

## Übersichtsreferat — Review Article

# Zur Größe der Irrtumswahrscheinlichkeit bei der positiven Vaterschaftsfeststellung anhand von Blutgruppenbefunden\*

K. HUMMEL

Hygiene-Institut der Universität Freiburg i. Br.

Eingegangen am 7. Juli 1970

### The Problem of Error in the Biostatistical Estimation of Fatherhood on the Basis of Blood Grouping Results

*Summary.* The error, estimated in the Bayes' theorem, which is realized when fatherhood is decided categorically for or against, will correspond to the error in reality, the more the gene frequencies for estimation of  $X$  and  $Y$  are realistic and the more the prior plausibilities used in the Bayes' formula correspond to reality. It is shown, that the prior plausibility 0.5, recommended by Essen-Möller, meet the practical requirements: the estimated error in all cases of illegitimacy was greater than in reality. The adequacy of the gene frequencies used for estimation of  $X$  and  $Y$  and those of the prior plausibilities in the formula of Bayes may be tested by comparison with empirical findings. The biostatistical basis however does not need to be confirmed by empirics.

*Key-Words:* Bayessesches Theorem — Biostatistik, Blutgruppengutachten — Essen-Möller-Verfahren — Irrtumswahrscheinlichkeit, positive Vaterschaftsfeststellung — Vaterschaftsfeststellung.

*Zusammenfassung.* Die unter Anwendung des Bayesschen Theorems erhaltene Irrtumserwartung bei kategorischer Entscheidung für bzw. gegen Vaterschaft stimmt mit der Irrtumsquote der Wirklichkeit um so besser überein, je wirklichkeitsbezogener die zur Berechnung von  $Y$  und  $X$  benutzten Genfrequenzen und die im Bayesschen Ansatz verwendeten a priori-Wahrscheinlichkeiten sind. Es wird gezeigt, daß die von Essen-Möller vorgeschlagene a priori-Wahrscheinlichkeit praktischen Anforderungen gerecht wird, indem die mit ihr erhaltenen Irrtumserwartungen (bei Unterhaltsklagen) stets größer sind als die in der Wirklichkeit sich einstellende Irrtumsquote. Mit empirischen Vergleichen kann die Adäquatheit der benutzten Genfrequenzen und der in den Bayesschen Ansatz eingehenden a priori-Wahrscheinlichkeiten getestet werden. Der biostatistische Ansatz selbst bedarf keiner Bestätigung durch die Empirie.

## I. Theoretische Grundlagen

Mit Hilfe von Blutgruppenbefunden läßt sich eine Vaterschaftswahrscheinlichkeit (VW) unter Benutzung des Bayesschen Theorems [1] errechnen, wobei die entscheidende Information von den Häufigkeiten  $Y$  (Vorkommen des Probandentyps unter Nichtvätern) und  $X$  (Vorkommen des Probandentyps unter

---

\* Herrn Professor Dr. R. Haas zum 60. Geburtstag gewidmet.

Vätern) ausgeht. Sollen *Einmannfälle* beurteilt werden, so verwendet man den von Essen-Möller [2] angegebenen Ansatz; bei *Mehrmannfällen* ist der Ansatz von Schulte-Mönting und Hummel [3] zu gebrauchen bzw. — wenn als sicher gelten kann, daß sich der wahre Vater unter den nichtausgeschlossenen Männern befindet — der Ansatz von Essen-Möller und Quensel [4].

Gleichheit vor dem Gesetz ist für die in Paternitätsprozesse verwickelten Männer nur dann gewährleistet, a) wenn sich die Gutachter zur Errechnung der serostatistischen VW einheitlich der gleichen *Methode* (welche optimal sein muß) bedienen, b) wenn sie der Berechnung dieselben *Genfrequenzen* zugrunde legen und c) wenn sie die Beurteilung nach gleichen *Maßstäben* vornehmen.

Eine *statistische Methode* zur positiven Vaterschaftsfeststellung ist dann für die Praxis *optimal*, wenn sie sich der Informationen  $Y$  und  $X$  bedient, wenn sie auf anerkannte statistische Verfahren zurückgeführt werden kann und wenn sie die vom Richter gestellten Fragen beantwortet. Es spricht nichts dagegen, daß die oben genannten drei Ansätze (nach EM, SMH und EMQ) diese Voraussetzungen erfüllen.

Als *Genfrequenzen* zur Berechnung der VW-Werte benutzt man am zweckmäßigsten *Mittelwerte* aus dem deutschen Raum.

Um nach gleichen *Maßstäben* zu beurteilen, sollten die Gutachter zu bestimmten VW-Werten stets gleiche verbale Prädikate gebrauchen. Zu einer einheitlichen Handhabung der statistischen Auswertung gehört auch eine verbindliche Normierung im „a priori-Bereich“: Die beiden Hypothesen Vaterschaft-Nichtvaterschaft sollten (wie schon Essen-Möller empfahl) eine gleichgroße Ausgangswahrscheinlichkeit haben ( $V = \bar{V}$ ). Des weiteren sollte die Ausgangswahrscheinlichkeit der Hypothese „Vaterschaft“ für jeden nichtausgeschlossenen Mann einer Mehrmannsache gleich groß sein ( $B_V = Z'_V = Z''_V \dots$ , vgl. [5]).

In den nachstehenden Formeln

$$W(i) = 1/(1 + Y_i/X_i) \text{ (Essen-Möller-Formel);}$$

$$W_M(i) = 1/\left[1 + Y_i/X_i \left(n + \sum_{\substack{j=1 \\ j \neq i}}^n X_j/Y_j\right)\right]$$

(Formel nach Schulte-Mönting und Hummel);

$$W'_M(i) = 1/\left(1 + \sum_{\substack{j=1 \\ j \neq i}}^n X_j Y_i / X_i Y_j\right)$$

(Formel im Sinne von Essen-Möller und Quensel)

sind diese Normierungen enthalten. Die erhaltenen  $W(i)$ - und  $W_M(i)$ -Werte besagen (in %), wie häufig Männer aus 100, im Blutgruppentyp gleichgelagerten Einmann- bzw. Mehrmannfällen die Erzeuger sind, wenn in den zur Blutgruppenbegutachtung gelangenden Akten der wahre Vater als Beihälter ebensohäufig nominiert wie nicht nominiert ist. — Die  $W'_M(i)$ -Werte (in %) geben an, wie häufig Männer aus 100 im Blutgruppentyp gleichgelagerten Mehrmannfällen die Erzeuger sind, wenn sich unter den in die Begutachtung einbezogenen Beihältern stets der wahre Vater befindet.

Würde man sämtliche Männer einer bestimmten VW als Erzeuger *verurteilen*, so würden sich unter diesen (bei den gegebenen normierenden Voraussetzungen) in der Häufigkeit des VW-Werts wahre Väter, in der Häufigkeit 1-VW Nichtväter befinden. — Würde bei allen Fällen einer bestimmten VW die *Klage zurückgewiesen*, so würden in der Häufigkeit des VW-Wertes wahre Väter einer Verurteilung entgehen, Nichtväter in einer Häufigkeit von 1-VW zu Recht nicht verurteilt.

Es ist anzustreben, daß die Größen  $W$ ,  $W_M$  und  $W'_M$  (bzw. deren Komplemente zu 1) mit den realistischen Zutreffens- bzw. Irrtumserwartungen möglichst gut übereinstimmen. Eine Übereinstimmung ist größenordnungsmäßig in der Tat gegeben. *Größere Differenzen* zwischen der rechnerischen und der realistischen Irrtumserwartung können durch folgende Umstände verursacht sein:

1. *Gültig für  $W$ ,  $W_M$  und  $W'_M$* : Die für die Berechnung der VW benutzten Genfrequenzen stimmen mit den für das Aktenmaterial zutreffenden nicht in wünschenswertem Maß überein.

2. *Gültig für  $W$ ,  $W_M$  und  $W'_M$* : Wenn zwischen der Kindesmutter und dem Beklagten (Zeugen) Blutsverwandtschaft besteht, stellt sich der Fall außerhalb der Stichprobenfälle, welche der Berechnung der VW zugrunde liegen.

3. *Gültig für  $W$  und  $W_M$* : Der Anteil der Aktenfälle, bei dem der wahre Vater mitgenannt ist, ist demjenigen nicht gleichzusetzen, bei dem der wahre Vater nicht genannt ist.

4. *Gültig für  $W'_M$* : Die Voraussetzung für die Zulässigkeit der Formel nach Essen-Möller und Quensel ist *nicht* erfüllt, weil der wahre Vater sich nicht mit Sicherheit unter den nominierten Beihältern befindet. Die Irrtumsquote ist damit in der Wirklichkeit größer als aus  $1-W'_M$  zu erwarten wäre.

5. *Gültig für  $W_M$  und  $W'_M$* : Die Häufigkeit wahrer Väter unter den „Beklagten“ ist nicht gleichzusetzen der unter „Zeugen“ (bei Unterhaltsklagen) bzw. nicht gleichzusetzen der unter „Klägern“ und „Zeugen“ (bei Ehelichkeitsanfechtungen). Die für diese Gruppen (Beklagte, Kläger, Zeugen) aufgrund der mit den Formeln nach SMH und EMQ erhaltenen Irrtumserwartungen können dann mit der Wirklichkeit in einem größeren Maß differieren.

Die unter 1.—5. aufgeführten Umstände<sup>1</sup> werden anschließend eingehender betrachtet.

*Zu 1. Die Häufigkeit der Phänotypen* erblicher Eigenschaften (entsprechend auch die der Genotypen) schwanken von Population zu Population. Innerhalb Deutschlands sind die Schwankungen jedoch — wenn man von kleineren Inzuchtgruppen absieht — nicht so groß, daß die VW-Werte dadurch ernsthaft verfälscht werden könnten (vgl. u. a. das Ergebnis eines Nord-Süd-Vergleichs [7]). Hinzu kommt, daß sich Abweichungen nach oben und unten innerhalb der einzelnen Systeme zu einem gewissen Teil kompensieren dürften.

Der Anteil *Fremdstämmiger* bei Vaterschaftsprozessen ist in den Genfrequenzen berücksichtigt, wenn diese tatsächlich aus *Aktenfällen* gewonnen wurden. Trotzdem verbleibt ein Rest an „Unschärfe“, einmal weil sich Fremdstämmige unter den Putativvätern häufiger als unter den Kindesmüttern befinden, zum

<sup>1</sup> Zu diesen kommt noch ein Einfluß durch Anwendung einer *nicht weiter differenzierten* Formalgenetik. Doch ist dieser so gering, daß auf ihn nicht weiter einzugehen ist (vgl. [6]).

anderen, weil bei Fremdstämmigen bestimmte charakteristische Merkmalsgruppen „gebündelt“ vorliegen und in dieser Bündelung weitergegeben werden. Der Wert *positiver* Beweise wird durch die genannten Umstände praktisch nicht angetastet; *negativen* Hinweisen wird man hingegen nicht das volle Gewicht beimessen können [6].

*Zu 2.* Bei *Verwandtschaft* zwischen dem Beklagten (Zeugen) und der Kindesmutter ist mit einer größeren Übereinstimmung zwischen den beiden Blutformeln zu rechnen, als im Mittel bei Nichtverwandtschaft. Für die Berechnung der VW bedeutet dies verminderte Differenzierungschancen: Je ähnlicher sich die Eltern sind, um so schwächer fallen die aus den Blutformeln sich ergebenden positiven und negativen Hinweise aus. *Falsche Hinweise* können sich aber aus Verwandtschaftsbeziehungen nicht ergeben, somit auch keine Fehlentscheidungen (zumindest nicht die fälschliche Annahme einer Vaterschaft bei Nichtvätern). Will man nicht dem (gut zu begründenden) Vorschlag folgen, Verwandtschaftsfälle wie alle übrigen zu behandeln, so kann man bei der Beurteilung diejenigen Systeme auslassen, in denen der Eventualvater mit der Kindesmutter übereinstimmt [6]. Die *Irrtumsquote in der Gesamtstatistik* wird im übrigen wegen der Seltenheit einer Verwandtschaft zwischen dem Putativvater und der Kindesmutter praktisch nicht beeinflusst.

*Zu 3.* Bei *Einmannsachen* (Unterhaltsklagen) kann man damit rechnen, daß Aktenfälle *ohne* benanntem Vater etwa ebensohäufig sind wie die *mit* benanntem Vater [8]. Die Irrtumserwartung  $1-W$  stimmt daher weitgehend mit der Wirklichkeit überein. — Bei *Mehrmannsachen* (Unterhaltsklagen) findet sich ein *höherer* Anteil Akten *mit* genanntem Vater als *ohne* diesen. Damit sind die mit  $1-W$  bzw.  $1-W_M$  angegebenen Irrtumserwartungen (bei kategorischer Annahme einer Vaterschaft) in Wirklichkeit kleiner. Die Irrtumswahrscheinlichkeit  $W$  bzw.  $W_M$  — bei kategorischer Annahme von *Nichtvaterschaft* — ist entsprechend in der Wirklichkeit größer.

*Ehelichkeitsanfechtungsfälle* zeigen andere  $V/\bar{V}$ -Relationen als Unterhaltsklagen. Bei *Einmannsachen* wird der Kläger als Erzeuger häufiger ausgeschlossen, als einer  $V/\bar{V}$ -Relation von 1 entspricht; die  $V/\bar{V}$ -Relation ist in der Wirklichkeit kleiner als 1, entsprechend ist die reale Irrtumsquote — bei kategorischer Annahme einer Vaterschaft des Klägers — größer als durch  $1-W$  angegeben [6]. — Bei *Zweimannsachen* befindet sich der Erzeuger meist unter den in die Begutachtung einbezogenen Männern. Damit ist die a priori gegebene VW für den nichtausgeschlossenen Mann stets größer als 0,5; die Irrtumswahrscheinlichkeit — bei kategorischer Annahme einer Vaterschaft des nichtausgeschlossenen Mannes — ist kleiner als durch  $1-W$  angegeben.

Ist weder der Kläger noch der Mehrverkehrszeuge von der Vaterschaft auszuschließen, und kann man — wie es in der Regel zulässig sein dürfte — davon ausgehen, daß sich der wahre Vater unter den beiden nichtausgeschlossenen Männern befindet, dann ist zur biostatistischen Auswertung die EMQ-Formel zu verwenden. Würde man die SMH-Formel benutzen, so wäre die durch  $1-W_M$  gegebene Irrtumserwartung im Verhältnis zur Wirklichkeit zu groß.

Stammt der Zweimannfall aus einer Drei- oder Mehrmannsache, so wird man sich eher zur SMH-Formel entschließen, weil es hier nicht mehr als voll gesichert

gelten kann, daß der wahre Vater mitbenannt ist. Stets wird in diesen Fällen die mit  $1 - W_M$  angegebene Irrtumserwartung für die Wirklichkeit zu groß sein, weil im Aktenmaterial häufiger der wahre Vater mitbenannt ist als nicht benannt.

*Zu 4.* Die Anwendung der EMQ-Formel ist nur dann zulässig, wenn gesichert ist, daß sich der Erzeuger unter den benannten Eventualvätern befindet. Es gehört nicht zur Aufgabe des Gutachters, sondern zu der des Richters, zu entscheiden, ob sich — den übrigen Beweismständen nach — der wahre Vater unter den benannten Beihältern befindet oder ob dies möglicherweise nicht der Fall ist. Der Gutachter sollte dem Gericht sowohl das Ergebnis der SMH-Formel als auch das der EMQ-Formel mitteilen und es dem Gericht überlassen, die Auswahl entsprechend dem gerade vorliegenden Fall zu treffen. Entscheidet sich das Gericht für den SMH-Wert, so würde eine Differenz der Irrtumserwartung  $1 - W_M$  zur Irrtumsquote der Wirklichkeit im wesentlichen auf eine von 1 abweichende  $V/\bar{V}$ -Relation zurückzuführen sein. Entscheidet das Gericht sich für den EMQ-Wert, so ließe eine Nichtübereinstimmung mit der statistischen Wirklichkeit darauf schließen, daß im einen oder anderen Fall die Formel nicht hätte angewandt werden dürfen (weil der wahre Vater nicht mitbenannt war).

*Zu 5.* Die Normierung  $B_V = Z_V$  (Häufigkeit von Vätern unter „Beklagten“ gleich der unter „Zeugen“), bzw.  $K_V = Z_V$  (Häufigkeit von Vätern unter „Klägern“ gleich der unter „Zeugen“), wie sie für Mehrmannfälle in die SMH- und EMQ-Formel eingeht, nimmt nur dann Einfluß auf die Irrtumswahrscheinlichkeit (bei Unterhaltsklagen), wenn diese für Beklagte bzw. Kläger und Zeugen getrennt ermittelt wird.  $B_V$  ist in aller Regel größer als  $Z_V$ ,  $K_V$  kleiner als  $Z_V$ . Die sich aus  $W_M$  oder  $W'_M$  ergebende Irrtumserwartung ( $1 - W_M$ ;  $1 - W'_M$ ) ist dementsprechend für „Beklagte“ größer als die Irrtumsquote in der Wirklichkeit; für „Kläger“ ist sie kleiner.

Die hier angestellten Betrachtungen über die nach den Formeln von EM, SMH und EMQ ermittelte Irrtumserwartung im Vergleich zur realistischen — bei kategorischer Entscheidung zugunsten der „Nullhypothese“ (= Vaterschaft) — haben folgendes ergeben:

a) In praktisch jedem Begutachtungsfall ist eine Differenz zwischen der rechnerisch (aus den Formeln nach EM, SMH und EMQ) erhaltenen Irrtumswahrscheinlichkeiten und der tatsächlichen Irrtumsquote zu erwarten.

b) Die Differenz kommt zu einem kleinen Teil dadurch zustande, daß die zur Auswertung benutzten Genfrequenzen mit den Genfrequenzen der Populationen, aus denen die Beteiligten stammen, nicht völlig übereinstimmen; die Abweichungen halten sich im allgemeinen in Grenzen und werden sich — als Plus-Minus-Abweichungen — z. T. ausgleichen.

c) Bedeutungsvoller sind Einflüsse, welche von effektiven a priori-Verhältnissen ausgehen, die vom Normierungswert 0,5 abweichen. Ist der Normierungswert im Vergleich zur Realität zu nieder, so wird eine höhere Irrtumserwartung vorgetäuscht als sie in Wirklichkeit besteht; für die Begutachtungspraxis läßt sich dies noch in Kauf nehmen; bedenklicher erscheint es, wenn der Normierungswert zu hoch liegt und sich damit eine im Vergleich zur Wirklichkeit zu geringe Irrtumserwartung ergibt. Das letztere trifft nur für Ehelichkeitsanfechtungen zu, bei denen als einziger möglicher Erzeuger der Kläger benannt ist.

Für die Begutachtungspraxis hat man sich bei diesen Fällen daran zu erinnern, daß die Ehe unter dem besonderen Schutz des Gesetzgebers steht. Somit erscheint es gerechtfertigt, auch diese Diskrepanz unberücksichtigt zu lassen: Würde man eine realistische a priori-Wahrscheinlichkeit einführen, so könnte es geschehen, daß mit dieser die VW eines *jeden* Klägers (in Einmannsachen) so niedrig zu liegen käme, daß beim Richter stets Zweifel an der Vaterschaft wachgerufen würden.

d) Die Normierung  $B_V = Z_V$  bzw.  $K_V = Z_V$ , wie sie in die SMH- und EMQ-Formel für Mehrmannfälle eingeht, bezieht sich auf die Irrtumserwartung innerhalb der Gruppe der Beklagten bzw. Kläger und Zeugen. Die Realität weicht von dieser Normierung in aller Regel erheblich ab. Die Abweichung bleibt aber ohne Auswirkung, solange man — wie dies meist geschieht — die Irrtumswahrscheinlichkeit auf *alle* Männer bezieht (ohne jeweils zu berücksichtigen, ob es sich um Beklagte, Zeugen oder Kläger handelt).

## II. Empirie zur Kontrolle der biometrischen Erwartungswerte

Wie im ersten Teil ausgeführt, gibt eine Vaterschaftswahrscheinlichkeit, berechnet auf der Basis des Bayesschen Theorems unter Verwendung möglichst wirklichkeitsentsprechender Genfrequenzen und a priori-Wahrscheinlichkeiten eine Zutreffenserwartung an, welche mit der tatsächlichen Zutreffensquote, bezogen auf ein realistisches Aktenmaterial, weitgehend übereinstimmen muß.

Da die Zutreffenswahrscheinlichkeit in keinem Fall 1 betragen kann, sondern immer mehr oder weniger unter 1 liegt, ist stets mit einer Irrtumsquote zu rechnen. Diese realisiert sich, wenn kategorisch zugunsten der Hypothese *Vaterschaft* entschieden wird und einige Nichtväter fälschlich als Väter verurteilt werden, oder (komplementär), wenn kategorisch zugunsten der Hypothese *Nichtvaterschaft* entschieden wird und wahre Väter zu Unrecht nicht verurteilt werden.

Der Blutgruppengutachter wird in concreto mit einer Realisierung der Irrtumserwartung konfrontiert, wenn bei der Ausdehnung eines Blutgruppengutachtens auf weitere Erbsysteme oder durch die anthropologische Begutachtung ein vorher als Erzeuger höchst verdächtiger Mann (mit  $W \geq 90\%$ ) von der Vaterschaft ausgeschlossen bzw. wenn seine Erzeugerschaft wenigstens als sehr unwahrscheinlich befunden wird.

Wie häufig es vorkommt, daß durch eine *erweiterte Blutgruppenuntersuchung* serostatistische Hinweise wiederlegt werden, ist nicht bekannt. Doch sind in der Literatur einige Vergleiche serostatistischer und *anthropologischer* Ergebnisse mitgeteilt [9, 10, 11]. Sie ergaben im ganzen eine gute Übereinstimmung. Nur in wenigen Fällen harmonierte das anthropologische Schlußurteil nicht mit dem serostatistischen. Schade und Scholz [11] fanden unter 65 Männern aus *Einmannfällen* (Beklagte und Kläger) mit  $W = 90\%$  2, die im anthropologischen Gutachten als Erzeuger „sehr unwahrscheinlich“ waren; in 2 weiteren Fällen blieb das anthropologische Resultat im „Unentschieden“, bei eher weniger wahrscheinlicher Vaterschaft. In 61 Fällen fiel das anthropologische Ergebnis gleichsinnig mit dem serostatistischen aus.

In eigenen Untersuchungen (zusammen mit Schwarzfischer und Schmidt [9]) erwiesen sich 7 von 157 Männern aus Einmannfällen mit  $W \geq 90\%$  im anthropologischen Gutachten als Nichtväter (Vaterschaft „sehr unwahrscheinlich“ bis „praktisch ausgeschlossen“). In 22 Fällen war nach dem anthropologischen Urteil

die Vaterschaft „möglich“, in 128 Fällen „wahrscheinlich“ bis „praktisch erwiesen“. — Nicht alle 7 „Widerspruchsfälle“ können als solche voll gewertet werden. So war bei 4 von diesen Mehrverkehr erwiesen; die Mehrverkehrer konnten aber (aus verschiedenen Gründen) nicht in die Blutgruppenbegutachtung einbezogen werden. In einem weiteren Fall pflog die Kindesmutter hWg. Nur in 2 von den 7 Fällen konnte kein Mehrverkehr namhaft gemacht werden.

Nach ständiger Rechtsprechung ist bei Fällen mit nachgewiesenem Mehrverkehr, aber nicht greifbaren Mehrverkehrszeugen (oder bei zu vielen Mehrverkehrern) eine VW von mindestens 99% erforderlich, um die Vaterschaft eines Mannes positiv festzustellen. Einer anthropologischen Begutachtung bedarf es dann in der Regel nicht mehr [12]. Keiner der Männer der 5 Fälle mit nicht greifbaren oder zu zahlreichen Mehrverkehrern erreichte  $W = 99\%$ ; der höchste Wert war  $W = 98,5\%$ . Dementsprechend war für keinen dieser Fälle die Voraussetzung für eine vollgültige serostatistische Entscheidung zugunsten der Hypothese Vaterschaft gegeben. Alle 5 Fälle waren daher regelentsprechend der anthropologischen Begutachtung zuzuführen.

Faßt man die Einmannfälle von Schade und Scholz [11] sowie die eigenen [9] zusammen, so erhält man 4 unter 222 beteiligten Männern mit  $W \geq 90\%$ , welche aufgrund der anthropologischen Begutachtung als Nichtväter anzusehen sind. Sie wären wohl alle zu Unrecht als Väter verurteilt worden, wenn der Richter allein aufgrund des serostatistischen Resultats (Vaterschaft mit  $W \geq 90\%$  wenigstens „wahrscheinlich“) geurteilt hätte. Wie stimmt diese Irrtumsquote mit der rechnerischen Erwartung überein? Rund 6% der „Nichtväter“ eines Gutachtenmaterials von 909 Einmannfällen [die 909 Nichtväter waren durch eine (ausnahmsweise 2) Blutgruppeneigenschaft als Erzeuger ausgeschlossen worden] wiesen  $W \geq 90\%$  auf. Geht man davon aus, daß a) das anthropologische Gutachten *alle Nichtväter* erfaßt und b) daß *alle* Entscheidungen dieser Expertise fehlerfrei sind (Annahmen, die zwar nicht eigentlich zutreffen, statistisch jedoch tragbar sind), dann ist die Relation der Fälle *ohne* Vater zu denen *mit* Vater nicht, wie bei Essen-Möller,  $Q = \bar{V}/V = 1$ , sondern  $Q = 0,45$  (3% Nichtväter mit  $W \geq 90\%$  bei  $Q = 1$ ;  $4/222 = 1,8\%$  Nichtväter mit  $W \geq 90\%$  bei  $Q = 0,43$ ). Die a priori gegebene VW beträgt nicht 0,5, sondern 0,7. Dieser Wert liegt zwischen demjenigen von Einmannsachen ( $\text{apr}W = 0,5$ ) und dem von Zweimannsachen ( $\text{apr}W = 0,83$ ) des serologischen Aktenmaterials [8].

Nun läßt sich mit 4 Fällen keine vertrauenswürdige Statistik aufstellen. Doch dürften auch mit den kleinen Zahlen folgende Aussagen möglich sein:

1. Es liegt kein Grund vor zur Annahme, die Irrtumsquote bei der Entscheidung von Einmannfällen sei in der Wirklichkeit größer als durch  $1 - W$  der Essen-Möller-Formel angegeben.

2. Es gibt keinen Grund zur Annahme, die a priori gegebene VW für Männer von Einmannfällen des anthropologischen Aktenmaterials unterscheide sich merklich von der für Männer des serologischen.

Es ist weiter nicht anzunehmen, die  $W$ -Werte der einer anthropologischen Begutachtung zugeführten Fälle würden im Mittel höher ausfallen als diejenigen der serologischen Aktenfälle (wie dies von Grote [13] sowie Schade und Scholz [11] angenommen wird): Die Tatsache, daß sämtliche dem Anthropologen zugeleiteten

Männer serologisch nicht auszuschließen sind, wirkt sich als solche nicht auf die Höhe der  $W$ -Werte aus<sup>2</sup>. Die  $W$ -Werte werden nur dann beeinflußt, wenn die Fälle (vom Richter) vorsortiert werden. Mit einer solchen Vorsortierung ist in praxi allerdings stets zu rechnen, weil die Gerichte Fälle mit zweifelhafter Vaterschaft des Beklagten (oder Klägers), z. B. wegen niederer VW, *bevorzugt* dem anthropologischen Gutachter zuweisen. Als Folge der Sortierung sind nun aber nicht höhere, sondern niedrigere  $W$ -Werte zu erwarten. Durch Anreicherung von Nichtvaterfällen kann dabei eine  $\text{apr}W$  von 0,5 sogar möglicherweise unterschritten werden; die Irrtumsquote der Wirklichkeit würde damit größer als durch  $1 - W$  angegeben. Wie jedoch der Vergleich der aus Modellterzetten erhaltenen Irrtumserwartung (3%, bei einer  $\text{apr}W$  von 0,5) mit der bei der anthropologischen Begutachtung festgestellten Fehlerquote ( $4:222 = 1,8\%$ ) wahrscheinlich macht, ist mit einer Unterschreitung der  $\text{apr}W$  von 0,5 in praxi nicht zu rechnen.

Auch für *Mehrmannfälle* wurden Vergleiche zwischen serostatistischen und anthropologischen Resultaten angestellt. Schade und Scholz [11] fanden unter 93 Zweimannfällen 19, bei denen einer der Männer  $W \geq 90\%$  erhielt, während der andere unter diesem Wert blieb. Vier Fälle werden von den Autoren herausgestellt, bei denen der serostatistisch wahrscheinlichere Mann (mit  $W \geq 90\%$ ) im anthropologischen Gutachten als Erzeuger praktisch ausgeschlossen wurde, während der weniger wahrscheinliche als Vater erkannt wurde. Doch können nicht alle 4 Fälle als echte Irrtumsfälle angesehen werden. Nur ein Fall wäre — unter den derzeit praktizierten Beurteilungskriterien — serostatistisch entschieden worden: Wertet man nach der EMQ-Formel aus (wobei angenommen wird, daß sich der wahre Vater unter den beteiligten Männern befindet), so ist nach einem früher gemachten Vorschlag [6], der sich an eine Vorschrift des schwedischen Rechts anlehnt, für eine serostatistische Entscheidung bei Zweimannfällen Voraussetzung, daß die Logarithmen der „kritischen Werte“ der beiden Männer um wenigstens eine Log-Stufe differieren. Des weiteren ist für den als Erzeuger in Frage kommenden Mann wenigstens  $W = 90\%$  zu verlangen. In 3 der von Schade und Scholz mitgeteilten „Irrtumsfällen“ liegt die Log-Differenz (wie mit den umgerechneten  $W$ -Werten festzustellen war) *unter* 1. Nur in einem Fall war sie größer als 1 (Beklagter  $W = 60 - 70\%$ ; Zeuge  $W = 96 - 97\%$ ;  $\lg 9,728 - 8,555 = 1,173$ ). Dieser eine Fall hätte serostatistisch entschieden werden können und hätte, wenn der Richter dem Gutachter gefolgt wäre, zu einem Fehlurteil geführt. — Im übrigen wäre der Fall auch mit der SMH-Formel — mit  $W = 90\%$  — biostatistisch zu entscheiden gewesen.

Im eigenen Material (ausgewertet zusammen mit Schmidt und Schwarzfischer [10]) fanden sich unter 15 serostatistisch (unter Anwendung der SMH-Formel) entschiedenen und anthropologisch überprüften Zweimannfällen 10, bei denen das Urteil des anthropologischen Gutachtens gleichsinnig zu dem des serostatistischen ausfiel. Bei 4 Fällen blieb das anthropologische Urteil im Un-

<sup>2</sup> Irreführend wirken hierzu die Ausführungen über  $Q'$  in [7], S. 31. Die Korrektur  $Q'$  (für Einmannsachen) hat nur dann ihre Berechtigung, wenn  $Y$  die Merkmalsfrequenz bei nichtausschließbaren Nichtvätern angibt. In den gebräuchlichen Tabellen bedeutet  $Y$  aber die Merkmalshäufigkeit in der Bevölkerung. Damit ist als  $\text{apr}W$  nicht  $Q'$ , sondern  $Q$  zu verwenden. Im übrigen ist — nach unserem heutigen Wissen — an derselben Stelle unrichtig, daß die  $\text{apr}W$  mit der Zahl der beteiligten Männer *abnimmt*; in Wirklichkeit ist sie bei Mehrmannsachen *größer* als bei Einmannsachen [8].



entschieden. *Ein* Fall wurde anthropologisch konträr entschieden, d. h. der serostatistisch als Vater „wahrscheinliche“ Mann erwies sich anthropologisch als Nichtvater, und der serostatistisch als Vater „unwahrscheinliche“ Mann wurde anthropologisch als Vater festgestellt.

Der Umfang der vorliegenden Vergleichsuntersuchungen ist zu gering, um verwertbare Schlüsse ziehen zu können. Es besteht aber kein Grund zur Annahme, daß die tatsächliche Irrtumsquote bei Mehrmannfällen mit der statistischen Irrtumserwartung (bei Benutzung der EMQ- bzw. SMH-Formel) praktisch nicht übereinstimmt. Vor allem ist nicht mit einer *höheren* Irrtumsquote in der Wirklichkeit zu rechnen.

Da Mehrmannfälle — im Gegensatz zu Einmannfällen — durch das Gericht für die anthropologische Begutachtung nicht vorsortiert sind, ist im übrigen nicht damit zu rechnen, daß die a priori-Wahrscheinlichkeit (wegen einer geänderten  $\bar{V}/V$ -Relation) beeinflußt wird, ebensowenig damit, daß sich das Mittel der  $W$ -Werte verschiebt.

Die Treffsicherheit der nach dem Bayesschen Ansatz erhaltenen Zutreffens- bzw. Irrtumswahrscheinlichkeiten kann man auch an Modellterzetten prüfen. Als *echte* Terzetten kann man Familienterzetten einsetzen; verwendbar sind auch forensische Fälle mit anthropologisch „praktisch erwiesener“ Vaterschaft des Putativvaters [9]. *Falsche* Terzetten gewinnt man entweder durch Vätertausch an echten Terzetten (von Hand oder über den Computer), oder man benutzt Ausschlußterzetten, bei denen man zur biostatistischen Auswertung das abschließende Blutgruppensystem wegläßt. Man berechnet die mittlere Ausschlußwahrscheinlichkeit für die beteiligten Mutter-Kind-Terzetten (vgl. [8, 14]) und konstruiert mit dieser eine 1:1-Relation der Väter-Nichtväterfälle *vor* der Blutgruppenbegutachtung (z. B. sind — bei einer mittleren Ausschlußwahrscheinlichkeit für Nichtväter von 0,85 — zu 15 Nichtvaterfällen 100 Vaterfälle zuzumischen). Dann ermittelt man die prozentuale Besetzung der einzelnen  $W$ -Bereiche (z. B. Bereiche  $W = 50 \dots 90\%$ ;  $90 \dots 95\%$ ;  $95 \dots 99\%$ ) und vergleicht mit der Besetzung, wie sie sich beim ausgewerteten Aktenmaterial findet. Würden sich dabei größere Differenzen finden, so könnten mehrere Ursachen verantwortlich gemacht werden. Einmal könnten die Genfrequenzen der Modellterzetten mit denen des Aktenmaterials nicht harmonieren (z. B. wegen soziologischer oder anderer herkunftsmäßiger Differenzen); oder es differieren die zur biometrischen Auswertung benutzten Genfrequenzen mit denen des Aktenmaterials oder/und denen der Modellterzetten. Schließlich führt auch eine nicht übereinstimmende a priori-Wahrscheinlichkeit zu Differenzen.

In der Literatur sind bis jetzt keine Ergebnisse von Vergleichsuntersuchungen mit Modellterzetten bekannt geworden. Jedoch sind lg  $Y/X$ -Verteilungen bei Vater- und Nichtvaterfällen sowie bei Gerichtsfällen aufgestellt, die zu Vergleichen herangezogen werden können [9].

Die hier angestellten Betrachtungen über den Wert empirischer Untersuchungen zur Kontrolle der Verlässlichkeit serostatistischer Aussagen haben zu folgendem Ergebnis geführt:

1. Es gibt keinen idealen empirischen Test, mit dem verlässlich geprüft werden kann, wie gut sich die statistische Irrtumserwartung in der Wirklichkeit realisiert.

2. Die Häufigkeit der empirisch gefundenen Widersprüche zwischen dem serostatistischen und anthropologischen Resultat steht nicht in Gegensatz zur Erwartung; sie wird vor allem nicht größer sein als erwartet. Voraussetzung für eine möglichst geringe Widerspruchsquote (bei noch ausreichender „Ausbeute“) ist, daß bei Ein- und Mehrmannfällen erst von  $W=90\%$  an (bzw. nach der SMH-Formel von  $W_M(1)=90\%$  an) ein positiver Hinweis als verwertbar gilt, bzw. daß — bei Verwendung des EMQ-Prinzips — für den als Vater verdächtigen Mann neben  $W_M(1)=90\%$  der  $\lg Y/X$ -Wert um wenigstens 1 Log-Stufe niedriger als der der anderen Männer liegt.

3. Die Tatsache, daß Gerichte trotz hoher VW (absolut und relativ) von Männern aus Ein- und Mehrmannfällen die Entscheidung in ausgewählten Fällen dem anthropologischen Gutachter überlassen, zeigt, daß Umstände *außerhalb* der Blutgruppenexpertise (einschließlich biostatistischer Auswertung) ausreichend stark sein können, um auch deutliche Hinweise auf die Vaterschaft zu erschüttern. Damit wird offenbar, daß die in der Wirklichkeit durch den Spruch der Gerichte realisierte Irrtumsquote weit geringer sein wird, als es der statistischen Irrtumserwartung entspricht.

4. Es besteht die Möglichkeit, die Besetzung einzelner  $W$ -Bereiche, wie sie beim biostatistisch ausgewerteten Aktenmaterial erhalten wird, mit der Besetzung durch Modellterzetten zu vergleichen. Würden sich bei einem solchen Vergleich — der bisher noch nicht vorgenommen wurde — größere Unterschiede ergeben, so kämen mehrere Ursachen in Frage. Eine Auskunft darüber, welche Ursachen beteiligt sind und in welchem Maß, kann der Vergleichstest aber nicht geben.

### Literatur

1. Bayes (comm. R. Price): An essay towards solving a problem in the doctrine of chances. Phil. Trans. **53**, 370 (1763).
2. Essen-Möller, E.: Die Beweiskraft der Ähnlichkeit im Vaterschaftsnachweis; theoretische Grundlagen. Mitt. Anthropol. Ges. (Wien) **68**, 9 (1938).
3. Schulte-Mönting, J., Hummel, K.: Über die Berechnung der Vaterschaftswahrscheinlichkeit bei Fällen mit mehr als einem im Blutgruppengutachten nicht ausgeschlossenen Mann. I. Mitteilung: Theoretische Grundlagen. Z. Immun.-Forsch. **138**, 295 (1969).
4. Essen-Möller, E., Quensel, C. E.: Zur Theorie des Vaterschaftsnachweises auf Grund von Ähnlichkeitsbefunden. Dtsch. Z. Ges. gerichtl. Med. **31**, 70 (1939).
5. Hummel, K.: Eine weitere Art A-priori-Wahrscheinlichkeit zur Ermittlung der Irrtumserwartung aus Vaterschaftsplaussibilitäten bei Mehrmannfällen. Z. Immun.-Forsch. **140**, 115 (1970).
6. — Biostatistischer Vaterschaftsbeweis. In: Lehrbuch der Gerichtlichen Medizin (Hrsg. A. Ponsold), 3. Aufl. Stuttgart: Thieme 1967.
7. — Die serologische Begutachtung mit biostatistischer Auswertung. In: Die medizinische Vaterschaftsbegutachtung mit biostatistischem Beweis (Hrsg. K. Hummel). Stuttgart: G. Fischer 1961.
8. — Schmidt, V.: Berechnung realistischer A-priori-Wahrscheinlichkeiten zur Verwendung bei der biostatistischen Vaterschaftsbegutachtung. Z. Immun.-Forsch. **139**, 407 (1970).
9. — Schwarzfischer, F., Schmidt, V.: Vergleich der nach Essen-Möller berechneten Vaterschaftswahrscheinlichkeit einschließlich „Zuordnung“ mit anthropologischen Resultaten bei Einmannfällen. Z. Rechtsmedizin **67**, 27 (1970).
10. — Schmidt, V., Schwarzfischer, F.: Die Berechnung der Vaterschaftswahrscheinlichkeit bei Fällen mit mehr als einem im Blutgruppengutachten nicht ausgeschlossenen Mann. 2. Mitteilung: Praktische Durchführung. Ergebnis einer Reihenuntersuchung im Vergleich zu anthropologischen Resultaten. Z. Immun.-Forsch. **139**, 50 (1970).

11. Schade, H., Scholz, W.: Problematische Beispiele zur positiven serologischen Vaterschaftsfeststellung. *Z. Rechtsmedizin* **67**, 64 (1970).
12. Beitzke, G.: Bewertung der Vaterschaftsbeweise. In: *Lehrbuch der Gerichtlichen Medizin* (Hrsg. A. Ponsold), 3. Aufl. Stuttgart: G. Thieme 1967.
13. Grote, W.: Empirische Häufigkeiten positiver und negativer Aussagen über die Vaterschaft aus Blutuntersuchungen nach der Essen-Möller-Methode. *Anthrop. Anz.* **30**, 258 (1968).
14. Hummel, K., Ihm, P., Schmidt, V.: 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. ges. gerichtl. Med.* **66**, 97 (1969).

Professor Dr. K. Hummel  
Hygiene-Institut der Universität  
D-7800 Freiburg i. Br.  
Hermann-Herder-Straße 11