

08. März 1993

# INFORMATIK BIOMETRIE und EPIDEMIOLOGIE

## IN MEDIZIN UND BIOLOGIE

Offizielles Organ  
der Deutschen Gesellschaft für  
Medizinische Informatik,  
Biometrie und Epidemiologie (GMDS) e.V.

Gustav Fischer Verlag Stuttgart  
Verlag Eugen Ulmer Stuttgart

**Band 24**  
**Heft 1/1993**  
ISSN 0934-9235

# INFORMATIK BIOMETRIE und EPIDEMIOLOGIE

IN MEDIZIN UND BIOLOGIE

Gustav Fischer Verlag Stuttgart  
Verlag Eugen Ulmer Stuttgart

Schriftleitung:  
Prof. Dr. Hans Geidel, Stuttgart  
Prof. Dr. W. van Eimeren, Neuherberg  
Prof. Dr. Walter Lehmacher, Hannover  
Dipl.-Math. Joachim Vollmar, Mannheim

## Herausgeber

P. Bauer (Köln) · W. van Eimeren (Neuherberg) · H. Geidel (Stuttgart) · I. Guggenmoos-Holzmann (Berlin)  
R. Haux (Heidelberg) · W. Lehmacher (Hannover) · J. Michaelis (Mainz) · H. Rundfeldt (Hannover)  
W. Stucky (Karlsruhe) · H. Thöni (Hohenheim) · J. Vollmar (Mannheim) · H.-E. Wichmann (Wuppertal)

## Wissenschaftlicher Beirat

R. Blomer (Frankfurt) · U. Ferner (Basel) · H. Haußmann (Hohenheim) · W. Köhler (Gießen)  
W. Maurer (Basel) · M. Precht (Freising) · O. Richter (Braunschweig) · M. Schumacher (Freiburg)  
E. Sonnemann (Trier) · H.-D. Unkelbach (Geisenheim) · H. F. Utz (Hohenheim)

## Inhaltsverzeichnis 1/1993

Editorial		1
Aspekte der Fragebogengestaltung für eine automatische Datenerfassung in epidemiologischen Studien	L. Kreienbrock/ R. Ostermann	2
Epidemiologische Methoden und Wege der Risikoabschätzung	H. Becher/K. Steindorf	14
Zertifikat Medizinische Informatik Durchführungsrichtlinien		28
Buchbesprechungen/Bookreviews		39

### Redaktion:

**Schriftleiter** (verantwortlich im Sinne des Presserechts): Prof. Dr. H. Geidel, c/o Verlag Eugen Ulmer, Postfach 700561, 7000 Stuttgart 70, Tel. 47 28 80.

Verantwortlich für die Mitteilungen der Deutschen Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie e.V.: Dipl. rer. pol. F. Stobrawa, Herbert-Lewin-Straße 1, 5000 Köln 41, Tel. 4 00 42 56

**Verlag:** Vertrieb und Werbung: Gustav Fischer Verlag GmbH & Co. KG, 7000 Stuttgart 70, Wollgrasweg 49, Tel. (07 11) 45 80 30, Postgirokonto Stuttgart 13 556, Stuttgarter Bank 45 290. Herstellung und Anzeigen: Verlag Eugen Ulmer GmbH & Co., Wollgrasweg 41, 7000 Stuttgart 70, Tel. (07 11) 45 07-0, Telex 7-23 634, Postscheckkonto Stuttgart 74 63-700, Zürich 80-47072, Wien 1083.662 Deutsche Bank AG, Stuttgart, Kto. 14/76 878, Südwestbank AG, Stuttgart, Kto. 21 000, Herstellung Siegfried Märkisch. Verantwortlich für die Anzeigen: Dieter Boger, Gudrun Müller, Tel. (07 11) 45 07-147, z. Z. ist die Anzeigenpreisliste Nr. 9 gültig. Anzeigenschluß: am 20. der Monate Januar, April, Juli, Oktober.

**Druck:** Druckhaus »Thomas Müntzer« GmbH, Neustädter Straße 1-4, O-5820 Bad Langensalza, Telefon (0 36 03) 23 21-23 23

Die in der Zeitschrift veröffentlichten Beiträge sind urheberrechtlich geschützt. Alle Rechte, insbesondere das der Übersetzung in fremde Sprachen, vorbehalten. Kein Teil dieser Zeitschrift darf ohne schriftliche Genehmigung des Verlages in irgendeiner Form — durch Fotokopie, Mikrofilm oder andere Verfahren — reproduziert oder in eine von Maschinen, insbesondere von Datenverarbeitungsanlagen, verwendbare Sprache übertragen werden. Auch die Rechte der Wiedergabe durch

Vortrag, Funk- und Fernsehsendung, im Magnettonverfahren oder ähnlichem Wege bleiben vorbehalten. Fotokopien für den persönlichen und sonstigen eigenen Gebrauch dürfen nur von einzelnen Beiträgen oder Teilen daraus als Einzelkopien hergestellt werden.

Informatik, Biometrie und Epidemiologie in Medizin und Biologie ist eine deutsch/englischsprachige Zeitschrift, die theoretische und anwendungsbezogene Arbeiten aus dem Gesamtgebiet der Informatik, Biometrie und Epidemiologie in Medizin und Biologie enthält. Die Zeitschrift veröffentlicht sowohl Beiträge über neu entwickelte Ansätze als auch über neue Anwendungsmöglichkeiten bekannter Methoden. Zusätzlich soll dem tutoriellen Aspekt durch Übersichtsartikel Rechnung getragen werden. Der Bezug zur Medizin oder Biologie darf dabei jedoch nicht vernachlässigt werden. Jedes Heft enthält u. a. Editorial, Übersichten zu bestimmten Themenkomplexen, Originalarbeiten, Kurzbeiträge, Stellungnahmen, Fallbeispiele, Informationen über Programme und Programmsysteme, Informationen zu Veranstaltungen und Neuerscheinungen auf dem Buchmarkt.

Die Zeitschrift wird zur Zeit referiert in den »Biological Abstracts« und »Current Index to Statistics«.

Informatik, Biometrie und Epidemiologie in Medizin und Biologie erscheint vierteljährlich.

Der Bezugspreis beträgt jährlich DM 273,- (Einzelheft DM 81,-) zuzüglich Versandkosten. Vorzugspreis für Mitglieder der Deutschen Region der Internationalen Biometrischen Gesellschaft DM 198,- zuzüglich Versandkosten. Bestellungen nehmen jede Buchhandlung sowie die Verlage entgegen.

# INFORMATIK BIOMETRIE und EPIDEMIOLOGIE

IN MEDIZIN UND BIOLOGIE

---

## Bd. 24 (1993) Inhaltsverzeichnis

ALESCH, F.; KORN, A., PAPPATERRA, J.; GANGLBERGER, J.: CASA – ein Computer-Assistierter mehrdimensionaler Stereotaxie-Atlas .....	57
BECHER, H.; STEINDORF, K.: Epidemiologische Methoden und Wege der Risikoabschätzung .....	14
BRINKER, T.; DIETZ, H.: Computersimulation zur Durchführung experimenteller Untersuchungen des Liquorsystems .....	129
ELLSÄSSER, K.-H.; KÖHLER, C. O.: Shared Care: Konzept einer verteilten Pflege – Kurz- und langfristige Perspektiven in Europa.....	188
FEDOROV, V. V.; HACKL, P.; MÜLLER, W. G.: Estimation and Experimental Design for Second Kind Regression Models .....	134
HAMPL, J.; STEINMETZ, A.; ALBERT, F.; WINKLER, J.; KUNZE, S.: Belegleser und strukturierte Dokumentationsbögen: Ein in den Routinebetrieb integrierbares System zur halbautomatischen Erfassung medizinischer Daten ...	94
HOFESTÄDT, R.: Modelling molecular counters with stochastic cellular grammars .....	169
KÖNIG, A.: Integration eines Interpreters in ein Labor-EDV-System zur flexiblen Nachbearbeitung von Laboraufträgen und Verbesserung der Plausibilitätskontrolle.....	178
KREIENBROCK, L.; OSTERMANN, R.: Aspekte der Fragebogengestaltung für eine automatische Datenerfassung in epidemiologischen Studien.....	2
LEINER, F.; HAUX, R.; GLÜCK, E.: Systematische Planung Medizinischer Basisdokumentationen im Krankenhaus.....	199
LIPINSKI, H.-G.; STRUPPLER, A.: Wissensbasierte funktionelle Neurochirurgie	50
MÄRZ, U.; HAMBURGER, C.; WIESE, U. H.: Beschleunigte Befund- und Behandlungs-Datenerfassung sowie Arztbriefherstellung in der Neurochirurgie mit Hilfe einer relationalen Datenbank .....	73
MÄRZ, U.; SCHMID, R.; PIEPGRAS, A.; REULEN, H.-J.: Diagnosen und OP-Dokumentation in der Neurochirurgie mit der Datenbank „CODIAG“ ....	87
MEYER, M.; HORBACH, L.; LANG, Ch.; WENTZ, B.: Erstellung einer strukturierten Diagnosenstatistik durch Nutzung der Struktur des Diagnoseschlüssels für Neurologie und Neurochirurgie .....	110
MIETH, M.; BRIGL, B.; GLÜCK, E.; HAUX, R.: Zur automatischen Indexierung von Diagnosen nach SNOMED und ihrer Anwendbarkeit in der Neurochirurgie.....	101

SAWINSKI, R.; HAMPL, J.; WERNER, R.; WINKLER, J.; WINTER, A.; HAUX, R.; KUNZE, St.: OP-Register und Arztbriefschreibung mit MDVS . . .	81
SCHERB, H.; BRÜSKE-HOHLFELD, I.: The Exact Confidence Curve . . . . .	152
SCHMITT, H. P.: Computer-aided Classification of the Malignancy of Astrocytomas with the Classifier TESTAST268 for the Choice of Therapy and as a Standardized Basis for Comparison of Results from Multi-Center Studies . . . . .	43
SCHMÜCKER, P.; DUJAT, C.; HAMPL, J.; STEINMETZ, A.; HOLLE, R.: Optische Archivierung von Krankenunterlagen in der Neurochirurgie Universitätsklinik Heidelberg . . . . .	123
SCHWERDTFEGER, K.; SCHULTE, S.; SCHWARZ, W.; STEUDEL, W. I.: Ein EDV-gestütztes Patienten-Management-System . . . . .	65
Zertifikat Medizinische Informatik – Durchführungslinien . . . . .	28
Editorial . . . . .	1,41
Nachruf: Arthur Linder . . . . .	121
Buchbesprechungen . . . . .	39, 116, 164, 214

Institut für Medizinische Biometrie  
 und Medizinische Informatik  
 der Universität Freiburg  
 Abt. Medizinische Biometrie u. Statistik  
 Stefan-Meier-Straße 26  
 79104 Freiburg

## Editorial

Sicher wird die meisten Leser die mit diesem ersten Heft des Jahres 1993 vollzogene Namensänderung unserer Zeitschrift nicht überrascht haben. Seit der im letzten Jahr vollzogenen Übereinkunft zwischen Verlag und gmds hat die Zeitschrift nicht nur einen erheblich weiteren Leserkreis gefunden, sondern ihr Profil in Medizinischer Informatik verstärkt.

Der in diesem Heft abgedruckte Beitrag über das Zertifikat „Medizinische Informatik“ signalisiert dabei neben dem wissenschaftlichen auch das – begrenzte – berufspolitische Engagement unserer Herausgeberschaft.

Entgegen den Befürchtungen, die gelegentlich laut wurden, hat dieser Gestaltwandel nicht zulasten anderer Themen stattgefunden. Dies wird auch nicht durch die jetzt verstärkte Berücksichtigung epidemiologischer Themen anders. Bei gleich hohem Qualitätsanspruch unserer Gutachter wird die Konsequenz für den Leser lediglich ein reichhaltigeres Angebot sein, liege nun des Lesers wissenschaftlicher Interessens-Schwerpunkt mehr im Bereich der Biometrie, Informatik oder Epidemiologie und sein bevorzugtes Anwendungsfeld in der Medizin oder Biologie!

Nur ein ausreichend weites Themenspektrum kann bei Qualitätsanspruch einerseits aus wirtschaftlicher Sicht hinreichend viele Leser und andererseits aus der Sicht wissenschaftlichen Einflusses genügend viele qualifizierte Autoren anziehen.

Wilhelm van Eimeren



## **Aspekte der Fragebogengestaltung für eine automatische Datenerfassung in epidemiologischen Studien<sup>\*)</sup>**

L. Kreienbrock, R. Ostermann

### **Zusammenfassung**

*In den letzten Jahren ist das Leistungsniveau von Scannern immer größer geworden. Insbesondere ist das Niveau der dort verwandten Texterkennungssoftware mittlerweile sehr hoch. Somit besteht nun auch bei epidemiologischen Studien die Möglichkeit, das gegebene Datenmaterial (Fragebögen) mit Hilfe eines Scanners und der (zumeist integrierten) Texterkennungssoftware zu erfassen. Im Rahmen dieses Beitrages sollen einige Aspekte der „scannerfreundlichen“ Fragebogengestaltung diskutiert werden. Dabei sollen u. a. einige einfache Regeln deskriptiver Statistik auf ihre Anwendbarkeit bezüglich dieses neuen Datenerfassungsmediums überprüft werden.*

### **Schlüsselwörter**

*automatische Datenerfassung, Fragebogengestaltung, epidemiologisches Management*

### **Summary**

*In recent years the power of scanners grows up more and more the level of the used software for textrecognition is very high. Therefore the possibility is given to use this modern technology to record data for epidemiological studies. In this paper some aspects of a “scanner-friendly” questionnaire are discussed. Some simple and basic rules of descriptive statistics are proved for their use concerning this new technology.*

### **Key Words**

*automatical data entry, formation of questionnaires, epidemiological management*

---

<sup>\*)</sup> Vortrag und Demonstration „Statistical Computing '91“, Schloß Reisenburg

## 1 Einleitung

In vielen Bereichen der Wissenschaft werden heutzutage moderne Technologien zur effizienten Bearbeitung vorliegender Probleme eingesetzt. Mit der zunehmenden Technisierung der (wissenschaftlichen) Arbeit ergibt sich die Frage, ob und in welchem Maße technische Hilfsmittel eingesetzt werden sollen, um wissenschaftliche Fragestellungen insgesamt effizient zu lösen.

Ein besonderes Problemfeld in diesem Zusammenhang und damit eine besondere Herausforderung stellen die modernen Datenverarbeitungstechnologien dar. Der technische Standard ist mittlerweile so hoch, daß Arbeitsplätze mit hochleistungsfähigen Systemen ausgerüstet werden können, ohne übliche Kostengrenzen zu sprengen. Durch diese Technisierung der Arbeitsplätze erweitert sich in der Regel das wissenschaftliche Leistungsdenken in dem Sinne, daß durch bessere technische Hilfsmittel entweder bessere wissenschaftliche Leistungen oder schneller die „alten“ Leistungen erbracht werden sollen.

Hierbei wird in der Regel allerdings nicht bedacht, daß im Ablauf wissenschaftlicher Arbeit ein vielschichtig verknüpftes, zum Teil interdisziplinäres Schema der Arbeit berücksichtigt werden muß, so daß eine Vereinfachung eines Teils der Arbeit nicht zwangsläufig das Gesamtkonzept vereinfacht. Als Beispiel hierfür mag die Einführung technischer Hilfsmittel dienen, die in der Regel auch die Bereitstellung entsprechenden technischen Hilfspersonals bzw. eine gewisse Lern- und Einarbeitungszeit zur effizienten Nutzung des Systems erfordert. In diesem Zusammenhang sei auf die Einführung des Personal Computers hingewiesen. Damit ein PC sinnvoll die wissenschaftliche Arbeit unterstützen kann, ist es unumgänglich, daß der Benutzer dieses PC's grundlegende Betriebssystemkenntnisse besitzt. Für das statistische Arbeiten mit einem PC ist es ebenfalls erforderlich, den korrekten syntaktischen Umgang der Statistik-Software zu beherrschen.

Der nachfolgende Beitrag beschäftigt sich mit einem solchen praktischen Aspekt wissenschaftlicher Arbeit. Im Rahmen epidemiologischer Studien, die als höchst interdisziplinäres Forschungsfeld angesehen werden müssen, sollen hierbei Aspekte der Fragebogengestaltung für eine automatische Datenerfassung im Vordergrund stehen.

Hierbei wird von der Vorstellung ausgegangen, daß die inhaltliche und formale Gestaltung von Erhebungsmaterialien nicht voneinander zu trennen sind, so daß etwa die Formulierung einer Frage (incl. entsprechender Antwortvorgaben) nicht nur den Prozeß der anschließenden Datenerfassung, sondern auch der nachfolgenden Analysen und Interpretationen beeinflusst. Damit ist eine direkte Verknüpfung der üblicherweise getrennt behandelten Bereiche sogenannter wissenschaftlicher und technischer Arbeit gegeben und entsprechende Wechselwirkungen zu berücksichtigen.

Bezogen auf die Gesamtkonzeption einer epidemiologischen Studie stellt sich dieser Aspekt der Verknüpfung einzelner Arbeitskomponenten vereinfacht wie in Abbildung 1 dar.

Wenn im folgenden die Fragebogengestaltung im Zusammenhang mit einer Automatisierung der Datenerfassung behandelt wird, so ist im Sinne des in Abbildung 1 aufgeführten Schemas die inhaltliche und technische Gestaltung von Fragebögen in eine der ersten Planungsphasen zuzuordnen, die Datenerfassung den Bereichen der Datengewinnung und -aufbereitung.

Die Aussage, ob und in welchem Maße eine automatische Datenerfassung einsetzbar ist, sollte sich deshalb stets aus dem hier kurz skizzierten Gesamtkonzept ergeben. Im folgenden werden wir dennoch einige grundlegende Aspekte zu diesem Themenkomplex aufzeigen. Dabei wollen wir nicht auf spezielle Technologien von Lesegeräten oder Scannern eingehen, denn die technische Weiterentwicklung dieser Geräte ist so rasant, daß eine heute gemachte Aussage schon in kürzester Zeit veraltet wäre. Dennoch haben

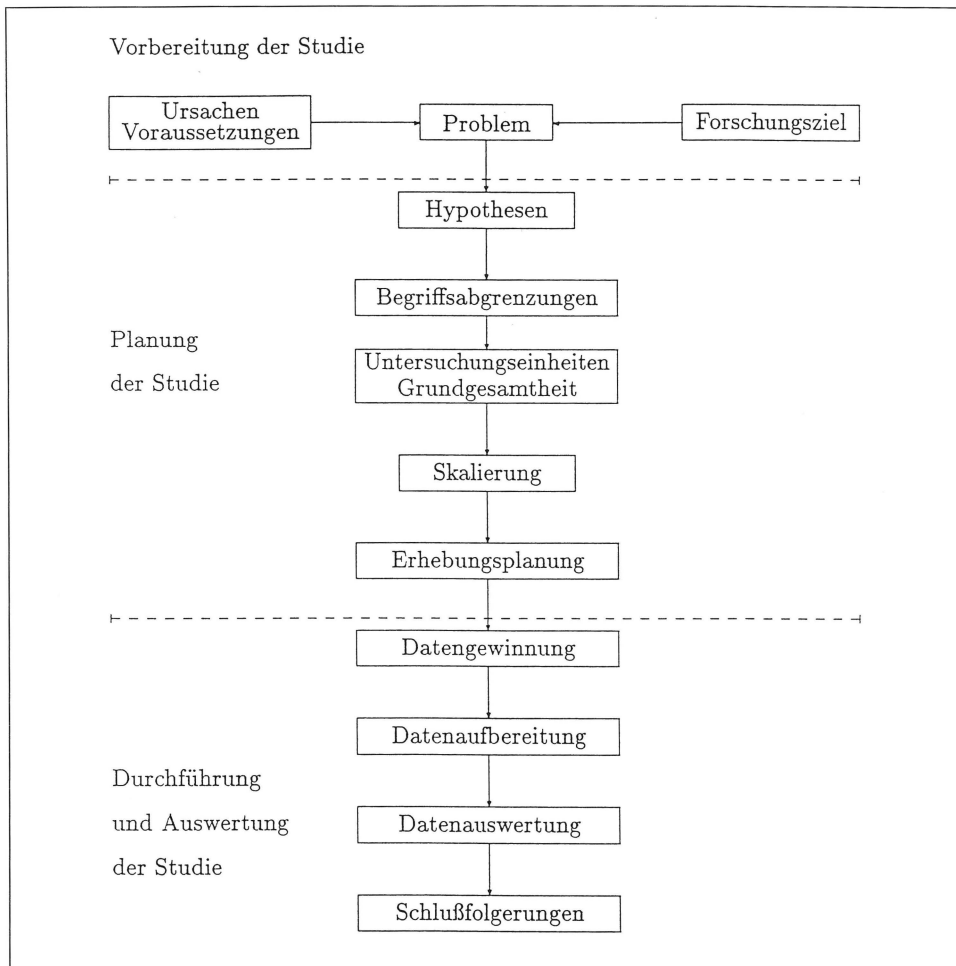


Abbildung 1: Ablaufschema einer epidemiologischen Studie (modifiziert nach KREIENBROCK 1989)

sich unsere hier gemachten Überlegungen an dem heutigen Stand der Technik orientiert und sind auch davon beeinflusst worden.

Hauptaugenmerk soll vielmehr auf einige Grundprinzipien der genannten Schnittstellen gelegt werden. Hierzu werden zunächst Informationsquellen und -geber der epidemiologischen Forschung sowie einige Strukturaspekte von Datenerfassungsdokumenten behandelt. Hieran kann sich dann eine Bewertung der potentiellen Einsatzmöglichkeiten automatischer Datenerfassung in der Epidemiologie anschließen.

## 2 Informationsquellen und Informationsgeber der Epidemiologie

Die Epidemiologie hat nach Definition der WHO die Beschreibung und Analyse der Verteilungen und Ursachen von Krankheiten in (menschlichen) Populationen zur Aufgabe. Damit zählt die Epidemiologie zu einem hochinterdisziplinären Forschungsbe-



reich, in dem sämtliche Disziplinen moderner Forschung vertreten sind. Hierzu gehören unter anderem Medizin, Statistik, Natur-, Sozial- und Ingenieurwissenschaften und viele weitere (Teil-) Disziplinen. Je nach konkretem Forschungsgegenstand können somit sämtliche modernen Forschungsrichtungen und -mittel in der Epidemiologie vertreten sein.

Um den Bereich der Datenerfassung zu charakterisieren sind demzufolge unterschiedlichste Informationen zur Beantwortung einer epidemiologischen Fragestellung einzuholen. Hier soll deshalb zunächst unterschieden werden, in welcher Form Informationen eingeholt werden, d. h. welche Informationsquellen prinzipiell zu berücksichtigen sind, und wer diese Informationen ursächlich zusammenstellt, d. h. wer die Informationsgeber der Epidemiologie darstellen.

### Informationsquellen

Geht man im Sinne einer analytischen Epidemiologie davon aus, daß das Hauptinteresse eine Untersuchung in der Prüfung einer Ursache-Wirkungs-Beziehung zwischen einem Risikofaktor und einer Krankheit (unter Berücksichtigung störender Einflußgrößen) ist, so lassen sich Informationsquellen sowohl für den Bereich der Risiken und Krankheiten eines einzelnen Menschen, wie auch im Bereich sekundärstatistischer Sammelstatistiken identifizieren. Damit ergeben sich insgesamt drei Gruppen von Informationsquellen der Epidemiologie.

#### – Primärstatistiken am Individuum

Hierzu zählen sämtliche Daten, die im Rahmen von konkret durchzuführenden Studien an Einzelpersonen erhoben werden. Die nutzbaren Erfassungsinstrumente sind vor allem Fragebögen und Untersuchungs- sowie Meßprotokolle, aber auch etwa Patientenakten, die im normalen Klinikbetrieb anfallen können und für eine Studie aufbereitet werden.

#### – Primärstatistiken aus der individuellen Umwelt

Zur Umwelt eines Individuums sind die Parameter zu zählen, die als Risikofaktoren auf dessen Krankheit wirken können, diesem direkt zuzuordnen sind, aber nicht direkt in Verbindung mit einem zu untersuchenden Individuum stehen. Damit sind etwa Daten zur Luft-, Wasser- und Bodenverschmutzung, Messungen an Arbeitsplätzen, soziologische Daten aus dem Wohnumfeld und vieles mehr als Informationsquelle denkbar.

#### – Sekundärstatistiken

Neben den Informationen, die eine direkte Zuordnung zum Individuum innerhalb der zu untersuchenden Population zulassen, werden weitere den Ursachen und Voraussetzungen zuzuordnenden Informationen für die Erforschung eines epidemiologischen Sachverhaltes notwendig sein. Aus diesem Grund bedient man sich häufig auch bereits bestehender Daten, sei es aus administrativen oder auch aus interpretatorischen Gründen. Hier sind vor allem allgemein gesellschaftliche oder Umweltdaten zu nennen, wie sie etwa von Meldeämtern, Versicherungsträgern, statistischen und anderen Ämtern oder Meßinstituten zur Verfügung gestellt werden können.

Die Quellen epidemiologischer Daten sind damit sehr unterschiedlich, so daß dies bei einer automatischen Datenerfassung Berücksichtigung finden muß. Dies gilt allerdings in noch größerem Maße für die Personen und Institutionen, die die Informationsbeschaffung durchführen.

## Informationsgeber

Insbesondere dann, wenn wie bei multizentrischen Studien primärstatistische Daten in unterschiedlichen Datenströmen von verschiedensten Personen und Institutionen ermittelt werden, stellt sich nicht nur unter dem Aspekt einer möglichen Automatisierung der Datenerfassung die Frage nach der Art der Informationsgeber. Im wesentlichen sind folgende Informationsgeber innerhalb epidemiologischer Studien anzutreffen.

- (geschultes) Interviewpersonal  
Immer dann, wenn hochspezialisiertes Wissen zu einem Forschungsgegenstand mit Hilfe einer Befragung erhoben werden muß, so wird Interviewpersonal eingesetzt

Abbildung 2: Beispiel für ein Interview durch geschultes Interviewpersonal

a: standardisierter Fragenteil

b: nicht standardisierte Gesprächssituation

<b>(a) 2. Was haben Sie damals .... geraucht</b>				
	ja	nein	w.n	k.A
Zigaretten	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 8	<input type="radio"/> 9
selbstgedrehte Zigaretten	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 8	<input type="radio"/> 9
Zigarren	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 8	<input type="radio"/> 9
Zigarillos	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 8	<input type="radio"/> 9
Pfeife	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 8	<input type="radio"/> 9
<b>und wieviel waren das pro</b>	<b>Tag</b>	<b>Woche</b>	<b>Monat</b>	
Zigaretten	/ _ / _ / _ /	/ _ / _ / _ /	/ _ / _ / _ / Stück	
selbstgedrehte Zigaretten		/ _ / _ / _ /	/ _ / _ / _ / Gramm	
Zigarren	/ _ / _ / _ /	/ _ / _ / _ /	/ _ / _ / _ / Gramm	
Zigarillos	/ _ / _ / _ /	/ _ / _ / _ /	/ _ / _ / _ / Stück	
Pfeife		/ _ / _ / _ /	/ _ / _ / _ / Gramm	
<small>(entf. = 777, w.n. = 888, k.A. = 999)</small>				
<b>3. Haben Sie damals Inhalter bzw. Lungenzüge gemacht ?</b>				
	Zigarette	Zigarre	Zigarillo	Pfeife
ja, eher tief	<input type="radio"/> 1	<input type="radio"/> 1	<input type="radio"/> 1	<input type="radio"/> 1
ja, eher flach	<input type="radio"/> 2	<input type="radio"/> 2	<input type="radio"/> 2	<input type="radio"/> 2
nein, nur gepafft	<input type="radio"/> 3	<input type="radio"/> 3	<input type="radio"/> 3	<input type="radio"/> 3
entfällt	<input type="radio"/> 7	<input type="radio"/> 7	<input type="radio"/> 7	<input type="radio"/> 7
weiß nicht	<input type="radio"/> 8	<input type="radio"/> 8	<input type="radio"/> 8	<input type="radio"/> 8
keine Angaben	<input type="radio"/> 9	<input type="radio"/> 9	<input type="radio"/> 9	<input type="radio"/> 9
<small>(falls kein Zigarettenraucher --&gt; weiter mit Frage 6)</small>				
<b>(b) 19. Beschreiben oder skizzieren Sie kurz, wo Ihr Hauptwohnraum bzw. Schlafraum lag.</b>				
<small>(Probing: Zeichnen Sie einen einfachen Grundriß oder beschreiben Sie uns den Weg in diese Räume (z.B. rechts hinter der Eingangstür usw.).)</small>				

Lippenzyanose	<input type="checkbox"/> ja	<input type="checkbox"/> nein
Husten z.Zt.	<input type="checkbox"/> ja	<input type="checkbox"/> nein
Dyspnoe	<input type="checkbox"/> ja	<input type="checkbox"/> nein
Atemfrequenz in Ruhe	<input type="text"/> <input type="text"/> pro Minute	
Anhusteversuch	<input type="checkbox"/> normal <input type="checkbox"/> positiv <input type="checkbox"/> stark positiv (Auswurf)	
Atemgeräusch	.....	
Pulsfrequenz	<input type="text"/> <input type="text"/> <input type="text"/> pro Minute	

Abbildung 3: Beispiel für einen vom medizinischen Personal auszufüllenden Untersuchungsbogen

werden müssen, das speziell für dieses Forschungsvorhaben ausgebildet und geschult ist. Diese Form der Befragung ermöglicht dann einerseits eine detaillierte Datenerhebung (vgl. z. B. Abbildung 2a), andererseits führt sie zu einer komplexen Datenstruktur, die ein hohes Maß an nicht standardisierten Daten enthält (vgl. z. B. Abbildung 2b).

– medizinisches und technisches (Hilfs-)Personal

Viele epidemiologische Studien nutzen die Einrichtungen des Gesundheitswesens und deren Personal, um gegebene Fragestellungen zu beantworten. Hier ist es dann unabdingbar einfache Datenerfassungsmedien zu benutzen, um das Personal nicht zeitlich zu überfordern, denn in epidemiologischen Studien werden Datenerfassungstätigkeiten oftmals zusätzlich zur eigentlichen Arbeit erbracht (z. B. Führen einer studienrelevanten Patientenliste oder standardisierter Anamnesebogen, vgl. Abbildung 3).

– Probanden

Die Selbstbeobachtung bzw. die eigene Beschreibung von epidemiologisch relevanten Daten wird häufig von den Teilnehmern einer Studie selbst durchgeführt (vgl. z. B. Abbildung 4a und b). Hier gilt wie in vielen anderen Forschungsfeldern auch, daß Sprache und Antwortvorgaben vornehmlich am Erfahrungshorizont des Probanden orientiert sein müssen, um Spezifität, Sensitivität und Reliabilität der Daten zu gewährleisten.

– Listen, Akten und elektronische Datenträger

Relevante Informationen zu epidemiologischen Sachverhalten findet sich oftmals

**1. Wie lautet die Adresse Ihrer Wohnung**

Ort : \_\_\_\_\_  
 Kreis : \_\_\_\_\_  
 Postleitzahl : \_\_\_\_\_  
 Straße : \_\_\_\_\_ Nr. \_\_\_\_\_  
 Telefon : \_\_\_\_\_

---

**2. Wohnen Sie in einem**

<sub>1</sub> Mehrfamilienhaus  
<sub>2</sub> Einfamilienhaus  
<sub>3</sub> Sonstiges (Wohnwagen o.ä.)

*Falls Sie in einem Mehrfamilienhaus wohnen, fahren Sie mit Frage 3 fort, sonst lassen Sie bitte die nächste Frage aus und gehen zu Frage 4.*

Abbildung 4: Beispiel für Fragebogen, der durch Probanden ausgefüllt wird  
 a: (oben) einfacher Fragebogen  
 b: (rechts) Selbstbeobachtung

auch auf bereits existierenden physikalischen Datenträgern (z. B. Patientenakten eines Krankenhauses) bzw. auf elektronischen Datenträgern unterschiedlichster Art. Die enthaltenen Informationen sind dabei sowohl von speziellem Inhalt wie auch im Rahmen von Routineerfassungen gesammelt worden, so daß Daten mit Individualcharakter wie auch hoch standardisiertes

Datenmaterial hier zur Verfügung stehen kann, welches vom einfachen Text (z. B. Arztbrief) bis hin zur komplexen Datenbankstruktur auf Mainframe-Computern reicht.

Wieviel Stunden im Freien gewesen?

Wie haben Sie sich gesundheitlich gefühlt?

- letzte Nacht?	- und heute tagsüber?
<input type="checkbox"/> gut	<input type="checkbox"/> gut
<input type="checkbox"/> ziemlich gut	<input type="checkbox"/> ziemlich gut
<input type="checkbox"/> mittelmäßig	<input type="checkbox"/> mittelmäßig
<input type="checkbox"/> schlecht	<input type="checkbox"/> schlecht
<input type="checkbox"/> sehr schlecht	<input type="checkbox"/> sehr schlecht

Luftnot?  
 keine  leicht  mittel  schwer

Husten?  
 kein  leicht  stark  Attacke

### 3 Strukturaspekte von Datenerfassungsdokumenten in der Epidemiologie

Wie die Beispiele des Abschnittes 2 gezeigt haben, liegen epidemiologisch relevanten Daten unterschiedlichste Strukturen und Quellen zu Grunde. Hierbei sind nicht alle in gleichem Maße zu einer automatischen Datenerfassung geeignet, da die Zeichenerkennung ein gewisses Maß an Standardisierung und Exaktheit voraussetzt. Um den Einsatz von Scannern oder analogen technischem Gerät (z. B. Belegleser mit 7-Segment-schrifterkennung) im Rahmen der epidemiologischen Forschung zu prüfen, ist es deshalb

erforderlich grundsätzliche Strukturaspekte von Datenerfassungsdokumenten nochmals kurz zu notieren.

- codierte vs. uncodierte Fragen  
Vorgegebene Codes (vgl. z. B. Abbildung 2a, Frage 2) bzw. uncodierte Fragen (vgl. z. B. Abbildung 2a, Frage 1) können sich entscheidend auf eine Zeichenerkennung auswirken, denn im ersten Fall ist eine einfache Markierung, dagegen im zweiten Fall ein (hand-)schriftlicher Eintrag im Datenerfassungsdokument vorzunehmen.
- offene vs. geschlossene Fragen  
Der oben angesprochene Erkennungsaspekt wirkt sich noch drastischer aus, wenn (wie bei offenen Fragen) nicht bekannt ist welche Antwortmöglichkeiten bestehen. So ist etwa die in Abbildung 2b gezeigte Frage ein Beispiel von vollständiger Individualität und somit nicht für eine Automatisierung geeignet.
- stetige vs. diskrete Merkmale  
Ein analoger Aspekt ergibt sich bei (theoretisch) stetigen Merkmalen. Hier ist zu fragen, ob für das epidemiologische Forschungsziel nicht eine Diskretisierung ausreichend ist, so daß die Zahl möglicher Antwortkategorien eingeschränkt wird. So sind z. B. etwa Größen wie Blutdruck, Körpergröße oder eine Schadstoffexposition häufig durchaus in kategorisierter Form von höherem Aussagewert als eine stetige Meßgröße.
- Skalenniveau  
Insbesondere die Frage nach Spezifität, Sensitivität und Reliabilität eines Befragungsgegenstandes läßt häufig die Diskussion entstehen, ob ein möglichst hohes Skalenniveau nicht doch nur eine Vortäuschung einer angestrebten Genauigkeit ist. Will man etwa im Rahmen einer epidemiologischen Untersuchung den Gesundheitszustand eines Probanden charakterisieren und stehen hierzu unterschiedlichste medizinische und biologische Parameter zur Verfügung, so sind diese durchaus nicht immer gleich gerichtet und die Beschreibung des Gesundheitszustandes erfolgt nur noch auf ordinalem oder gar dichotomen Niveau. In diesem Sinne ist zu fragen, ob nicht schon auf der Ebene der Datenerfassung mit diesem Skalenniveau gearbeitet werden kann, was die automatische Datenerfassung erleichtern wird.

Diese prinzipiellen Strukturaspekte bei der Datenerfassung gelten zwar im wesentlichen für sämtliche empirische Untersuchungen, jedoch kompliziert sich diese Betrachtung im Rahmen der Epidemiologie durch zwei hier besonders eklatant auftretende Phänomene, die für die automatische Datenerfassung von besonderer Bedeutung sind.

- Filter und Verzweigungen  
Da im Rahmen einer zu prüfenden Ursache-Wirkungs-Beziehung ein hohes Maß an Vernetzung zwischen den einzelnen zu betrachtenden Faktoren und den damit assoziierten Daten besteht, kann jede Datenstruktur einen Filter darstellen, so daß in den Datenerfassungsdokumenten Sprünge und Verzweigungen entstehen, die bei einer automatischen Datenerfassung besondere Berücksichtigung finden müssen. Eine zu komplexe Verzweigung kann u. U. die Aussagekraft eines einzelnen ausgefüllten Fragebogens mindern (zu komplex für den Probanden).
- Medizin vs. Statistik  
Die Epidemiologie ist als interdisziplinäres Forschungsfeld im wesentlichen als eine Schnittstelle zwischen einer statistischen („Massenverhalten“) und einer medizinischen Betrachtungsweise („Individualbetrachtung“) anzusehen. Von diesem Blickpunkt aus ergibt sich trotz eines hohen Maßes an Strukturierung stets auch ein großer Anteil von offenen Beschreibungen, die wie oben bereits erwähnt für eine automatische Datenerfassung nicht oder noch nicht geeignet sind.

#### 4 Einsatzmöglichkeiten automatischer Datenerfassung in der Epidemiologie

Automatische Datenerfassungsgeräte haben mittlerweile einen so hohen technischen Stand erreicht, der es möglich erscheinen läßt, diese technischen Medien nicht nur im Rahmen einer industriellen oder wirtschaftlichen Nutzung, sondern auch im Wissenschaftsbetrieb einzusetzen.

Im Gegensatz zu großindustriellen Einsatzgebieten wie z. B. der automatischen Erfassung von Lotteriedeckungen oder Wertscheinen, Bankbelegen, Frachtpapieren etc., die in hohem Maße standardisiert sind und in extrem großen Stückzahlen von mehreren (hundert-)tausend meist durch Belegleser mittlerweile (fast) fehlerfrei erfaßt werden (vgl. z. B. WEIDELICH 1991), ist bei der epidemiologischen Datenerfassung wie oben dokumentiert von anderen Rahmenbedingungen auszugehen.

Neben der beschriebenen Informations- und Strukturvielfalt sind insbesondere die Stückzahlen gering; dagegen liegen individualspezifische Daten häufig in großen Mengen vor. So werden etwa in Fall-Kontroll-Studien zur Ätiologie seltener Krankheiten oftmals nur wenige hundert Probanden untersucht oder befragt, an diesen Individuen aber hunderte von Merkmalen erfaßt.

Damit ist die automatische Datenerfassung als Teil des „Epidemiological Management“ anzusehen und es entsteht hier insbesondere die Frage nach der Effizienz des Einsatzes. Diese Effizienz wird in der Regel von zwei Aspekten bestimmt.

##### – Datenqualität

Die Datenqualität bzw. die Anzahl in einem zu analysierenden Datensatz erhaltener Fehler wird bei herkömmlicher Datenerfassung von verschiedensten Komponenten bestimmt. Hierzu zählen Falschangaben des Probanden, Fehler der Datenerfasser, bei der Codierung und der eigentlichen Datenübertragung von beispielsweise Papierdokumenten in Dateneingabemaschinen. Trotz Plausibilitätsprüfungen, Doppelingaben und Rücksprachen ist bei epidemiologischen Studien davon auszugehen, daß der Fehleranteil je nach Erhebungsgegenstand, Informationsquelle und -geber etwa bei 1 bis 5% liegen kann.

Als entscheidender Faktor zur Steuerung der Datenqualität kann der Aufwand verstanden werden, der zur Verbesserung der Datenqualität geleistet werden muß.

##### – Forschungsressourcen

Ein limitierender Faktor für die Fehlerrecherche stellen in der Regel die Ressourcen dar, die für ein Forschungsvorhaben zur Verfügung stehen. Im Sinne der oben aufgeführten Datenqualitätsaspekte zählen hierzu Verständlichkeit für den Probanden (empirisches Know How), Ausbildungs- bzw. Motivationsstand der Datenerfasser (Schulungs- bzw. Motivationsaufwand), Codierungskontrolle durch Experten, Qualifikation und Zeitaufwand der Dateneingabe sowie der Aufwand an Programmierung von Plausibilitäts- und Dateneingabesystemen.

Im Sinne eines „Epidemiological Management“ kann die Entscheidung über den Einsatz einer automatisierten Datenerfassung an obigen Gesichtspunkten orientiert werden.

Kann eine Fehlerquote für Datenqualität definiert werden, so stellt sich zunächst nur die Frage, ob ein automatisches System diese erfüllen kann. Ist dies der Fall kann eine einfache Investitions- und Finanzierungsrechnung (fixe Investitionskosten + Arbeitsstunden) die Entscheidung für oder gegen ein automatisches oder ein herkömmliches System herbeiführen.

Zum jetzigen Zeitpunkt (Sommer 1992) ist absehbar, daß eine solche Vergleichsrechnung teilweise schon jetzt zu Gunsten einer automatisierten Datenerfassung ausgehen kann. Dies gilt insbesondere dann, wenn nicht nur ein Forschungsvorhaben als Grundlage der

Entscheidung dient, sondern ähnliche Studien in naher Zukunft weiterhin durchgeführt werden.

Die technischen Grundlagen von Scannern erlauben zwar heute immer noch kein Lesen von Handschriften, was offene und uncodierte Fragen von einer automatischen Erfassung ausschließt, aber immer dann, wenn Fragen vorcodiert gestellt werden können, ist ein automatisches Lesen mit Fehlerquoten, die der manuellen Eingabe gleichen, schon jetzt durchführbar. So ist unter Verwendung von sogenannten „Blindfarben“ oder Templates, die beim Scanvorgang nicht mitgelesen werden, es z. Zt. möglich etwa 1500 bis 1800 identische Seiten eines Fragebogens mit durchschnittlicher Informationsdichte pro Stunde einzulesen.

Im Rahmen der Arbeitstagung „Statistical Computing '91“ wurde unter Leitung der Autoren ein Pre-Praxis-Test durchgeführt. Dabei wurden die Teilnehmer angehalten, den Fragebogen aus Abbildung 5b auszufüllen. Anschließend wurden diese ausgefüllten Fragebogen direkt mit einem Scanner eingelesen. Zum Erstaunen zahlreicher Teilnehmer dieser Tagung war die Erkennungsrate sehr hoch, obwohl einige sich bewußt bemüht hatten, den Fragebogen falsch bzw. mehrdeutig auszufüllen.

Die technische Entwicklung von Hard- und Software ist dabei so rasant, daß diese gerade getroffene Aussage bei Erscheinen dieses Beitrages schon veraltet sein kann. Die hier getroffenen Aussagen bzgl. der Handschrifterkennung gelten mit leichten Abschwächungen auch für einen Belegleser mit 7-Segmentschrifterkennung.

Für die Ausgestaltung von Datenerfassungsmaterial hat dieser technische Stand nur wenige Konsequenzen. Die Abbildung 5a zeigt etwa einen Fragebogen „klassischer“ Art, der nach Befragung zunächst codiert und zur Auswertung dann manuell eingegeben

Abbildung 5: Fragebogen mit unterschiedlichen Skalenniveaus  
a: normale Version

<b>Bitte sorgfältig ausfüllen.</b>	
<b>Zutreffendes bitte ankreuzen bzw. ausfüllen</b>	
1. Geschlecht	.....
2. Welches statistisches Software-Produkt benutzen Sie am häufigsten?	.....
3. Wie sind Sie mit diesem Produkt zufrieden? Bitte Schulnote angeben	.....
4. Wieviele verschiedene Prozeduren (Verfahren) benutzen Sie durchschnittlich pro Woche aus 2. ?	..... Prozeduren
5. Wieviele Minuten benutzen Sie durchschnittlich pro Woche die Software aus 2. ?	..... Minuten
6. Haben Sie Verbesserungsvorschläge für die Software aus 2. ?	
1. ....	
2. ....	
3. ....	

Bitte sorgfältig ausfüllen.  
Zutreffendes bitte ankreuzen.

1. Geschlecht  w  m

2. Welches statistische Softwareprodukt benutzen Sie am häufigsten?  
 BMDP  ISP  P-STAT  S  
 SAS  SPSS  Sonstiges

3. Wie sind Sie mit diesem Produkt zufrieden?  
Bitte Schulnote angeben.  
 1 sehr gut  2 gut  3 befried.  
 4 ausrchn.  5 mangelh.

4. Wieviele verschiedene Prozeduren (Verfahren) benutzen Sie durchschnittlich pro Woche aus 2.?  
 0  1-2  3-5  
 6-10  mehr als 10

5. Wieviele Minuten benutzen Sie durchschnittlich pro Woche die Software aus 2.?  
 Hunderter  000  100  200  300  
 400  500  600  700  
 800  900  
 Zehner  00  10  20  30  
 40  50  60  70  
 80  90  
 Einer  0  1  2  3  
 4  5  6  7  
 8  9

Abbildung 5: Fragebogen mit unterschiedlichen Skalenniveaus  
b: scannergeeignete Version

werden muß. Abbildung 5b zeigt die „Scanner-Version“ der gleichen Fragestellung. Unter Berücksichtigung weniger Bedingungen kann somit mit gleichem Aufwand ein Erfassungsmaterial erstellt werden, das ohne großen Personalaufwand eingegeben werden kann.

Hervorzuheben sei noch die Umwandlung von offenen Fragen in geschlossene, wie dies in Abbildung 5 dokumentiert ist. So war etwa auch die Frage 5 nach der durchschnittlichen Nutzungsdauer in eine geschlossene Frage umformbar; die Einzelinformationen können mittels Konvertierungssoftware wieder in die ursprüngliche Information überführt werden.



## 5 Abschließende Bemerkungen

Als Resümee dieser Betrachtungen lassen sich deshalb folgende Schlußfolgerungen ziehen.

- Offene Fragen und Beschreibungen lassen sich z. Zt. nicht automatisch erfassen.
- Für geschultes Personal ist die z. Zt. vorhandene Technik dann einsetzbar, wenn das Datenerfassungsmaterial (mit wenig Aufwand) entsprechend technischer Vorgaben umgestaltet wird. Damit ist eine erhebliche Zeit- und Kostenersparnis bei der Dateneingabe zu erreichen.
- Hilfspersonal, das nicht ausreichend geschult ist, kann nur bei einfachsten Fragen mit der Technik konfrontiert werden, so daß der Einsatz hier nur eingeschränkt möglich ist.
- Kurzfristige Änderungen am Inhalt und Layout des Erfassungsmaterials lassen sich schnell selbst vornehmen (etwa mit Hilfe eines Textverarbeitungs-systems oder eines Formulargenerators in Verbindung mit einem Laserdrucker).
- Befragungen sollten nach automatisch lesbaren und nicht lesbaren Teilen separiert werden, damit teilweise eine Automatisierung ermöglicht wird.
- Automatisch eingelesene Daten lassen sich mit Hilfe einer Konvertierungssoftware (Softwarefilter) in die gewünschten Formatstrukturen der jeweiligen Datenbank- bzw. Statistik-Software umwandeln.

Ausgehend von Erfahrungen über einen Antwortbias bei Verwendung von Antwortvorgaben bei geschlossenen Fragen (SCHWARZ et al. 1989) kann man bei der automatischen Datenerfassung in Erwägung ziehen, jedem Befragten einen „persönlichen“ Fragebogen zu erstellen. Dies könnte durch Permutation (Nominalskala) bzw. durch Inversion (Ordinalskala) von Antwortvorgaben geschehen. Aber auch die Permutation von Fragesequenzen, soweit diese nicht durch Filterfragen in eine Abhängigkeitsstruktur eingebunden sind, läßt sich in diese Überlegungen einbeziehen. Dies gilt vor allem für die immer mehr an Bedeutung gewinnenden epidemiologischen Routine- oder Monitoringuntersuchungen, bei denen über Jahre ein Erhebungsaufwand entsteht.

Als abschließende Bemerkung möchten wir darauf hinweisen, daß dieser Beitrag nicht als Aufruf zu einem weiteren Personalabbau zu Gunsten von technischem Gerät verstanden werden soll. Vielmehr ist unser Anliegen diese Hilfsmittel so einzusetzen, daß vorhandene Personalressourcen (nicht nur innerhalb der Epidemiologie) mehr für wissenschaftliche Inhalte als für technische Details genutzt werden. Beim jetzigen Stand der Technik erscheint dies möglich und wünschenswert.

## Literatur

- HOLM, K. (1975): Die Befragung, Band 1: Der Fragebogen. Francke, München  
 KAMMERMEIER, P. (1991): Scannen und Drucken. Addison-Wesley, Bonn/München  
 KREIENBROCK, L. (1989): Stichprobenverfahren. Oldenbourg, München/Wien  
 KURZWEIL/DISCOVER (1989): Kurzweil/Discover User's Guide. Xerox Imaging Systems, Inc.  
 SCHWARZ, N., HIPPLER, H.-J., NOELLE-NEUMANN, E. (1989): Einflüsse der Reihenfolge von Antwortvorgaben bei geschlossenen Fragen. ZUMA-Nachrichten 25, p. 24–38  
 WEDELICH, F. (1991): Zeichen erkennen. Bild der Wissenschaft 8/91, p. 92–93

# Epidemiologische Methoden und Wege der Risikoabschätzung

Heiko Becher, Karen Steindorf

## Zusammenfassung

*Risikoabschätzungen basieren auf epidemiologischen und/oder experimentellen Daten. Die Klasse der epidemiologischen Studien, die für eine Risikoabschätzung geeignet sein können, wird beschrieben und eine formale Definition relevanter Begriffe, wie das Risiko, die Dosis und verwandte Maßzahlen, wird vorgenommen. Die Verwendung moderner Regressionsverfahren zur Schätzung von Dosis-Wirkungs-Beziehungen wird erläutert. Wir beschreiben einen häufig verwendeten populationsbezogenen Parameter, das sogenannte Unit Risk, und entwickeln Methoden zu seiner Schätzung. Der gesamte Weg der Risikoabschätzung wird an dem Beispiel des Kanzerogens Arsen verdeutlicht: Die Auswahl geeigneter Studien, die Anpassung von Dosis-Wirkungs-Kurven und die Schätzung des Unit Risk. Ein Ausblick auf weitere Forschungsansätze wird gegeben.*

## Schlüsselwörter:

Risikoabschätzung; Epidemiologie; Dosis-Wirkungs-Beziehungen; Unit Risk; Arsen.

## Summary

*Risk assessment is based on epidemiologic and/or experimental data. The class of epidemiologic studies that may be used for risk assessment is described and a formal definition of relevant terms, such as risk, dose and related measures is given. The use of modern regression methods for obtaining dose-response-relationships is demonstrated. We describe a parameter which has become popular in recent past, the so-called Unit Risk, and present some estimates. The whole process of risk assessment is exemplified using the carcinogen arsenic: The selection of appropriate studies, the fit of dose-response models and unit risk estimation. Directions for further research are indicated.*

## Key Words:

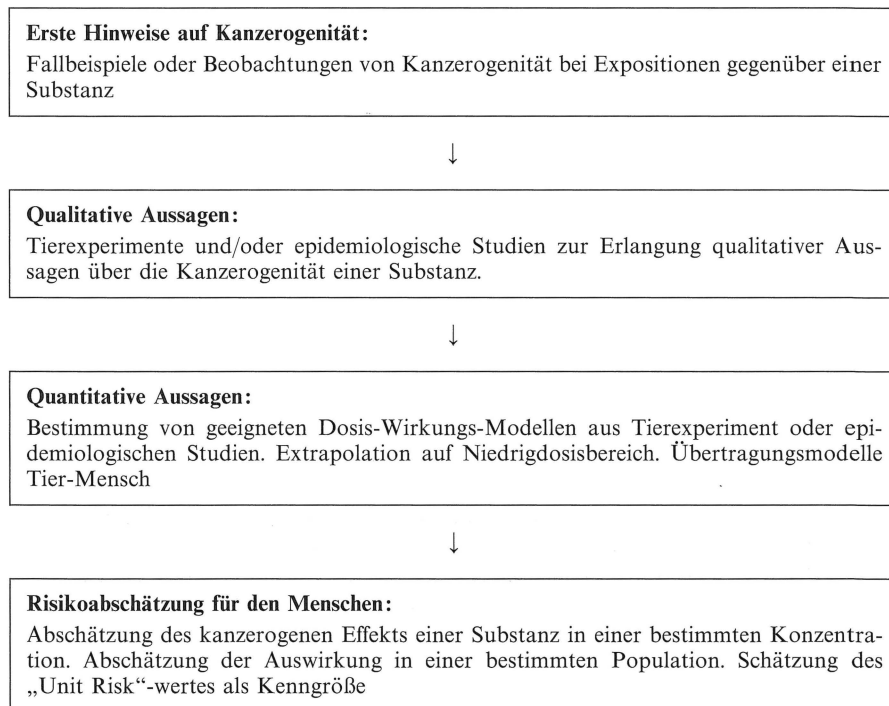
Risk Assessment; Epidemiology; Dose-Response; Unit Risk; Arsenic

## 1. Einleitung

Risikoabschätzung ist eine Methode, Risiken zu untersuchen, um sie besser vermeiden oder reduzieren zu können (WILSON & CROUCH, 1987). Was sind Risiken und wie können sie untersucht werden? Kann man Risikoabschätzung benutzen, um eine Rangfolge kanzerogener Gefahrstoffe anzugeben? In diesem Beitrag soll aufgezeigt werden, wie in der Epidemiologie aus einer geeigneten Studie Kenngrößen abgeleitet werden können, die die Gefährdung einer Population durch eine Substanz beschreiben. Dies kann zum Beispiel die Gefährdung durch eine Umweltbelastung in bestimmter Höhe sein, die für die Festlegung von Grenzwerten untersucht werden soll. Ebenso relevant kann die Untersuchung der Gesundheitsgefährdung durch Lebensstilfaktoren wie Tabakkonsum sein. Der methodische Weg soll theoretisch dargestellt und an einem Beispiel verdeutlicht werden. Es gibt einige Monographien im angelsächsischen Sprachraum, die zur Ergänzung und Vertiefung herangezogen werden können (GORDIS, 1988; MCCOLL, 1987; MOOLGAVKAR, 1990). Krebserzeugende Stoffe stehen in diesem Beitrag im Vordergrund, die Resultate gelten aber sinngemäß auch bei anderen Endpunkten als Krebs. Der Weg zur Risikoabschätzung kann durch folgende Graphik dargestellt werden.

Am Anfang stehen oft Fallbeispiele, in denen Ärzte aufgrund einer Beobachtung in ihrer Praxis den Verdacht äußern, daß eine bestimmte Exposition oder – synonym – eine bestimmte Belastung die Krankheit ihres Patienten hervorgerufen haben kann. Für eine – zunächst qualitative – Absicherung einer solchen Vermutung sind Studien erforder-

Abbildung 1: Der Weg zur Risikoabschätzung



aus: WAHRENDORF & BECHER (1990)

lich, bei denen es sich um tierexperimentelle Arbeiten oder um epidemiologische Studien handeln kann. Nicht überall wird versucht, darüberhinaus eine quantitative Bewertung von Substanzen vorzunehmen. So beschränkt sich die International Agency for Research on Cancer (IARC) mit Sitz in Lyon/Frankreich derzeit auf eine rein qualitative Klassifikation der Substanzen hinsichtlich ihrer Kanzerogenität in eine der vier Gruppen (IRAC, 1987):

- sicher kanzerogen für den Menschen (Gruppe 1)
- wahrscheinlich oder möglicherweise kanzerogen (Gruppen 2a und 2b)
- nicht klassifizierbarer Stoff (Gruppe 3)
- vermutlich nicht kanzerogen für den Menschen (Gruppe 4).

Für eine quantitative Risikoabschätzung muß versucht werden, aus geeigneten epidemiologischen oder tierexperimentellen Studien Dosis-Wirkungs-Beziehungen zwischen der Substanz und dem beobachteten Effekt herzuleiten. Toxikologie und Epidemiologie können sich hierbei wirkungsvoll ergänzen. Ist es gelungen, aus den vorhandenen Daten eine Dosis-Wirkungs-Beziehung aufzustellen, kann versucht werden, die Auswirkung einer Substanz in einer bestimmten Konzentration auf eine bestimmte Population abzuschätzen. Es liegen nicht immer epidemiologische Daten vor, die zur Risikoabschätzung geeignet sind, so daß Tierexperimente die Basis der Entscheidung sind. Beruht die Kenntnis der Dosis-Wirkungs-Beziehung auf tierexperimentellen Daten, so muß ein geeignetes Übertragungsmodell Tier-Mensch angewendet werden. Die Wahl eines solchen Modells ist ein „sehr unsicherer Prozeß in der Extrapolation vom Tier zum Menschen“ (KALDOR & DAY, 1985). Einen pragmatischen Ansatz bietet die sogenannte „Comparative Potency Method“ (WICHMANN, 1990), bei der angenommen wird, daß die relative Wirksamkeit zweier Substanzen im Tierexperiment und beim Menschen gleich ist. Der Endpunkt der Risikoabschätzung, dem dann das Risikomanagement (Risk-Management) folgt, besteht häufig in der Angabe einer bestimmten Kenngröße, z. B. des Unit Risk.

Im weiteren werden vor allem die epidemiologischen Aspekte einer Risikoabschätzung erläutert. Weitergehende Betrachtungen zu tierexperimentellen Untersuchungen gehen über den Rahmen des vorliegenden Artikels hinaus und finden sich z. B. in ROLLER et al. (1992) oder WICHMANN & IHME (1992). Es soll zunächst versucht werden, einige Kriterien für epidemiologische Studien zu diskutieren, die zur Risikoabschätzung benutzt werden können. Danach wird besonders auf den Begriff des Risikos und Methoden bei der Wahl der Dosisvariablen eingegangen. In einem weiteren Abschnitt wird ein Parameter vorgestellt, der in den letzten Jahren zunehmend an Bedeutung gewonnen hat, das Unit Risk. Nach einer mathematisch exakten Definition wird auf Schätzverfahren für das Unit Risk eingegangen. In einem folgenden Abschnitt wird anhand eines Beispiels – der Risikoabschätzung für Arsen – der gesamte Weg der Risikoabschätzung dargestellt.

## 2. Epidemiologische Studien zur Risikoabschätzung

Epidemiologische Untersuchungen können in deskriptive und analytische Studien unterteilt werden. Während in deskriptiven Studien, z. B. Korrelationsstudien, Hinweise auf das Vorhandensein bestimmter Risikofaktoren gegeben werden können, erlauben es analytische Studien, den Effekt eines Faktors auf eine Erkrankungswahrscheinlichkeit zu schätzen und dabei auch Störgrößen zu berücksichtigen. Als wesentliche Studientypen sind *Fall-Kontroll-Studien* und *Kohortenstudien* zu nennen, die fast ausschließlich Basis der Risikoabschätzungen sind. Beide Studientypen haben ihre spezifischen Vor- und Nachteile, wobei hier nur auf die entsprechende Literatur verwiesen werden soll (BRESLOW & DAY, 1980, 1987; ROTHMAN, 1986; AHLBOHM & NORELL, 1991).

Nicht jede Fall-Kontroll-Studie bzw. jede Kohortenstudie ist zur quantitativen Risikoabschätzung geeignet. Neben allgemeinen Qualitätskriterien für epidemiologische Studien müssen insbesondere Daten über die Höhe und den Verlauf der Schadstoffbelastung vorliegen. Es gibt keine klar formulierten objektiven Maßstäbe oder eine festgelegte Rangfolge von Kriterien, nach denen die Qualität einer epidemiologischen Studie eindeutig bestimmt werden kann. Für jeden Einzelfall ist eine spezielle Betrachtung erforderlich, so daß bisweilen je nach Sichtweise ein und dieselbe Studie völlig unterschiedlich beurteilt wird. Daraus ergibt sich, daß Unterschiede in der Bewertung der Basisdaten auch zu Kontroversen bei der Beurteilung der daraus resultierenden Risikoabschätzung führen. Einige allgemeine Kriterien, deren Beurteilung im konkreten Fall schwierig sein kann, und deren Behandlung über den Rahmen dieses Beitrags hinausgehen würde, sind:

- Die Studienpopulation ist geeignet ausgewählt worden,
- die erforderlichen Daten wurden vollständig und hinreichend verzerrungsfrei erhoben,
- der Einfluß wichtiger Störfaktoren (confounder) konnte angemessen berücksichtigt werden.

Zu allgemeinen Aspekten bei Planung und Durchführung analytischer epidemiologischer Studien kann hier nur auf die Literatur verwiesen werden. (BRESLOW & DAY, 1980, 1987; ROTHMAN, 1986; AHLBOHM & NORELL, 1991).

Expositionsspezifische Kriterien lassen sich leichter formulieren, es ist jedoch ebenso schwierig zu beurteilen, ob sie als erfüllt gelten können:

- Es liegen ausreichende Informationen über die Höhe und den zeitlichen Verlauf der Schadstoffbelastung vor,
- es kann mit hinreichender Sicherheit ausgeschlossen werden, daß ein beobachteter Effekt nicht durch eine Mischexposition mit anderen Substanzen hervorgerufen wurde.

Im allgemeinen ist es schwierig zu entscheiden, wann die Information über die Höhe und den zeitlichen Verlauf der Belastung als „ausreichend“ bezeichnet werden kann. Eine Dosisbestimmung, die in ihrer Präzision mit der von Tierexperimenten vergleichbar ist, liegt generell nicht vor. Bisweilen ist es so, daß die Angabe einer Dosis nicht mehr als ein „best guess“, also die plausibelste Schätzung, ist. In diesen Fällen müssen die daraus resultierenden Risikoabschätzungen mit besonderer Vorsicht interpretiert werden. Beispiele für Dosisbestimmungen in epidemiologischen Studien werden in Abschnitt vier angegeben.

### 3. Der Risikobegriff

In der Literatur zur Risikoabschätzung wird häufig davon ausgegangen, daß der Terminus „Risiko“ keiner zusätzlichen Erklärung bedarf. Bisweilen werden verbale Definitionen gegeben, die aber jegliche mathematische Exaktheit vermissen lassen. Es soll daher in diesem Kapitel versucht werden, den Begriff des Risikos aus der Sicht derjenigen Fachrichtungen herzuleiten, auf denen die Risikoabschätzung basiert: der Statistik bzw. Mathematik, der Epidemiologie und der Toxikologie.

Der Risikobegriff ist eng an den Begriff der Wahrscheinlichkeit  $P$  gekoppelt. Zur exakten Definition ist es nötig, zunächst einige Begriffe aus dem Bereich der Überlebenszeitanalyse einzuführen (siehe z. B. KALBFLEISCH & PRENTICE, 1980). Sei  $T$  eine positive, stetige Zufallsvariable für die Überlebenszeit,  $S(t) = P(T > t)$  die Überlebensfunktion (survival function) für das Alter  $t$  und

$$\lambda(t) = \lim_{\Delta \rightarrow 0} \frac{P(t < T < t + \Delta \mid T > t)}{\Delta}$$

die sog. Hazard-Rate für das Alter  $t$ . Das *Risiko*, oder die *Wahrscheinlichkeit*, daß eine Person im Zeitintervall  $(0, t)$  stirbt, wird mit  $P(t)$  bezeichnet. Der Zusammenhang zwischen  $P(t)$ ,  $S(t)$  und  $\lambda(t)$  ergibt sich mit (BRESLOW & DAY, 1980, p. 51)

$$P(t) = 1 - S(t) = 1 - \exp\left(-\int_0^t \lambda(u) du\right).$$

Oft ist das Zielereignis nicht „Tod“, sondern speziell „Tod durch eine bestimmte Krankheit  $D$  (disease)“, oder auch „Auftreten einer Krankheit  $D$ “ von Interesse. In diesem Fall muß geeignet berücksichtigt werden, daß sog. „konkurrierende Todesursachen“ (competing risks) auftreten können. Dazu benötigt man die krankheitsspezifische Hazard-Rate

$$\lambda_D(t) = \lim_{\Delta \rightarrow 0} \frac{P(t < T < t + \Delta, D | T > t)}{\Delta}$$

(vgl. ROBINS & GREENLAND, 1989). Das Risiko, daß eine Person im Zeitintervall  $(0, t)$  an einer Krankheit  $D$  stirbt,  $P(t, D)$ , ist

$$P(t, D) = \int_0^t \lambda_D(u) S(u) du.$$

Diese Gleichung ergibt sich direkt aus der Tatsache, daß

$$\lambda_D(u) = \frac{dP(t, D)}{du} \frac{1}{S(u)} \quad \text{und somit auch} \quad \lambda_D(u) S(u) = \frac{dP(t, D)}{du}$$

gilt. Durch Integration erhält man das obige Ergebnis.

Betrachten wir nun einen Faktor  $X$ . Wenn man davon ausgehen kann, daß mit dem Faktor  $X$  die Wahrscheinlichkeit für  $D$  erhöht ist, spricht man oft direkt von einem Risikofaktor  $X$ . Das *Zusätzliche Risiko* oder *Exzess-Risiko*  $ER(t, D, x)$  ist definiert als die zusätzliche Wahrscheinlichkeit  $P$  für das Auftreten der Krankheit  $D$  unter dem Faktor  $X$  mit Wert  $x$  im Zeitraum  $(0, t)$  (cumulative hazard), d. h.

$$\begin{aligned} ER(t, D, x) &= P(t, D | X = x) - P(t, D | X = 0) \\ &= \int_0^t \lambda_D(u | X = x) S(u | X = x) du - \int_0^t \lambda_D(u | X = 0) S(u | X = 0) du. \end{aligned}$$

Bei der Berechnung wird für  $t$  bisweilen die Lebensspanne 70 Jahre willkürlich festgelegt (WHO, 1987); in anderen Berechnungen wird die gesamte Lebensdauer betrachtet. Man spricht dann von dem *Lebens-Exzess-Risiko* (excess lifetime risk)

$$ER(D, x) = \int_0^{\infty} \lambda_D(u | X = x) S(u | X = x) du - \int_0^{\infty} \lambda_D(u | X = 0) S(u | X = 0) du.$$

Sowohl die Überlebensfunktion als auch die Hazard-Rate sind spezifisch für eine bestimmte Population, so daß das Exzess-Risiko immer populationsspezifisch ist.

Wie bereits angedeutet, kann unter dem Zielereignis der Tod an dieser Krankheit (Mortalität) oder das Auftreten einer bestimmten Krankheit (Inzidenz) gemeint sein.

Unter der Bedingung, daß der Faktor  $X$  die Erkrankungswahrscheinlichkeit nicht senkt, liegt das so definierte Exzess-Risiko immer zwischen 0 und 1. Das Exzess-Risiko wird häufig im Widerspruch zu den obigen Definitionen abgekürzt mit „Risiko“ bezeichnet. Meist wird aus dem Kontext heraus die jeweilige Bedeutung erkennbar. Es soll aber noch einmal deutlich gemacht werden, daß das Exzess-Risiko nicht zu verwechseln ist

mit der bedingten Wahrscheinlichkeit für das Auftreten der Krankheit  $D$  unter der Exposition gegenüber dem Faktor  $X$ . Falls kein Zusammenhang zwischen  $D$  und  $X$  besteht, ist  $ER = 0$ , da die Wahrscheinlichkeit für die Krankheit unter den Exponierten genauso groß ist, wie unter den nicht Exponierten. Dies bedeutet selbstverständlich nicht, daß die Wahrscheinlichkeit für die Krankheit unter der Exposition null ist.

#### 4. Der Dosisbegriff

Die in der Toxikologie und Epidemiologie verwendeten Dosis-Wirkungs-Beziehungen hängen stark von dem verwendeten Maß für die Dosis ab. Der Dosisbegriff in der Epidemiologie soll deshalb näher erläutert werden, vgl. auch WICHMANN (1990).

Risikoabschätzungen werden für alle möglichen Arten von Expositionen durchgeführt: Berufliche oder Umweltbelastung mit Chemikalien, Strahlenbelastungen, Ernährungsfaktoren, Tabakkonsum, etc. Entsprechend dieser großen Spannweite gibt es verschiedene gebräuchliche Ansätze, aus der Exposition bzw. Belastung durch eine Substanz über eine bestimmte Zeit in einer bestimmten Konzentration oder Menge die Dosis (-variable) zu berechnen, die dann in die Dosis-Wirkungs-Analyse einfließt (CHECKOWAY et al., 1989). Zu diesen gehören z. B.

- die kumulative Belastung, aufsummiert über den gesamten Belastungszeitraum (Einheit z. B.  $\mu\text{g}/\text{m}^3 \times \text{Jahre}$ ),
- die durchschnittliche Belastung, wie sie z. T. beim Abschätzen des Lungenkrebsrisikos bei Zigarettenrauchen verwendet wird (Einheit z. B. Zigaretten/Tag),
- die maximale Belastung während des Untersuchungszeitraums (Einheit z. B.  $\mu\text{g}/\text{m}^3$ ),
- die Strahlenbelastung bei strahlenbelasteten Personen (Einheit z. B. rad).

Es muß sorgfältig beachtet werden, für welche Dosisvariable eine Dosis-Wirkungs-Beziehung angegeben ist. Insbesondere bei Umweltrisiken besteht häufig die Frage nach dem Risiko unter einer lebenslangen konstanten Exposition einer Substanz in der Atemluft in einer bestimmten Konzentration, d. h. die Dosisvariable in der Dosis-Wirkungs-Analyse ist hier die Konzentration in der Atemluft. Als Basis zur Schätzung dieser Kenngröße liegen allerdings meist Daten mit einem anderen Expositionsmuster vor, so daß entsprechende Umrechnungen erforderlich sind. Werden z. B. Daten über die mittlere Arbeitsplatzbelastung verwendet, so ist diese Umrechnung relativ einfach (Beispiel in Abschnitt 7). Die Umrechnung einer Dosis-Wirkungs-Beziehung, bei der das absolute oder relative Risiko auf der Basis der kumulativen Belastung angegeben ist, wird ebenfalls in Abschnitt 7 erläutert. Extremfälle sind einmalige Spitzenbelastungen, wie z. B. Exposition bei einem Arbeitsunfall oder – als bekanntes Beispiel – die strahlenbelastete Kohorte von Hiroshima und Nagasaki, die intensiv untersucht wurde, um Strahlenrisiken zu schätzen (BEIR, 1990).

Die Erfassung der Dosis kann auf verschiedene Arten erfolgen, wobei je nach Substanz einer bestimmten Methode der Vorzug zu geben ist. Meist ist man allerdings nicht in der Lage, mehrere Datenquellen zu haben, so daß ein wechselseitiger Vergleich und damit eine Art Validierung nicht möglich ist. Folgende Datenquellen zur Dosiserhebung sind in der Praxis relevant:

*Dosis aus Fragebogenangaben:* Bei Rauchgewohnheiten können die in einer persönlichen Befragung erhobenen Angaben geeignet sein, die damit verbundenen Risiken quantitativ abzuschätzen (z. B. JEDRYCHOWSKI et al., 1992).

*Dosis aus Arbeitsplatzkonzentrationsmessungen:* Die Schadstoffbelastung, der ein Arbeiter ausgesetzt war, wird häufig mit Hilfe von Konzentrationsmessungen geschätzt. Durch Multiplikation mit der Dauer der Beschäftigung erhält man die kumulative Belastung (z. B. ENTERLINE et al., 1987).

*Dosis aus Krankenakten:* Bei der Untersuchung von Risiken nach Strahlentherapie kann

es möglich sein, die individuellen Strahlendosen direkt aus Krankenakten zu entnehmen (z. B. BOICE et al., 1987).

*Dosis aus biologischen Markern:* Bei Schadstoffen mit langer Halbwertszeit im Körper (Dioxin) oder bei Untersuchung von Kurzeffekten (Schadstoffkonzentration im Urin) können biologische Marker herangezogen werden (z. B. FLESCHE-JANYS et al., 1992; RIBOLI et al., 1990).

Ein weiteres Problem bei der Dosisbeurteilung betrifft den Zeitraum von Exposition bis zum Ausbruch der Krankheit, die sogenannte Latenzzeit. Latenzzeiten sind nicht exakt bestimmbar, insbesondere bei Krebs können sehr lange Zeitspannen zwischen Exposition und Diagnose der Krankheit liegen. Unter Berücksichtigung der Latenzzeit werden nur die Expositionen betrachtet, die eine bestimmte Zeit zurückliegen, da die in jüngster Vergangenheit erfolgte Exposition möglicherweise keinen (oder geringeren) Einfluß auf die Krankheit hat.

## 5. Dosis-Wirkungs-Beziehungen in der Epidemiologie

Neben dem Exzess-Risiko spielt in der Epidemiologie das *Relative Risiko* ( $RR$ ) eine wichtige Rolle, da nicht in allen Typen von analytischen Studien, z. B. nicht in Fall-Kontroll-Studien, absolute Risiken schätzbar sind. Das relative Risiko ist definiert als

$$RR(D, x) = \frac{P(D | X = x)}{P(D | X = 0)},$$

also als das Verhältnis der Erkrankungswahrscheinlichkeiten mit bzw. ohne Exposition. Verwandte Maßzahlen sind das Odds Ratio oder das standardisierte Mortalitätsverhältnis (standardized mortality ratio, SMR). Oft wird die Annahme gemacht, daß das relative Risiko nicht vom Alter abhängt. Es kann dann auch als der Quotient der Hazard-Raten und damit als Ratenverhältnis (rate ratio) bezeichnet werden.

Möglichkeiten, diesen Wert in epidemiologischen Studien zu schätzen, sind in der entsprechenden Literatur beschrieben (z. B. BRESLOW & DAY, 1980, 1987). Eine Methode zur Schätzung einer solchen Dosis-Wirkungs-Beziehung aus Fall-Kontroll-Studien ist die logistische Regression. Dieses Regressionsmodell hat die Form  $\text{logit } P(D | x) = \alpha + \beta x$ . Dabei sind  $\alpha$  und  $\beta$  die zu schätzenden Regressionsparameter, und  $\text{logit } (p)$  bezeichnet die logistische Transformation  $\text{logit } (p) = \log(p/(1 - p))$ . In diesem Modell wird ein exponentieller Anstieg des relativen Risikos mit der Dosis  $x$  der Form  $RR(x) = \exp(\beta x)$  angenommen. Es können auch in Erweiterung dieses Modells beliebige andere funktionale Zusammenhänge zwischen der Dosis und dem relativen Risiko geschätzt werden.

Einen entsprechenden regressionsanalytischen Ansatz zur Schätzung einer Dosis-Wirkungs-Beziehung aus einer Kohortenstudie bieten die sogenannte Poisson-Regression (BRESLOW & DAY, 1987) und die Cox-Regression (COX, 1972). Bei der Poisson-Regression wird durch geeignete Gruppierung der verschiedenen Expositionen die Zahl der auftretenden Erkrankungs- bzw. Todesfälle in Abhängigkeit von der Dosis modelliert, wohingegen bei der Cox-Regression die individuelle Dosis direkt berücksichtigt wird. Auch hier ist eine Erweiterung zur Analyse beliebiger funktionaler Zusammenhänge möglich.

Häufig liegt auch der Fall vor, daß lediglich eine mittlere Dosis für die exponierte Gruppe bekannt ist. Da somit nur eine Schätzung des relativen Risikos für einen Dosiswert möglich ist, kann streng genommen nicht von einer „Dosis-Wirkungs-Beziehung“ gesprochen werden. Dennoch läßt sich in diesem Fall eine Risikoabschätzung durchführen, wenn auch mit eingeschränkter Aussagekraft.



## 6. Risikoabschätzung für Populationen – das Unit Risk

Durch die Angabe einer Dosis-Wirkungs-Beziehung der Form  $RR(x) = f(x)$  ist ein Merkmal der Risikoabschätzung bereits erfüllt, da so ein funktionaler Zusammenhang zwischen der Dosis und dem erzielten Effekt, hier angegeben durch das relative Risiko, vorliegt. Mit der Bestimmung des relativen Risikos ist das Problem allerdings in der Regel nicht gelöst, da aus gesundheitspolitischer Sicht die Angabe von absoluten Risiken wichtig ist. Das relative Risiko gibt nur bedingt einen Hinweis darauf, da die Prävalenz der Krankheit dabei nicht einfließt.

Um eine Rangfolge der Toxizität oder Kanzerogenität bestimmter Substanzen aufstellen zu können, wird die Dosis-Wirkungs-Funktion auf eine Maßzahl reduziert, indem das Risiko auf eine bestimmte Dosis bezogen wird. Dabei hat sich bei Umweltschadstoffen in der Atemluft eingebürgert, das Exzess-Risiko anzugeben, das mit einer konstanten Konzentration von  $1 \mu\text{g}/\text{m}^3$  verbunden ist. Das mit einer lebenslangen konstanten Exposition dieser Höhe verbundene Exzess-Risiko wird als *Unit Risk* bezeichnet. Fehlende exakte Definitionen für den Begriff des Risikos haben jedoch dazu geführt, daß in der Literatur das Unit Risk häufig nur heuristisch eingeführt wird. Eine Definition im Sinne des Risikobegriffs in Abschnitt drei findet man in EPA (1984), CRUMP & ALLEN (1987), oder WAHRENDORF & BECHER (1990) mit

$$\text{Unit risk} = P(D \mid \text{konstante Exposition } 1 \mu\text{g}/\text{m}^3) - P(D \mid \text{keine Exposition}).$$

Es gibt andere verbale Definitionen, z. B. von der WHO (1987): *Unit risk = risk associated with a lifetime exposure to  $1 \mu\text{g}/\text{m}^3$*  oder des LAI (Länderausschuß für den Immissionschutz) (1992): *Unit risk = Schätzwert des Risikos pro Doseinheit*, die inhaltlich das gleiche aussagen sollen. Mit der Wahl der Konzentration  $1 \mu\text{g}/\text{m}^3$  ist eine implizite Einschränkung verbunden, da diese Einheit im wesentlichen nur für Luftschadstoffe sinnvoll ist. Auch hier gibt es z. B. mit Radon – einem radioaktiven Edelgas, das als Risikofaktor für Lungenkrebs gilt – eine Ausnahme, da hier die Einheit Becquerel/ $\text{m}^3$  angemessen ist. Für Substanzen im Boden oder im Trinkwasser müssen entsprechende andere Einheiten gewählt werden.

Wie bereits erwähnt, sind sowohl die Überlebensfunktion als auch die Hazard-Funktion spezifisch für eine bestimmte Population. Als Konsequenz folgt daraus, daß das Unit Risk derselben Substanz in verschiedenen Populationen verschieden sein kann. Dies mag zunächst verwundern, wird aber durch folgende Überlegung plausibel: Hat ein bestimmter anderer Risikofaktor, z. B. Rauchen, in verschiedenen Populationen eine unterschiedliche Prävalenz, ist zwar das *relative Risiko* für den interessierenden Faktor  $X$  in beiden Populationen gleich, aber die Wahrscheinlichkeiten für die Krankheit mit oder ohne Exposition gegenüber  $X$  sind größer und damit auch das *absolute Risiko*. In der praktischen Umsetzung der ermittelten Unit-Risk-Werte wird implizit angenommen, daß die einzelnen Risiken additiv sind (LAI, 1992). Auf der anderen Seite gibt es Evidenz dafür, daß der gemeinsame Effekt verschiedener Substanzen auf das relative Risiko nicht additiv, sondern multiplikativ ist, und damit auch nicht von einer Additivität der Unit-Risk-Werte ausgegangen werden kann. Für kleine relative Risiken, wie sie hier vorliegen, sind die Unterschiede jedoch minimal, so daß die Methode des LAI gerechtfertigt ist.

Die nachfolgend vorgestellten Methoden zur Schätzung dieses Parameters verwenden die Mortalitätsraten der betrachteten Population, um das Grundrisiko abzuschätzen.

*Methode 1* (EPA, 1984; CRUMP & ALLEN, 1986):

Es seien  $l_i$ ,  $i = 1, \dots, n$ , die jährlichen Sterberaten der zugrundeliegenden Population für die Altersgruppe  $i$  (5-Jahres Intervalle) für alle Todesursachen und  $d_i$ ,  $i = 1, \dots, n$ , die

entsprechenden Sterberaten für die betreffende Krankheit  $D$ . Es ist üblich, Sterberaten pro 100000 Personen anzugeben (BECKER et al., 1984). Es gilt also

$$d_i = \frac{\text{Anzahl Todesfälle pro Jahr in Altersgruppe } i}{\text{Anzahl Personen in Altersgruppe } i} \times 100000.$$

Die Wahrscheinlichkeit für ein Individuum, im  $i$ -ten 5-Jahres-Altersintervall an der Krankheit  $D$  zu sterben, unter der Bedingung, daß es zu Beginn dieses Intervalls noch lebt, ist dann  $5(d_i/100000)$ . Für das letzte Intervall (Alter 85 und mehr) ist diese Wahrscheinlichkeit approximativ  $d_{18}/l_{18}$ . Damit wird die Wahrscheinlichkeit  $P_0 = P(D \mid \text{keine Exposition})$  geschätzt durch

$$P_0 = \sum_{i=1}^{17} \frac{5d_i}{100000} \prod_{j=1}^{i-1} \frac{1 - 5l_j}{100000} + \frac{d_{18}}{l_{18}} \prod_{j=1}^{17} \frac{1 - 5l_j}{100000}.$$

Wir gehen davon aus, daß in der Dosis-Wirkungs-Analyse das relative Risiko in Abhängigkeit von der kumulativen Belastung geschätzt wurde, daß also  $\widehat{RR}(x)$  vorliegt. Dann kann analog zu  $P_0$  die Sterbewahrscheinlichkeit bei einer Exposition unter einer konstanten lebenslangen Konzentration von  $1 \mu\text{g}/\text{m}^3$  berechnet werden. Dazu muß aus der durchschnittlichen Konzentration  $1 \mu\text{g}/\text{m}^3$  auf eine entsprechende kumulative Belastung geschlossen werden. Ausgehend von den in den Bevölkerungsstatistiken üblichen 5-Jahres-Altersintervallen ergibt sich die kumulative Exposition für das  $i$ -te Altersintervall aus  $x_i = [5i-2.5] \times 1 \mu\text{g}/\text{m}^3 \times \text{Jahre}$ .

Die Wahrscheinlichkeit  $P(D \mid \text{konstante lebenslange Konzentration von } 1 \mu\text{g}/\text{m}^3)$  kann somit geschätzt werden durch

$$P_1 = \sum_{i=1}^{17} \frac{5d_i^*}{100000} \prod_{j=1}^{i-1} \frac{1 - 5l_j^*}{100000} + \frac{d_{18}^*}{l_{18}^*} \prod_{j=1}^{17} \frac{1 - 5l_j^*}{100000}.$$

mit  $d_i^* = d_i \times \widehat{RR}(x_i)$  und  $l_i^* = l_i + d_i^* - d_i$ . Die Schätzung des Unit Risk ergibt sich somit als  $\widehat{UR} = P_1 - P_0$ .

### Methode 2

Eine einfachere Methode zur Schätzung des Unit Risk wird in WHO (1987) benutzt. Dieses Verfahren ist anwendbar, wenn innerhalb der Studienpopulation keine unterschiedlichen Belastungskategorien vorliegen, sondern von der Angabe einer durchschnittlichen (kumulativen) Exposition auszugehen ist. Bezeichnet  $RR$  das relative Risiko (bzw.  $SMR$ ) in der exponierten Studienpopulation,  $P_0$  die Wahrscheinlichkeit für die Krankheit ohne Exposition bis zum Alter 70 Jahre (background lifetime risk) in der Zielpopulation,  $x$  die durchschnittliche lebenslange Exposition in der Studienpopulation, dann kann das Unit Risk geschätzt werden durch

$$\widehat{UR} = P_0(\widehat{RR} - 1)/x.$$

Bei dieser Schätzung wird implizit ein linearer Zusammenhang zwischen der durchschnittlichen Belastung und dem relativen Risiko vorausgesetzt. Dies ist insofern sinnvoll, als nur die durchschnittliche lebenslange Exposition in der gesamten Studienpopulation als bekannt vorausgesetzt wird. Dies sieht man leicht, indem man die lineare Funktion  $RR(x) = 1 + \beta x$  in die allgemeine Darstellung des Exzess-Risikos  $ER(D, x) = P_0 \times RR(x) - P_0$  einsetzt. Für das Unit-Risk mit  $x = 1$  ergibt sich daraus die obige Schätzformel.

### Methode 3

Eine weitere Methode, die einen Kompromiß der beiden obigen Methoden darstellt, wird in WAHRENDORF & BECHER (1990) für Kohortenstudien vorgestellt und läßt sich

unmittelbar auf Fall-Kontroll-Studien erweitern. Es wird vorausgesetzt, daß eine Aufteilung der exponierten Untersuchungspopulation in  $k$  Gruppen, aufgeteilt nach der Höhe der Exposition, vorliegt. Dann werden für die einzelnen Belastungsgruppen separate Unit-Risk-Schätzungen nach der WHO-Formel vorgenommen und zu einem gewichteten Mittelwert nach folgender Formel zusammengefaßt:

$$\widehat{UR} = P_0 \sum_{i=1}^k \frac{\widehat{RR}_i - 1}{x_i} \frac{w_i}{w_1 + \dots + w_k}.$$

Dabei stellen analog zur vorherigen Methode  $RR_i$  das relative Risiko der exponierten Population in der  $i$ -ten Gruppe,  $x_i$  die entsprechende durchschnittliche lebenslange Belastung dieser Studienpopulation und  $w_i/(w_1 + \dots + w_k)$  ein geeignet gewähltes Gewicht dar. In WAHRENDORF & BECHER (1990) werden für  $w_i$  die gruppenspezifischen Personenjahre  $PY_i$  gewählt, es sind aber auch andere Gewichte z. B. die erwarteten Anzahlen  $E_i$  möglich. Für alle Studiendesigns bietet sich generell auch die Gewichtung mit der reziproken geschätzten Varianz des gruppenspezifischen relativen Risikos an, d. h.  $w_i = (\widehat{Var}(\widehat{RR}_i))^{-1}$ . Die genannten Vorschläge stellen lediglich Ad-hoc-Ansätze für die Wahl der Gewichtsfunktionen dar und lassen sich weiter verfeinern.

Mit Hilfe dieser Methoden kann – je nach Datenlage – das Unit Risk aus einer epidemiologischen Studie geschätzt werden. Die statistische Variabilität dieser Schätzungen ist noch nicht untersucht und ist Gegenstand eines Forschungsvorhabens der Autoren. Zusätzliche Probleme, die bei der Risikoabschätzung einer Substanz auftreten, werden anhand des folgenden Beispiels erläutert.

## 7. Ein Beispiel – Risikoabschätzung für Arsen

Arsen ist ein in der Natur weit verbreitetes Element, bei dem eine Emission in die Umwelt hauptsächlich durch Metallverhüttung erfolgt. Der Nachweis einer kanzerogenen Wirkung von Arsen (Tumoren der Atemwege) wurde im wesentlichen durch epidemiologische Kohortenstudien bei Arbeitern in Kupferhütten erbracht. In diesen Studien konnte bei Bestimmung der Dosis auf vorhandene Messungen der Arsenkonzentration in der Atemluft zurückgegriffen werden. Angaben zum Beschäftigungsbereich eines Arbeiters, dem Zeitraum der Beschäftigung und der Arsenkonzentration wurden dann benutzt, um die individuelle kumulative Dosis zu schätzen. Jede dieser Studien hat bei der Dosisbestimmung ihre spezifischen Merkmale, auf die hier nicht weiter eingegangen werden soll. Zu weiteren Details siehe BECHER & WAHRENDORF (1992) und WAHRENDORF & BECHER (1990), jedoch wurden alle hier genannten Studien von verschiedenen Organisationen (EPA, WHO, IARC) für geeignet angesehen, einen signifikanten Beitrag zur Bewertung von Arsen zu leisten.

Die Daten erlaubten eine Schätzung von Dosis-Wirkungs-Beziehungen mit den in Abschnitt 5 beschriebenen Methoden. Die in diesen Studien hergeleiteten Dosis-Wirkungs-Beziehungen haben die Form  $RR(x) = f(x)$ , wobei  $x$  die kumulative Belastung am Arbeitsplatz in der Einheit  $\text{g/m}^3 \times \text{Jahre}$  bezeichnet. Zur Umrechnung in die entsprechende Umweltbelastung wird angenommen, daß die berufliche Belastung an 8 Stunden/Tag und 240 Tagen/Jahr vorliegt. Der sich damit ergebende Faktor ist  $8/24 \times 240/365 = 0,22$ , d. h. eine Konzentration von  $1 \text{ mg/m}^3$  am Arbeitsplatz entspricht einer konstanten Konzentration von  $0,22 \text{ mg/m}^3$  in der Atemluft.

Die Wahl des funktionalen Zusammenhangs zwischen der Dosis und dem relativen Risiko ist schwierig, da weder aus medizinischer Sicht einem bestimmten Modell der Vorzug zu geben ist, noch aus statistischer Sicht ein bestimmtes Modell herausragt (EPA, 1984; WAHRENDORF & BECHER, 1990). Tabelle 1 gibt einen Überblick über Schätzungen

von Dosis-Wirkungs-Beziehungen aus den wichtigsten Kohortenstudien, die zur quantitativen Risikoabschätzung für Arsen-Exposition verwendet werden können. Die Streubreite der Unit-Risk-Schätzungen erklärt sich daraus, daß ihnen

- verschiedene Kohortenstudien,
- verschiedene Publikationen zu einer Kohorte (unterschiedliche Dauer des follow-up, Subgruppenanalysen)
- verschiedene Expositionsdaten
- unterschiedliche Annahmen zur Latenzzeit (Nichtberücksichtigen der Exposition in den letzten m Jahren)
- die Verwendung verschiedener Dosis-Wirkungs-Modelle

zugrundeliegen. Diese Schätzungen basieren alle auf dem ersten Schätzverfahren, bei dem also die Dosis-Wirkungs-Beziehungen direkt einfließen.

Zum Vergleich geben wir ein Resultat unter Verwendung des Schätzverfahrens der WHO. Für eine Untergruppe der Takoma-Kohorte schätzten ENTERLINE et al. (1987) bei einer kumulativen Belastung am Arbeitsplatz von  $12,555 \text{ mg/m}^3 \times \text{Jahre}$  ein SMR von 332,1, bei 23 beobachteten und 6,93 erwarteten Todesfällen von Tumoren der Atemwegsorgane. Die Umrechnung in eine mittlere lebenslange (70 Jahre) Umweltbelastung liefert den Wert  $x = 12,555 \text{ mg/ml}^3 \times 8/24 \times 240/365 \times 1/70 = 39,3 \text{ } \mu\text{g/m}^3$ . Für  $P_0$  (Tumoren der Atemwegsorgane, Männer, BRD) ergibt sich ein Wert von 0,053 und daraus folgt  $UR = 0,053 (3,321 - 1)/39,3 = 3,1 \times 10^{-3}$ . Dieser Wert liegt voll im Bereich der Werte, die mit dem aufwendigeren Schätzverfahren (Methode 1) erzielt wurden.

Tabelle 1: Dosis-Wirkungs-Beziehungen und Unit Risk Schätzer für Arsen

Kohorte	Autor	Dosis-Wirkungs-Beziehung	Unit Risk Schätzer $\times 10^{-3}$
Anaconda	LUBIN et al. (1981)	$1 + 0,0001175x$	0,97
		$1 + 0,0002003x$	1,7
		$1 + 0,0003389x$	2,8
	WELCH et al. (1982)	$1 + 0,0003173x$	2,6
		$\exp(0,0001142x)$	0,96
	BRESLOW et al. (1983)	$\exp(0,000038x)$	0,32
	LEE-FELDSTEIN (1983, 1986)	$1 + 0,0003754x$	3,1
		$\exp(0,0002748x)$	2,4
	ENTERLINE et al. (1987)	$1 + 0,04062x^{0,3843}$	9,8
	WAHRENDORF & BECHER (1990)	$\exp(0,00004874x)$	0,41
$\exp(0,0003416x^{1/2})$		1,7	
Takoma	ENTERLINE et al. (1987)	$1 + 0,0003228x$	2,7
		$1 + 0,04897x^{0,3499}$	9,7
		$1 + 0,04857x^{0,4081}$	13,0
		$1 + 0,0007359x$	6,8
		$1 + 0,0005237x$	4,3
8 versch. Kupferhütten	ENTERLINE et al. (1987)	$1 + 0,0003812x$	3,1
		$\exp(0,001x)$	9,5
		$\exp(0,00081x)$	7,5
		$\exp(0,0439x^{1/2})$	30,0
		$\exp(0,0093x^{1/2})$	4,8
Michigan	OTT et al. (1974)	$1 + 0,0003629x$	3,0
		$1 + 0,0009224x$	7,5

x Arsenexposition in  $\mu\text{g/m}^3 \times \text{Jahre}$  (kumulative Arbeitsplatzbelastung)

aus: WAHRENDORF & BECHER (1990)

Um aus der Gesamtheit der Ergebnisse eine gemeinsame Schätzung für das Unit Risk ableiten zu können, gilt es, Kriterien für die Güte eines Einzelergebnisses festzulegen. Zur Beurteilung der Dosis-Wirkungs-Beziehungen sollte, falls möglich, dabei sowohl ein statistisches Entscheidungskriterien als auch die biologische Plausibilität berücksichtigt werden. Die Endbewertung des LAI war  $4,0 \times 10^{-3}$  als gerundeter Wert. Die WHO legte ebenfalls einen Wert von  $4,0 \times 10^{-3}$  fest, die EPA einen Wert von  $4,3 \times 10^{-3}$ . Streng genommen läßt sich eine Unabhängigkeit verschiedener Unit-Risk-Schätzungen nur dann annehmen, wenn ihnen unterschiedliche Daten bzw. Studien zugrundeliegen. Dies ist bei den einzelnen Schätzungen derselben Studie nicht der Fall. In WAHRENDORF & BECHER (1990) wurden daher die Mediane der Schätzungen innerhalb der einzelnen Studien betrachtet, und von diesen wiederum der Mittelwert berechnet. Dieses ist ein pragmatisches Vorgehen, das den Autoren angemessen erschien. Um zu zeigen, daß die Angabe eines Schätzwertes problematisch ist, schreiben WAHRENDORF & BECHER (1990, S. 109) „Aus wissenschaftlicher Sicht halten wir es jedoch für angemessener, ein Intervall für das Unit-Risk von Arsen anzugeben, welches die Streubreite der Schätzungen dokumentiert, die sich alleine aufgrund unterschiedlicher Modellannahmen ergeben können. Das Intervall ( $10^{-3}$ ,  $10^{-2}$ ) kann nach den vorliegenden Analysen als ein Maß für diesen Schwankungsbereich angesehen werden.“

## 8. Diskussion

Für eine Bewertung von Substanzen ist die quantitative Risikoabschätzung unverzichtbar. Dabei benutzen nationale und internationale Gremien (EPA, WHO, LAI) ungeachtet der methodischen Einschränkungen die Erkenntnisse aus der Epidemiologie, sofern geeignete Studien vorliegen. Dies ist für viele Substanzen nicht zutreffend, so daß dann toxikologische Ergebnisse als ausschließliche Basis für Entscheidungen herangezogen werden müssen. Sofern jedoch gute epidemiologische *und* experimentelle Daten vorliegen, wird ersteren eine größere Aussagekraft eingeräumt. Radon oder Arsen sind nur zwei Beispiele dafür. Es gibt auch Substanzen, für die nur epidemiologische Daten zur Verfügung stehen (IARC, 1987). Die quantitative Betrachtung von Lebensstilfaktoren bezieht sich naturgemäß nur auf epidemiologische Daten. Quantitative Risikoabschätzung ist wichtig, um eine sachliche Diskussion führen zu können, welche Risiken relevant sind und welche nicht. Viele Risiken werden unbewußt als lebensbedrohlich eingestuft, weil Menschen sich ihnen ausgeliefert fühlen, wie z. B. Unfälle im Luftverkehr. In Wirklichkeit sind diese Risiken verschwindend gering im Vergleich zu denen, die man vermeintlich selbst „im Griff hat“ und damit selbst vermeiden kann (SLOVIC, 1987). Dazu zählen zum Beispiel Lebensstilfaktoren wie Rauchen, Ernährung oder Autofahren.

Quantitative Risikoabschätzung basiert häufig auf Daten mit limitierter Aussagekraft. Damit ist die alleinige Angabe eines Punktschätzers, also einer einzigen Zahl, irreführend, da so für den Laien eine nicht vorhandene Präzision vorgetäuscht wird. Während es bereits schwierig ist, die statistische Variabilität dieses mehrstufigen Schätzprozesses analytisch zu untersuchen, entziehen sich andere Unsicherheitsfaktoren, die grob mit Güte Merkmalen zu kennzeichnen sind, einer soliden statistischen Behandlung. Arsen ist hier ein vergleichsweise „gutes“ Beispiel: verschieden große Studie stehen zur Verfügung, es gibt Dosis-Abschätzungen, die von unabhängigen Experten als hinreichend valide angesehen werden, und es ergeben sich aus einzelnen Studien und verschiedenen Publikationen der selben Studie Schätzwerte, die, wenn sie auch nicht gleich sind, aber mit Ausnahme weniger Extremwerte in einem akzeptablen Intervall liegen. Für andere Substanzen ist die Datenlage weniger befriedigend.

Es gibt andere Konzepte zur Beschreibung von Risiken, deren Vor- und Nachteile im Vergleich zu dem Unit-Risk-Ansatz noch nicht hinreichend geklärt sind. Das sogenannte „loss-of-lifetime“-Konzept (ROBINS & GREENLAND, 1991), bei dem versucht wird, die mit einer bestimmten Exposition verbundene Lebenszeitverkürzung zu schätzen, ist hier als mögliche Alternative zu nennen. Alle Verfahren sollen für regulatorische Entscheidungen wissenschaftlich fundierte Grundlagen liefern. Epidemiologen sind gefordert, bestmögliche Daten zu gewinnen. Statistiker sind gefordert, durch die Wahl besserer Analysemethoden möglichst präzise Resultate vorzulegen. Sie müssen aber auch auf deren Variabilität und damit den Grenzen der Interpretierbarkeit mit aller Deutlichkeit hinweisen.

### Acknowledgement

Wir bedanken uns bei zwei Gutachtern für eine große Anzahl konstruktiver Kommentare zu der ersten Version dieses Artikels.

Gefördert vom Umweltbundesamt, Berlin, Forschungsnr. 11606089

### Literatur

1. AHLBOHM, A., NORELL, S. (1991): Einführung in die moderne Epidemiologie, BGA-Schriften 2/91, MMV Medizin Verlag München
2. BECHER, H., WAHRENDORF, J. (1992): Metalle/Arsen. In: H. E. Wichmann; H.-W. Schlipkötter; G. Fülgraff (Eds) Handbuch der Umweltmedizin. Ecomed, Landsberg; Kap. VI—3
3. BECKER, N., FRENTZEL-BEYME, R., WAGNER, G. (1984): Krebsatlas der Bundesrepublik Deutschland. Springer-Verlag, Berlin
4. BEIR, V. (1990): Health Effects of Exposure to Low Levels of Ionizing Radiation. National Academy Press, Washington
5. BOICE, J. D., BLETTNER, M., KLEINERMANN, R. A., STOVALL, M., MOLONEY, W. C., ENGHOLM, G., AUSTIN, D. F., BOSCH, A., COOKFAIR, D. L., KREMENTZ, E. T., et al. (1987): Radiation Dose and Leukemia Risk in Patients Treated for Cancer of the Cervix. *JNCI*, **79**, 1295—1311.
6. BRESLOW, N. E., DAY, N. E. (1980): Statistical Methods in Cancer Research Vol. I — The Analysis of Case-Control Studies. IARC Scientific Publications No. 32
7. BRESLOW, N. E., DAY, N. E. (1987): Statistical Methods in Cancer Research Vol. II — The Design and Analysis of Cohort Studies. IARC Scientific Publications No. 82.
8. BRESLOW, N. E., LUBIN, J. H.; MAREK, P., LANGHOLZ, B. (1983): Multiplicative Models and Cohort Analysis. *J Am Stat Assoc*, **78**, 1—12
9. CHECKOWAY, H., PEARCE, N., CRAWFORD-BROWN, D. J. (1989): Research Methods in Occupational Epidemiology. Oxford University Press. New York
10. COX, D. R. (1972): Regression Models and Life Tables (with discussion). *J. R. Stat. Soc. B*, **34**, 187—220
11. CRUMP, K. S., ALLEN, B. C. (1986): Quantitative Assessment of Carcinogenic Hazards Using Epidemiological Data. In: R. S. McColl (Ed) Environmental Health Risks: Assessment and Management University of Waterloo Press, 133—158
12. ENTERLINE, P. E., MARSH, G. M., ESMEN, N. A., HENDERSON, V. L., CALLAHAN, C. M., PAIK, M. (1987): Some Effects of Cigarette Smoking, Arsenic, and SO<sub>2</sub> on Mortality Among US Copper Smelter Workers. *J Occup Med*, **29**, 831—8
13. ENVIRONMENTAL PROTECTION AGENCY, USA (1984): The Carcinogen Assessment Group's Final Risk Assessment on Arsenic. EPA-600/8-83-02IF
14. FLESCH-JANYNS, D., BERGER, J., KONIETZKO, J., MANZ, A., NAGEL, S., WAHRENDORF, J. (1992): Quantification of Exposure to Dioxins and Furans in a Cohort of Workers of a Herbicide Producing Plant in Hamburg, FRG. *Chemosphere* **25**, 2
15. GORDIS, L. (Ed.) (1988): Epidemiology and Health Risk Assessment. Oxford University Press

16. INTERNATIONAL AGENCY FOR RESEARCH ON CANCER (1987): Overall Evaluations of Carcinogenicity: An Updating of IARC Monographs Volumes 1 to 42. IARC Monogr Eval Carcinog Risk Chem Hum, Suppl. 7, 1–440
17. JEDRYCHOWSKI, W., BECHER, H., WAHRENDORF, J., BASA-CIERPIALEK, Z., GOMOLA, K. (1992): Effect of Tobacco Smoking on Various Histological Types of Lung Cancer. *J Cancer Res Clin Oncol*, **118**, 276–282
18. KALBFLEISCH, J. D., PRENTICE, R. L. (1980): *The Statistical Analysis of Failure Time Data*. Wiley, New York
19. KALDOR, J.; DAY, N. (1985): *The Use of Epidemiological Data for the Assessment of Human Cancer Risk (with comments)*. Banbury Report 19: Risk Quantitation and Regulatory Policy, 79–87
20. LAI (Länderausschuß für Immissionsschutz) (1992): *Beurteilungsmaßstäbe zur Begrenzung des Krebsrisikos durch Luftverunreinigungen*. Ministerium für Umwelt, Raumordnung und Landwirtschaft des Landes Nordrhein-Westfalen (Hrsg.)
21. LEE-FELDSTEIN, A. (1983): Arsenic and Respiratory Cancer in Humans: Follow-up of Copper Smelter Employees in Montana. *JNCI*, **70**, 601–10
22. LEE-FELDSTEIN, A. (1986): Cumulative Exposure to Arsenic and its Relationship to Respiratory Cancer Among Copper Smelter Employees. *J Occup Med*, **28**, 296–302
23. LUBIN, J. H., POTTERN, L. M., BLOT, W. J., TOKUDOME, S., STONE, B. J., FRAUMENI, J. F. (1981): Respiratory Cancer Among Copper Smelter Workers: Recent Mortality Statistics. *J Occup Med*, **23**, 779–784
24. MCCOLL, R. S. (Ed.) (1987): *Environmental Health Risks: Assessment and Management*. University of Waterloo Press
25. MOOLGAVKAR, S. H. (Ed.) (1990): *Scientific Issues in Quantitative Cancer Risk Assessment*. Birkhäuser, Boston
26. OTT, M. G., HOLDER, B. B., GORDON, M. D. (1974): Respiratory Cancer and Occupational Exposure to Arsenicals. *Arch Environ Health*, **29**, 250–255
27. RIBOLI, E., PRESTON-MARTIN, S., SARACCI, R., HALEY, N. J., TRICHOPOULOS, D., BECHER, H., BURCH, D., FONTHAN, E. T. H., GAO, Y., JINDAL, S. K. et al. (1990): Exposure of Nonsmoking Women to Environmental Tobacco Smoke: a 10-Country Collaborative Study. *Cancer Causes Control*, **1**, 243–252
28. ROBINS, J. M., GREENLAND, S. (1989): Estimability and Estimation of Excess and Etiologic Fractions. *Statistics in Medicine*, **8**, 845–859
29. ROBINS, J., GREENLAND, S. (1991): Estimability and Estimation of Expected Years of Life Lost Due to a Hazardous Exposure. *Statistics in Medicine*, **10**, 79–93
30. ROLLER, M., CSICSAKY, M., POTT, F. (1992): Methoden der Risikoabschätzung aus Daten experimenteller Kanzerogenitätsuntersuchungen. *Zentrbl Hyg Umweltmed*, **192**, 479–493
31. ROTHMAN, K. J. (1986): *Modern Epidemiology*. Little and Brown, Boston
32. SLOVIC, P. (1987): Perception of Risk. *Science*, **236**, 280–285
33. WAHRENDORF, J., BECHER, H. (1990): Quantitative Risikoabschätzung für ausgewählte Umweltkanzerogene. Umweltbundesamt Berlin (Ed). UBA-Bericht, E. Schmidt Verlag Berlin, **1/90**, 1–205
34. WELCH, K., HIGGINS, I., OH, M., BURCHFIEL, C. (1982): Arsenic Exposure, Smoking, and Respiratory Cancer in Copper Smelter Workers. *Arch Environ Health*, **37**, 325–335
35. WICHMANN, H. E. (1990): Grundsätze zur Ableitung umweltbezogener Grenzwerte aus epidemiologischen und toxikologischen Untersuchungen. In: G. Giani, R. Repges (Eds.) *Biometrie und Informatik – neue Wege zur Erkenntnisgewinnung in der Medizin*. Medizinische Informatik und Statistik, **71**, 30–38
36. WICHMANN, H. E., IHME, W. (1992): Quantitative Abschätzung von Risiken durch chemische Noxen. In: H. E. WICHMANN, H. W. SCHLIPKÖTER, G. FÜLGRAFF (Eds.) *Handbuch der Umweltmedizin*. Ecomed, Landsberg; Kap. III – 1.5.1
37. WILSON, R., CROUCH, E. A. C. (1987): Risk Assessment and Comparisons: An Introduction. *Science*, **236**, 267–270
38. WHO (1987): *Air Quality Guidelines for Europe*. WHO regional publications. European series No. 23

Anschrift der Verfasser: Heiko Becher und Karen Steindorf, Deutsches Krebsforschungszentrum, Abteilung Epidemiologie, Im Neuenheimer Feld 280, W-6900 Heidelberg

## **Zertifikat Medizinische Informatik Durchführungsrichtlinien**

Herausgegeben von der Deutschen Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie e. V. (GMDS) und der Gesellschaft für Informatik e. V. (GI)

Erarbeitet und überarbeitet von den Mitgliedern der Zertifikatskommission für das Zertifikat Medizinisch Informatik (September 1992):

Kurt Böhm, Carl-Theo Ehlers (Vorsitzender), Uwe Feldmann, Jörg Hofmann † (federführend), Peter Kandzia, Rüdiger Klar, Claus O. Köhler (federführend), Albrecht Neiß, Rudolf Thurmayr

Diese dritte Fassung löst die von J. Möhr herausgebene zweite Auflage der Durchführungsrichtlinien aus dem Jahre 1979, erschienen im Schattauer-Verlag als Heft 2 der Schriftenreihe der Deutschen Gesellschaft für Medizinische Dokumentation, Informatik und Statistik e. V., ab.

### **§ 1 Berufsbild Medizinische Informatik**

Die Medizinische Informatik umfaßt die systematische Verarbeitung von Informationen in der Medizin (Humanmedizin) und im Gesundheitswesen.

Im Mittelpunkt steht die Modellierung von informationsverarbeitenden Systemen mit dem Ziel, diese zu beschreiben, zu analysieren, zu realisieren, zu betreiben und zu bewerten. Es kommen Methoden der

- Medizinischen Informatik,
- Informatik,
- Medizinischen Biometrie,
- Epidemiologie und
- Wirtschaftswissenschaften

zur Anwendung.

Der Medizinische Informatiker (Zertifikatsinhaber) muß die Aufgaben der Medizinischen Informatik mit wissenschaftlichen Methoden bearbeiten können. Er muß Fähigkeiten zur verantwortlichen Planung, Implementierung und Leitung des operationellen Betriebes bei der Lösung von Vorhaben (gemäß Gegenstandskatalog „Medizinische Informatik und Informatik“ des Ökologie-Kurses des Medizinstudiums) in Einrichtungen der Gesundheitsversorgung besitzen. Dazu soll er Mitarbeiter auf dem Gebiet der Medizinischen Informatik aus- und weiterbilden können.



## § 2 Inhalt und Zweck des Zertifikats

Mit dem Zertifikat wird ein Ausbildungsstandard für das Fachgebiet Medizinische Informatik im Sinne des § 1 definiert. Es bescheinigt den Inhaber die Qualifikation, eine leitende Stellung auf dem Gebiet der Medizinischen Informatik zu übernehmen.

Voraussetzungen zum Erwerb des Zertifikats sind (s. § 6):

- ein adäquates Eingangsstudium,
- eine operationelle Qualifikation und
- ggfs. eine zum Inhalt des Eingangsstudium komplementäre Fortbildung.

Die nachstehenden §§ 3 bis 9 legen die Bedingungen für den Erwerb des Zertifikats im einzelnen fest. Nähere Erläuterungen finden sich im Anhang.

## § 3 Trägergesellschaften

Trägergesellschaften sind die Deutsche Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie e. V. (GMDS) und die Gesellschaft für Informatik e. V. (GI).

Die GMDS verwaltet und archiviert alle Unterlagen und führt die Stammrolle der Zertifikatsinhaber.

## § 4 Zertifikatskommission

Für die Durchführung des Verfahrens ist die von den beiden Trägergesellschaften gebildete gemeinsame Zertifikatskommission (im Sinne des § 11 Abs. 3 der GMDS-Satzung) zuständig. Sie wird bei der Deutschen Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie e. V. angesiedelt.

Die Zertifikatskommission besteht aus je zwei Vertretern der nachstehend genannten Fachrichtungen:

- Medizinische Informatik,
- Informatik,
- Medizin,
- Medizinische Biometrie und Wirtschaftswissenschaften.

Für die Bestellung sind die nach der Satzung der GMDS bzw. der GI hierfür vorgesehenen Gremien zuständig. Die Mitglieder und Amtsinhaber werden auf 4 Jahre bestellt. Wiederholte Bestellungen sind zulässig.

Aus dem Kreis der Mitglieder werden der Vorsitzende und der stellvertretende Vorsitzende gewählt und von den satzungsmäßig zuständigen Gremien der Trägergesellschaften bestätigt.

## § 5 Der Vorsitzende

Der Vorsitzende führt die Geschäfte nach Maßgabe der Beschlüsse der Zertifikatskommission und dieser Durchführungsrichtlinien. Zu seinen Aufgaben gehören insbesondere:

- a. Schriftverkehr,
- b. Vorbereitung und Leitung der Kommissionssitzungen,
- c. Durchführung der Kommissionsbeschlüsse.

Der stellvertretende Vorsitzende unterstützt den Vorsitzenden bei seinen Aufgaben und vertritt ihn insbesondere im Falle der Verhinderung.

## § 6 Antrag

Der Antrag ist formlos zu stellen. Die beizufügenden Unterlagen umfassen (s. auch Anhang):

- Lebenslauf und ein aktuelles Lichtbild
- Darstellung des beruflichen Werdegangs des Bewerbers.
- Unterlagen zum Nachweis des Abschlusses eines adäquaten Eingangsstudiums,
- Verzeichnis der Veröffentlichungen, Projekte usw. nach den Fachgebieten (siehe § 1) geordnet,
- ggfs. Nachweis der komplementären Fortbildung,
- Stellungnahme eines Zertifikatsinhabers oder eines Professors für Medizinische Informatik zum Antrag,
- Beleg-Exemplare von einigen repräsentativen Veröffentlichungen, Berichten o. ä.,
- Quittung über die Einzahlung der Bearbeitungsgebühr.

Alle Unterlagen mit Ausnahme des Lichtbildes und der Belegexemplare sind 10-fach einzureichen.

Mit der Einreichung des Antrags unterwirft sich der Antragsteller diesen Durchführungsrichtlinien.

Die Unterlagen werden vertraulich behandelt.

Für Mitglieder von GMDS oder GI gilt eine reduzierte Bearbeitungsgebühr (siehe Anhang).

## § 7 Eröffnung des Verfahrens

Der Vorsitzende prüft die Unterlagen auf Erfüllung der Voraussetzungen zur Eröffnung des Verfahrens.

Voraussetzung für die Eröffnung des Verfahrens sind

- a. ein abgeschlossenes wissenschaftliches Hochschulstudium (in der Regel Informatik, Medizin, Medizinische Informatik),
- b. eine mindestens fünfjährige, einschlägige und voll berufliche, praktische Tätigkeit auf dem Gebiet der Medizinischen Informatik im Sinne der operationellen Qualifikation nach § 2 dieser Zertifikatsordnung nach Abschluß des Grundstudiums,
- c. eine komplementäre Fortbildung in den Fachrichtungen (s. § 8).

Bei Absolventen eines Studiengangs Medizinische Informatik (z. B. Universität Heidelberg/Fachhochschule Heilbronn, Universität Hildesheim) entfällt die Forderung nach komplementärer Fortbildung.

## § 8 Beurteilungsrichtlinien

### Inhaltliche Definition der operationellen Qualifikation

Die operationelle Qualifikation wird durch die Befähigung zum Lösen von Problemen im Gesundheitswesen mit Methoden der in § 1 genannten Fachgebiete erreicht. Sie kann durch Nachweis erfolgreich abgeschlossener Projekte (Dokumentation, betriebsinterne Berichte) oder wissenschaftliche Publikationen aus den nachstehenden beispielhaft genannten Tätigkeitsgebieten belegt werden:

- Informationshaltung und -auswertung,
- Rechnergestützte Organisation von Gesundheitsversorgungssystemen,
- Medizinisch-technische Informatik.

Als Eingangsvoraussetzung im Sinne von § 6 gilt die Tätigkeit in wenigstens einem der genannten Arbeitsgebiete. Die Erteilung des Zertifikats setzt Kenntnisse der Grundlagen in den anderen Gebieten voraus.

Nähere Angaben zu den Tätigkeitsgebieten siehe Anhang.

#### **Inhaltliche Definition der Fortbildung**

Eine Fortbildung soll zur absolvierten Hochschulausbildung komplementäres Grundlagenwissen enthalten. Es wird eine Fortbildung gefordert, die neben der Medizinischen Biometrie und den Wirtschaftswissenschaften vor allem die

- Medizinische Informatik,
- Medizin und
- Informatik

enthält. Die Fortbildung muß gewährleisten, daß Grundkenntnisse in allen fünf Fachrichtungen gemäß § 1 vorhanden sind. Näheres siehe Anhang.

#### **Beurteilungsverfahren**

Die Zertifikationskommission prüft, ob die eingereichten Unterlagen für die geforderte Qualifikation in den Fachgebieten ausreichen. Für Wissensstoff, der nicht im absolvierten Hochschulstudium enthalten ist, erfolgt die Beurteilung auf Grund eingereicherter Nachweise über komplementäre Fortbildungsmaßnahmen, durch Beurteilung der eingereichten Publikationen oder die Nachweise über durchgeführte Projekte. Die Breite der Erfahrung und des Wissens gilt als nachgewiesen, wenn beispielsweise in Publikationen mehrere unabhängige Aspekte des Fachgebiets Medizinische Informatik behandelt werden oder der Bewerber an der Lehre auf dem Fachgebiet beteiligt ist.

#### **§ 9 Anerkennung**

Die Entscheidung über die Anerkennung erfolgt in einer Sitzung der Kommission.

Eine Ablehnung muß begründet sein und kann ggfs. Empfehlungen zur Vervollständigung der Voraussetzungen enthalten.

Die Ablehnung ist vom Vorsitzenden der Zertifikatskommission zu unterzeichnen und dem Antragsteller zuzustellen.

Gegen die Ablehnung der Zertifikatskommission kann der Antragsteller innerhalb von drei Monaten einen begründeten Einspruch an die Präsidenten von GMDS oder GI richten.

Die Präsidenten von GMDS und GI erteilen das Zertifikat „Medizinisch Informatik“.

#### **§ 10 Aberkennung**

Das Zertifikat kann aberkannt werden, wenn die Erteilung aufgrund falscher Voraussetzungen erfolgte.

#### **§ 11 Inkrafttreten**

Diese Durchführungsrichtlinien treten mit Wirkung vom 1. 1. 1993 in Kraft.

Damit verlieren alle früheren Richtlinien ihre Gültigkeit.

## Anhang

### 1. Allgemeines

Maßgebend für den Erwerb des Zertifikats „Medizinisch Informatik“ sind die vorstehend abgedruckten Richtlinien zum Erwerb des Zertifikats sowie die Beschlüsse der Zertifikatskommission.

Die hier im Anhang vorgelegten praktischen Hinweise sollen dem Bewerber bei der Zusammenstellung des Antrages helfen. Sie ergänzen die in den Richtlinien festgelegten Anforderungen inhaltlich und sollen den Bewerber in die Lage versetzen, selbständig zu prüfen, in welchem Umfang die Voraussetzungen für den Erwerb des Zertifikats erfüllt sind.

Die Präsidien der Gesellschaft für Informatik (GI) und der Deutschen Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie (GMDS) waren von Beginn an der Meinung, daß an die Vergabe des Zertifikats Medizinische Informatik hohe Anforderungen zu stellen seien. Es bescheinigt dem Inhaber die Qualifikation, eine leitende Stellung im Bereich der Medizinischen Informatik zu übernehmen.

Wünschenswert ist, daß Zertifikatsinhaber und Vorgesetzte ihre Mitarbeiter auf die Möglichkeit des Zertifikaterwerbs hinweisen. Je früher der Entschluß das Zertifikat angestrebt wird, desto zielgerichteter und schneller können die Anforderungen erfüllt werden. Dies kann durch Fortbildungskurse, Selbststudium und Mitarbeit in geeigneten Projekten geschehen.

Bei der Formulierung des Antrags ist nach den Richtlinien (§§ 6, 7, 8) vorzugehen.

Die Zusammenstellung der geforderten Nachweise soll vollständig, systematisch, knapp, übersichtlich und nach den fünf Fachgebieten

- Medizinische Informatik,
- Medizin,
- Informatik,
- Medizinische Biometrie und
- Wirtschaftswissenschaften

gegliedert sein.

Kurse und Seminare sind nach Art und Umfang sowie durch eine Bescheinigung der erfolgreichen Teilnahme zu belegen. Die Teilnahme an entsprechenden Fortbildungsveranstaltungen der Fachgesellschaften und ihrer Fortbildungsakademien wird empfohlen.

### 2. Voraussetzungen einer Bewerbung

#### 2.1. Operationelle Tätigkeit

1. Die Richtlinien sehen eine **mindestens** 5-jährige operationelle Tätigkeit nach Abschluß des Studiums vor. Diese Forderung ist nicht automatisch durch eine 5-jährige Berufserfahrung erfüllt.

Hat der Bewerber ein Zweitstudium in dieser Zeit betrieben, war er nur halbtags tätig oder liegen Unterbrechungen aus anderen Gründen vor, so wird jeweils nur ein Teil dieser Zeit als operationelle Tätigkeit anerkannt.

Es wird weiterhin von der Kommission sehr sorgfältig auf eine einschlägige und hinreichend breite Tätigkeit während dieser Zeit geachtet.

Wer den überwiegenden Teil seiner operationellen Tätigkeit nur in einem Spezialgebiet wie etwa EKG-Auswertung oder Labordatenverarbeitung ausgeübt hat, wird mit

dieser Praxis keine Aussicht auf Anerkennung im Sinne dieses Zertifikats haben. Vor einem Antrag sollte der Interessent daher z. B. durch Teilnahme an anderen Projekten seine Erfahrungen verbreitern und/oder in einem anderen Institut hospitieren.

2. Der Bewerber muß, über theoretische Beiträge hinaus, im Gesundheitswesen praktisch tätig gewesen sein und sich darin bewährt haben. Dieser Nachweis bedingt im allgemeinen eine aktive Beteiligung an wenigstens einer klinischen Studie.
3. Da mit der Vergabe des Zertifikats auch zum Ausdruck gebracht werden soll, daß der Bewerber für eine leitende Tätigkeit qualifiziert ist, wird darauf geachtet, daß diesbezügliche Erfahrungen auch vorliegen. Der Bewerber sollte etwa hinreichend lange und erfolgreich z. B. als Projektleiter gearbeitet und dabei Erfahrungen gesammelt und die Probleme von Koordination, Finanzierung und Menschenführung kennengelernt und gelöst haben.
4. Nähere Angaben zur operationellen Qualifikation, Informationshaltung und -auswertung:
  - Medizinische Dokumentation,
  - Betrieb von Informationssystemen im Gesundheitswesen,
  - Planung und Auswertung medizinischer Untersuchungsreihen,
  - Datenschutz.

Die medizinische Dokumentation umfaßt dabei die inhaltliche Definition von Dokumentationsvorhaben unter Berücksichtigung medizinischer Zielsetzungen sowie Auswahl und Anwendung von Dokumentationssystemen und Kodierungsverfahren.

Der Betrieb von Informationssystemen im Gesundheitswesen schließt die Lösung der Probleme bei der algorithmischen Realisierung, insbesondere die Erfüllung von Benutzeranforderungen durch Entwurf und Realisierung von Erfassungs-, Speicherungs-, Auswertungs- und Darstellungsreihen, wegen der engen methodischen Beziehung zu den Informationssystemen mit ein. Die operationelle Qualifikation sollte möglichst die intensive Mitarbeit an mindestens einer klinischen Prüfung von der Versuchsplanung über Datenerfassung und -speicherung bis zur Auswertung enthalten. Eine auf die Anwendung biometrischer Verfahren beschränkte Tätigkeit ist jedoch nicht als operationelle Qualifikation mit Tätigkeitsschwerpunkt „Informationshaltung und -auswertung“ anzusehen.

Rechnergestützte Organisation von Gesundheitsversorgungssystemen:

Dieses Tätigkeitsgebiet wird im Sinne sowohl der betriebswirtschaftlichen Auffassung von Organisation als auch der Systemorganisation verstanden und umfaßt Tätigkeiten in

- Gesundheitsversorgungssystemen oder
- Rechenzentren mit entsprechenden Aufgaben.

Die Organisation umfaßt die Analyse, Planung und Optimierung der Betriebsabläufe durch Rechnereinsatz in Gesundheitsversorgungssystemen, wie z. B. Praxis, Klinik, betriebsärztlicher Dienst oder öffentlichem Gesundheitswesen und betrifft die Personal-, Finanz- und Sachmittelverwaltung. In dieser Hinsicht ist die Kenntnis der beteiligten Instanzen sowie deren Kommunikationskanäle, Zielsetzungen und Erfordernisse notwendig.

Auf Rechenzentren angewandt betrifft sie gegebenenfalls deren Finanzplanung, Auswahl und Beschaffung der Betriebsmittel (Personal, Organisations- und Dokumentationsmittel, Netzentwicklung, Verwaltung, Hardware und Software) sowie die Optimierung der Rechenzentrumsleistung einschließlich Datensicherung und -schutz.

Medizinisch-technische Informatik:

Hierzu zählen Tätigkeiten auf dem Gebiet der Bildverarbeitung, der Prozeßsteuerung und Biosignalverarbeitung, die Bearbeitung technischer Aspekte der Benutzerschnittstelle bzw. der Rechnerkopplung sowie der Netzwerkprobleme, wobei diese Tätigkeiten

auch die Planung und die Integration der Problemlösung in ein medizinisches System beinhalten müssen.

## 2.2. Fortbildung

Die für den Erwerb des Zertifikats relevanten Fächer einer Fortbildung sind:

- Medizinische Informatik,
- Informatik,
- Medizin,
- Medizinische Biometrie,
- Wirtschaftswissenschaften.

Möglichkeiten der Fortbildung sind z. B.:

- Besuch von Vorlesungen des komplementären Studienfachs während des Studiums (Nebenfach),
- Durchführung eines Aufbaustudiums,
- Besuch von Kursen und Seminaren, die zunehmend von Fachgesellschaften (z. B. GMDS, GI, BVMI, DVMD, Biometrische Gesellschaft) angeboten werden.
- Selbststudium im Kontakt mit einem Fachvertreter (im Bedarfsfall kann der Kontakt über die Zertifikatskommission vermittelt werden).

## 2.3. Sonderfälle

### 2.3.1. Diplom-Informatiker der Medizin

Im Studiengang Medizinische Informatik (z. B. Universität Heidelberg/Fachhochschule Heilbronn und Universität Hildesheim) sind die oben angeführten Fächer integrale Bestandteile des Studiums. Deshalb betrachtet die Kommission die gestellten Anforderungen in Hinsicht auf die Ausbildung als erfüllt. Die Kommission wird bei Bewerbern mit diesem Studienabschluß die geforderte operationelle Qualifikation besonders sorgfältig prüfen.

### 2.3.2. Dipl.-Informatiker mit Nebenfach Medizin

Wurde während des Informatik-Studiums das Nebenfach Medizin oder Medizinische Informatik gewählt, so sind in der Regel neben der operationellen Qualifikation nur noch ausreichende Kenntnisse in den Fächern Biometrie und Wirtschaftswissenschaften nachzuweisen. Grundsätzlich wird bei allen Anträgen, gleichgültig ob ein Regelfall vorliegt oder nicht, eine Einzelfallprüfung durchgeführt.

## 3. Gebühren

Der Nachweis über die Einzahlung der Bearbeitungsgebühr (z. Z. DM 150,– für Mitglieder der GI/GMDS bzw. DM 200,– für Nichtmitglieder auf das Konto-Nr. 0001601822 der GMDS bei der Deutschen Apotheker- und Ärztebank in Köln. Bankleitzahl 370606 15) ist in Kopie (einfach) dem Antrag beizufügen.

## 4. Unterlagen

Alle Unterlagen (Zeugnisse, Bestätigungen, Lebenslauf, beruflicher Werdegang, Literaturverzeichnis, Vorlesungsverzeichnis, Projektbeschreibungen etc.) sind dem formlosen Antrag in 10-facher Ausfertigung (Lichtbild und Publikationen nur einmal) beizufügen.

## 5. Literaturhinweise

Mit diesen Angaben sollen weniger bestimmte Bücher empfohlen als vielmehr Hinweise auf Inhalt und Umfang der Literatur zum Selbststudium gegeben werden. Zu jedem Fachgebiet sind relevante Zeitschriften angegeben.

### 5.1. Allgemeines

1. Berufsverband Medizinischer Informatiker (Hrsg.): Der Medizinische Informatiker. Heidelberg, Eigenverlag 1985
2. BRAUER, W., HAACKE, W., MÜNCH, S. (Hrsg.): Studien und Forschungsführer Informatik. 2. Aufl., Berlin – Heidelberg – New York – London – Paris – Tokyo, Springer 1989
3. Bundesanstalt für Arbeit (Hrsg.): Medizinischer Informatiker/Medizinische Informatikerin. Blätter zur Berufskunde, 3 – IA 04, Bielefeld, Bertelsmann 1991 (2. Auflage)
4. DAY, R. A.: How to write and publish a scientific paper. 2nd edit., Isi Press, Philadelphia 1979
5. HAUX, R., DUDECK, J., GAUS, W., LEVEN, F.-J., KUNATH, H., MICHAELIS, J., PRETSCHNER, D. P., SONNTAG, H. G., THURMAYR, R., WOLTERS, E.: GMDS-Empfehlungen zur Ausbildung in Medizinischer Informatik. Biometrie und Informatik in Medizin und Biologie 22 (1991) 180 – 197
6. HAWKINS, C., SORGI, M.: Research: How to plan, speak, and write about. Springer, Berlin – Heidelberg – New York – Tokyo 1985
7. TRAMPISCH, H. J. et al. (Hrsg.): Studien- und Forschungsführer Medizinische Informatik, Biometrie und Epidemiologie. GMDS, Köln 1992

Zeitschrift:

Informatik, Biometrie und Epidemiologie in Medizin und Biologie (Fischer, Stuttgart)

### 5.2. Medizinische Informatik

Systematische Verarbeitung von Informationen in der Medizin; Nutzen der Medizinischen Informatik für die Medizin; Modelle des Informationsflusses; Definition der Aufgaben von rechnergestützten Informationssystemen; Analyse und Gestaltung des sozialen Umfeldes; Methoden der Planung, Entwicklung und Auswahl von medizinischen Anwendungssystemen; Signal- und Bildverarbeitung in der Medizin; Ziele, Arten, Aufbau, Retrieval, Medizinische Dokumentation und Schlüsselsysteme in der Medizin, ihre Ziele und Strukturen; Arten und Funktionen medizinischer Register; Methoden der Qualitätssicherung ärztlichen Handelns; Anwendungssysteme in Krankenhäusern, Arztpraxen, betriebsärztlichen Dienst etc.; Daten- und Wissensbanken in der Medizin; Modelle der medizinischen Entscheidungsfindung; medizin-technische EDV-Systeme z. B. für EKG, EEG etc.; Patienten-Überwachungssysteme; Computersimulation von Krankheitsfällen und Pathomechanismen; computergestütztes Lernen zur Diagnostik und Therapie bestimmter Leitsymptome; Grundkonzepte zum Datenschutz im Gesundheitswesen; Maßnahmen zur Gewährleistung des Datenschutzes, der ärztlichen Schweigepflicht und des Persönlichkeitsschutzes der Betroffenen.

1. BAKKER, A. R., EHLERS, C. TH., BRYANT, W. E. (eds.): Hospital Information Systems: Scope-Design-Architecture. North-Holland. Amsterdam 1992
2. BALL, M. J., HANNAH, K. J. et al. (Eds.): Nursing Informatics. Springer, Berlin – Heidelberg – New York – London – Paris – Tokyo 1988

3. BALL, M., DOUGLAS, J. V., O'DESKY, R. I., ALBRIGHT, J. W. (eds.): Health Information Systems — A Practical Guide. Springer, New York et al. 1991
4. BOSCH, M., KÖHLE, M., KÖHLER, C. O., SCHAEFER, O. P., SCHLAEFER, K., STAUFER, D., VOGELSANG, A., WEBER, W.: Computer in Arztpraxis. (Herausgegeben von Köhler, C. O., Schaefer, O. P.), 3. überarb. Aufl., Bd. 4 Schriftenreihe zur Informationsverarbeitung im Gesundheitswesen (Hrsg.: Köhler, C. O., Maurer, C.), excomed, Landsberg 1991
5. KÖHLER, C. O. (Hrsg.): Dokumentation und Information — Handbuch für Klinik und Praxis. Lose-Blatt-Sammlung, ecomed. Landsberg ab 1980
6. PUPPE, F.: Diagnostische Problemlösungen in Expertensystemen. Springer, Berlin — Heidelberg et al. 1987
7. SHORTLIFFE, E., PERREAULT, L. E. (eds.): Medical Informatics — Computer Application in Health Care. Addison-Wesley, Bonn-Reading (Mass.) 1990
8. WINGERT, F.: Medical Informatics. Springer, New York et al. 1981

Zeitschriften:

Methods of Information in Medicine (Schattauer, Stuttgart—New York)

Medical Informatics (Taylor & Francis, London)

### 5.3. Informatik

Erforderlich sind grundlegende Kenntnisse über die Methoden der Informatik, das Werkzeug der Informatikers — den Computer — sowie die Fähigkeit zur informatikspezifischen Arbeitsweise bei der Formalisierung, Strukturierung, Modellierung und Lösung von Anwendungsproblemen. Grundlegende Kenntnisse der Informationsverarbeitung, Automaten- und Sprachmodelle, Algorithmen, Datenbank- und Informationssysteme, Software-Engineering, Rechnerarchitektur, Simulation.

1. BAUER, F. L., GOOS, G.: Informatik — Eine einführende Übersicht. Springer, Berlin — Heidelberg — New York — London — Paris — Tokyo (2. Bd.) 1984, 1991
2. BODE, A., HÄNDLER, W.: Rechnerarchitektur. Band I und II, Springer, Berlin — Heidelberg — New York — London — Paris — Tokyo 1983
3. GOLDSCHLAGER, L., LISTER, A.: Informatik — Eine moderne Einführung. Hanser, München — Wien 1990
4. HOPCROFT, J. E., ULLMANN, J. D.: Einführung in die Automatentheorie, Formale Sprachen und Komplexitätstheorie. Addison-Wesley, Bonn-Reading (Mass.) et al. 1990
5. REMBOLD, U. (Hrsg.): Einführung in die Informatik. Hanser, München — Wien 1987
6. SCHLAGETER, G., STUCKY, W.: Datenbanksysteme — Konzepte und Modelle. Teubner, Stuttgart 1983
7. SCHMIDT, B.: Systemanalyse und Modellaufbau, Fachbereiche Simulation 1. Springer, Berlin — Heidelberg — New York — London — Paris — Tokyo 1985
8. TANENBAUM, A. S.: Computer-Netzwerke. Wolfram, Attenkirchen 1990
9. UNGERER, T.: Innovative Rechnerarchitekturen — Bestandsaufnahme, Trends, Möglichkeiten. McGraw-Hill, Hamburg 1989
10. VETTER, M.: Aufbau betrieblicher Informationssysteme. Teubner, Stuttgart 1987
11. VETTER, M.: Strategie der Anwendungssoftware-Entwicklung. Teubner, Stuttgart 1988

Zeitschriften:

Wirtschaftsinformatik (Vieweg, Braunschweig — Wiesbaden)

Informatik Forschung und Entwicklung (Springer, Berlin — Heidelberg et al.)



## 5.4. Medizin

Medizinische Terminologie, Struktur und Funktionen des menschlichen Organismus (Anatomie und Physiologie), allgemeine Pathologie, diagnostische und therapeutische Maßnahmen einschließlich Labor- und Röntgendiagnostik bzw. Strahlentherapie, Grundlagen der Inneren Medizin und Chirurgie.

1. ARNOLD, M., BRAUER, H.-P., DENEKE, J. F. V., FIEDLER, E.: Der Beruf des Arztes in der Bundesrepublik Deutschland. Deutscher Ärzte-Verlag, Köln 1984
2. BLEYL, U., DÖHNERT, G., HÖPKER, W. W., HOFMANN, W.: Allgemeine Pathologie. Heidelberger Taschenbücher Nr. 163, Springer, Berlin – Heidelberg – New York 1976
3. BÜHLMANN, A. A., FROESCH, E. R.: Pathophysiologie. 5. Aufl., Springer, Berlin – Heidelberg – New York 1989
4. FALLER, A.: Der Körper des Menschen. Einführung in Bau und Funktion. 11. Aufl., Thieme, Stuttgart 1988
5. GANONG, W. F.: Lehrbuch der Medizinischen Physiologie. (Übersetzung von W. Auerswald). 4. Aufl., Springer, Berlin – Heidelberg – New York 1974
6. HOFFMANN-LA-ROCHE, URBAN & SCHWARZENBERG (Hrsg.): Roche Lexikon Medizin. 2. Aufl., Urban & Schwarzenberg, München – Wien-Baltimore 1987
7. KAUFMANN, W., LÖHR, G. W.: Pathophysiologie. 4. Aufl., Thieme, Stuttgart 1992
8. MÖRICKE, K. D., BETZ, E., MERGENTHALER, W.: Biologie des Menschen. 12. Aufl., Quelle & Meyer, Heidelberg 1989
9. PSCHYREMBEL, W.: Klinisches Wörterbuch. de Gruyter, Berlin – New York 1990
10. SCHMIDT, R. F. (Hrsg.): Grundriß der Sinnesphysiologie. Heidelberger Taschenbücher. Springer, Berlin – Heidelberg – New York 1977

Zeitschriften:

Deutsche Medizinische Wochenschrift (Thieme, Stuttgart)  
Strahlentherapie (Urban & Vogel, München)

## 5.5. Medizinische Biometrie

Modellbildung in der Medizin, Versuchsplanung, statistische Verfahren, statistische Auswertungssysteme.

1. ADAM, J.: Einführung in die medizinische Biometrie. Fischer UTB, Jena – Stuttgart 1992
2. BRESLOW, N. E., DAY, N. E.: Statistical Methods in Cancer Research. Vol. I, The Analysis of Case-Control Studies, (IARC Scientific Publications No. 32), IARC, Lyon 1987
3. BRESLOW, N. E., DAY, N. E.: Statistical Methods in Cancer Research. Vol. II, Design and Analysis of Cohort Studies, (IARC Scientific Publications No. 82), IARC, Lyon 1987
4. FRIEDMAN, L. M., FURBERG, C. D., DEMETS, D. L.: Fundamentals of Clinical Trials. Wright, Boston – Bristol – London 1980
5. HARTUNG, J.: Lehr- und Handbuch der angewandten Statistik. Oldenbourg, München – Wien 1985
6. HEINECKE, A., HULTSCH, E., REPGES, R.: Medizinische Biometrie, Biomathematik und Statistik, Springer, Berlin – Heidelberg – New York – Tokyo 1992

7. KINDER, H.-P., OSIUS, G., TIMM, J.: Statistik für Biologen und Mediziner. Vieweg, Braunschweig 1982
8. POCOCK, S. J.: Clinical Trials — A Practical Approach. Wiley, Chischester — New York — Brisbane — Toronto — Singapore 1983
9. RASCH, D.: Biometrisches Wörterbuch. Deutsch, Thun-Frankfurt 1988
10. SCHWARTZ, D., FLAMANT, R., LELLOUCH, J.: Clinical Trials. Academic Press, London — New York — Toronto — Sidney — San Francisco 1980
11. SACHS, B.: Angewandte Statistik, Anwendungen statistischer Methoden. 6. Aufl., Springer, Berlin — Heidelberg — New York — Tokyo 1984
12. VICTOR, N., SCHÄFER, H., NOWAK, H.: Arzneimittelforschung nach der Zulassung, Bestandsaufnahme und Perspektiven. Springer, Berlin — Heidelberg — New York — London — Paris — Tokyo 1991

Zeitschriften:

Controlled Clinical Trials (Elsevier, New York)

Statistics in Medicine (Wiley, Chichester)

Biometrics (Biometric Soc., Alexandria, VA)

## 5.6. Wirtschaftswissenschaften

Aufbau- und Ablauf-Organisation, Kostenrechnung, Management.

1. EICHHORN, P., LÖSCH, A. (Hrsg.): Krankenhäuser als Unternehmen. Nomos, Baden-Baden 1979
2. EICHHORN, S. (Hrsg.): Handbuch Krankenhaus-Rechnungswesen. Gabler, Wiesbaden 1983
3. EICHHORN, S.: Krankenhausbetriebslehre. Theorie und Praxis des Krankenhausbetriebes. (2 Bände, 3. Aufl.). Kohlhammer, Stuttgart — Berlin — Köln — Mainz 1975, 1976
4. EIFF, V. W. (Hrsg.): Handbuch Krankenhaus-Management. ecomed, Landsberg 1979
5. LAUX, E. (Hrsg.): Wirtschaftliches Krankenhaus. Management, Planung, Rechnungswesen. Prüfung. WiBERA-Fachschr. NF Bd. 9. Kohlhammer, Stuttgart — Berlin — Köln — Mainz 1980
6. MEYER, M. (Hrsg.): Krankenhausplanung — Die Lösung medizinisch-ökonomischer Probleme der Praxis mit Methoden der Systemforschung. Fischer, Stuttgart — New York 1979
7. MÜLLER, H. W. (Hrsg.): Führungsaufgaben im modernen Krankenhaus. (2. Aufl.). Kohlhammer, Stuttgart — Berlin — Köln — Mainz 1983
8. PURZER, K.: Das Rechnungswesen der Krankenhäuser. Boorberg, München 1978
9. SCHICKE, R. K.: Ökonomie des Gesundheitswesens. Vandenhoeck und Ruprecht, Göttingen 1981
10. WÖHE, G.: Einführung in die Allgemeine Betriebswirtschaftslehre. (17. Aufl.). Vahlen, München 1990

Zeitschriften:

Operations Research (Operations Research Soc., Baltimore, MD)

Krankenhaustechnik (ecomed, Landsberg)

## BUCHBESPRECHUNGEN/BOOKREVIEWS

RASCH, D., GUIARD, V., NÜRNBERG, G.

**Statistische Versuchsplanung.**

Einführung in die Methoden und Anwendung des Dialogsystems CADEMO.

Stuttgart: Gustav Fischer 1992.

386 S., 159 Abb., 26 Tab., kt. DM 89,-, ISBN 3-437-40247-1.

Zu dem nunmehr seit einigen Jahren kommerziell erhältlichen System CADEMO für die statistische Versuchsplanung gibt es jetzt ein einführendes Lehrbuch. Das Buch ist in zwei Hauptteile gegliedert. Um statistischen Laien die Lektüre zu erleichtern, wurde der Band um eine kurze Einführung in die Wahrscheinlichkeitsrechnung erweitert. Außerdem liegt eine Demonstrationsdiskette bei, die eine Vorstellung von der Anwendung des Systems vermittelt.

Der erste Hauptteil im Umfang von knapp 180 Seiten erläutert anhand von zahlreichen Beispielen die Grundlagen der statistischen Versuchsplanung. Die einzelnen Kapitel befassen sich mit den Grundprinzipien der Versuchsplanung: Wiederholen, Blockbildung, Randomisierung, erläutern häufig verwendete Versuchsanlagen und die Ideen der optimalen Versuchsplanung. Schwerpunkt ist die Präsentation ausgewählter Methoden zur Bestimmung des Stichprobenumfanges. Das abschließende Kapitel des ersten Hauptteils verweist auf Probleme der Modellwahl und auf die Robustheit statistischer Methoden.

Dieser erste Teil hat den Charakter eines Lehrbuchs, wenn auch zugeschnitten für Aufgaben, die man mit CADEMO bearbeiten kann. Durch die systematische Zusammenstellung der in diesem System verwendeten Methoden wird eine Lücke in der bisher vorliegenden Programmdokumentation geschlossen. Manche Vorschläge des Systems, die bisher nur mittels langwierigem Quellstudium nachvollziehbar waren, sind jetzt transparent. Wohl unter Rücksichtnahme auf den Lehrbuchcharakter wurde allerdings auf eine vollständige Offenlegung der verwendeten Heuristiken und Strategien verzichtet.

Der zweite Hauptteil von etwa 150 Seiten beschreibt kurz den Aufbau und ausführlich die Anwendung des CADEMO-Systems. Die Gliederung entspricht jener der Architekturkomponenten des Programms in Modulkomplexe und Moduln. Dabei wird in allgemein anwendbare Modulkomplexe (z. B. „Beurteilung von Varianzen“) und in Modulkomplexe für spezielle Anwendungsgebiete (z. B. „Feldversuche“) unterschieden. Die Autoren haben sich für eine anschauliche Darstellung anhand von Bildschirm- und Konsultationsbeispielen entschieden. Im Unterschied zum ersten Teil ist dieser wie ein Handbuch geschrieben und vor allem für Personen gedacht, die entweder CADEMO oder neuere Moduln des Systems nicht kennen.

Die im Titel verwendete Bezeichnung Dialogsystem mag etwas ungewöhnlich erscheinen. Es war jedoch gewiß eine kluge Entscheidung der Autoren, sich der wenig fruchtbaren Diskussion zu entziehen, ob CADEMO nun ein „Expertensystem“ sei oder nicht, indem sie diesen Begriff vermeiden.

Das Buch ist von besonderem Nutzen für Anwender von CADEMO, die näheres über die dem System zugrundeliegenden Methoden wissen möchten. Es bietet gleichzeitig potentiellen Interessenten die Möglichkeit, sich ausführlich über CADEMO zu informieren. Der erste Teil des Buches kann überdies Wissenschaftlern verschiedenster Fachrichtungen mit eher geringen Statistikkenntnissen als Einführung in die Grundbegriffe der Versuchsplanung empfohlen werden.

U. Haag, Heidelberg

## TABLE OF CONTENTS

Regression models for an angular response	N. I. Fisher and A. J. Lee	665
Linear regression analysis based on Buckley–James estimating equation	J. S. Lin and L. J. Wei	679
Blocked arteries and multivariate regression	D. F. Percy	683
Testing the effect of treatment in experiments with correlated binary outcomes	D. Zucker and J. Wittes	695
Testing goodness of fit of a multinomial model against overdispersed alternatives	B. S. Kim and B. H. Margolin	711
Multivariate methods for clustered binary data with multiple subclasses, with application to binary longitudinal data	B. Rosner	721
A parametric family of correlation structures for the analysis of longitudinal data	A. Muñoz, V. Carey, J. P. Schouten, M. Segal, and B. Rosner	733
Modelling balanced longitudinal data: Maximum likelihood estimation and analysis of variance	A. Forcina	743
Exchangeability in multivariate Markov chain models	A. Gottschau	751
Sequential monitoring for comparison of changes in a response variable in clinical studies	M. C. Wu and K. K. G. Lan	765
Use of external rates in nested case–control studies with application to the International Radiation Study of Cervical Cancer Patients	D. C. Thomas, M. Blettner, and N. E. Day	781
Semiparametric estimation of random effects using the Cox model based on the EM algorithm	J. P. Klein	795
A statistical distribution with an unbounded hazard function and its application to a theory from demography	D. Zelterman	807
Robust procedures for testing equality of covariance matrices	P. C. O'Brien	819
Proportional hazards model with covariates subject to measurement error	T. Nakamura	829
Optimal administration of dual screening tests for detecting a characteristic with special reference to low prevalence diseases	S. Geisser and W. Johnson	839
Estimating the probability of toxicity at the recommended dose following a Phase I clinical trial in cancer	J. O'Quigley	853
Estimating the total number of distinct species using presence and absence data	S. A. Mingoti and G. Meeden	863
Sample size calculations for normal variates under binomial censoring	J. R. Skalski	877
Shorter Communications		
Testing homoscedasticity in a two-way table	G. S. Mudholkar and I. C. Sarkar	883
An algorithm for the mean and variance of the noncentral hypergeometric distribution	J. Liao	889
The efficiency of a weighted log-rank test under a percent error misspecification model for the log hazard ratio	D. M. Zucker	893
A robust line transect method	S. T. Buckland and B. J. Turnock	901
The Consultant's Forum		
A GLM approach to quantal response models for mixtures	C. Cox	911
Starting $D$ -optimal designs for batch kinetic studies of enzyme-catalyzed reactions in the presence of enzyme deactivation	F. X. Malcata	929
Marginal models for correlated binary responses with multiple classes and multiple levels of nesting	B. F. Qaqish and K.-Y. Liang	939
Maximum likelihood estimation for interval-censored data using a Weibull-based accelerated failure time model	P. M. Odell, K. M. Anderson, and R. B. D'Agostino	951
Reader Reaction: On estimating standardized risk differences from odds ratios	K. F. Yu	961
Response	S. Greenland and P. W. Holland	963
Reader Reaction: A note on modelling the covariance structure of repeated measurements	P. S. Gill	965
Correspondence		969
Book Reviews		973
Correction		979
Papers to be published		981
The Biometric Society		983