

Aus der Prosektur des Bezirkskrankenhauses „Am Sund“ Stralsund
(Prosektor: Dr. med. H. GROSSE)

Über „Berksons Fallacy“ und die Selektion durch den Tod

Von

HANS GROSSE

(Eingegangen am 21. November 1963)

Trotz zahlreicher kritischer Stimmen (CORNET, BERKSON, MAINLAND, GROSSE, KRAUS, KOLB, VAN DER LINDEN, CORNFIELD, PARRISH und EBERLY, NIELSEN und KESSING, OPDAHL) hat sich die Methodik der Sektionsstatistik nicht verbessert. Immer wieder erscheinen in medizinischen Zeitschriften und Büchern Arbeiten, welche behaupten, durch Vergleich von Sektionsreihen gleichen Alters und Geschlechtes „mathematisch exakt“ Häufigkeitsunterschiede bewiesen zu haben. Auch die den Medizinern beratend zur Seite stehenden Statistiker haben sich kaum mit dem störenden Einfluß der Absterbegeschwindigkeit befaßt.

CORNET sagte 1899, daß nur bei Bezug auf die Lebenden gleichen Alters und Geschlechtes Häufigkeitsschwankungen der Tuberkulose oder anderer Erkrankungen bewiesen werden können, niemals aber beim Vergleich von Sektionsreihen allein. 1946 wies BERKSON auf die trügerischen Prozentsatzunterschiede bei Krankheitskombinationen hin, und 1953 prägte MAINLAND den Terminus „Berksons Fallacy“ für solche Trugschlüsse.

In USA haben KRAUS, CORMFIELD, PARRISH und EBERLY diese Erkenntnisse propagiert, in Schweden VAN DER LINDEN und im deutschen Schrifttum kamen GROSSE wie KOLB unabhängig voneinander zu ähnlichen Einwänden gegen die traditionelle Methodik. KOLB kritisierte die Formel nach PFAUNDLER und v. SEHT, und GROSSE suchte anhand von verschiedenen Syntropien im Sektionsgut das vorliegende Problem darzustellen.

1. Selektion der Verstorbenen

So sollen hochgradige Formen der Arteriosklerose und schwere Folgen derselben (Apoplexie und Herzinfarkt) bei Tuberkulose, Lebercirrhose, Carcinom, Sarkom und Hämblastose seltener sein, als der Erwartung entspricht. Maligne Geschwülste sollen bei Tuberkulose, bei Lues, bei Diabetes mellitus und bei Ulcus gastroduodenale seltener sein, als der Erwartung entspricht. — Würde ein Tyrann alle Lebenden einer bestimmten Altersklasse töten und sezieren lassen, so würde er „signifikant seltener“ eine schwere Arteriosklerose oder einen Herzinfarkt finden als unter den gleich alten „spontan“ Verstorbenen. 1958 hat GROSSE die Skleroseintensität der Krebsfälle mit der der Unfall-, Mord- und Selbstmordfälle verglichen und keine Unterschiede gefunden. Auch PARRISH und EBERLY hielten 1962 die Skleroseintensität der Unfalltoten für repräsentativer als die der sog. Kontrollsektionen und fanden keinen Unterschied zwischen dem Arteriosklerosegrad der Unfall- und der Lebercirrhosesektionen.

Während sich BERKSON, CORMFIELD, PARRISH und EBERLY u.a. darauf beschränkt haben, die Unzulässigkeit des Vergleiches gleich alter Sektionsreihen

darzustellen, wenn sie „Berksons Fallacy“ unterliegen, hat GROSSE versucht, diesen störenden Einfluß zu eliminieren. Dagegen ist einzuwenden, daß zu viele verschiedene nicht exakt zu analysierende Störungen die Auslese des Sektionsgutes beeinflussen. Wenn man aber in einer relativ kleinen Stadt wie Stralsund Gelegenheit hat, einen relativ großen Anteil der Verstorbenen zu sezieren und bei Ausdehnung der Sektionen auf fragwürdige Todesfälle keine grundsätzlichen Modifikationen der Resultate erhält, liegt es nahe, die eliminierbaren Störungen zu definieren.

Unter einer Million lebender Männer im Alter von 50—55 Jahren kommen pro Jahr etwa 10000 Todesfälle vor. Unter dieser Million finden sich etwa 5000 (= 0,5%) Krebsfälle und 3000 (= 0,3%) aktive Tuberkulosen. Von den Krebsfällen versterben pro Jahr etwa 2000 (= 40%) und von den Tuberkulosekranken etwa 120 (= 4%). Wieviel Kombinationsfälle von Krebs und Tuberkulose sind nun unter den Lebenden und unter den Verstorbenen zu erwarten? Die traditionelle Methodik der Sektionsstatistik geht von dem Prinzip aus, daß auch unter 100 Tuberkulosekranken 20% Krebsfälle zu erwarten seien, da die Sektionsfälle ohne Tuberkulose etwa 20% Krebsfälle aufweisen. Das wären also 24 Kombinationsfälle. Diese Zahl kann aber nicht richtig sein, weil unter einer Million Lebenden *nur 15 Kombinationsfälle* vorkommen können. 15 lebende Kombinationsfälle können niemals pro Jahr 24 Todesfälle stellen. Mit anderen Worten, bei einer Tuberkulose-Morbidität der Lebenden (mit wie ohne Krebs) von 0,3% kann die Tuberkulose-Mortalität der Krebsfälle nicht 1,2% betragen.

Von den 15 lebenden Kombinationsfällen werden etwa 40% am Krebs versterben (also 6) und von den restlichen 9 werden etwa 4% zusätzlich an Tuberkulose sterben. Es kommen also 6 (oder 6,36) Todesfälle auf 15 lebende Kombinationsfälle (Tabelle).

Man kann die Erwartungswerte für die Kombinationsfälle also nicht durch einfachen Prozentsatzvergleich oder Vierfeldertafelanalyse bestimmen, sondern muß die Absterbequoten der untersuchten Krankheitszustände pro Jahr kennen. Dafür geben die Lebensversicherungen und die klinischen Daten nur annähernde Werte, denn die klinischen Angaben über Erkrankung A sind meist zu sehr bereinigt, durch Weglassen der ungünstigen A-Fälle einerseits und die am Leben bleibenden, fälschlich diagnostizierten A-Fälle andererseits. Auch unter den Angaben der Lebensversicherungen finden sich die diagnostizierten A-Fälle allein, es fehlen die nicht erfaßten A-Fälle, und die fälschlich als A-Fälle diagnostizierten verbessern das Resultat, da sie länger am Leben bleiben.

Die Tabelle (1—6) enthält weitere Beispiele für die fehlerhafte Methodik der traditionellen Sektionsstatistik.

Man kann diese Berechnung wesentlich vereinfachen, indem man die Absterberate der Kombinationsfälle derjenigen der ernsteren Erkrankung gleichsetzt. Dies ist deshalb möglich, weil in der Sektionsstatistik im allgemeinen alle Altersklassen zusammengefaßt werden. Denn die in den früheren Altersklassen zusätzlich verstorbenen Kombinationsfälle vermindern den Überlebendenbestand. Mit anderen Worten, die Kombinationsfälle werden zum Teil etwas früher ins Grab sinken, fehlen aber dann in den nächstfolgenden Altersklassen.

Setzt man nun die Absterberate der Gesamtpopulation gleich 1, so kann man die der untersuchten Erkrankungen als Vielfache davon berechnen. Nun

Tabelle

	Lebende		Tote/Jahr	%		
1	1000000 5000 3000 15	Gesamtpopulation Krebsfälle Tuberkulosefälle Kombinationsfälle	10000 2000 120 6	100 20 1,2 0,06	% der Carcinomfälle 0,3	% der Tuberkulosefälle 5
				24	0,24	1,2
						20
2	1000000 6000 7000 42	Gesamtpopulation Herzinfarktfälle Tuberkulosefälle Kombinationsfälle	20000 2000 1000 14	100 10 5 0,07	% der Herzinfarktfälle 0,7	% der Tuberkulosefälle 1,4
				100	0,5	5
						10
3	1000000 10000 15000 150	Gesamtpopulation Krebsfälle Diabetesfälle Kombinationsfälle	20000 4000 600 60	100 20 3 0,3	% der Carcinomfälle 1,5	% der Diabetesfälle 10
				120	0,6	3
						20
4	1000000 10000 6000 60	Gesamtpopulation Krebsfälle Herzinfarktfälle Kombinationsfälle	20000 4000 2000 40	100 20 10 0,2	% der Carcinomfälle 1	% der Herzinfarktfälle 2
				400	2	10
						20
5	1000000 4000 3000 12	Gesamtpopulation Herzinfarktfälle Tuberkulosefälle Kombinationsfälle	10000 1000 120 3	100 10 1,2 0,03	% der Herzinfarktfälle 0,3	% der Tuberkulosefälle 2,5
				12	0,12	1,2
						10
6	1000000 6000 1000 6	Gesamtpopulation Herzinfarktfälle Lebercirrhosefälle Kombinationsfälle	20000 2000 200 2	100 10 1 0,01	% der Herzinfarktfälle 0,1	% der Lebercirrhosefälle 1
				20	0,1	1
						10

Die umrandeten Werte entsprechen den Postulaten der traditionellen Methodik.

braucht nur noch der traditionelle Erwartungsprozentsatz oder die traditionelle Erwartungszahl an Kombinationsfällen durch die Absterberate jener Erkrankung geteilt zu werden, welche die weniger ernste Prognose hat. Niemals darf durch die Absterbequote der Erkrankung mit der infausteren Prognose dividiert werden. Dieses Verfahren ist auch für jene Fälle exakt, bei denen eine ernstere Krankheit mit einem Zustand durchschnittlicher Lebenserwartung (Irisfarbe, Haarfarbe, Geburtsmonat, Sterbemonat usw.) in Beziehung gesetzt wird. Denn hier ist durch die Absterberate 1 zu teilen, mit anderen Worten, das traditionelle Verfahren exakt.

Würde man nun die Berechnung der Fehlergrenzen anschließen, so würde man eine Exaktheit vortäuschen, welche niemals erreicht werden kann. Denn es ist die exakte Höhe der Absterbequote der gleich alten Population zwar sehr leicht zu erhalten, aber die der an den einzelnen Leiden Verstorbenen sehr schwer. Schließlich werden in diese Klinik diese, in eine andere Klinik andere Fälle eingewiesen und die Auslese des Sektionsgutes ist nur dann zu eliminieren, wenn

annähernd alle Todesfälle seziert werden. Es ist also in sektionsfreudigeren Populationen leichter, die Sektionsfälle als repräsentativ für alle Verstorbenen anzusehen, während in sektionsfeindlichen Populationen die Bedenken gegen die Auswahl des Sektionsgutes größer sein müssen. Wir erhalten aber bei Division des traditionellen Vergleichswertes durch die Absterbequote der weniger ernsten Erkrankung einen Näherungswert, welcher den tatsächlichen Verhältnissen weit gerechter wird, als die traditionelle Methodik es vermag. Bei allen rein zufälligen Kombinationen von Zuständen mit erhöhter Absterbequote muß diese angeblich „hochsignifikante“ negative Korrelationen beweisen, welche nichts als eine Illusion darstellen, bedingt durch „Berksons Fallacy“. Ebenso muß sie bei annähernd gleichen Hundertsätzen der Kombinationsfälle und der sog. Kontrollfälle gleichen Alters und Geschlechtes aleatorische Zusammenhänge, also rein zufällige Kombinationen annehmen, obwohl „latente positive Korrelationen“ (GROSSE) vorliegen.

Bei der Beachtung der Absterberelationen kommen wir aber in einen gewissen Gegensatz zu BERKSONS Ausgangsposition. BERKSON hat auch die positiven Korrelationen im Sektionsgut, wie zwischen Diabetes und Cholelithiasis sowie zwischen Rauchen und Lungenkrebs als vorgetäuscht angesehen und zwar durch die Auslese des Sektionsgutes. Wenn man mit GROSSE annimmt, daß die Auslese des Sektionsgutes vor allem durch den Tod erfolgt und sich diese Annahme bei Ausdehnung der Sektionen auf die außerhalb der Kliniken Verstorbenen bestätigt, so werden alle positiven Korrelationen im Sektionsgut noch entschiedener und intensiver, als bei der traditionellen Berechnung vermutet werden konnte.

2. Selektion der lebenden Patienten

Manche glauben, die Selektion durch den Tod zu eliminieren, indem sie ihre Untersuchungen auf die lebenden Patienten begrenzen. So vergleicht ELKELESS die Aortenverkalkung bei krebskranken und krebsfreien Patienten. Nun sind aber unter den Patienten, welche wegen ihres Krebses zum Arzt kommen, sicher nicht mehr Arteriosklerosen als unter den Lebenden der Population. Jedoch sind unter den Patienten, welche nicht wegen eines Krebses zur Klinik oder zur Röntgenuntersuchung kommen, wesentlich mehr mit verkalkter Aorta als unter der gleich alten Population. Es ist also zu erwarten, daß Krebskranke signifikant seltener eine Aortenverkalkung aufweisen als krebsfreie Patienten.

PFAUNDLER und v. SEHT fanden unter 546 luischen und 1737 tuberkulösen Kindern nur vier Kombinationsfälle, es waren also 6% der insgesamt 28090 behandelten Kinder tuberkulös, aber nur 0,7% der luischen Kinder wiesen tuberkulöse Veränderungen auf.

Hier darf keineswegs auf eine negative Syntropie zwischen Tuberkulose und Lues geschlossen werden, denn ein Bias liegt hier vor. Man müßte die Tuberkulosefrequenz unter den luischen Kindern mit der unter den gleich alten lebenden Kindern der Gesamtpopulation vergleichen, nicht aber mit den klinisch beobachteten Fällen, unter denen sich die Tuberkulosekranken gesetzmäßig anreichern. Die Formel PFAUNDLERS und v. SEHTS wurde also ohne Beachtung CORNETS aufgestellt.

Zusammenfassung

Obwohl CORNET 1899 den Beweis führte, daß Häufigkeitsschwankungen der Tuberkulose und anderer Erkrankungen niemals durch Vergleich von Sektionsreihen gleichen Alters und Geschlechtes bewiesen werden können, wird auch heute noch gegen diese Erkenntnis verstoßen. Dabei täuscht die Berechnung der Fehlerrgrenzen eine Exaktheit vor, welche im Sektionsgut niemals erreicht werden kann. Durch „Berksons Fallacy“ werden scheinbar signifikante negative Korrelationen zwischen jenen Krankheitszuständen vorgetäuscht, deren Absterberate wesentlich höher liegt als die der gleichen Gesamtpopulation. Die durch den Einfluß der unterschiedlichen Absterbequoten bedingten Störungen lassen sich nach GROSSE eliminieren. Die Absterbequote der Kombinationsfälle ist nicht größer als die Summe der Absterbequoten der beiden miteinander kombinierten Erkrankungen. Meist genügt es, den traditionellen Erwartungsprozentsatz der Kombinationsfälle durch die Absterbequote derjenigen Erkrankung zu teilen, welche die günstigere Prognose hat.

Concerning Berkson's Fallacy and Selection by Death

Summary

Although CORNET in 1899 showed, that fluctuations in the frequency of tuberculosis and other diseases could never be proven by comparing autopsy groups of the same age and sex, this knowledge even today is rejected. An accuracy which can never be attained in autopsy material is falsely ascribed to the calculation of the ranges of error. With "Berkson's Fallacy" apparently significant negative correlations are thought to exist among those disease states where the mortality rates are essentially higher than those mortality rates of the entire population of the same age. The disturbances introduced by the influences of different mortality quotas may be eliminated according to GROSSE. The mortality quota of the combination-cases is not greater than the sum of the mortality quotas of both diseases when these are combined. Usually it is sufficient to divide the traditionally anticipated percentage of the combination-cases by the mortality quota of that disease which has the more favorable prognosis.

Literatur

- BERKSON, J.: Limitations of the application of fourfold table analysis to hospital data. *Biomet. Bull.* **2**, 47 (1946).
- CORNET, G.: Die latenten Herde der Tuberkulose und die Tuberkulindagnostik im Lichte neuerer Forschung. *Berl. klin. Wschr.* **1904**, 359, 392.
- CORNFIELD, J.: Principles of research. *Amer. J. ment. Defic.* **64**, 240 (1959).
- ELKELESS, A.: Calcified atherosclerosis and cancer. *Brit. J. Cancer* **13**, 403 (1959).
- GROSSE, H.: (1) Sind unsere sektionsstatistischen Methoden exakt? *Virchows Arch. path. Anat.* **330**, 192 (1957).
— (2) Arteriosklerose und Krebs. *Z. Krebsforsch.* **62**, 519—526 (1958).
- KADAS, J.: Beiträge zum Verhältnis von Krebs und Arteriosklerose. *Zbl. allg. Path. path. Anat.* **102**, 450 (1961).
- KOLB, P.: (1) Zum Syntropie-Index von PFAUNDLER und v. SEHT. *Z. Kinderheilk.* **80**, 50 (1957).
— (2) Syntropie-Untersuchungen in Sektions- und Klinikmaterial. *Med. Klin.* **1963**, 1839.

- KRAUS, A. S.: The use of hospital data in studying the association between a characteristic and a disease. *Publ. Hlth Rep. (Wash.)* **69**, 1211 (1954).
- LINDEN, W. VAN DER: Some biological traits in female gallstone-disease patients. *Acta chir. scand.*, Suppl. 269 (1961).
- , u. A. TRAUBE: Arterioscleros och cancer i obduktionsmaterial. *Svenska Läk.-Tidn.* **56**, 3536 (1959).
- MAINLAND, D.: The risk of fallacious conclusions from autopsy data on the incidence of disease with application to heart disease. *Amer. Heart J.* **45**, 644 (1953).
- NIELSEN, L., and V. KESSING: Levercirrose og arteriosclerotisk hjertelidelse. *Ugeskr. Læg.* **124**, 33 (1962).
- OPSAHL, N.: On hefthyppigheten blandt sinnsyke. *Norsk Mag. Lægevidensk.* **94**, 77 (1933).
- PARRISH, H. M., and A. L. EBERLY: Negative association of coronary atherosclerosis and chronic alkoholism, a statistical fallacy. *J. Indiana med. Ass.* **54**, 347 (1961).
- PFAUNDLER, M., u. L. v. SEHT: Über Syntropie von Krankheitszuständen. *Z. Kinderheilk.* **300**, 100 (1921).

Dr. H. GROSSE, Bezirkskrankenhaus „Am Sund“, Stralsund, Billrothstr. 22