

Aus dem Universitätsinstitut für gerichtliche Medizin Köln  
(Direktor: Prof. Dr. R. MANZ)

## Statistische Betrachtungen zur Trunkenheitsdiagnose und zur Anwendbarkeit der Nystagmusprobe\*

Von

EBERHARD LOMMER

Mit 5 Textabbildungen

Unsere Gerichte ermitteln einen Trunkenheitsgrad aus der Blutalkoholkonzentration (BAK), den Zeugenbeweisen *und* dem sog. Arztprotokoll. Manche Richter haben sich dabei ein empirisches Urteil über die diagnostischen Qualitäten der regelmäßig mit Blutabnahmen befaßten Ärzte gebildet. Für den Sachverständigen ergibt sich daraus die Aufgabe, gelegentlich Arztbefunde diskutieren zu müssen, die der BAK nicht zu entsprechen scheinen. Eine reibungslose Zusammenarbeit ist auf die Dauer jedoch nur möglich, wenn der Aussagewert des Protokolls in seinen Einzelangaben und in seiner Schlußdiagnose für den örtlichen Gebrauch — wie im Interesse der allgemeinen Erkenntnis — *objektiv* geprüft wird.

Unsere darauf abgestellten Untersuchungen bestärken die latente Forderung, den Fragebogen von entbehrlichen Fragen zu entlasten (WAGEMANN<sup>11</sup>) und dafür eine von den üblichen Proben unabhängige, mit der BAK funktional verknüpfte diagnostische Methode einzuführen. Eine solche Methode sollte ganz auf die Verhältnisse eines Polizeireviers abgestimmt, überall und ohne Hilfsmittel des Laboratoriums durchführbar und vom Willen des Untersuchten weitgehend unabhängig sein, wie bisher nur die in Köln auf Vorschlag von TASCHEN<sup>8, 9</sup> eingeführte *Nystagmusprobe*. Sie wurde in einigen Tausend Fällen angewandt und hat sich auch anderwärts in der dort geübten Form bewährt (SCHULTE und ROTH<sup>6</sup>). Ein Vergleich mit RAUSCHKE<sup>5</sup> ist wegen der starken Abweichung vom Original nicht möglich, seine relativ kleine Versuchszahl läßt weder Normalverteilung noch Gesetzmäßigkeit erkennen, so daß seine Ablehnung dieser Variante verständlich wird.

Wir wollen zeigen, daß die Beurteilung von Trunkenheitsproben und die Diagnose selbst statistisch begründete Entscheidungsakte sind (KOLLER<sup>3</sup>), daß man also mit bestimmbar streuungen rechnen *muß*. Durch statistischen Vergleich der üblichen, von verschiedenen Ärzten

\* Auszugsweise vorgetragen auf dem Kongreß für gerichtliche und soziale Medizin Zürich, 13. 9. 58.

ausgeführten Proben soll der Einfluß subjektiver Faktoren auf diese Schätzungsmethoden geprüft werden. Die Bearbeitung der Ergebnisse mit der Nystagmusprobe soll deren grundsätzliche Brauchbarkeit aufzeigen, Auskunft über die Reproduzierbarkeit geben und die Frage der Zuordnung von BAK und Nystagmusdauer in mathematischer Formulierung anschnneiden.

Hierzu wurden alle Befunde aus Köln-Stadt 1956, die von 4 Ärzten M, N, F, T und einer Gruppe Chirurgen K erstellt wurden (ohne

die von Gelegenheitssachverständigen), statistisch ausgewertet. Die Reproduzierbarkeit wurde an den Befunden des 1957 eingearbeiteten Arztes P geprüft. Die Untersucher standen vergleichbaren Teilmassen von Trunkenheitsfällen gegenüber (s.  $M_x$  in Tabelle 2). Da der Mittelwert unseres Materials von 1955—1957 konstant bei  $1,5\text{‰}$  blieb, sind auch diese Befunde miteinander vergleichbar.

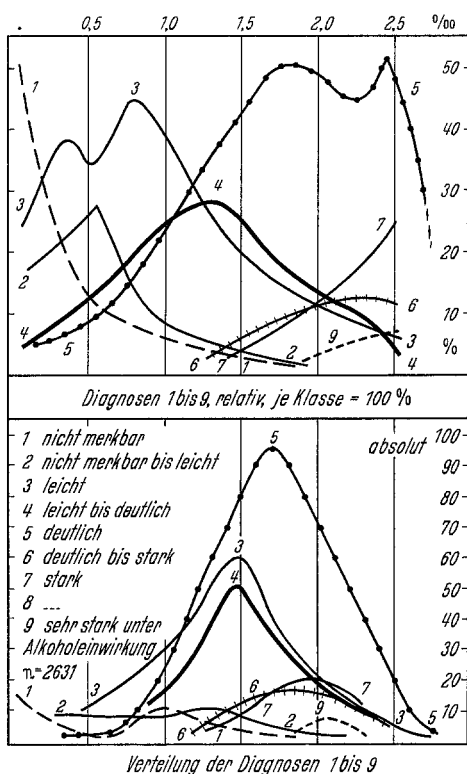


Abb. 1. Oben: Diagnosen 1—9, relativ, je Klasse = 100 %; unten: Verteilung der Diagnosen 1—9

### *Der diagnostische Ausdruck für einen Trunkenheitsgrad*

Die Ausgangssituation war die Beobachtung einer unbefriedigenden Zuordnung von Arztdiagnose und BAK. Sie ist deshalb zuerst zu besprechen. Wir boten mit Rücksicht auf eine detaillierte Angabe für das Ge-

richt und das ärztliche Bemühen um subtile Diagnosen sowie zur Verhütung unklarer Eigenschöpfungen eine Skala von 10 diagnostischen Ausdrücken zur Wahl an. Für die rechnerische Bearbeitung wurden sie mit einer Zensur — vgl. Abb. 2 — versehen. Die absolute Verteilung der Diagnosen folgt aus Abb. 1 unten, wo aus Raumgründen die Gesamtmasse, das bekannte Polygon mit dem Gipfel bei  $1,53\text{‰}$  ohne einen zweiten Gipfel bei etwa  $0\text{‰}$ , nicht dargestellt wurde. Die rechnerischen Ergebnisse finden sich in Tabelle 1. Dort und mit Abb. 1 oben ist der Versuch gemacht worden, der tatsächlichen diagnostischen Leistung

und der Streuung näherzukommen. Die starke Streuung vereitelt eine differenzierte Diagnostik; die bei den Kurven 3 und 5 erkennbare Doppelgipfligkeit läßt uneinheitliche Ausdeutung der Diagnosenvorschläge „leicht“ bzw. „deutlich“ vermuten. Praktisch werden nur leichte, mittlere und grobe Trunkenheitsgrade unterschieden, wie aus der in beiden Darstellungen sichtbaren Bevorzugung der Ausdrücke „leicht“, „leicht bis deutlich“ und „deutlich“ hervorgeht. Diese Ausdrücke sind verhältnismäßig unverbindlich. Die Darstellung bestätigt die bekannte, auch unter den Bedingungen der polizeilichen Untersuchung erhaltene, breite Variation möglicher Trunkenheitsbilder. Analoge Folgerungen lassen sich bereits aus den Kurven 0, H, L und M bei LAVES<sup>4</sup> ziehen.

Die Rechnung nach Tabelle 1 rechts bestätigt noch einmal die starke Überlappung bei der Wahl eines bestimmten diagnostischen Ausdrucks.

 Tabelle 1. Diagnosen *M, N, F, T* (1956) und *P* (1957)

Dia- gnosen ( <i>y</i> )	Absolut				Relativ			
	An- zahl	Verteilung	$M_x$	$\sigma_x$	Maxima aus Abb.1 oben	Planimetrisch bestimmte Streubereiche der relativen Verteilung. Es liegen etwa		
						68 %	95 %	99 %
						der Fälle in den Bereichen		
(1)	109	0—2,35	0,84	0,56	gegen 0	0—0,5	0—1,35	0—1,95
(2)	129	0—2,25	0,95	0,57	0,55	0,1—0,8	0—1,4	0—1,8
(3)	575	0—2,75	1,31	0,50	0,35, 0,80	0,2—1,4	0—2,0	0—2,7
(4)	446	0—2,75	1,45	0,43	1,30	0,7—1,9	0,2—2,4	0—2,6
(5)	1000	0—2,85	1,70	0,43	1,85, 2,45	1,3—2,4	0,4—2,7	0—2,8
(6)	154	1,15—2,65	1,89	0,35	2,30	1,6—2,4	1,3—2,5	
(7)	180	1,25—2,85	1,98	0,38	gegen 2,6	1,8—2,5	1,2—2,6	
(8)	5	1,55—2,55	2,05					
(9)	32	1,45—2,65	2,13	0,30	gegen 2,5	2,0—2,6		
(10)	1	2,65						
(1—10)	2631	0,00—2,85	1,53	0,54	Angaben in ‰ BAK ( <i>x</i> )			

Die Verteilung nähert sich weitgehend einer Normalverteilung, denn von 2631 Diagnosen liegen 63,23% im Bereich  $\pm 1\sigma_y$ , 92,13% innerhalb  $\pm 2\sigma_y$  und 98,54% innerhalb  $\pm 3\sigma_y$ . Die Gaussssche Verteilung gibt 68,27, 95,45 bzw. 99,73% Wahrscheinlichkeit dafür an, daß ein herausgegriffener Wert in das betreffende Streuungsintervall fällt (WEBER<sup>12</sup>). Statt mit etwa 3 muß bei versierten Ärzten noch mit etwa 15 Fehldiagnosen pro 1000 gerechnet werden. Bei Gelegenheitssachverständigen und unter schwierigen Verhältnissen müssen echte Fehldiagnosen noch häufiger werden (vgl. Tabelle 2,  $K_{56}: \sigma_x, \sigma_y$ ).

In Abb. 2 ist die aus den absoluten Zahlen gewonnene Streuung von *y* in diagnostischen Einheiten über den Klassen von *x* (BAK) vereinfacht dargestellt. Die Regressionsgerade  $R_y$  liegt auffallend flach,

weil in der Praxis offenbar gravierende Diagnosen vermieden werden. Infolge der starken Streuung finden sich bei geringer BAK im  $+3\sigma$ -Bereich und bei hoher BAK im  $-3\sigma$ -Bereich statistisch „richtige“ Diagnosen, die mit der landläufigen Vorstellung nicht in Einklang zu bringen sind. Das bekannte Entlastungsvorbringen, eine hohe BAK sei fehlerhaft, weil die Diagnose günstig laute, wie der Vorhalt, bei ungünstiger Diagnose sei eine niedrige BAK unwahrscheinlich, ist also unberechtigt. Auch eine enorme Vermehrung der Beobachtungszahl wird hieran nichts ändern. Die Versager bei niedriger BAK veranlassen gerade zur elastischen Handhabung der 1,5 $^0$ / $_{00}$ -Grenze. SEIFERT<sup>7</sup> stellte sie bezüg-

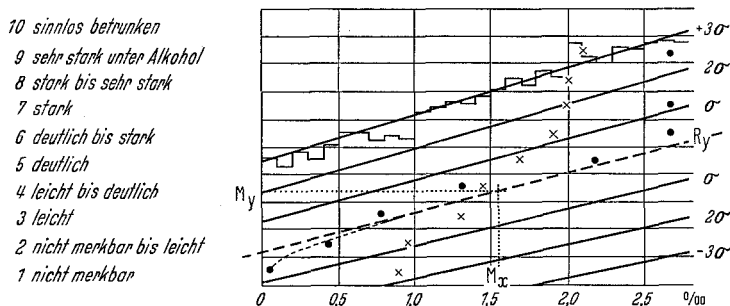


Abb. 2. Streuung von 2631 Diagnosen mit den Schwerpunkten für die diagnostischen Ausdrücke.  $\times$  absolut;  $\bullet$  relativ; ---- arithmetisches Mittel

lich der nordischen Rechtsprechung heraus, übergang aber a. a. O. die Anwendbarkeit der §§ 2, 71 StVZO bezüglich der Probanden, die ein unverhältnismäßig gutes klinisches Bild boten.

In der bisher besprochenen Gesamtverteilung ist die Leistung mehrerer Ärzte enthalten. Einzelnen betrachtet finden sich auch bei ihnen gewisse Unterschiede, die in subjektiven Momenten begründet sein müssen. Die Korrelationskoeffizienten der Tabelle 2 sind auffallend niedrig, obwohl es sich um erfahrene Untersucher handelt. Die Unterschiede bei  $r$  sind für P/T, M/T und F/T signifikant, für F/N, M/N und P/N zweifelhaft, im übrigen keine echten Differenzen (WEBER<sup>13</sup>). N und T hielten wir für die sichersten Diagnostiker. Bei der Chirurgengruppe K ist die besondere Schwierigkeit der Untersuchung Verletzter am niedrigen  $r$  und hohen  $\sigma$  zu erkennen. Die Steigung der Regressionskonstanten  $R_y$  zeigt bei M das umfangreichste Repertoire an Diagnosen an, T und F sind wesentlich zurückhaltender gewesen.

Eine Diagnose ist also mehr, als man es bisher gewohnt war, als statistischer Entscheidungsakt zu betrachten. Auch versierte Ärzte kommen bei einer scheinbar so einfachen Feststellung wie des Trunkenheitsgrades zu merklichen (individuellen?) Abweichungen. Es ist also zu fragen, ob die Diagnosen überwiegend intuitive Arztleistungen waren

Tabelle 2. *Diagnostische Leistung der einzelnen Untersucher (Korrelationsrechnung)*

		Untersucher						Gesamt
		M <sub>56</sub>	N <sub>56</sub>	F <sub>56</sub>	T <sub>56</sub>	K <sub>56</sub>	P <sub>57</sub>	
Anzahl	<i>n</i>	431	556	187	680	270	507	2631
$\%_{00}$ BAK	<i>M<sub>x</sub></i>	1,61	1,52	1,47	1,51	1,51	1,55	1,53
Mittel	<i>M<sub>y</sub></i>	4,68	4,02	4,23	4,52	4,16	5,78	4,33
(diagnostische Einheit)								
Streuung	$\sigma_x$	0,51	0,53	0,51	0,56	0,63	0,54	0,54
	$\sigma_y$	1,53	1,60	1,06	1,13	1,89	1,40	1,50
Korrelationskoeffizient	<i>r</i>	0,57	0,43	0,61	0,35	0,46	0,55	0,53
		$\pm 0,10$	$\pm 0,10$	$\pm 0,14$	$\pm 0,10$	$\pm 0,14$	$\pm 0,10$	$\pm 0,04$
Regressionskonstante	<i>R<sub>x</sub></i>	0,193	0,144	0,299	0,173	0,156	0,213	0,190
	<i>R<sub>y</sub></i>	1,72	1,32	1,27	0,69	1,39	1,45	1,47

Klassengröße  $x = 0,1\%_{00}$ ,  $y = 1$  diagnostische Einheit.

oder ob sie durch die üblichen Trunkenheitsproben gefördert oder etwa ungünstig beeinflußt worden sind.

Tabelle 2 läßt überdies erkennen, daß man sich bei der späteren Benutzung von Arztangaben vor Gericht weniger auf die Gesamtleistung des nachträglich gebildeten Teams, in der die individuellen Unterschiede verwischt sind, als viel mehr auf den einzelnen Untersucher einstellen muß. Insoweit ist es berechtigt, wenn sich die Richter ein Urteil über den „Wert“ der Diagnosen eines bestimmten Arztes zu bilden suchen.

### *Grobe Trunkenheitszeichen als Anhalt für die Schätzung eines Trunkenheitsgrades*

Die breite Diskussion des Ausfalls der Proben bei sog. Grenzfällen vor Gericht und die Erörterungen des vorigen Abschnittes forderten zu einer Prüfung der Leistungsfähigkeit der Proben, wie sie schon in LAVES a. a. O. S. 198 versucht ist, auf. Der Ausfall der Proben wird überwiegend nach den Alternativen *bestanden*: *nicht bestanden* registriert. Die Verteilungen kommen einer (allerdings sehr flachen) Normalverteilung nahe, so daß der Zweizeilenkorrelationskoeffizient berechnet werden durfte. Es wurde hier mit Klassen von  $x = 0,2\%_{00}$  gerechnet, weil die Werte ohnehin mit den anderen Tabellen nicht vergleichbar sind, auch wurden nur die Befunde 1956 ohne Chirurgen ausgewertet. Die gegenüber der Diagnosenzahl in Tabelle 2 fehlenden Angaben sind auf den Vordrucken nicht ausgefüllt. Die hier auftretenden negativen Koeffizienten besagen Nichtbestehen der betreffenden Probe bereits unterhalb der BAK des Mittels  $M_x$  der Ausgangsverteilung, die Verteilung der Alternative „nicht bestanden“ war stets geringer besetzt.

Tabelle 3 läßt die sehr unterschiedliche Bewertung der angeführten Proben durch die einzelnen Untersucher erkennen. Beim Zeigeversuch,

Tabelle 3. *Zweizeilenkorrelationskoeffizient einiger Trunkenheitsproben*

	T	F	N	M	Gesamt
Gang	+ 0,40 ± 0,10	+ 0,39 ± 0,20	+ 0,47 ± 0,10	+ 0,57 ± 0,10	+ 0,216 ± 0,015
Blindgang	+ 0,37 ± 0,10	- 0,41 ± 0,19	- 0,55 ± 0,09	- 0,35 ± 0,13	+ 0,197 ± 0,068
Aufhebversuch	+ 0,33 ± 0,03	- 0,57 ± 0,21	+ 0,30 ± 0,11	+ 0,29 ± 0,14	+ 0,466 ± 0,063
Zeigerversuch	- 0,31 ± 0,04	- 0,37 ± 0,20	+ 0,29 ± 0,21	+ 0,65 ± 0,09	+ 0,125 ± 0,073
Sprechweise	+ 0,79 ± 0,05	+ 0,29 ± 0,21	+ 0,16 ± 0,11	+ 0,48 ± 0,10	+ 0,415 ± 0,061
Benennen	+ 0,37 ± 0,10	+ 0,39 ± 0,20	+ 0,22 ± 0,11	+ 0,43 ± 0,12	+ 0,343 ± 0,065
Urteilsfähigkeit	+ 0,41 ± 0,10	+ 0,29 ± 0,21	+ 0,48 ± 0,10	+ 0,47 ± 0,12	+ 0,387 ± 0,060

den T und F bei niedriger, M erst bei höherer BAK nicht bestanden finden, wird wohl verschiedene Technik angewandt. In der Gesamtverteilung werden Blindgang und Zeigerversuch am frühesten nicht mehr bestanden. Die Übereinstimmung mit der Auffassung bei LAVES a. a. O. S. 199 ist aber nur teilweise: nach Abb. 4 sind diese Proben zu empfindlich, um etwas über den Trunkenheitsgrad aussagen zu können. Abb. 3

Tabelle 4. *Umfang der Verteilungen zu Tabelle 3*

	T <sub>ss</sub>	F <sub>ss</sub>	N <sub>ss</sub>	M <sub>ss</sub>	Gesamt
Gang . . .	646	170	520	403	1739
Blindgang .	647	175	473	384	1679
Aufhebversuch .	627	167	492	389	1675
Zeigerversuch	638	164	477	397	1676
Sprechweise	643	171	552	402	1768
Benennen .	668	177	548	410	1803
Urteilsfähigkeit	665	175	556	407	1803

demonstriert die Verteilung einiger bestandener (ausgezogen) bzw. nicht bestandener (punktirt) Proben, wobei die Fallzahl jeweils = 100 gesetzt ist. Die Ausfälle weisen ein Defizit gegenüber der Erwartung auf, das auch in Abb. 4 deutlich wird. Hier wird sinnfälliger, daß die brauch-

barsten Ergebnisse mit den Proben erreicht werden können, die der Untersuchte nicht als solche erkennt: Gang, Urteilsfähigkeit, Benennen und Sprechweise. Die Beurteilung der Sprechweise ist in der Großstadt sicher schwierig, weil neben das ursprüngliche Idiom zahlreiche fremde Mundarten getreten sind. Wenn ein Großteil sicher als betrunken zu bezeichnender Personen „vor den kritischen Augen des Polizeiarztes“ auch die letztgenannten Proben noch besteht, liegt die Vermutung nahe, daß eine willensmäßige Einstellung auf die Untersuchung für deren Dauer vielfach noch möglich ist. Die besprochenen Proben können also

der Wahl eines geeigneten Ausdrucks für einen bestimmten Trunkenheitsgrad nur wenig förderlich sein, der Arzt ist auf seinen klinischen Blick angewiesen. Es wäre allerdings vom Sachverständigen unärztlich gehandelt und würde vom Laien nicht verstanden werden, wenn man vom Verhalten des Untersuchten ganz absehen und einzig Schlüsse aus der BAK ziehen wollte. Wenn aber selbst versierten Ärzten — wie

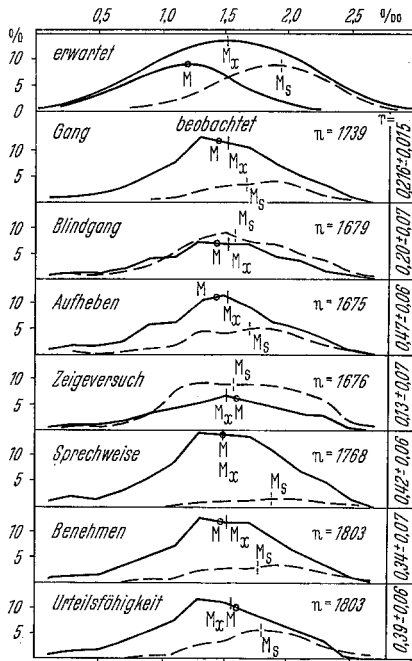


Abb. 3

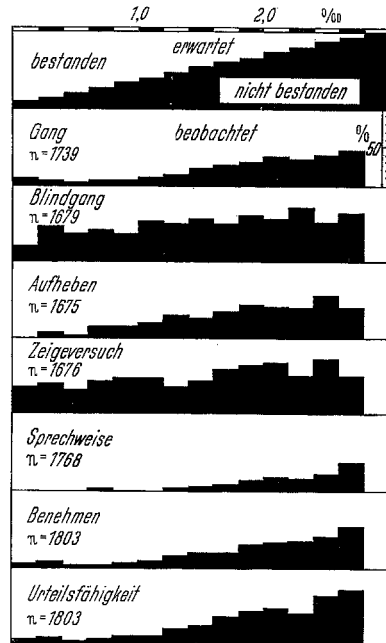
Abb. 3. Grobe Trunkenheitsproben  $n = 100 \%$ 

Abb. 4

Abb. 4. Grobe Trunkenheitsproben je Klasse ( $0,2 \text{ ‰}$ ) =  $100 \%$ 

die Abb. 3 und 4 eindringlich zeigen — alkoholbedingte Ausfälle entgehen oder verborgen werden können, wie so häufig bei Laienbeobachtung, müssen grob schätzende Methoden verlassen oder mindestens durch den Versuch einer objektiven Messung ergänzt werden. Der Versuch konnte nur im Routinebetrieb zu gültigen Ergebnissen kommen. Damit war der Verzicht auf jeden Apparat verbunden.

#### *Die Dauer des Dreh-Nachnystagmus als Funktion der Blutalkoholkonzentration*

Als TASCHEN<sup>10</sup> 1954 in Kiel die Nystagmusprobe im Film demonstrierte, besaßen wir für die Zuordnung der Nystagmusdauer zur BAK experimentell gewonnene Daten, die bereits gestatteten, ein arithmetisches

Mittel zu berechnen. Die damalige Kurve wurde nicht veröffentlicht, weil wir ein Absinken bei den höchsten ‰-Werten (wie bei SCHULTE und ROTH<sup>6</sup>, vgl. auch Abb. 5), das wir für den Ausdruck einer einsetzenden Lähmung hielten, statistisch nicht sichern konnten. Mit

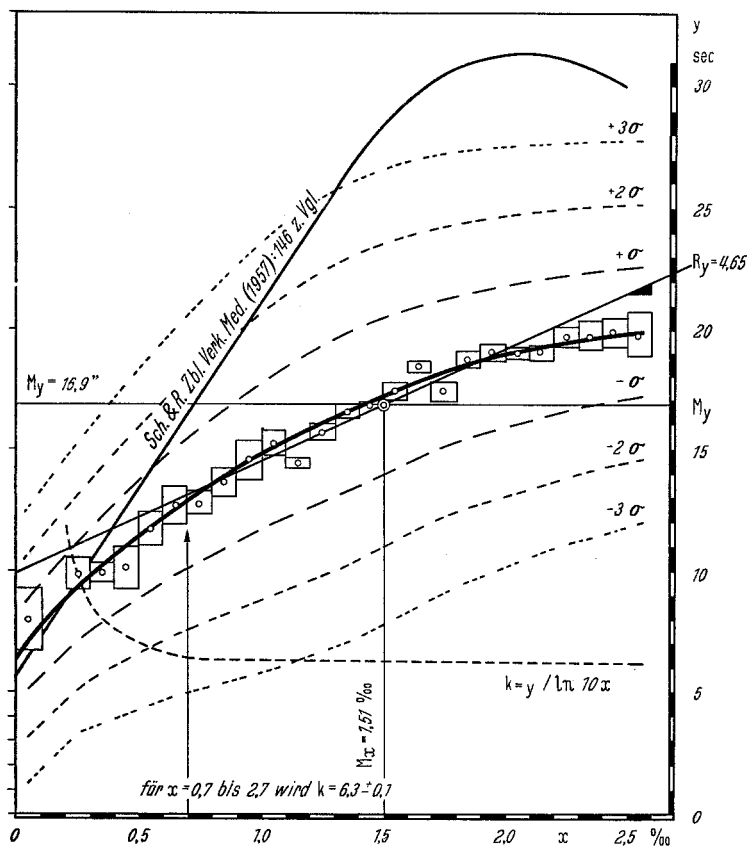


Abb. 5. Nystagmusdauer  $y$  als Funktion der Blutalkoholkonzentration  $x$ .  $y = k \cdot \ln 10 x$

wachsender Fallzahl ist diese Erscheinung geschwunden. Bei gleichbleibender Technik (s. Anhang) gewannen T 1955—1957 und F 1956 bis 1957 zusammen 1723 Messungen. Die Ergebnisse beider Untersuchungen deckten sich in kurvenmäßiger Darstellung und in der Rechnung so gut, daß wir sie als *eine* statistische Masse behandeln und der Konstruktion der Abb. 5 zugrunde legen durften. Von den 1723 Messungen fallen 65,58, 90,95 bzw. 99,36% in die Bereiche  $\pm 1,2$  bzw.  $3\sigma_y$ , und 11=0,64% eindeutig fehlerhafte Werte außerhalb. Die Verteilung kommt mithin einer Gauss'schen Verteilung sehr nahe. Aber auch über den einzelnen Klassen von  $x$  sind analoge Verteilungen der Nystagmus-



dauer  $y$  zu beobachten. Die Konstruktion der Mittelwertskurve erschien daher erlaubt. Die theoretischen Schwankungen der Kolonnenmittel sind als Rechtecke eingezeichnet. Die Kurve stellt die Nystagmusdauer als Funktion der BAK dar, um die sich die bei jeder Messung unvermeidlichen Streuungsbereiche gruppieren. Die Funktion ist nicht linear, die Berechnung des zugehörigen mathematischen Ausdrucks nach der Interpolationsmethode NEWTONS<sup>2</sup> und der Differenzmethode stieß zunächst auf Schwierigkeiten. Wir fanden schließlich, daß sie im Bereich von  $0,7^0_{00}$  bis Ende einer *logarithmischen* Kurve von der Form

$$y = k \cdot \ln 10x$$

entspricht, wenn  $k = 6,3 \pm 0,1$  ist. In diesem Bereich sind auch die Differenzen der Kolonnenmittel statistisch nach WEBER<sup>13</sup> gesichert. Da mit verfeinerter Methodik durchgeführte Untersuchungen darauf hinweisen, daß die Erscheinung nicht einheitlich ist, muß daran gedacht werden, daß der Anfangsteil der Kurve tatsächlich einer anderen Gesetzmäßigkeit gehorcht. Solche Funktionen sind aus der Technik bekannt. Der Zweck der Untersuchung ist aber bereits erreicht, wenn der *forensisch interessierende Teil definiert* ist. Es ist schwierig, für die geringen BAK eine ausreichende Fallzahl aus der Praxis zu erhalten, denn die eingehenden Proben entsprechen nicht mehr der Verteilung, wie sie Verf. noch 1937 für BUHTZ<sup>1</sup> zeichnen konnte. Die Erkenntnis einer einfachen logarithmischen Abhängigkeit ist unseres Erachtens ein starkes Argument für die Einführung der Probe auf breiter Basis.

Die zur Auffindung der Beziehung führende Kurve  $k = y/\ln 10x$  könnte durch Einführung von  $100x$  noch etwas verbessert werden. Dann nimmt jedoch die interessierende Gleichung eine unbequeme Form an und es verschwinden 2 Auf- oder Zufälligkeiten:  $k$  hat die Größenordnung des Normalwertes von  $y$  beim Nüchternen und ist das 10fache des Korrelationskoeffizienten.

An der unteren Grenze des forensisch interessierenden Bereichs ist die mittlere Nystagmusdauer gegenüber der Norm etwa verdoppelt ( $0,8^0_{00}:14$  sec), bei den höchsten BAK etwa verdreifacht. Die praktische Anwendbarkeit hängt wesentlich vom Streuungsmaß und der Reproduzierbarkeit durch verschiedene Untersucher ab, was wir auf dem Wege der Korrelationsrechnung prüften.

Obwohl nach WEBER<sup>14</sup> bei nichtlinearer Verbundenheit besser das Korrelationsverhältnis gebildet werden sollte, hielten wir die im Hinblick auf die nahezu normale Verteilung auch erlaubte Errechnung des Korrelationskoeffizienten zu  $r = 0,63 \pm 0,04$  und der Regressionskonstanten  $R_y = 4,65$  für zweckmäßiger\*. Das Korrelationsverhältnis ist weniger bekannt, es kann nur zwischen 0 und  $\pm 1$  liegen und ist stets

\*  $R_y$  zeigt das Fortschreiten um  $n$  sec Nystagmusdauer an, wenn  $x$  um die Einheit ( $1,0^0_{00}$ ) wächst.

größer als  $r$ , wodurch leicht Fehldeutungen möglich werden. Der Vergleich der beteiligten Ärzte (ohne N, der die Probe kaum angewandt hat) ergibt sich aus Tabelle 5.

Die Mittelwerte  $M_y$  liegen bei den verschiedenen Untersuchern mit Ausnahme der unter besonders schwierigen Verhältnissen prüfenden Krankenhausärzte bei 16—17 sec, das angebotene Material ist vergleichbar ( $M_x$ ). Die Streuung  $\sigma_y$  hängt offenbar von der Übung ab, die bei P geringer als bei T und F ist. Die Unterschiede in den Korrelationskoeffizienten sind ohne Kenntnis der Urtafel nicht ohne weiteres zu

Tabelle 5. *Nystagmusprobe (Korrelationsrechnung)*

	T 1955—1957 F 1956—1957	M 1956	K 1956	Gesamt	P 1957 zum Vergleich
$n$ . . .	1723	289	100	2112	467
$M_x$ $^{\circ}/_{00}$	1,51	1,59	1,53	1,52	1,49
$M_y$ sec	16,9	16,2	13,7	16,3	16,5
$\sigma_x$ $^{\circ}/_{00}$	$\pm 0,535$	$\pm 0,466$	$\pm 0,519$	$\pm 0,529$	$\pm 0,522$
$\sigma_y$ sec	$\pm 3,9$	$\pm 4,8$	$\pm 4,7$	$\pm 4,8$	$\pm 6,1$
$r$ . . .	$0,63 \pm 0,04$	$0,3 \pm 0,16$	$0,63 \pm 0,18$	$0,47 \pm 0,05$	$0,42 \pm 0,12$
$R_x$ . .	0,085	0,028	0,070	0,051	0,036
$R_y$ . .	4,65	3,03	5,74	4,34	4,89

Klassengröße  $x = 0,1^{\circ}/_{00}$ ,  $y = 1$  sec.

verstehen. Auf der Zähltafel für M sind eigenartigerweise nur die Sekunden 0, 5, 8, 10, 12, 15, 20, 25 (und einigemal 30) über die ganze Skala von 0— $2,8^{\circ}/_{00}$  ganz abweichend von den diagonalen Feldern der übrigen Tafeln dichter besetzt. Hier kann nur schematisiert oder ohne Stoppuhr gearbeitet worden sein. Bei P sind in der Zeit der Einarbeitung — die bei allen verhältnismäßig lang war! — durch rasche Wiederholung der Probe oder ungleichmäßige Reizsetzung Zeiten beobachtet worden, die aus dem Rahmen unserer Norm herausfallen. Mangels entsprechender Kennzeichnung waren sie aber von der statistischen Bearbeitung nicht auszunehmen, haben also die Streuung erheblich beeinflußt. Trotzdem liegt die Mittelwertskurve von P nur 0,5 sec niedriger als die von T und F. Wenn sich aber eine solche Übereinstimmung ergibt, darf man annehmen, daß die Werte von T und F im Prinzip bestätigt worden sind. Von den 467 Messungen  $P_{57}$  liegen 56,9, 84,0 bzw. 96,5% in den Bereichen  $\pm 1,2$  bzw.  $3\sigma_y$  und sind 3,5% sicher grob fehlerhaft. Eine neue Berechnung unter Ausschluß dieser 3,5% hätte keine wesentliche Verbesserung im Sinne einer Annäherung an die Gauss'sche Kurve gebracht. Wir beobachten jetzt bei P eine viel geringere Streuung in den täglichen Eingängen.

In Abb. 5 wurde die Kurve von SCHULTE und ROTH<sup>6</sup> zum Vergleich übernommen. Sie geht von unserem Nüchternwert aus, steigt aber viel

steiler an, so daß man eine stärkere Reizsetzung als nach der Originalvorschrift annehmen muß. Sie haben die Fehlerquellen, die auch in unserer Statistik noch nicht auszumerzen waren, bereits gekannt. Es sind dies vorzugsweise: nicht streng axiale Drehung des Probanden, verschiedene Drehbeschleunigung und -verzögerung und Zweifel über das Ende des ohne Hilfsmittel beobachteten Augenrucksens. Für weitere Untersuchungen ist also eine kritische Bemerkung der Untersucher zur Durchführbarkeit der Probe unerlässlich, denn sie prüfen *nicht* unter den Bedingungen des Trinkversuchs mit willigen und eingeübten Probanden.

Trotz dieser erheblichen Erschwernis haben wir zur Orientierung die 570 Eingänge eines Vierteljahres 1956, die Nystagmusmessungen nebst Diagnosen enthalten, mit Hilfe der Korrelationsrechnung verglichen. Es mußte dabei eine von den mitgeteilten Tabellen abweichende Klassengröße gebildet werden, die bezüglich der Diagnosen je 2 Einheiten der Abb. 2 zusammengefaßt, weil vorläufig nicht mehr als etwa folgende Zuordnung der Nystagmuszeiten zur Trunkenheitsdiagnose vertretbar erschien (Tabelle 6).

Tabelle 6. *Vorläufiger Versuch einer Zuordnung von Nystagmusdauer und Trunkenheitsdiagnose*

Vereinfachte Diagnose	Vorläufig zugeordnete Nystagmusdauer (sec)	Entspricht Diagnose nach Abb. 2
I sehr leicht	bis 8	1 nicht merkbar 2 nicht merkbar bis leicht
II leicht	9—13	3 leicht 4 leicht bis deutlich
III mittel	14—17	5 deutlich 6 deutlich bis stark
IV stark	18 und 19	7 stark 8 stark bis sehr stark
V sehr stark	20 und mehr	9 sehr stark 10 sinnlos betrunken

Die Durchrechnung ergab für die Originaldiagnosen ( $n=570$ ):  $r=0,83 \pm 0,04$ ,  $M_x=1,49^{0/00}$  (Klassengröße  $0,1^{0/00}$ ),  $M_y=2,58$  vereinfachte Einheiten,  $\sigma_x=\pm 0,50$ ,  $\sigma_y=\pm 0,37$ ,  $R_y=0,61$ . Eine nachträgliche Diagnosenstellung allein auf Grund der Einteilung nach Tabelle 6 führte zu folgenden Ergebnissen:  $r=0,70 \pm 0,06$ ,  $M_x=1,49$ ,  $M_y=2,46$ ,  $\sigma_x=\pm 0,51$ ,  $\sigma_y=\pm 1,1$ ,  $R_y=1,51$ .

Die beträchtliche Änderung von  $R_y$  führt zu einer Regressionsgeraden, die in der wünschenswerten Lage diametral durch ein Feld wie in Abb. 2 verläuft, also eine bessere Zuordnung der Diagnosen zur BAK ergibt. Die Originaldiagnosen haben eine sehr flache Regressionsgerade, weil die unverbindlichen Diagnosen bevorzugt waren. Da aber unbereinigte Nystagmuszeiten verwendet werden mußten, ergab sich eine viel zu starke Streuung  $\sigma_y=\pm 1,1$  Einheiten. Man kann also vorläufig

nur als Tendenz feststellen, daß sich eine tatsächliche Verbesserung der diagnostischen Aussage erwarten läßt, wenn es gelingt, einwandfreie Nystagmuszeiten zu erhalten. Hierüber soll zu gegebener Zeit berichtet werden.

Unter der Voraussetzung einer Verteilungskurve mit höherem Exzeß — also einer schlankeren Glockenkurve der einer bestimmten Nystagmusdauer zugeordneten BAK — erwarten wir noch weitere Aufschlüsse bei Anwendung der Nystagmusprobe. Wir denken an Zusammenhänge von Nystagmusdauer und Restalkohol, frischer und alter Hirnschädigung, Übermüdung, Störungen seitens des Sehorgans und des Gleichgewichtssinnes. Bis jetzt haben wir bei den vereinzelt Fällen von Taubstummheit (2), Einäugigkeit (1) und vermutlicher Hirnerschütterung (etwa 50) keine auffällig abweichenden Nystagmuszeiten beobachtet. Einer der Untersucher (P) vermutet Zusammenhänge mit der vegetativen Ausgangslage.

Nachdem wir gesehen haben, daß bei Einhaltung der Technik grundsätzlich übereinstimmende Mittelwerte, wenn auch mit zufallsbedingten Meßfehlern in Form einer mehr oder weniger breiten Streuung, zu erzielen sind, glauben wir die Probe als Mittel zur Verfeinerung und Ergänzung der Trunkenheitsdiagnostik empfehlen zu dürfen. Man wird sich allerdings auch hier einer überspitzten Diagnostik enthalten müssen. Ebenso wie mit Hilfe der üblichen Trunkenheitsproben sind hier nicht 10 differenzierte Diagnosen, sondern nur etwa 5, wie in Tabelle 6, möglich. Wenn das Rüstzeug des Polizeiarztes um eine vom Willen des Untersuchten unabhängige Probe, die mit den üblichen Proben wenig gemeinsam hat, ergänzt wird, darf man aus der Trias: *Intuitives Erfassen*, *Schätzungsmethoden* und *messende Nystagmusprobe* eine Verbesserung der diagnostischen Leistung erwarten. Auf diese Verbesserung kommt es aber bei der Überzahl routinemäßig zu bewältigender Gerichtsverfahren im Interesse aller Beteiligten entscheidend an.

### Zusammenfassung

Der Ausfall der üblichen Trunkenheitsproben ist dem Trunkenheitsgrad nicht in wünschenswertem Maße zugeordnet. Die auftretenden Streuungen entsprechen der breiten Variation vorkommender Trunkenheitsbilder. Verhältnismäßig brauchbar sind die Proben, die vom Untersuchten nicht als solche erkannt werden (Gang, Benehmen, Sprechweise, Urteilsfähigkeit). Blindgang und Zeigerversuch fördern die Wahl eines diagnostischen Ausdrucks nicht. Der Untersucher ist von subjektiven Schätzungen und intuitiver Erfassung der Situation abhängig. Es werden seltener gravierende Diagnosen gestellt, als man erwarten könnte. Die Drehnachnystagmusprobe nach TASCHEN ist unabhängig vom Willen des Untersuchten. Die gemessenen Zeiten unterliegen wie jede dia-

gnostische Entscheidung der zufallsbedingten Streuung. Sie sind der Blutalkoholkonzentration korrelativ zugeordnet und reproduzierbar. Für den Bereich von 0,7—2,8‰ Blutalkoholkonzentration wird eine mathematische Formulierung für den Verlauf der Mittelwertskurve angegeben. Eine Diagnosenstellung allein mit Hilfe der Nystagmusprobe zeigt bereits einen optimalen Verlauf der Regressionsgeraden für die Zuordnung von diagnostischem Ausdruck und Blutalkoholkonzentration. Es wird empfohlen, nur etwa 5 verschiedene Grade von Trunkenheit zu diagnostizieren.

*Die praktische Durchführung der Nystagmusprobe nach TASCHEN<sup>9</sup>*

Nach Unterrichtung über den Gang der Probe wird die aktiv mitwirkende Versuchsperson vom Untersucher in 10 sec 5mal exakt um die Körperlängsachse gedreht. Der Kopf soll keine Kreisbewegung machen, die Füße drehen auf der Stelle. Beim Anhalten soll das Gesicht möglichst wenig in entgegengesetzten Drehsinn geraten. Mit dem Anhalten gibt der Untersucher die Stoppuhr frei, der Untersuchte fixiert einen in Sehweite gehaltenen Finger oder Bleistift. Wenn der grob- bis mittelschlägige Nystagmus nicht mehr erkennbar ist, wird gestoppt. Eine Wiederholung soll erst nach längerer Pause erfolgen. Die Augen dürfen nicht geschlossen werden.

### Literatur

- <sup>1</sup> BUHTZ, G.: Der Verkehrsunfall, S. 103. Ferdinand Enke 1938. — <sup>2</sup> KÖRWIEN, H.: Graphisches Rechnen, S. 175. Leipzig 1949. — <sup>3</sup> KOLLER, S.: Statistik — ärztliche Erfahrung — und der Einzelfall. Therap. Ber. **30**, 77 (1958). — <sup>4</sup> LAVES-BITZEL-BERGER. Der Straßenverkehrsunfall, S. 196. Ferdinand Enke 1956. — <sup>5</sup> RAUSCHKE, J.: Alkohol-Nystagmus und Leistungsschädigung. Medizinische **1958**, 460. — <sup>6</sup> SCHULTE, K., u. H. ROTH: Drehnystagmus und Blutalkoholkonzentration. Zbl. Verkehrs-Med. **3**, 144 (1957). — <sup>7</sup> SEIFERT, P.: Nachdenkliches zur Praxis der verkehrsmedizinischen Alkoholbegutachtung in Deutschland. Ärztl. Lab. **4**, 60 (1958). — <sup>8</sup> TASCHEN, B.: Nystagmusprobe zur Feststellung der Alkoholbeeinflussung. Kriminalistik **8**, 180 (1954). — <sup>9</sup> TASCHEN, B.: Eine einfache Nystagmusprobe zur Feststellung der Alkoholbeeinflussung. Med. Mschr. **25** (1955). — <sup>10</sup> TASCHEN, B.: Nystagmusprüfungen zur Feststellung der Alkoholbeeinflussung. Dtsch. Z. ges. gerichtl. Med. **44**, 436 (1955). — <sup>11</sup> WAGEMANN, E.: Narrenspiegel der Statistik, 2. Kap. München: Lehnen 1950. — <sup>12</sup> WEBER, E.: Grundriß der biologischen Statistik, S. 100. Jena: Gustav Fischer 1948. — <sup>13</sup> WEBER, E.: a. a. O., S. 181. — <sup>14</sup> WEBER, E.: a. a. O., S. 223.

Dr. EBERHARD LOMMER, Universitätsinstitut für gerichtliche Medizin,  
Köln, Zulpicher Straße 47