Z Rechtsmed 85, 189–197 (1980)

Zeitschrift für
Rechtsmedizin

Distribution of Likelihood Ratio in Relation to the Exclusion Probability

Kayoko Minakata and Minoru Asano

Department of Legal Medicine, Hamamatsu University, School of Medicine, Hamamatsu, 431-31, Japan

Summary. The relation between the exclusion probability (E) and the paternity probability is derived by assuming the distributions of logarithm of paternity likelihood ratio, $\log(Y/X)$ for true fathers and unexcluded non-fathers as the normal distributions.

Under this assumption the value $\log(1-E)$ is equal to the mean of the mean value for true fathers (a) and that for unexcluded non-fathers (b), i.e., $\log(1-E)=(a+b)/2$. This relation holds quite well for the various actual distributions of $\log(Y/X)$ of father-child combinations and those of father-mother-child combinations using 14 blood group systems. Therefore, the derived relation is found to be a convenient way to deduce one of the three quantities (E, a, b) from the remaining two quantities in the actual distributions.

Key words: Paternity probability – Exclusion probability – Distribution of likelihood ratio

Zusammenfassung. Die Relation zwischen Ausschlußwahrscheinlichkeit (E) und der Vaterschaftswahrscheinlichkeit wird von der Annahme abgeleitet, daß die logarithmische Verteilung des Wahrscheinlichkeits-Verhältnisses, $\log(Y/X)$, für echte Väter und nicht-ausgeschlossene Nicht-Väter als eine Normalverteilung anzusehen ist.

Unter dieser Voraussetzung ist der Wert $\log(1-E)$ gleich dem Mittel aus den Mittelwerten für echte Väter (a) und nicht-ausgeschlossene Nicht-Väter (b): $\log(1-E)=(a+b)/2$. Die Gleichung läßt sich ebenso auf die verschiedenen Verteilungen des $\log(Y/X)$ von Vater-Kind- und von Vater-Mutter-Kind-Kombinationen bei Verwendung von 14 Blutgruppensystemen anwenden. Somit kann die Gleichung als eine gute Methode zur Ableitung einer der drei Größen (E, a, b) aus den jeweils verbleibenden zwei in der betreffenden Verteilung benutzt werden.

Schlüsselwörter: Vaterschaftswahrscheinlichkeit – Ausschlußwahrscheinlichkeit – Wahrscheinlichkeits-Verhältnis

0044/3433/80/0085/0189/$ 1.80

In the diagnosis of paternity, the exclusions with safe genetic markers can be accepted as conclusive evidences. Such exclusions, however, are not always expected for every non-fathers, giving rise to ambiguity in the diagnosis. To distinguish between true fathers and unexcluded non-fathers, it is necessary to calculate the paternity probability, $X/(X+Y)$ which is so called Essen-Möller value [1]. Here X is the frequency of the accused man's phenotype among true fathers for a given mother-child couple and Y is the frequency of the corresponding phenotype among non-fathers. Furthermore, the paternity probability gives the conclusive assignment if it is sufficiently high, as it is in most cases where extensive series of genetic tests are used [2].

There were several reports which compare the exclusion probability (E) with the paternity probability in each case of the diagnosis [3,4]. Salmon plotted the paternity probabilities vs. the exclusion probabilities in 48 cases of diagnosis of paternity [5]. Hummel et al. investigated several cases in HLA system where the exclusion probabilities were fairly high but the paternity probabilities were relatively low [6].

For the theoretical investigation of the distribution of frequencies for X and Y, Essen-Möller used first a binomial distribution which is an approximation of a normal distribution [1,7]. Hummel assumed the distribution of $\log(Y/X)$ to be a normal distribution in his theoretical studies on the classification of a putative man among possible fathers [3,8]. Recently, we reported several statistical studies on the diagnosis of the estimated paternity probability from blood group findings of his relatives by using a Monte Carlo method [9,10].

As expected reasonably, the frequencies X and Y vs. $\log(Y/X)$, expressed by histograms, are approximated to be normal curve if a number of blood group systems are examined and sufficiently large numbers of possible values of Y/X are plotted in the distribution. Therefore, the normal curve can be used with a sufficiently high accuracy to solve problems pertaining the distributions for X and Y.

In this paper, a general relation between the exclusion probability and the paternity probability is derived by assuming the normal curves for X and Y. The relation is verified by the actual distributions of three cases of children (duo), those of three cases of mother-child couples (trio).

Theoretical Calculation

The shape of the distribution for true fathers X and that for non-fathers Y are schematically illustrated in Fig. 1. The ordinate y stands for the relative frequency X (or Y). The abscissa x stands for the value of $\log(Y/X)$,

$$x = \log(Y/X) \tag{1}$$

If the distribution of the relative frequency of true fathers X is assumed to be a normal distribution, following equations can be written,

$$X = (2\pi s^2)^{-\frac{1}{2}} \exp[-(x-a)^2/2s^2] \tag{2}$$

$$\int_{-\infty}^{\infty} Y \mathrm{d}x = 1 - E \tag{3}$$

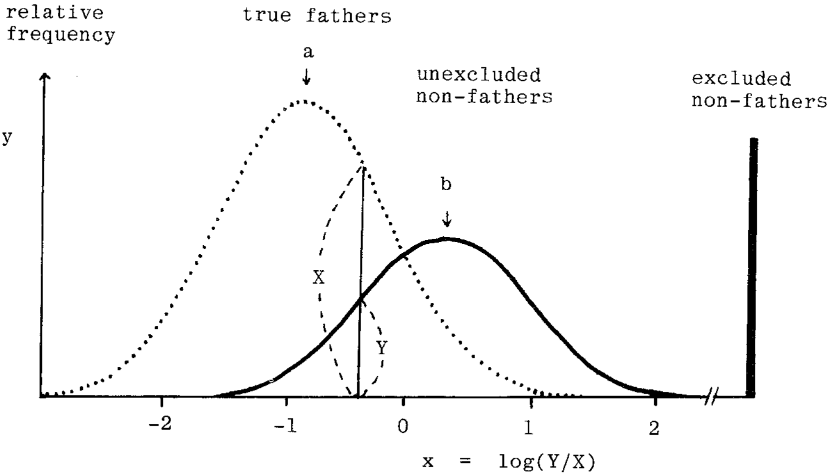

Fig. 1. Schematic distributions of relative frequencies of $\log(Y/X)$ for true fathers and non-fathers. Since the exclusion probability for non-fathers is assumed to be E, the integral of unexcluded non-fathers Y is $(1-E)$. a and b are mean values for true fathers and unexcluded non-fathers, respectively

where

a = mean value of $\log(Y/X)$ for the distribution of true fathers,
s = standard deviation of the distribution for true fathers,
E = exclusion probability of non-fathers.

Substituting equation (2) into (1), Y is expressed as

$$Y=10^{a+\frac{s^2}{2\log e}}(2\pi s^2)^{-\frac{1}{2}}\exp[-\{x-(a+\frac{s^2}{\log e})\}^2/2s^2] \tag{4}$$

Integrating Eq. (4) with respect to x and comparing with Eq. (3), we have

$$10^{a+\frac{s^2}{2\log e}}=1-E \tag{5}$$

Therefore, Eq. (4) is rewritten as

$$Y=(1-E)(2\pi s^2)^{-\frac{1}{2}}\exp[-(x-b)^2/2s^2] \tag{6}$$

where

$$b=a+\frac{s^2}{\log e} \tag{7}$$

Equation (6) means that the distribution of unexcluded non-fathers Y is also a normal distribution with mean value b and standard deviation s, which is the same as that of X. Also, it turns out from Eq. (7) that the difference between two peaks a and b depends only on the variances (s^2). Combining Eqs. (5) and (7), we have

$$(a+b)/2=\log(1-E) \tag{8}$$

Equation (8) expresses the relation between the mean of the two peaks (a and b) and the exclusion probability (E).

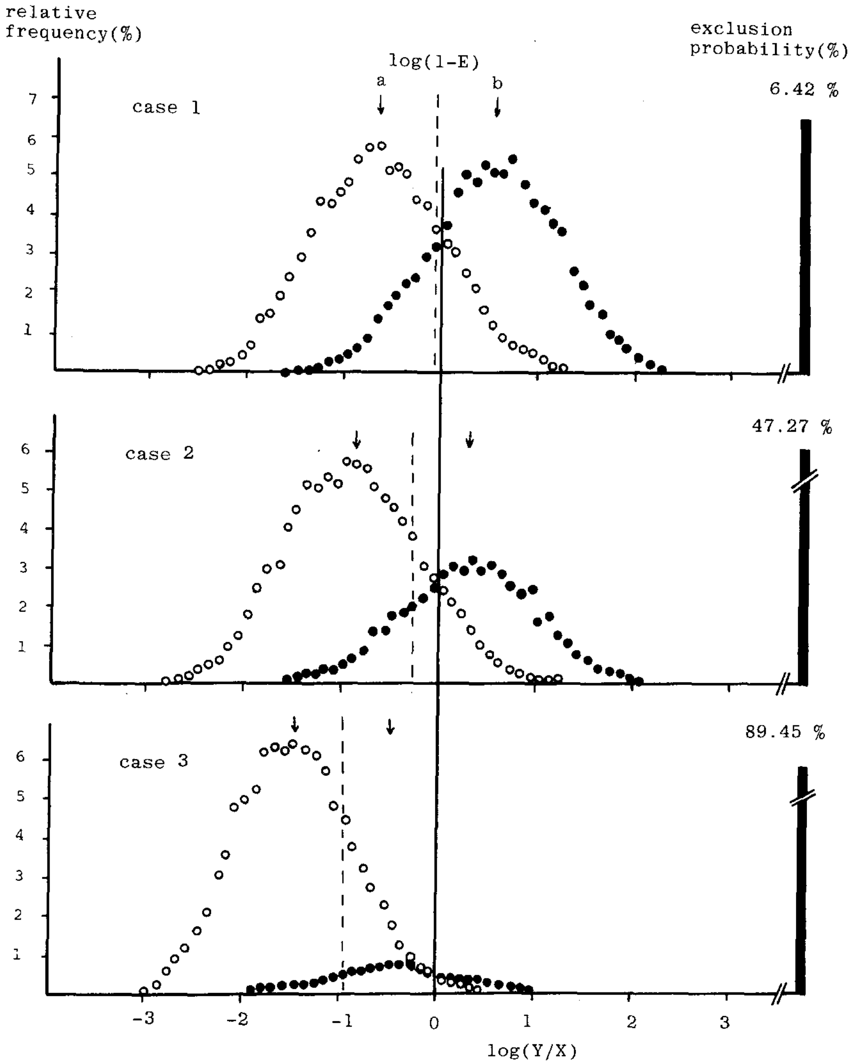

Fig. 2. Distributions of duo. The relative frequencies of log(Y/X) for 10^4 true fathers and 10^4 non-fathers concerning three kinds of children. The phenotypes of the children are as follows; *case 1:* A, MNs, CcDEe, P+, Fy(ab), Jk(ab), Gm(1, 16, 21), Inv(−1), Gc 2−1, Hp 2−1, PGM 2−1, AcP(AB), GPT 2−1, EsD 2−1; *case 2:* A, MNs, CcDEe, p+, Fy(ab), Jk(ab), Gm(1, 16, 21), Inv(−1), Gc 2−1, Hp 2−1, PGM 2−1, AcP(AB), GPT 2−1, EsD 2; *case 3:* A, MNs, CcDEe, P+, Fy(b), Jk(ab), Gm(1, 16, 21), Inv(−1), Gc 2−1, Hp 2−1, PGM 2−1, AcP(AB), GPT 2−1, EsD 2. The relative frequency of log(Y/X) for true fathers and unexcluded non-fathers are designated in open circles (○) and black circles (●), respectively. Those for excluded non-fathers are expressed in black sticks. The arrows indicate the mean positions for true fathers *a* and those for unexcluded non-fathers *b*. The values log($1-E$) are designated by chain lines. The class-interval of log(Y/X) is 0.1. The relative frequencies which are greater than 0.001 are drawn in every distribution

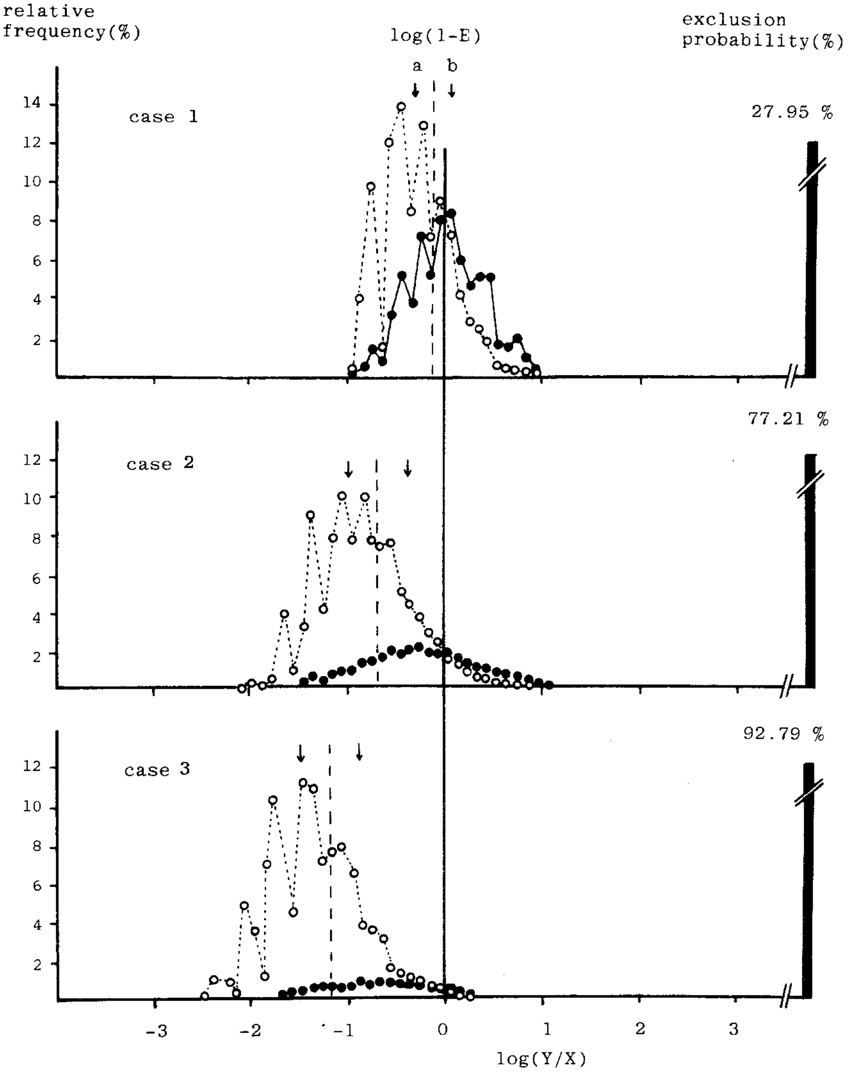

Fig. 3. Distributions of trio. The relative frequencies of log(*Y*/*X*) for 10^5 true fathers and 10^5 non-fathers concerning three kinds of mother-child couples. The phenotypes of the mothers and children are as follows; *case 1:* mother: A, MNs, CcDEe, P+, Fy(ab), Jk(b), Gm(1, 16, 21), Inv(−1), Gc 2−1, Hp 2−1 PGM 2−1, AcP(AB), GPT 2−1, EsD 2−1; child: A, MNs, CcDEe, P+, Fy(ab), Jk(b), Gm(1, 16, 21), Inv(−1), Gc 2−1, Hp 2−1, PGM 2−1, AcP(AB), GPT 2−1, EsD 2−1; *case 2:* mother: A, MNs, ccDEe, P+, Fy(a), Jk(b), Gm(1, 16, 21), Inv(−1), Gc 2−1, Hp 2−1, PGM 2−1, AcP(B), GPT 1, EsD 2−1; child: A, MNs, CcDEe, P−, Fy(a), Jk(b), Gm(1, 16), Inv(−1), Gc 2−1, Hp 2, PGM 2−1, AcP(B), GPT 1, EsD 2−1; *case 3:* mother: A, MNs, ccDEe, P−, Fy(a), Jk(b), Gm(1, 16, 21), Inv(−1), Gc 2−1, Hp 2−1, PGM 2−1, AcP(B), GPT 1, EsD 2−1; child: A, MNs, CcDEe, P+, Fy(a), Jk(b), Gm(1, 16), Inv(−1), Gc 2−1, Hp 2, PGM 2−1, AcP(B), GPT 1, EsD 2−1

Distribution Curve

To examine how the derived formula holds for the actual distributions, we calculated the distributions of $\log(Y/X)$ for various cases. Following 14 systems with Japanese gene frequencies are used in the calculation [9, 11]; ABO (A, B, O), MNSs, Rh (C, c, D, E, e), P (P+ non P–), Duffy (Fy^a, Fy^b), Kidd (Jk^a, Jk^b), Gm (1, 2, 5, 16, 21), Inv (1, non-1), Gc (1, 2), Hp (1, 2), PGM (1, 2), AcP (A, B), GPT (1, 2), EsD (1, 2).

The distributions of duos and those of trios are defined as the distributions of $\log(Y/X)$ of possible fathers concerning a fixed child and those concerning a fixed mother-child couple, respectively [3, 8, 9, 10]. The corresponding exclusion probabilities are determined by the phenotypes of the child and those of the mother-child couple, respectively.

Duos

The distributions of three cases of children are shown in Fig. 2. For the distribution of true fathers, 10^4 men are chosen based on the phenotype of each child. The products of Y/X of 14 systems are calculated and arranged in the order of the values of $\log(Y/X)$, which is divided by the interval of 0.1 in logarithmic scale. The relative frequencies of true fathers are designated by open circles (○), as shown in Fig. 2. For the distribution of non-fathers, 10^4 men are chosen at random in the general populations. Values of Y/X are calculated for unexcluded men. The relative frequencies of unexcluded non-fathers are designated by black circles (●) and those of excluded non-fathers are designated by black sticks, as shown in Fig. 2.

Trios

The distributions of three cases of mother-child couples are shown in Fig. 3. 10^5 true fathers are chosen based on the phenotypes of each mother-child couple and 10^5 non-fathers are chosen at random in the general populations.

Results and Discussion

Exclusion probabilities, mean values, and variances of the distributions calculated for 14 systems are listed in Table 1. The relative area under the distribution curves for $\log(Y/X) < -1$ (or >1) is also listed in Table 1 [9, 10]. These values represent the overall features of the distributions quantitatively. In Figs. 2 and 3, mean values a and b are designated by arrows and the values of $\log(1-E)$ are expressed by chain lines.

The distributions of duos can be approximated as normal curves in any cases of exclusion probabilities (Fig. 2). Variances of the two curves (s_t^2 and s_n^2) for duos are almost the same as shown in Table 1, in agreement with the theoretical prediction expressed in Eqs. (2) and (6) by assuming normal curves.

The distributions of trios are rather indented (Fig. 3). Comparing the distributions of duos with those of trios, the distributions of duos are found to resemble more closely to normal curves than those of trios. Since the number of samplings examined for trios is 10^5, which is ten times larger than that of duos (10^4), this phenomenon is not due to the insufficiency of the number of samplings, but may be due to the following fact: In cases of duos the phenotypes of fathers are restricted only by the phenotype of children, whereas in cases of trios the phenotypes of fathers are restricted not only by the phenotype of the children but also by that of the mothers. The fact that variances of trios are smaller than those of duos also indicates the smaller freedoms in cases of trios.

Table 1. Parameters of the distributions

	Case		Exclusion (%)	Mean value a or b	Variance s_t^2 or s_n^2	Relative area under the curve $\log(Y/X)$		$(a+b)/2$	$\log(1-E)$
						<-1 (%)	>1 (%)		
Duo	(1)	true fathers		−0.629	0.4722	30.60	0.98	−0.050	−0.028
		non-fathers	6.42	0.529	0.4668	1.35	23.61		
	(2)	true fathers		−0.884	0.4817	44.46	0.42	−0.291	−0.278
		non-fathers	47.27	0.303	0.4590	1.55	7.94		
	(3)	true fathers		−1.465	0.3643	77.75	0.01	−0.978	−0.977
		non-fathers	89.45	−0.499	0.3802	2.09	0.09		
Trio	(1)	true fathers		−0.326	0.1128	0.00	0.03	−0.159	−0.142
		non-fathers	27.95	0.009	0.1430	0.00	0.41		
	(2)	true fathers		−0.994	0.2064	40.33	0.00	−0.676	−0.642
		non-fathers	77.21	−0.358	0.2398	2.21	0.01		
	(3)	true fathers		−1.474	0.1886	78.95	0.00	−1.183	−1.142
		non-fathers	92.79	−0.892	0.2128	2.98	0.00		

a = Mean value of the distribution of $\log(Y/X)$ for true fathers
b = Mean value of the distribution of $\log(Y/X)$ for unexcluded non-fathers
s_t^2 = Variance of the distribution for true fathers
s_n^2 = Variance of the distribution for unexcluded non-fathers

Although some distribution patterns deviate from normal curves, the validity of Eq. (8) is examined for each case. In Table 1 the two quantities $(a+b)/2$ and $\log(1-E)$ are compared, since the two quantities should be equal if Eq. (8) holds. Table 1 shows that the difference between the two quantities is at most 0.04, indicating that Eq. (8) holds quite well for various actual distributions.

Equation (8) also implies the following relation; if the exclusion probability (E) increases, the mean value of the two peaks, $(a+b)/2$ becomes smaller, that is, the mean of the distributions gives high paternity probability. This tendency is confirmed in Figs. 2 and 3.

By using the difference between the value of $\log(Y/X)$ of a putative man and the mean value of true father a (or that of non-fathers b), Hummel elucidated the fraction of true fathers (or non-fathers) having higher paternity probability than the putative man [3, 8]. According to his method, the mean value of true fathers a and that of non-fathers b are necessary to find such a fraction of true fathers and that of non-fathers, respectively. Equation (8) can be applied to such situation; that is, if one of the two peaks (a or b) is calculated, the other can be obtained easily by using the relation (8).

Hoppe proposed an index of paternity using the exclusion probability (E) [12, 13, 14, 15]. If only "exclusion" is considered in the group where the same number of true fathers and non-fathers are collected, the part of true fathers is 1 and that of unexcluded non-fathers is $(1-E)$. Hoppe's index of paternity is defined as the ratio of the above two quantities, $(1-E)/1=1-E$. By this definition the paternity probability becomes as $1/\{1+(1-E)\}=1/(2-E)$. However, the unexcluded men, irrespective of their values of Y/X, are treated in the same way in Hoppe's procedure. The exclusion probability (E) has no relation with the distribution pattern of Y/X. Our study reveals the relation between E and the distribution pattern of Y/X, since $\log(1-E)$ represents the mean of the two distributions of $\log(Y/X)$, as illustrated in Figs. 2 and 3.

The distribution pattern of $\log(Y/X)$ gives us more precise information as to the statistical tendency of the paternity probability than E does; the latter quantity is only a certain representative value for large numbers of possible fathers. The distribution pattern provides an adequate method to know the efficiency of the diagnosis [9, 10]. For example, larger fractions of putative men can be diagnosed as true father ($\log(Y/X)<-1$) or non-father ($\log(Y/X)>1$) in the distribution of case 1 in Fig. 2 than that of case 1 in Fig. 3, although E in the former case is smaller than that in the latter case.

In general, it has been accepted as a qualitative tendency that the paternity probability $X/(X+Y)$, often used in the paternity diagnosis [1], increases with increasing E. The quantitative relation between $X/(X+Y)$ and E, however, has not been derived yet. The present work gives a simple relation between the characteristic values (a, b) of paternity probability and E by using the distribution for X and Y versus $\log(Y/X)$.

(The computer calculation has been performed on a Facom 230-75 at Nagoya University Computation Center)

References

1. Essen-Möller E, Quensel CE (1939) Zur Theorie des Vaterschaftsnachweises auf Grund von Ähnlichkeitsbefunden. Dtsch Z Gerichtl Med 3:70
2. Wecht CH (1976) Legal Medicine Annual: Paternity actions—A matter of opinion or a trial of the blood? Appleton-Century-Crofts, New York, pp 237
3. Hummel K (1974) Biostatistical opinion of parentage. Biostatistische Abstammungsbegutachtung. Tabellenband I. Fischer, Stuttgart
4. Hummel K (1974) Zur Blutgruppenbegutachtung in Vaterschaftsprozessen. Zentralbl Jugendrecht u Jugendwohlf 511
5. Salmon D (1978) Indice de paternité: application à un échantillon de 48 trios père présumé-mère-enfant. Rev Fr Transfus Immunohematol 21:239
6. Hummel K, Conradt C (1978) Differenzen zwischen Vaterschafts- und Ausschlußwahrscheinlichkeit aus HLA-Befunden. Z Rechtsmed 81:217
7. Hartmann G, Mueller B, Rittner Ch, Steffens Ch, Wehner HD (1975) Vaterschaftsfragen. Gerichtl Med. Springer, Berlin Heidelberg New York
8. Hummel K, Ihm P, Schmidt V (1969) Beurteilung einer nach der Formel von Essen-Möller gefundenen Vaterschaftswahrscheinlichkeit im Hinblick auf die gegebene Mutter-Kind-Konstellation: Beschreibung der Verfahren; Tabellen und Graphik. Dtsch Z Gerichtl Med 66:97
9. Asano M, Minakata K, Hattori H (1980) General formulas of the estimated likelihood ratio in the diagnosis of paternity of a deceased putative father. Z Rechtsmed 84:125
10. Asano M, Minakata K, Hattori H (1980) Diagnosis of paternity for cases without mother and without both mother and putative father on the basis of blood group findings from the relatives. Z Rechtsmed 84:135
11. Asano M, Minakata K, Hattori H (1979) Calculation of likelihood ratio *Y/X* in diagnosis of paternity using computer methods. Z Rechtsmed 82:263
12. Hoppe HH (1957) Zum Wert der Blutgruppenuntersuchung für den Vaterschaftsprozeß: Berechnung zur Ausschließung und Wahrscheinlichkeit der Vaterschaft. Blut 3:195
13. Sachs V, Hoppe HH (1960) Die Feststellung der Vaterschaft auf Grund blutgruppenserologischer Reaktionen. Dtsch Z Gerichtl Med 50:9
14. Fiedler H, Hoppe HH, Pettenkofer HJ (1968) Ein neuer Weg zum „positiven Vaterschaftsbeweis" über die statistische Auswertung serologischer Befunden. Bundesgesundheitsbl 129
15. Hummel K (1969) Zum biostatistischen Vaterschaftsnachweis aufgrund von Blutgruppenbefunden: Eine allgemeine Betrachtung, zugleich eine Stellungnahme zur Methode von Sachs und Hoppe sowie von Fiedler, Hoppe und Pettenkofer. Bundesgesundheitsbl 379

Received March 5, 1980