considerably, e. g. l. c. (15).

When estimating reference intervals by parametric methods in clinical chemistry one often takes into consideration only two types of distribution, i. e. the normal and the log-normal one (15). If the data do not fit the normal distribution because of positive skewness, the log-normal one will be used in spite of sometimes considerable deviations, for example see l. c. (2, 4). The unjustified assumption that the data are log-normal distributed, results in an upper ref-

erence limit that is obviously too high in comparison with the maximal reference value. If the normal and the log-normal distribution have low levels of significance in the goodness-of-fit test, then often the seldom used truncated normal distribution may be suitable for fitting the data; for example see IgA, transferrin or haptoglobin (table 6).

The reason why the distribution of some serum proteins is "cut off" on the left side may be explored for the case of haptoglobin (fig. 2). The histograms f, g and h show that the data follow fairly well the shape of a normal distribution on the right side of the modus. If haptoglobin were to be normal distributed, the values should extend on the left into the negative range. Clinical laboratory data, however, cannot be negative; even a zero or very low haptoglobin value is pathological and, for example, a sign of haemolysis. Therefore the distribution of reference values of a healthy population must be cut off on the left side at a value  $c > 0$ . The value  $c$  may be estimated by  $\hat{c}$ , which gives the best fit of the data in the *Kolmogorov-Smirnov* test.

Reed et al. (15) also found that haptoglobin follows neither a normal nor a log-normal distribution. Therefore they proposed a non parametric estimation of the reference interval. We checked the fit of the haptoglobin values presented by these authors to the truncated normal distribution.  $\alpha_F$  was only 0.05. A somewhat better significance level ( $\alpha_F = 0.22$ ) resulted from the fitting of our haptoglobin data, if the non-standardized, crude values were used. The distribution into significantly differing reference sample groups and the subsequent standardization and combining of the data are obvious reasons for the relatively good fit ( $\alpha_F = 0.37$ ) to the truncated normal distribution.

The estimation of the underlying distribution by means of the pooled and standardized reference values is based on the assumption that one type of distribution underlies all subgroups or reference sample groups of a special protein. This assumption can be hardly proved graphically, because of the small sample sizes, as shown earlier (fig. 2e—g). The comparison of the significance levels  $\alpha_F$  of the goodness-of-fit test for the considered distributions of the different subgroups and reference sample groups of a protein showed that, in most cases, the finally accepted distribution had the highest  $\alpha_F$ -values. As an example, these values for the subgroups of haptoglobin are presented in fig. 2. Taking into account that the truncated distribution and the log-normal distribution are essentially different only in the upper and lower range, it is understandable that

for some subgroups the alternative distribution showed a somewhat better fit than the usually prevailing type of distribution.

The square root transformation  $\sqrt{x_i + d}$  (19) applied to our data did not result in a better fit, compared to the levels of significance of the distributions selected according to table 6.

The existing goodness-of-fit tests require too many items to give reliable information of the fit in the lower and upper 2.5% of a distribution. With  $N = 100$  and a level of significance of 0.20, the difference between the theoretical distribution and the sample distribution may be nearly 10% in the lower and upper end. Thus other tests must be developed for that purpose.

#### Appendix A

To test the hypothesis that a set of  $k$  subgroups of the reference values of a serum protein is equally distributed, the two samples *Kolmogorov-Smirnov* test (12) is appropriate. The test has to be applied to all pairs of subgroups of a protein. It is a distribution-free test, which does not require the assumption of equality in variance or of normal distribution as other tests do.

Two criteria were applied to the selection of suitable sets:

1. The reference sample groups should consist of as many subgroups as possible.
2. Only sets with  $\alpha_G > 0.20$  were accepted.

The second criterion and the testing in pairs need further explanation:

The null hypothesis of the two samples *Kolmogorov-Smirnov* test is that both subgroups are equally distributed. As an erroneous combination of subgroups may cause greater damage than keeping them apart, a fairly high level of significance has to be chosen. The distributions of the reference sample groups of a protein are supposed to be of the same type, only differing in the parameters  $\mu_i$  and  $\sigma_i$ . This assumption is based on the fact that in most cases the goodness-of-fit test of *Kolmogorov-Smirnov* gives the best fit to the same type of distribution for a specified protein. A small difference  $|\mu_i - \mu_j|$  and  $|\sigma_i - \sigma_j|$  giving a  $\alpha_G > 0.2$  may be acceptable, because of the advantages of greater sample size after combining subgroups  $i$  and  $j$ . With increasing sample size  $N$ , the test is more sensitive to smaller differences between two subgroups for a given  $\alpha_G$ . With increasing  $N$ , however, the confidence intervals decrease, so that the combining of subgroups is not so important.



Testing together with a third subgroup  $k$ , e.g.  $\mu_i < \mu_j < \mu_k$  may result in a significance level  $\alpha_G > 0.2$ , though subgroups  $i$  and  $k$  are too different. Thus, testing in pairs and a fairly high error of the first kind,  $\alpha_G > \alpha_{\min} = 0.2$ , were chosen.

The following describes the procedure in more detail: The *Kolmogorov-Smirnov* test has to be made for all pairs of subgroups. The pair with the greatest level of significance  $\alpha_G$  forms a set. The pair with the next greatest  $\alpha_G$  may form a second set or, if one of its subgroups belongs to the existing set, all possible pairs of this new subgroup and the subgroups from the existing set have to be checked. If every  $\alpha_G$  is greater than  $\alpha_{\min}$  the new subgroup may enlarge the existing set. If no further enlargement is possible, the procedure is finished.

As an example the procedure is carried out for IgM ( $\alpha_{\min} = 0.2$ ):

$$\alpha_{1,2} = 0.10; \alpha_{1,3} = 0.08; \alpha_{1,4} = 0.27; \alpha_{1,5} = 0.77$$

$$\alpha_{2,3} = 0.00; \alpha_{2,4} = 0.06; \alpha_{2,5} = 0.03$$

$$\alpha_{3,4} = 0.03; \alpha_{3,5} = 0.08$$

$$\alpha_{4,5} = 0.92$$

The first set is formed by the subgroups 4 and 5; the next greatest  $\alpha_G$  is  $\alpha_{1,5} = 0.77$ . As 5 is already in a set with 4, the value  $\alpha_{1,4}$  has to be checked:

$$\alpha_{1,4} = 0.27 > 0.2.$$

As all values  $\alpha_G$  are greater than  $\alpha_{\min}$ , the combination (1 + 4 + 5) forms the enlarged set. In this example  $\alpha_G^*$  (see legend of table 5) is  $\alpha_{1,4} = 0.27$ .

The remaining  $\alpha_{i,j}$  are too small. Therefore the procedure is finished with three reference sample groups: (1 + 4 + 5), (2), (3).

## Appendix B

The normal distribution truncated on the left side at  $c$  is defined by the following function

$$F(x) = \frac{\Phi\left(\frac{x-\mu}{\sigma}\right) - \Phi\left(\frac{c-\mu}{\sigma}\right)}{1 - \Phi\left(\frac{c-\mu}{\sigma}\right)}$$

where  $\Phi$  is the standard normal distribution. Its mean and variance are derived from the corresponding values  $\mu$  and  $\sigma^2$  of the normal distribution:

$$E(x) = \mu + \sigma \cdot \theta$$

$$\text{Var}(x) = \sigma^2 \left(1 + \frac{c-\mu}{\sigma} \cdot \theta - \theta^2\right)$$

with

$$\theta = \frac{\varphi\left(\frac{c-\mu}{\sigma}\right)}{1 - \Phi\left(\frac{c-\mu}{\sigma}\right)}$$

and  $\varphi$  as density function of the normal distribution.

If  $c$  is known, the estimation of  $\mu$  and  $\sigma$  in  $F(x)$  can be made by an iterative algorithm, using  $\bar{x}$  and  $s^2$  of the sample

$$0) \quad \mu_0 = \bar{x}, \quad \sigma_0 = s = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2}$$

and repeated calculation of part k, until changes of  $\mu_k$  and  $\sigma_k$  are sufficiently small

$$k) \quad \theta_k = \frac{\varphi\left(\frac{c-\mu_{k-1}}{\sigma_{k-1}}\right)}{1 - \Phi\left(\frac{c-\mu_{k-1}}{\sigma_{k-1}}\right)}$$

$$\sigma_k = \frac{s}{\sqrt{1 + \frac{c-\mu_{k-1}}{\sigma_{k-1}} \theta_k - \theta_k^2}}$$

$$\mu_k = \bar{x} - \sigma_k \cdot \theta_k.$$

To find the best estimation  $\hat{c}$  of  $c$ , and consequently the corresponding set  $(\hat{c}, \hat{\mu}_c, \hat{\sigma}_c)$ , a discrete set of values  $c < x_{\min}$  was taken to calculate the set  $(c, \mu_c, \sigma_c)$ , using the above given estimator for each  $c$ . The set  $(c, \mu_c, \sigma_c)$ , which gives the best fit in the *Kolmogorov-Smirnov* test was taken as estimator  $(\hat{c}, \hat{\mu}_c, \hat{\sigma}_c)$ .

## References

- Liappis, N. (1980) *Klin. Pädiatr.* 192, 370–378.
- Liappis, N., Beeser, H., Fritsche, Ch. & Hildenbrand, G. (1981) *Lab. Med.* 5, 73–78.
- Joliff, C. R., Cost, K. M., Stivins, P. C., Grossman, P. P., Nolte, C. R., Franco, S. M., Fijan, K. J., Fletcher, L. L. & Shriner, H. C. (1982) *Clin. Chem.* 28, 126–128.
- Liappis, N., Beeser, H. & Fritsche, Ch. (1983) *Lab. Med.* 7, 66–70.
- Beckman Instruments, Information papers ADS-7, ICS 1-N, ICS 1-M.
- Laurell, C.-B., Kullander, S. & Thorell, J. (1967) *Scand. J. Clin. Lab. Invest.* 21, 337–343.
- Lyngbye, J. & Krøll, J. (1971) *Clin. Chem.* 17, 495–500.
- Weeke, B. & Krasilnikoff, P. A. (1972) *Acta Med. Scand.* 192, 149–155.
- Dotchev, D., Liappis, N. & Hungerland, H. (1973) *Clin. Chim. Acta* 44, 431–435.
- Richtlinien zur Blutgruppenbestimmung und Bluttransfusion, pp. 30–34, Deutscher Ärzte-Verlag, Köln (1980).
- Dixon, W. J. (1953) *Biometrics* 9, 74–89.

12. Büning, H. & Trenkler, G. (1978) Nichtparametrische statistische Methoden, pp. 133–138, de Gruyter, Berlin–New York.
13. Solberg, H. E. (1981) Statistical treatment of collected reference values and determination of reference limits. In: Reference Values in Laboratory Medicine (Gräsbeck, R. & Alström, T. eds.), pp. 193–205, John Wiley & Sons, Chichester.
14. Elveback, L. R. & Taylor, W. F. (1969) *Ann. N. Y. Acad. Sci.* 161, 538–548.
15. Reed, A. H., Henry, R. J. & Mason, W. B. (1971) *Clin. Chem.* 17, 275–284.
16. Miller, L. H. (1956) *J. Am. Stat. Assoc.* 51, 111–121.
17. Lilliefors, H. W. (1967) *J. Am. Stat. Assoc.* 62, 339–402.
18. Paul, H. A. (1984), working paper, Institut für Statistik und Mathematische Wirtschaftstheorie, Universität Augsburg.
19. Harris, E. K. & DeMets, D. L. (1972) *Clin. Chem.* 18, 605–612.

Dipl. Chem. Dr. med. Werner Behr  
Institut für Laboratoriumsmedizin  
Zentralklinikum Augsburg  
Stenglinstraße 2  
D-8900 Augsburg