EDV in Medizin und Biologie

EDP in Medicine and Biology

Zeitschrift für

Computeranwendungen
bei
Versuchsplanung,
Datensammlung,
biometrischer Auswertung

in
klinischer Medizin,
experimenteller Medizin,
Veterinärmedizin,
Agrarwissenschaften und
Biologie

Herausgegeben von:
H. Fink, Wuppertal
G. Fischbeck, München
H. Geidel, Stuttgart
G. Grabner, Wien
H. Rundfeldt, Hannover

unter Mitarbeit von:
H.-J. Lange, München
N. Victor, Gießen
G. Vinek, Linz
E. Weber, Heidelberg
F. X. Wohlzogen, Wien

Band 17
Heft 1/2/1986

ISSN 0300-8282

Gustav Fischer Verlag Stuttgart · Verlag Eugen Ulmer Stuttgart

EDV in Medizin und Biologie
EDP in Medicine and Biology

Versuchsplanung, Datensammlung und biometrische Auswertung in Biologie, klinischer und experimenteller Medizin und Agrarwissenschaften

Schriftleiter:
Prof. Dr. H. Geidel, Stuttgart

Gustav Fischer Verlag, Stuttga
Verlag Eugen Ulmer, Stuttgart

Redaktion: Herausgeber: Prof. Dr. med. H. Fink, BAYER AG, Pharma Forschungszentrum, Abt. Dokumentation und Biometrie, Friedrich-Ebert-Straße, 5600 Wuppertal 1; Prof. Dr. G. Fischbeck, Institut für Pflanzenbau und Pflanzenzüchtung der Technischen Universität München, 8050 Freising-Weihenstephan; Prof. Dr. H. Geidel, Rechenzentrum der Universität Hohenheim, Postfach 700562, 7000 Stuttgart 70; Prof. Dr. med. G. Grabner, Ordinariat für Gastroenterologie und medizinische Computerwissenschaften, Garnisongasse 13, A-1097 Wien; Prof. Dr. H. Rundfeldt, Institut für Statistik und Biometrie der Tierärztlichen Hochschule Hannover, Bischofsholer Damm 15, 3000 Hannover.
Schriftleiter (verantwortlich im Sinne des Presserechts): Prof. Dr. H. Geidel, Rechenzentrum der Universität Hohenheim, 7000 Stuttgart 70 (Hohenheim), Postfach 700562, Tel. (07 11) 45 01 / 21 30

Verlag: Vertrieb und Werbung: Gustav Fischer Verlag GmbH & Co KG, 7000 Stuttgart 70, Wollgrasweg 49, Tel. (07 11) 45 50 38. 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 Sabine Laib. Verantwortlich für den Anzeigenteil: Dieter Boger, z. Z. ist die Anzeigenpreisliste Nr. 6 gültig. Anzeigenschluß: Am 10. der Erscheinungsmonate Februar, Mai, August, November.

Druck: Johannes Illig, Buchdruck Offsetdruck, 7320 Göppingen-Ursenwang, Daimlerstr. 13.

Mit der Annahme eines Manuskriptes erwerben die Verlage das ausschließliche Verlagsrecht für die Dauer eines Jahres nach Erscheinen des Beitrags, und zwar auch für etwaige spätere Vervielfältigungen durch Nachdruck oder durch andere Verfahren wie Photokopie, Mikrokopie, Xerographie u. a., sowie den Vergütungsanspruch gegen gewerbliche Unternehmen, die einzelne Vervielfältigungen zum innerbetrieblichen Gebrauch herstellen. Die Verlage sind berechtigt, das Vervielfältigungsrecht an Dritte zu vergeben und die Lizenzgebühren, die durch Überlassung des Vervielfältigungsrechts an Dritte entstehen, geltend zu machen. Lizenzgebühren dieser Art werden hälftig zwischen den Verlagen und dem Verfasser geteilt. Soweit die Herstellung einzelner Vervielfältigungen gewerblichen Zwecken dient, richtet sich die Vergütung sowie deren Verteilung nach dem zwischen dem Börsenverein des Deutschen Buchhandels und dem Bundesverband der Deutschen Industrie abgeschlossenen Rahmenabkommen. Die Gebühr ist an die Inkassostelle für Photokopiergebühren beim Börsenverein des Deutschen Buchhande e. V., Frankfurt a. M., Postfach 3914, zu entrichten. Erfolgt die Entrichtung der Gebühr durch Wertmarken der Inkassostelle, so ist für jedes kopierte Blatt eine Marke im Betr von DM –,30 zu verwenden.

Die Zeitschrift »EDV in Medizin und Biologie« erscheint viermal jährlich. Der Bezugspreis b trägt ab Band 16 jährlich DM 248.– (Einzelheft DM 75.–) zuzüglich Versandspesen. Vorzug preis für Mitglieder der Deutschen Gesellschaft für Medizinische Dokumentation und Statist DM 194.– zuzüglich Versandspesen. Bestellungen nehmen jede Buchhandlung sowie die Ve lage entgegen.

Redaktionelle Mitteilungen: »EDV in Medizin und Biologie« bringt Originalarbeiten ur Übersichtsartikel aus dem Gesamtgebiet der Anwendung von Computern bei der Planur von Versuchen und Erhebungen, bei der Sammlung von Informationen und bei der B arbeitung und Auswertung der Daten einschließlich der Anwendung biometrischer M thoden, jedoch keine Arbeiten über kaufmännische Datenverarbeitung und Literatu dokumentation. Die Arbeiten erscheinen in deutscher oder englischer Sprache, mit ku zen Zusammenfassungen in beiden Sprachen. Manuskripte sind an die Herausgeber richten. Es werden nur unveröffentlichte Arbeiten aufgenommen. Anstelle eines Honora erhalten die Verfasser kostenlos 50 Sonderdrucke. Weitere Sonderdrucke können z sätzlich gegen Berechnung geliefert werden. Der Text soll so knapp gefaßt sein, daß d Umfang einer Arbeit 20 Schreibmaschinenseiten nicht überschreitet. Die Manuskrip blätter sind einseitig mit doppeltem Zeilenabstand zu schreiben. Manuskripte könne nur zum Satz gegeben werden, wenn sie vollständig sind, d. h. alle Abbildungen, Abb dungs-Unterschriften sowie die deutsche und englische Zusammenfassung enthalte Am Ende jedes Beitrags sind der Name des Autors (Vor- und Zuname, Titel und gegeb nenfalls Dienststellung, Institut, Firma usw.) und seine genaue Adresse anzugeben. D Zahl der Abbildungen soll auf das unbedingt Notwendige beschränkt werden. Als Vorla werden reproduktionsfähige Zeichnungen und bei Fotografien scharfe Hochglanza züge erbeten. Farbige Abbildungen können nur bei Kostenersatz durch den Autor aufg nommen werden. Tabellen, Gleichungen und Fußnoten sind fortlaufend zu numeriere Die Literaturzitate sind entweder zu numerieren oder mit Namen und Jahresza »SCHULZ (1970)« zu kennzeichnen. Die Zeitschrift wird referiert in den »Biologic Abstracts«.

EDV in Medizin und Biologie
EDP in Medicine and Biology

INHALTSVERZEICHNIS/CONTENTS BAND/VOLUME XVII (1986)

Computeranwendungen/Computer applications

Statistische Verfahren/Statistical Methods

EDV in Medizin und Biologie **17** (1/2), 1–5, ISSN 0300-8282

WAMASTAT – ein System zur Benutzerführung bei der Auswertung klinischer Studien

B. Laminger, W. Dorda und Ch. Reichetzeder

Zusammenfassung

Die Verwendung von Statistiksystemen und -programmen verlangt vom Benutzer häufig ein mehr oder weniger großes EDV-Spezialwissen. Diese Anforderung stellt für einen Arzt eine oft unvertretbare zusätzliche Belastung dar. Aus diesem Grund wurde am Institut für Medizinische Computerwissenschaften in Wien ein Rahmensystem (WAMASTAT) entwickelt, das die Anwendung der vorhandenen Statistiksoftware ohne EDV-Spezialkenntnisse ermöglicht.

Summary

Today statistical analysis is one of the most important aids in medical science. The increasing number of experimental and epidemological, retrospective and prospective studies calls for fast, complete and mathematical adequate evaluation of data.

Besides self-written programs for special purposes, several purchasable program-packages for evaluation and graphic display of results (e.g. SAS, SAS/GRAPH, BMDP, GDDM, . . .) are made available by the Institute of Medical Computer-Sciences (IMC). We made the experience that many physicians would like to use those programs, but only a small group of physicians, which were already familiar with data-processing, actually used the possibilities. The knowledge required to use statistical systems – e.g. choosing the adequate program out of a number, the syntax of the query-languages etc. – prevents the use of those systems by physicians on a wide basis.

Therefore a frame system (WAMASTAT) has been developed at the IMC, which makes it possible to use the systems without detailed knowledge of data processing.

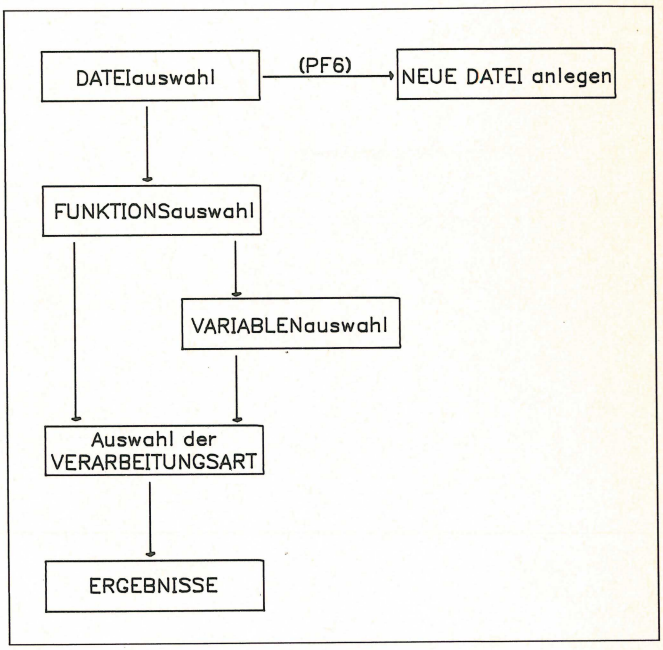
The main attributes of WAMASTAT are:
- *menu driven*
- *on-line help*
- *automatic activation of the programs*
- *automatic generation of syntactic correct statements*
- *automated display of the results on various mediums.*

1. Einleitung

Statistische Analysen gehören heute zu den wesentlichsten Hilfsmitteln in der medizinischen Forschung. Zunehmende experimentelle und epidemiologische, retrospektive und prospektive Untersuchungen fordern eine rasche, vollständige und mathematisch adäquate Auswertung der Daten.

Neben selbstentwickelten Programmen stehen am Institut für Medizinische Computerwissenschaften auch von verschiedenen Firmen angebotene Programmpakete zur Auswertung und Ergebnisdarstellung (SAS, SAS/Graph, BMDP, GDDM, . . .) zur Verfügung. Man stellte fest, daß viele Ärzte die

Abb. 1. Ablaufschema einer Auswertung im System WAMASTAT.



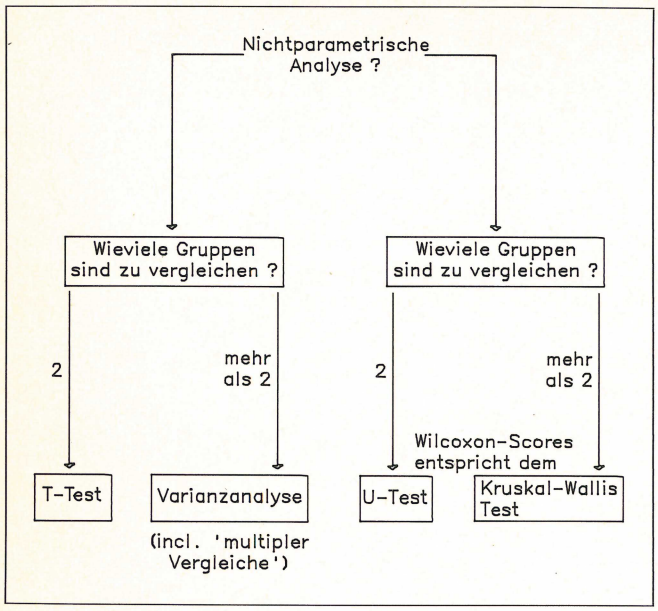


Abb. 2. Automatische Auswahl des adäquaten statistischen Verfahrens im System WAMASTAT.

gebotenen Möglichkeiten zwar gerne nutzen würden, jedoch nur einige EDV-Spezialisierte dies tun. Die Probleme bei der Auswahl der richtigen Programme eines oder mehrerer Statistiksysteme für eine bestimmte Problemstellung und die komplizierte Syntax der Abfragesprachen verhindern eine Verwendung dieser Systeme auf breiter Basis.

Um den Benutzern den Zugang zu den Statistikpaketen ohne EDV-Kenntnisse zu ermöglichen, wurde daher das System WAMASTAT entwickelt [1, 2, 3].

Neben einer Beschreibung des Systems soll ein Beispiel für seine Anwendung gegeben werden.

2. Beschreibung des Systems

Bei der Realisierung von WAMASTAT wurden folgende Richtlinien beachtet:

- Der Benutzer wird durch Auswahlschirme, die auf Knopfdruck weiter erläutert werden, bei der Formulierung der Anfrage unterstützt (»Menue-Technik«).
- Durch diesen interaktiven Definitionsprozeß werden vom System syntaktisch richtige Eingaben für die Statistikprogramme erstellt und diese werden automatisch aktiviert.
- Die Ausgabe der Ergebnisse sowie deren graphische Darstellung werden vom System WAMASTAT an den verwendeten Terminaltyp angepaßt.
- Das System bietet neben statistischen und graphischen Funktionen komfortable Datenerfassungskomponenten und verwaltet die Datenbestände äußerst benutzerfreundlich.
- Das System WAMASTAT ist einfach erweiterbar. Darüber hinaus sind noch nicht im System enthaltene Problemstellungen durch Absetzen eigener Befehle für die Statistikprogramme lösbar.

WAMASTAT wurde für das Betriebssystem CMS-VM/SP Rel. 2 entwickelt. Unterstützt werden im Rahmen des Systems Terminals der Typen IBM 3179, 3277, 3278 und 3279-2 und -3.

Für die Speicherung und Verwaltung der Datenbestände wird das Produkt SAS (Statistical Analysis System) verwendet.

Die Systemkomponenten von WAMASTAT gliedern sich in drei *Teile:*

- *Gemeinsame* Systemkomponenten für alle Benutzer. Dazu gehören das Rahmenprogramm, Programme für die Ablaufsteuerung, Definitionsmodule, Hilfsprogramme zur Dateiverwaltung und Ergebnisdarstellung/-ausgabe, Steuertabellen sowie Dateien, welche die Erläuterungen enthalten.
- *Benutzereigene* Datenbestände, Tabellen und vordefinierte Abläufe.
- *Schnittstellen* für die Übergabe von Anforderungen und Ergebnissen zwischen WAMASTAT und den Statistiksystemen.

3. Arbeiten mit dem System WAMASTAT

Die Daten aus klinischen Studien liegen üblicherweise in Tabellen vor und werden in WAMASTAT-Dateien gespeichert und mit Hilfe von WAMASTAT-Funktionen verarbeitet.

Die Abbildung 1 zeigt den grundsätzlichen Ablauf im System WAMASTAT:

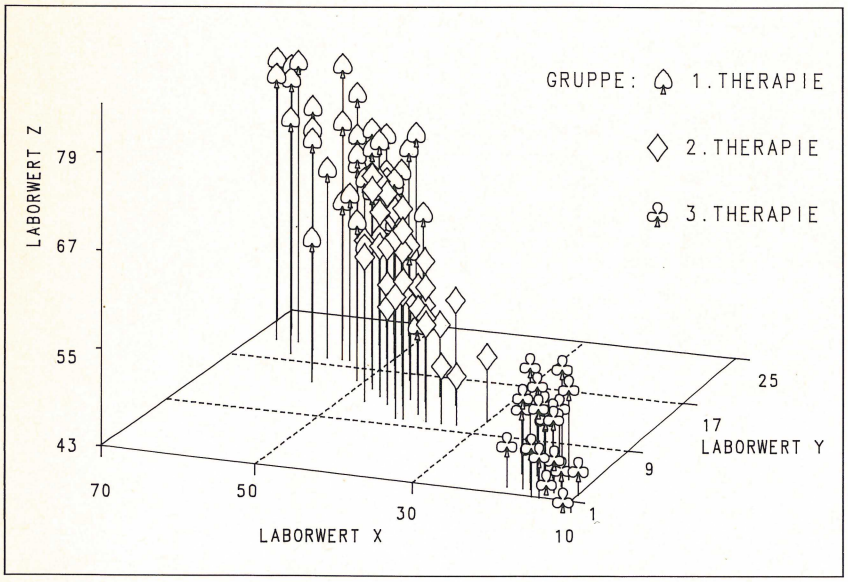


Abb. 3. Therapiestudie.

Abb. 4. Beispiel für einen Auswahlschirm im System WAMASTAT.

```
**********************************************************************

                          * * *  WAMASTAT  * * *                 DLR$VAR1
DORDA                                                             27/01/86
                                                                    12:38
Datei : THERAPIE - THERAPIESTUDIE

Funktion : GRUPPENVERGLEICH

         ==>  Waehlen Sie bitte eine VARIABLE aus  : *  14 *  <==

         Nr      Name         Variablenbeschreibung
         ------------------------------------------------------
  >>>>   011     NAME
  >>>>   012     ALTER
  >>>>   013     SEX          GESCHLECHT
  >>>>   014     GROESSE
  >>>>   015     LAB_X        LABORBEFUND X




                                                        Seite 002 von 002
  2-Info     8-STORNO      3/9-FUNKTIONSAUSWAHL       10-VOR      11-RUECK

**********************************************************************
```

1. Auswahl des zu bearbeitenden Datenbestandes
2. Auswahl der gewünschten Verarbeitung (Funktion)
3. Je nach gewählter Funktion eventuell weitere Präzisierung der Anforderung (z. B. Auswahl einer Variablen und der jeweils gewünschten Art der Verarbeitung)
4. Intern wird dann das entsprechende Programm aus den Statistiksystemen aktiviert und die jeweils gewünschten Anforderungen werden diesem übergeben (dieser Schritt läuft ohne Eingriff des Benutzers automatisch ab)
5. Darstellung und Ausgabe der Ergebnisse
6. Abhängig von der gewählten Funktion weiter bei 1., 2. oder 3.

Der Benutzer hat an jeder Stelle die Möglichkeit, zum vorhergehenden Schirm zurückzugehen oder den Dialog zu beenden.

Es sei noch festgestellt, daß jeder der angegebenen Arbeitsschritte durch einfache Auswahl auf dem entsprechenden Bildschirm erfolgt, so daß die Benutzung des Systems äußerst einfach und in wenigen Minuten erlernbar ist.

3.1 Anlegen einer neuen Datei

Zu Beginn einer Studie müssen die Variablen, welche dokumentiert werden sollen, definiert werden. Durch Tastendruck gelangt man in den entsprechenden Programmzweig (vgl. Abb. 1). Es wird vom System eine Bildschirmmaske angezeigt, auf welcher der Name der Studie (Dateinamen) sowie die Variablenbeschreibung (bestehend aus Variablenname, Variablentyp (Text/Zahl/Datum), Stellenzahl und wahlweise einer längeren Variablenbezeichnung) eingegeben werden müssen.

Danach wird durch die Taste PF1 die entsprechende Datei angelegt, und mit Hilfe der Funktion »Editieren der Datei (EINGABE von Beobachtungen)« kann sofort mit der Dateneingabe begonnen werden.

3.2 Überblick über die derzeit im System enthaltenen Funktionen

Die derzeit im System enthaltenen Funktionen können in vier Gruppen eingeteilt werden:

1. DATEIfunktionen
2. Statistik mit EINER Variablen (univariat)
3. Statistik mit ZWEI Variablen (bivariat)
4. Zusätzliche Funktionen.

Um einen Überblick über die Möglichkeiten des Systems zu geben, sind die bereits im Routinebetrieb stehenden Funktionen in Tabelle 1 aufgelistet; weitere Funktionen werden laufend in das System integriert.

Bei der Funktion »Editieren der Datei« werden vom System pro Beobachtung ein oder mehrere Bildschirmseiten für die

Tab. 1. Derzeit im System WAMASTAT realisierte Funktionen.

DATEIfunktionen
Editieren der Datei (EINGABE von Beobachtungen)
Ausgabe des Datei-Inhaltes
Kopieren der Datei
Neue Variablen zur Datei hinzufügen
Klassenbildung
Löschen einzelner Beobachtungen
Löschen einzelner Variablen
Löschen der Datei
Statistik mit EINER Variablen (univariat)
HÄUFIGKEITEN einer Variablen
MITTELWERT und STANDARDABWEICHUNG
Weitere univariate Statistik (MEDIAN ...)
HISTOGRAMM einer Variablen
Statistik mit ZWEI Variablen (bivariat)
KONTINGENZTAFELN, CHI-Quadrattest
T-PAARvergleich
GRUPPENvergleich
KORRELATIONSMATRIX
Lineare REGRESSION
PLOT zweier Variablen
Spezielle graphische Funktionen
PLOT von Mittelwerten, Median, ...
Zusätzliche Funktionen
Letzte Liste anzeigen

```
 ---------------------------------------------------------------
!                * * *  W A M A S T A T  * * *                  !
! WAMADEMO                                              29/01/86 !
! DLRVA1   - GRUPPENVERGLEICH                             15:52 !
! DATEINAME : THERAPIE - THERAPIESTUDIE                         !
 ---------------------------------------------------------------

ANALYSIS FOR VARIABLE ==> LAB_X <== CLASSIFIED BY VARIABLE  GRUPPE

ANALYSIS OF VARIANCE

LEVEL              N        MEAN          AMONG MS      WITHIN MS
                                            4954.5        693.658
1                 20       46.55                       +---------+
2                242       41.88          F VALUE      !  PROB>F !
3                 20       14.35             7.14      !  0.0002 !
4                 67       40.75                       +---------+


TUKEY'S STUDENTIZED RANGE (HSD) TEST FOR VARIABLE: LAB_X
COMPARISONS SIGNIFICANT AT THE 0.05 LEVEL ARE INDICATED BY '***'

                SIMULTANEOUS            SIMULTANEOUS
                   LOWER    DIFFERENCE     UPPER
  GRUPPE        CONFIDENCE    BETWEEN    CONFIDENCE
COMPARISON        LIMIT       MEANS       LIMIT

1     - 2        -11.146       4.673      20.493
1     - 4        -11.521       5.804      23.129
1     - 3         10.699      32.200      53.701     ***

2     - 1        -20.493      -4.673      11.146
2     - 4         -8.256       1.131      10.517
2     - 3         11.707      27.527      43.346     ***

4     - 1        -23.129      -5.804      11.521
4     - 2        -10.517      -1.131       8.256
4     - 3          9.071      26.396      43.721     ***

3     - 1        -53.701     -32.200     -10.699     ***
3     - 2        -43.346     -27.527     -11.707     ***
3     - 4        -43.721     -26.396      -9.071     ***

*********************************************************************
```

Abb. 5. Ergebnisse der Funktion Gruppenvergleich im System WAMASTAT.

Dateneingabe vorgeschlagen. Auf Wunsch können diese interaktiv modifiziert werden. Auch das Auffinden von Beobachtungen, welche spezielle Bedingungen erfüllen, wird mit dieser Funktion ermöglicht. Eine große Hilfe bei der Durchführung klinischer Studien bietet die Möglichkeit, interaktiv Einberufungsbriefe oder Formulare zu definieren, die, mit den jeweiligen Daten ergänzt, auf Knopfdruck für alle Patienten ausgedruckt werden können.

Die Funktion »Klassenbildung« ermöglicht, die verschiedenen Ausprägungen einer Variablen in Klassen zusammenzufassen.

In der Funktion »Weitere univariate Statistik« wird neben den Lokalisations- und Streuungsmaßen ein Test auf Normalverteilung berechnet und auf Wunsch durch Verteilungsplots ergänzt.

Beim »Histogramm einer Variablen« kann interaktiv die vom System vorgeschlagene Graphik in ihrem Design variiert werden. Durch die Wahl verschiedener Diagrammarten (Kreis-, Balken-, Sterndiagramm) können unterschiedliche Datenmerkmale hervorgehoben werden.

Die Funktion »Gruppenvergleiche« führt eine Reihe von Tests zum Vergleich zweier oder mehrerer Gruppen durch.

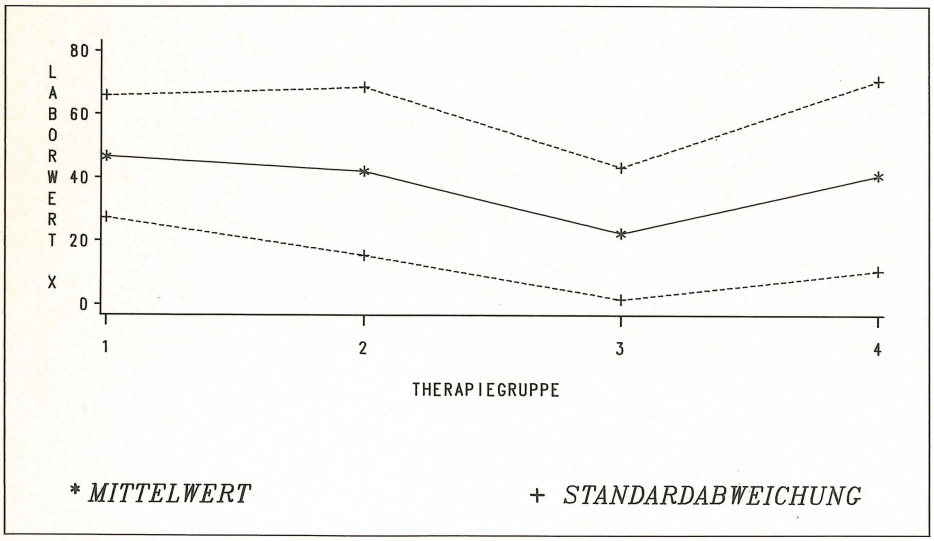


Abb. 6. Therapiestudie.

Welches der statistischen Verfahren zum Vergleich von Gruppen auf Grund der jeweiligen Fragestellung und Datenstruktur anzuwenden ist, wird vom System automatisch festgestellt (vgl. Abb. 2). Es wird also je nach Struktur der zu analysierenden Daten entweder ein T-Test oder U-Test bzw. eine Varianzanalyse (inkl. korrigierter multipler Vergleiche zwischen jedem Paar von Gruppen) oder ein Kruskal-Wallis-Test durchgeführt.

Mit der Funktion »Plot von Mittelwerten, Median, etc.« kann der zeitliche Verlauf einer Variablen in einem Patientenkollektiv graphisch dargestellt werden. Außerdem ist es möglich, den Vergleich der Mittelwerte bzw. der Mediane mehrerer Gruppen zu veranschaulichen.

3.3 Lösen spezieller Fragestellungen

Die meisten der üblichen Fragestellungen der medizinischen Statistik können mit Hilfe der derzeit realisierten WAMASTAT-Funktionen beantwortet werden. Fallweise treten aber spezielle statistische Problemstellungen auf. Da das System WAMASTAT die Daten intern in SAS-Dateien speichert, kann dann sofort – ohne jede Datenumspeicherung – auf das Statistiksystem SAS zurückgegriffen werden. Weiters kann mit Hilfe der SAS-Prozedur BMDP das Statistiksystem BMDP eingesetzt werden. Umgekehrt können mit Hilfe von SAS erstellte Datenbestände sofort vom System WAMASTAT verarbeitet werden. Wie ersichtlich, kann also zwischen der Verwendung von SAS und WAMASTAT hin- und hergewechselt werden: Eine im System WAMASTAT implementierte Fragestellung wird man wegen der Benutzerfreundlichkeit in diesem System durchführen, während man für spezielle Fragestellungen auf SAS und BMDP zurückgreift.

Ein spezieller Verarbeitungszweig erlaubt die Aktivierung individuell vordefinierter Abläufe: z. B. zur dreidimensionalen Darstellung des Zusammenhanges von Variablen (Abb. 3).

4. Ein Beispiel

Eine Datei enthält die Daten einer Therapiestudie. Es soll untersucht werden, ob die Therapie einen Einfluß auf den Laborwert X hat. Diese Fragestellung wird mit Hilfe der Funktion »Gruppenvergleich« gelöst.

Die dazu notwendigen Eingaben sind:
- Auswahl der Datei
- Auswahl der Funktion »Gruppenvergleich«
- Auswahl der Variablen (Therapieart, Laborwert X)

Die Auswahl kann mittels Lichtstift oder durch Eingabe einer Nummer erfolgen. Ein Beispiel für einen Auswahlschirm zeigt Abbildung 4. Abbildung 5 zeigt die Ergebnisse.

Vom System wurde festgestellt, daß vier Therapiegruppen zu vergleichen sind, und daher eine Varianzanalyse inklusiver multipler Vergleiche nach Tukey durchgeführt. Dadurch ist sofort ersichtlich, daß sich die dritte Therapiegruppe signifikant ($p < 0{,}05$) von den anderen im Laborwert X unterscheidet. Dies kann auch noch durch die Funktion »Plot von Mittelwert, Median, etc.« veranschaulicht werden (vgl. Abb. 6). Die Ergebnisse, die zuerst auf dem Bildschirm ausgegeben werden, können wahlweise auch sofort ausgedruckt werden.

5. Erfahrungen

Die Verwendung des Systems erfordert praktisch keine EDV-Kenntnisse. Zu jedem Eingabeschirm existiert auch eine ausführliche und jederzeit abrufbare Erläuterung der notwendigen und möglichen Eingaben, so daß der Benutzer die Handhabung von WAMASTAT sofort in der Praxis erlernen kann.

WAMASTAT wurde im Juni 1984 in die Produktion übernommen. Bereits mehr als 350 Ärzte benutzten das System im Dezember 1985. Das System wurde auch von jenen sehr positiv aufgenommen, die bereits Erfahrungen im Umgang mit den vorhandenen Statistikprogrammen hatten, da es eine wesentliche Erleichterung darstellt. Am erfreulichsten ist jedoch, daß Ärzte, die bislang von der EDV keinen Gebrauch machten, durch WAMASTAT dazu motiviert wurden und seither sehr intensiv mit dem System arbeiten.

Literatur

DORDA, W., B. LAMINGER, CH. REICHETZEDER, P. SACHS: Computereinsatz zur Analyse medizinischer Daten. Die Auswertungssysteme WAMAS und WAMASTAT. In GRABNER, G. (Hrsg.): WAMIS – Wiener Allgemeines Medizinisches Informationssystem. Springer-Verlag. Berlin, Heidelberg, New York, Tokyo, 1985, 251–291.

DORDA, W., B. LAMINGER, P. SACHS, CH. REICHETZEDER: Software-Systeme zur Auswertung medizinischer Daten. In GELL, G. und CH. EICHTINGER (Hrsg.): Medizinische Informatik 84, R. Oldenbourg-Verlag. Wien, München, 1984, 78–82.

REICHETZEDER, CH., B. LAMINGER, W. DORDA: Ein System zur Benutzerführung bei der Auswertung medizinischer Daten. In GELL, G. und CH. EICHTINGER (Hrsg.): Medizinische Informatik 84, Oldenbourg-Verlag. Wien, München, 1984, 83–87.

Eingegangen am 10. Februar 1986.
Anschrift der Verfasser: Dipl.-Ing. Brigitte Laminger, Dipl.-Ing. Wolfgang Dorda und Christian Reichetzeder, Institut für Medizinische Computerwissenschaften, Garnisongasse 13, A-1090 Wien/Österreich.

EDV in Medizin und Biologie **17** (1/2), 6–8, ISSN 0300-8282

Die Berechnung des Lebenserwartungsdefizits bei Tumorerkrankungen mit dem Statistikpaket SPSS

H. Kolles und O. Schmitt

Zusammenfassung

Das Lebenserwartungsdefizit ist ein Zeitmaß, das sich aus der Differenz der normalen Lebenserwartung im jeweiligen Alter und dem bei Vorliegen einer Tumorerkrankung tatsächlich erreichten Alter errechnen läßt. Es ist eine Größe, mit deren Hilfe die zeitlichen Auswirkungen lebenserhaltender Maßnahmen auf den weiteren Lebensverlauf beurteilt werden können. Den herkömmlichen statistischen Maßnahmen der Tumorbehandlung ist es insofern überlegen, als bei seiner Ermittlung sowohl die Altersverteilung der Patienten als auch die Dauer der Erkrankung berücksichtigt werden.

Im folgenden wird die Berechnung des Lebenserwartungsdefizits mit Hilfe des Statistikpakets SPSS beschrieben.

Summary

The deficiency of life expectancy is a time measure that can be calculated as the difference from normal life expectancy in the respective age group and the really reached age of patients with a malignant tumour disease. It is a parameter that allows to examine the effects of therapy in the development of the disease. With respect to the traditional statistical measures used in the treatment of tumour diseases it has many advantages because as well the age-distribution as the duration of the disease are considered.

In this paper the calculation of the deficiency of life expectancy by means of the statistical package SPSS is described.

Tabelle 1. Vergleich der Leistungsfähigkeit der beiden Größen »Lebenserwartungsdefizit« und »Überlebenszeit«.

Lebenserwartungsdefizit (LED)	Überlebenszeit (UEZ)
– kleines V (s/x̄)	– sehr großes V (über 100 %)
– Berücksichtigung der Altersverteilung	– keine Berücksichtigung
– Maß für die Malignität eines Tumors	– maßgeblich durch den Zeitpunkt der Tumordiagnose bestimmt
– geeignet zur Beurteilung von Rezidiven und Metastasen	– nicht geeignet bei langwierigen Tumorerkrankungen

Tabelle 2. Aufbau eines SPSS-Programmes zur Berechnung des LED. Fehlende Werte lassen sich durch die MISSING-VALUES-Anweisung aus den Rechnungen ausschließen.
Die Angabe von UEZ, KREBSALT und EX erfolgt hier in Monaten, die Angabe des LED in Jahren.
Die INPUT-MEDIUM-Anweisung wurde nicht spezifiziert. Beachte: Das Geburtsjahr kann 0 sein (z. B. 2. 3. 00), ohne daß der Fall als fehlend behandelt werden muß. Deshalb steht in der RECODE-Anweisung »Blank« und nicht »0«.

```
DATA LIST       FIXED GMONAT                1– 2
                      GJAHR                 3– 6
                      MDIAGPTu              7– 8
                      JDIAGPTU              9–12
                      TODMONAT             13–14
                      TODJAHR              15–18
                      SEX                  19
                      DIAG                 20–22
INPUT MEDIUM    CARD/TAPE/DISK/OTHER
N OF CASES      UNKNOWN
RECODE          GMONAT TO TODJAHR (BLANK = 9999)
COMPUTE         UEZ = (TODJAHR – JDIAGPTU)
                      *12 + (TODMONAT – MDIAGPTU)
COMPUTE         KREBSALT = (JDIAGPTU – GJAHR)
                *12 + (MDIAGPTU – GMONAT)
                STERBETAFELERGEBNISSE (s. Tabelle 3)
COMPUTE         LED = EX – UEZ/12
IF              (GMONAT = 9999 OR JDIAGPTU = 9999
                OR TODJAHR = 9999) LED = 9999
MISSING VALUES  LED (9999)
VAR LABELS      GMONAT, Geburtsmonat/
                GJAHR, Geburtsjahr/
                MDIAGPTU, Monat der Diagnose des
                  Primärtumors/
                JDIAGPTU, Jahr der Diagnose des
                  Primärtumors/
                TODMONAT, Todesmonat/
                TODJAHR, Todesjahr/
                SEX, Geschlecht/
                DIAG, Diagnose laut ICD-Code
                UEZ, Überlebenszeit
                KREBSALT, Alter bei Diagnose des
                  Primärtumors/
                LED, Lebenserwartungsdefizit/
                EX, Statistische Lebenserwartung im Alter x/
BREAKDOWN       TABLES = LED BY DIAG
*SELECT IF      (DIAG = 174)
CONDESCRIPTIVE LED
READ INPUT DATA
END INPUT DATA
FINISH
```

Problemstellung

Das Lebenserwartungsdefizit (LED) wurde bereits früher als ein Maß zur Beurteilung des Therapieerfolges bei Tumorerkrankungen vorgeschlagen (KROKOWSKI, 1979). Gegenüber den herkömmlichen statistischen Methoden hat diese Größe, die außerhalb des zeitlichen Verlaufs der Erkrankungen liegt, insofern Vorteile, als sie sowohl die Altersverteilung der Patienten mit der entsprechenden Lebenserwartung als auch die Dauer der Erkrankung selbst berücksichtigt.

Das LED errechnet sich folgendermaßen:

$LED = E_x - ÜZ$

E_x: statistische Lebenserwartung im Alter x, dem Alter bei Diagnose des Tumors

ÜZ: Überlebenszeit, Zeitpunkt von der Diagnose des Tumors bis Tod

Die Lebenserwartung im Alter x (E_x) kann den sog. amtlichen Sterbetafeln entnommen werden. Sie erscheinen jedes Jahr in abgekürzter Form und werden nach jeder Volkszählung in Form einer „allgemeinen Sterbetafel" neu herausgegeben. Dabei entscheidet das Datum der Diagnosestellung über die zugrundezulegende Sterbetafel.

Die weniger aufwendige Berechnung der herkömmlichen Maße (Überlebenszeit bzw. 5-Jahres-Überlebenszeit) machte die Einbeziehung von Sterbetafelergebnissen nicht erforderlich, weswegen sich auch bislang die klinische Anwendung des LED noch nicht durchgesetzt hat. Verglichen mit der Überlebenszeit (ÜZ) bietet es jedoch entscheidende Vorteile (Tab. 1), weswegen unseres Erachtens das LED klinisch breitere Anwendung finden sollte.

Im folgenden Beitrag wird ein einfaches Verfahren zur Berechnung des LED mit Hilfe des Statistikpakets SPSS (BEUTEL u. SCHUBÖ, 1983) vorgestellt. SPSS ist in Deutschland an fast allen Rechenzentren verfügbar. Es erfordert keine Programmierkenntnisse, so daß es auch klinisch relativ einfach anzuwenden ist.

Tabelle 3. Mögliche Integration von Sterbetafelergebnissen in das SPSS-Programm:

1. Im logischen Ausdruck, der bei der IF-Anweisung in Klammern steht, werden die Intervallgrenzen für die Einteilung des Alters bei Diagnose des Primärtumors festgesetzt.
   Im folgenden Fall werden die Intervalle in Monaten angegeben, um eine größere Präzision der Rechnungen zu erzielen.
2. Die Sterbetafel ordnet jedem Alter genau eine statistische Lebenserwartung E_x zu.

```
                                    1.          2.
                                    I----------I I-----I
IF      (KREBSALT GE  0 AND LT 12) EX = 793
IF      (KREBSALT GE 12 AND LT 24) EX = 806
IF      (KREBSALT GE 24 AND LT 36) EX = 795
IF      (KREBSALT GE 36 AND LT 48) EX = 784
.                       .           .
.                       .           .
.                       .           .
```

Methode

Bei der Berechnung des LED sind mehrere Einflußgrößen voneinander zu trennen, die als Variablen im SPSS-Programm berücksichtigt werden müssen: Geburtsdatum, Sterbedatum, Diagnosezeitpunkt des Primärtumors, Geschlecht, Art und Lokalisation des Primärtumors.

Mit Hilfe dieser Variablen werden zur Berechnung des LED folgende Intervalle definiert: das Erkrankungsalter, berechnet

Tabelle 4. Beispielausdruck zu BREAKDOWN.
Die verwendeten Daten stammen aus dem Saarländischen Krebsregister und sind Teil einer Studie zur ossären Metastasierung (Angabe des LED in Monaten).

Variable	Code	Value label	Sum	Mean	Std. Dev.	Variance	N
For entire population			9901.4151	17.5557	10.3364	106.8408	(564)
Diag.	151.	Magen	167.4166	16.7417	10.2900	105.8833	(10)
Diag.	153.	Dickdarm	166.4166	11.8869	7.4731	55.8468	(14)
Diag.	154.	Mastdarm	120.9167	12.0917	10.9507	119.9173	(10)
Diag.	156.	Gallenwege	29.5833	9.8611	1.9744	3.8981	(3)
Diag.	157.	Pankreas	14.4167	14.4167	0.0000	0.0000	(1)
Diag.	159.	Verdauungsorgane	39.4167	39.4167	0.0000	0.0000	(1)
Diag.	162.	Lunge	836.5832	17.7996	6.7417	45.4510	(47)
Diag.	164.	Mediastinum	18.9167	18.9167	0.0000	0.0000	(1)
Diag.	170.	Knochen primär	401.7499	36.5227	26.7638	716.3028	(11)
Diag.	172.	Hautmelanom	51.1667	51.1667	0.0000	0.0000	(1)
Diag.	174.	Mamma	6288.1656	17.6634	9.5177	90.5870	(356)
Diag.	179.	Uterus	126.0833	15.7604	7.5516	57.0266	(8)
Diag.	180.	Cervix uteri	443.3333	18.4722	9.0002	81.0040	(24)
Diag.	182.	Corpus uteri	93.0833	13.2976	4.7184	22.2636	(7)
Diag.	183.	Ovar	193.3333	13.8095	7.4472	55.4609	(14)
Diag.	184.	Weibl. Geschlechtsorg.	16.7500	8.3750	1.8267	3.3368	(2)
Diag.	188.	Harnblase	110.7500	15.8214	2.6803	7.1838	(7)
Diag.	189.	Niere	406.9166	14.0316	6.7062	44.9734	(29)
Diag.	190.	Auge	6.0833	6.0833	0.0000	0.0000	(1)
Diag.	193.	Schilddrüse	256.4166	19.7244	12.9219	166.9761	(13)
Diag.	194.	Endokrine Drüsen	26.8333	26.8333	0.0000	0.0000	(1)
Diag.	199.	Unbekannter Tumor	87.0833	29.0278	26.8143	719.0090	(3)

Total Cases = 564

Tabelle 5. Beispielausdruck zu CONDESCRIPTIVE (Angabe des LED in Monaten).

Mean	17.663	Std. error	0.504	Std. Dev.	9.518
Variance	90.587	Kurtosis	–0.384	Skewness	0.114
Range	56.583	Minimum	–13.167	Maximum	43.417
Sum	6288.166				

Valid observations – 356 Missing observations – 0

als Differenz aus Diagnosezeitpunkt des Primärtumors und dem Geburtsdatum, und die Überlebenszeit, berechnet als Differenz aus Todesdatum und Diagnosedatum des Primärtumors. Sie lassen sich im SPSS-Programm am zweckmäßigsten durch „COMPUTE"-Anweisungen ausdrücken (Tab. 2).

Die statistische Lebenserwartung im Alter x (E_x) läßt sich in das SPSS-Programm in Form von »IF«-Anweisungen integrieren (Tab. 3). Das LED wird als Differenz der statistischen Lebenserwartung im Alter x (E_x) und der Überlebenszeit als COMPUTE-Anweisung in das SPSS-Programm eingefügt, wobei auch die Möglichkeit besteht, fehlende Werte aus den Rechnungen zu eliminieren (Tab. 2).

Geht man von einem maximalen Alter bei Diagnosestellung des Primärtumors aus, so muß man bei allgemeinen Sterbetafeln die gleiche Anzahl von IF-Anweisungen, bei abgekürzten Sterbetafeln (z. B. mit 5-Jahres-Schritten) entsprechend weniger in das Programm einbauen.

Für eine derart große Anzahl von IF-Anweisungen muß in der Regel der Speichersatz, in dem diese Datentransformationen angewiesen werden, der »TRANSSPACE«, vergrößert werden. Dies geschieht am einfachsten durch die »ALLOCATE«-Anweisung bzw. bei der Siemens-Version durch die »SPACE«-Anweisung. Bei einer Datei mit mehreren Tumorarten empfiehlt sich zur Datenauswertung die Prozedur »BREAKDOWN«, bei der alle statistischen Mittelwerte sowie die Standardabweichungen des LED, für jede Tumorart getrennt, übersichtlich untereinander ausgegeben werden (Tab. 4).

Bei einer Datei mit nur einer Tumorart (z. B. Mamma-Karzinom) läßt sich die Prozedur »CONDESCRIPTIVE« verwenden, die die wichtigsten deskriptiven statistischen Kenngrößen für das LED berechnet (Tab. 5).

Literatur

BEUTEL, P., W. SCHUBÖ: SPSS 9-Statistik-Programm-System für die Sozialwissenschaften. Fischer, Stuttgart–New York (1983)

KROKOWSKI, E.: Krebsvorsorge und moderne Krebsdiagnostik. In: Deutsche Akademie für medizinische Fortbildung (Hrsg.), Neue Aspekte der Krebsbekämpfung. Thieme, Stuttgart (1979), S. 1–8

Eingegangen am 1. Oktober 1985.
Anschrift der Verfasser: PD Dr. med. O. Schmitt, Ltd. Oberarzt der Orthopädischen Universitätsklinik und Poliklinik, D-6650 Homburg/Saar.

EDV in Medizin und Biologie 17 (1/2), 9–11, ISSN 0300-8282

Welche Fallzahlen erfordert die Methode CRITLEVEL?
Ergebnisse einer Simulationsstudie

U. Abel und J. Berger

Zusammenfassung

CRITLEVEL ist eine explorative Methode zur Auffindung solcher Werte quantitativer prognostischer Faktoren, bei denen ein Prognosesprung eintritt. Als Folge der in der Methode enthaltenen Mehrfachtestung können leicht Fehlschlüsse auftreten. In der Arbeit werden die Resultate ausgedehnter Computersimulationen präsentiert, welche Anhaltspunkte für die zur Vermeidung dieser Fehlschlüsse erforderlichen Fallzahlen liefern.

Schlüsselwörter: CRITLEVEL, Prognose, Fallzahlen, Computersimulation

Summary

CRITLEVEL is an exploratory procedure designed for finding those levels of quantitative prognostic variables where a pronounced deterioration of prognosis occurs. Due to multiple testing implied by the technique, CRITLEVEL may lead to erroneous conclusions. The paper presents the results of extensive computer simulations examining the sample sizes required for safe application of the method.

Key words: CRITLEVEL, prognostic factor, sample size, computer simulation

I. Einleitung

CRITLEVEL (1) ist ein exploratives Verfahren zur Bestimmung solcher (»kritischer«) Werte quantitativer prognostischer Variablen V, bei deren Überschreitung eine ausgeprägte Verschlechterung der Prognose, ein »Prognosesprung«, eintritt. Das Verfahren besteht aus den folgenden drei Schritten:

1. Man wähle ein Intervall I von Werten der Variablen V, innerhalb dessen kritische Niveaus zu suchen sind. Ist m der kleinere der Stichprobenumfänge von Werten v, die unterhalb bzw. oberhalb von I liegen, so muß m groß genug sein, um sinnvolle Failure-time-Vergleiche zu gestatten.
2. Für alle Werte v innerhalb I werden gleich große Populationen $P_{\leq v,m}$ vs. $P_{>v,m}$ von je m Patienten mit Werten unterhalb bzw. oberhalb von v miteinander verglichen.
3. Die Resultate der Vergleiche (wir wollen hier annehmen: die p-Werte, die der Logrank-Test liefert) werden graphisch gegen V aufgetragen.

CRITLEVEL ist in der Praxis mehrfach angewandt worden (1, 2) und hat sich dabei recht gut bewährt. Gleichzeitig hat sich freilich herausgestellt, daß man auch bei der Benutzung von Nonsensevariablen V zuweilen deutliche Extrema der in Schritt 3 erzeugten Graphik erhält, die man fälschlicherweise als kritische Werte deklarieren würde. Die Höhe der Extrema nimmt in diesem Falle mit wachsendem Stichprobenumfang

Tab. 1. Mittelwerte (obere Ziffer) und Standardabweichungen (untere Ziffer) der in den beschriebenen Simulationen gewonnenen empirischen Verteilungen der Werte p_0 für verschiedene Konstellationen (n, m, Z).

n	m/n [%]	10	20	30	40
Z ~ U (0,1)	50	.0289	.0598	.1063	.2165
		.0291	.0626	.1018	.1547
	100	.0242	.0774	.1179	.1955
		.0307	.0755	.1185	.1501
	200	.0201	.0527	.0965	.1964
		.0225	.0511	.0859	.1567
	300	.0227	.0488	.0813	.1858
		.0242	.0526	.0795	.1450
Z ~ U (0,3)	50	.0406	.0741	.1245	.2029
		.0349	.0751	.1172	.1607
	100	.0263	.0472	.1044	.1978
		.0307	.0435	.1086	.1508
	200	.0205	.0524	.0939	.1788
		.0248	.0532	.0889	.1487
	300	.0196	.0514	.0974	.1783
		.0202	.0530	.0765	.1332
keine Zensierung	50	.0273	.0646	.0993	.2001
		.0321	.0757	.0886	.1572
	100	.0230	.0511	.1137	.1696
		.0345	.0498	.0982	.1439
	200	.0189	.0535	.1065	.1680
		.0232	.0546	.0975	.1284
	300	.0214	.0451	.1064	.1672
		.0271	.0492	.0999	.1548

Tab. 2. Aus den simulierten Verteilungen von p_0 geschätzte Wahrscheinlichkeit $\hat{W}$ (q) dafür, einen Wert $p_0 < q$ bei Anwendung der Methode CRITLEVEL zufällig zu erhalten, berechnet für verschiedene Konstellationen (n, m, Z) a. q = 0.05; b. q = 0.01; c. q = 0.005; d. q = 0.001

Tab. 2a $\hat{W}$ (0.05)

n	m/n [%]	10	20	30	40
Z ~ U (0,1)	50	.8187	.5282	.3587	.1255
	100	.8599	.4396	.3223	.1692
	200	.9184	.5738	.3307	.1716
	300	.8922	.6120	.4201	.1745
Z ~ U (0,3)	50	.6811	.4612	.3028	.1670
	100	.8390	.6206	.3527	.1685
	200	.9084	.5782	.3659	.1934
	300	.9306	.5867	.3119	.1696
keine Zensierung	50	.8273	.5197	.3350	.1850
	100	.8641	.5798	.3018	.2149
	200	.9234	.5708	.3310	.1860
	300	.8972	.6409	.3582	.2263

Tab. 2b $\hat{W}$ (0.01)

n	m/n [%]	10	20	30	40
Z ~ U (0,1)	50	.2656	.1775	.1419	.0381
	100	.3640	.1534	.1259	.0590
	200	.3832	.1801	.0936	.0611
	300	.3369	.2192	.1450	.0588
Z ~ U (0,3)	50	.1412	.1681	.1156	.0604
	100	.3305	.1993	.1339	.0592
	200	.3924	.1886	.1223	.0691
	300	.3694	.1937	.0779	.0522
keine Zensierung	50	.3167	.2064	.1035	.0718
	100	.4143	.1659	.0987	.0801
	200	.4231	.1892	.1142	.0580
	300	.3786	.2259	.1413	.0885

Tab. 2c $\hat{W}$ (0.005)

n	m/n [%]	10	20	30	40
Z ~ U (0,1)	50	.1619	.1233	.1078	.0274
	100	.2474	.1101	.0957	.0440
	200	.2440	.1204	.0636	.0460
	300	.2090	.1531	.1040	.0434
Z ~ U (0,3)	50	.0753	.1221	.0874	.0457
	100	.2196	.1326	.1004	.0443
	200	.2580	.1279	.0875	.0519
	300	.2240	.1317	.0508	.0375
keine Zensierung	50	.2089	.1532	.0724	.0553
	100	.2997	.1065	.0709	.0608
	200	.2828	.1294	.0831	.0418
	300	.2488	.1554	.1072	.0680

Tab. 2d $\hat{W}$ (0.001)

n	m/n [%]	10	20	30	40
Z ~ U (0,1)	50	.0712	.0709	.0722	.0172
	100	.1296	.0672	.0642	.0291
	200	.1109	.0646	.0359	.0306
	300	.0921	.0879	.0634	.0282
Z ~ U (0,3)	50	.0274	.0758	.0582	.0307
	100	.1113	.0703	.0659	.0293
	200	.1250	.0703	.0534	.0344
	300	.0925	.0726	.0269	.0233
keine Zensierung	50	.1050	.0978	.0428	.0382
	100	.1746	.0536	.0437	.0409
	200	.1398	.0721	.0522	.0260
	300	.1210	.0867	.0717	.0467

ab, steigt aber mit der Zahl der Vergleiche, d. h. mit der Größe des Intervalls I und der Dichte der Stichprobenwerte von V in I. Es gibt keine bequeme Abschätzung für den jeweiligen und kombinierten Einfluß der beiden Effekte.

Man kann versuchen, Fehleinschätzungen dadurch zu vermeiden, daß man nur solche Werte als kritisch bezeichnet, die sich auch bei Berücksichtigung der Multiplizität der in Schritt 2/3 durchgeführten Tests als solche erweisen, deren zugehöriger p-Wert also das korrigierte α-Niveau unterschreitet. Die üblichen Korrekturen für Mehrfachtests (3) wären jedoch zu konservativ, da die Prüfgrößen, die in CRITLEVEL für verschiedene Werte von V berechnet werden, hochgradig abhängig voneinander sind. Theoretische Untersuchungen des Problems, obwohl gewiß recht reizvoll, stehen noch gänzlich aus. Das Ziel der vorliegenden Arbeit war es daher, mit Hilfe von Computersimulationen praktische Anhaltspunkte für die Fallzahlerfordernisse bei der Benutzung von CRITLEVEL zu gewinnen.

II. Untersuchungsmethoden

Für jeden Simulationslauf wurden drei Zufallsvariable generiert:

1. Die theoretische Überlebenszeit Y; für sie wurde eine Exponentialverteilung mit Parameter 1 angenommen.
2. Die Zensierzeitverteilung Z; Z wurde als uniform auf [0,1] oder [0,3] und unabhängig von Y vorausgesetzt. Darüber hinaus wurden auch Simulationen für unzensierte Überlebenszeiten durchgeführt.
3. Die Verteilung der prognostischen Variablen V: hierfür wurde ebenfalls eine von Y, Z unabhängige Gleichverteilung auf [0,1] zugrunde gelegt.

Für die beobachtete Überlebenszeit S gilt S = min (Y, Z). Der zu erwartende Anteil a (Y, Z) an zensierten Daten ergibt sich in Abhängigkeit von Y und Z zu:

$$a(Y, Z) = {}_0\int^t \frac{x}{t} e^{-x}\, dx + {}_t\int^\infty e^{-x}\, dx$$ (mit t = 1 bzw. = 3, je nachdem, ob Z ~ U (0,1) oder Z ~ U (0,3)),

woraus man für Z ~ U (0,1) den Wert a(Y, Z) = 63,21 %, für Z ~ U (0,3) a(Y, Z) = 31,67 % erhält.

Es wurden Simulationen für die Gesamtstichprobenumfänge n = 50, 100, 200, 300 gerechnet. Für jeden Wert von n wurden Populationsvergleiche (Schritt 2/3 der Methode mit den Fallzahlen m = n (0.1, 0.2, 0.3, 0.4) durchgeführt. Wie man leicht nachrechnet, besteht das Verfahren CRITLEVEL für ein gegebenes Fallzahlpaar (m, n) aus n-2m Zwei-Gruppen-Vergleichen von Populationen $P_{\leq v,m}$, $P_{>v,m}$. Das Mini-

mum der p-Werte aus diesen Vergleichen werde mit p_0 bezeichnet. Die Wiederholungszahl der Simulationsläufe pro Konstellation (n, m, Z) betrug 100. Für jede Konstellation wurden die Parameter der aus den 100 Werten geschätzten Verteilung von p_0 berechnet und aufgezeichnet.[1)]

III. Resultate und Diskussion

Tabelle 1 enthält für jede der betrachteten Konstellationen den Mittelwert und die Standardabweichung der Werte p_0. Für die praktische Anwendung von CRITLEVEL relevant ist die Frage, wie hoch die Wahrscheinlichkeit dafür ist, einen p_0-Wert $< 0.05, 0.01, 0.005, 0.001$ zu erhalten. Zuverlässiger als aus der empirischen Verteilung von p_0 lassen sich diese Wahrscheinlichkeiten mittels Transformation auf die Normalverteilung schätzen. (Die Arcsin-Transformation erwies sich in diesem Zusammenhang als brauchbar.) Die errechneten Werte sind in Tabelle 2 dargestellt. Aus den beiden Tabellen lassen sich die folgenden Ergebnisse ablesen:

1. Für die Verteilung von p_0 ist das Verhältnis m/n maßgebend. Die durch die Erhöhung der Stichprobenumfänge m hervorgerufene Vergrößerung des Mittelwerts und der Standardabweichung von p_0 wird bei konstantem Verhältnis m/n durch die Zunahme der Anzahl der Vergleiche und die daraus resultierende Verringerung dieser Parameter recht genau kompensiert.
2. Das Ausmaß der Zensierungen spielt für die Ergebnisse keine erkennbare Rolle.
3. Wie Tabelle 2 lehrt, sollten Minima p_0, die über 0.005 liegen, für $m/n \leq 0.4$ (und somit vermutlich in jedem praktischen Anwendungsfall) nicht als kritisch angesehen werden, da sie in mehr als 5 % der Fälle zufällig auftreten. Werte von $p_0 \leq 0.005$ (besser noch: $p_0 \leq 0.001$) können als kritisch bezeichnet werden, sofern das Verhältnis m/n mindestens 40 % beträgt. Das Intervall I (s. Schritt 1 der Methode) sollte also äußerstenfalls aus den zentralen 20 Prozent der Stichprobenwerte von V bestehen. Weitergehende Fallzahlerwägungen sind, soweit dies aufgrund der vorliegenden Untersuchungen beurteilt werden kann, entbehrlich.

Abschließend sind noch einige Bemerkungen zur Wahl der Verteilung Y, Z, V angebracht. Die Annahme der Gleichverteilung von Z ist in den meisten klinischen Anwendungen nicht unrealistisch: auch scheint ja der Anteil der Zensierungen ohne Bedeutung für die Resultate zu sein. Die spezielle Form von V sollte, wie man sich leicht überlegt, belanglos sein, solange nur V von Y und Z unabhängig ist. Mit der Exponentialverteilungsannahme für Y werden eine Reihe von Anwendungsfällen (z. B. in der Onkologie) abgedeckt. Dennoch läßt sich natürlich der Einwand, daß es sich hier um einen Spezialfall handelt, nur schwer entkräften. Aus einigen mit anderen Verteilungen Y durchgeführten Simulationsläufen schließen wir jedoch, daß auch die Wahl von Y die Ergebnisse nicht wesenlich beeinflußt.

Literatur

1. ABEL, U., J. BERGER, H. WIEBELT (1984): CRITLEVEL: An exploratory procedure for the evaluation of quantitative prognostic factors. Meth. Inform. Med. **23,** 154–156.
2. ABEL, U. (1985): Aspekte des Krebsgeschehens aus biometrischer Sicht. Zur Veröff. angenommen in: Fortschr. d. Medizin.
3. SCHILLER, K., SONNEMANN, E. (1981): Tests zum multiplen Niveau α. Arbeitsberichte der Abteilung Statistik Nr. 10, Universität Dortmund.

[1)] Sämtliche Rechnungen wurden auf der IBM 3032 des DKFZ durchgeführt. Der gesamte CPU-Zeitbedarf betrug rund 200 000 Sekunden.

Eingegangen am 29. Januar 1986.
Anschrift der Verfasser: Dr. U. Abel, Deutsches Krebsforschungs-Zentrum (DKFZ), Abteilung Biostatistik, Postfach 10 19 49, D-6900 Heidelberg 1.

EDV in Medizin und Biologie 17 (1/2), 12–20, ISSN 0300-8282

Fuzzy modeling and reasoning in a medical diagnostic expert system*

K.-P. Adlassnig

Summary

CADIAG-2 (Computer-Assisted DIAGnosis) is a data-driven, rule-based fuzzy medical expert system developed for diagnostic screening and on-line consultation in a hospital. It is integrated into the system WAMIS (Wiener Allgemeines Medizinisches InformationsSystem [Vienna General Medical Information System]), the medical information system of the Vienna General Hospital. Extended clinical trials in the fields of gastroenterology and rheumatology have been conducted. First results obtained by testing about 500 clinical cases indicate the applicability of CADIAG-2 in this hospital setting.

This paper describes the main components and the formal concept of CADIAG-2. An example of a diagnostic process in the field of pancreatic diseases is provided for illustrative purpose.

Zusammenfassung

CADIAG-2 (Computerunterstützte Diagnose [Computer-Assisted DIAGnosis]) ist ein medizinisches Expertensystem für diagnostisches Screening und On-line-Konsultation im Krankenhaus. Fuzzy Produktionsregeln bilden die Grundlage der Wissensrepräsentation in diesem System. Der Diagnoseprozeß erfolgt datengetrieben, wobei die Patientendaten aus der Datenbank des medizinischen WAMIS (Wiener Allgemeines Medizinisches InformationsSystem) entnommen werden. Dies wird ermöglicht durch eine umfassende Integration des CADIAG-2 in das System WAMIS. CADIAG-2 wird derzeit in den Gebieten Gastroenterologie und Rheumatologie getestet. Die Ergebnisse der Auswertung der ersten 500 klinischen Fälle demonstrieren die Anwendbarkeit des Systems im Krankenhausbereich.

Die vorliegende Arbeit beschreibt das formale Konzept sowie die wesentlichen Komponenten von CADIAG-2. Ein illustratives Beispiel eines diagnostischen Prozesses aus dem Bereich der Pankreaserkrankungen ist beigefügt.

Schlüsselwörter:

Medizinisches Expertensystem CADIAG-2, Fuzzy Mengen und Fuzzy Logik, Medizinisches Informationssystem WAMIS, Interne Medizin.

1. Introduction

CADIAG-2[1] is a data-driven, rule-based fuzzy medical expert system, integrated into the medical information system WAMIS[2] of the Vienna General Hospital. The Vienna General Hospital is the teaching hospital of the University of Vienna Medical School and is constituted of about 70 medical clinics and institutes. At present, more than 1,200,000 case histories from about 700,000 patients are contained in its central patient data base. Both CADIAG-2 and WAMIS are on-line systems programmed in CICS/VS command level language and PL/1. They run in a time-sharing environment on an IBM 4381 Model P03 under DOS/VS, controlled by VM. For a description of WAMIS see [1, 2].

CADIAG-2 is at present being tested in this hospital setting. It is under retrospective evaluation at the 2nd Department for Gastroenterology and Hepatology (Director: Prof. Dr. G. Grabner) and the Hospital for Rheumatic Diseases of the Social Insurance Institute of Trade in Baden near Vienna (Director: Prof. Dr. G. Kolarz). The latter is associated with the University of Vienna Medical School for joint research. Until now, about 500 clinical cases, each case including up to 800 symptoms, signs, and laboratory test results, have been evaluated. These large symptom patterns consist of findings present, present to a certain degree, and definitely absent. Reports on these evaluations can be found in [3, 4].

CADIAG-2 is a successor of CADIAG-1 [4, 5]. Both systems employ two basic medical concepts for expressing medical associations between symptoms[3] and diseases: (1) the necessity of occurrence of a symptom with the disease; and (2) its sufficiency for concluding the disease. But whereas CADIAG-1 applies some qualitative categories of relationships between symptoms and diseases, expressed mathematically by means of first-order predicate calculus, CADIAG-2 uses two quantitative numbers to characterize the strengths of association between symptoms and diseases. Furthermore, while CADIAG-1 presupposes dichotomous decisions on the presence or absence of symptoms, CADIAG-2 allows uncertainty about these decisions and thus represents borderline symptoms, i.e., symptoms rendering a state between normal and pathological, more adequately. Fuzzy set theory [6] and fuzzy logic [6, 7] provide the necessary framework to formalize and manipulate certain as well as borderline symptoms, allow to propagate and aggregate this information, and to draw diagnostic conclusions. With the means of fuzzy logic, diag-

* This work was supported by a Max Kade Fellowship for Scientific Research and supported in part by the National Science Foundation NSF-IST Grant 8320416 and NASA Grant NCC2-275.

[1] CADIAG stands for *C*omputer-*A*ssisted *DIAG*nosis.

[2] WAMIS is the German acronym for *W*iener *A*llgemeines *M*edizinisches *I*nformations-*S*ystem (Vienna General Medical Information System).

[3] Throughout the text, the term »symptom« is considered to be synonymous with the terms »sign« and »laboratory test result«.

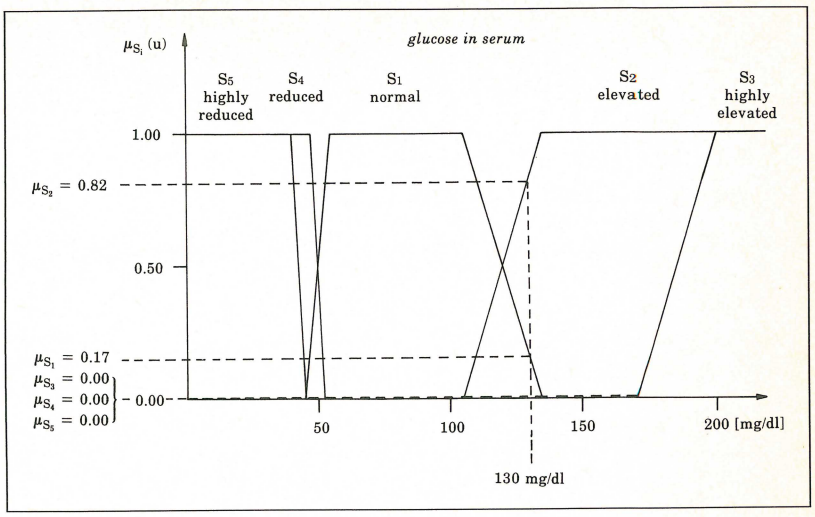


Figure 1. Determination of degrees of compatibility for a quantitative laboratory test result (serum glucose = 130 mg/dl) with the semantic medical concepts *serum glucose normal, elevated, highly elevated, reduced,* and *highly reduced.*

noses can be logically concluded with a certain degree of confidence. Additionally, a heuristic evidence aggregation function calculates a numerical support score for each diagnostic hypothesis. This score expresses the degree to which the given medical evidence supports the inferred diagnosis.

Medical expert systems whose design considerations influenced the conception of CADIAG-2 are CASNET [8, 9], MYCIN [10] (see also [11]), INTERNIST/CADUCEUS [12, 13], and EXPERT [14] (see also [15]). CADIAG-2's inference scheme has its roots in work published in [16, 17]. Further approaches to medical decision making applying fuzzy set theory, fuzzy logic, and possibility theory can be found in a survey on medical diagnosis and fuzzy subsets [18]. More recent work ist reported in [19–30].

The aim of this paper is to describe the main components and the formal concept of CADIAG-2. This is illustrated by examples of a CADIAG-2 description of a disease and a diagnostic process, both from the field of pancreatic diseases. Since a detailed and complete description of the mathematical basis of CADIAG-2 can be found elsewhere [3], this is not repeated here, but instead, emphasis is put on the description of the principal mechanisms of CADIAG-2 and their application in a hospital environment.

2. Medical expert system CADIAG-2

2.1. *Knowledge representation*

2.1.1. *Patient data*

In WAMIS, the stored patient records usually contain patient data on a detailed observational level, i.e., detailed history items, signs from physical examinations, quantitative laboratory test results, etc. An interface program, called patient data fuzzy interpreter, accesses the given medical data and transfers them to the CADIAG-2 system. During this step, medical information is abstracted and aggregated and converted into a representation commonly used in diagnostic discourse. This information can then be processed by the fuzzy inference engine.

An example for an abstracted symptom is *elevated glucose level,* set according to the result of the glucose test and the definition of *elevated.* The formal modeling of semantic medical concepts such as *elevated* that considers their inherent uncertainty, i.e., their gradual transition to adjacent medical concepts, is based on fuzzy set theory (see [6]). Adopting this theory, every symptom is considered to be a fuzzy set. Fuzzy sets are defined by fuzzy membership functions that assign to every symptom a degree of membership, expressing the degree of compatibility of a measured concrete value with the semantic concept under consideration. These membership functions are determined by the physician according to numerical ranges for normal and pathological findings, where the physician also indicates the transition zone. At present, about 400 membership functions for about 100 laboratory tests, where the test results are abstracted into classes such as *normal, elevated, highly elevated, reduced,* and *highly reduced,* are included into the patient data fuzzy interpreter. Most of them will be additionally adjusted according to sex and age of the patient during the actual diagnostic process. Figure 1 shows with an example how the degrees of compatibility of a quantitative laboratory test result are determined.

An example for an aggregated symptom is *limited motion in the hand joint,* defined according to physician's degree measurements of the *radial, ulnar, dorsal,* and *volar motion of the hand joint.* These detailed degree measurements are part of the medical documentation in WAMIS. Measurements like this and many other symptoms are simplified as binary symptoms in the WAMIS system. In these cases, the fuzzy logical concept automatically coincides with Boolean logic[4], a subtheory of fuzzy logic. The patient data fuzzy interpreter contains at present about 900 aggregation functions.

Once the abstracted and aggregated symptoms together with their degrees of compatibility have been transferred to the CADIAG-2 system, the degrees of compatibility can be altered by the physician according to his subjective perception of the case. By doing this, even symptoms defined as binary in the documentation system WAMIS can obtain intermediate values.

In more formal terms, degrees of compatibility $\mu_{S_i}(u_k) \in [0,1]$ between a symptom S_i and a measured or observed concrete value u_k, which is a member of the universe of discourse U of the respective symptom S_i, i.e., $u_k \in U$ ($U = \{u\}$), are determined by the patient data fuzzy interpreter and can be altered by the physician consulting CADIAG-2. A degree of zero means *no* and unity *full* compatibility with the meaning of the symptom S_i.

[4] Actually, Kleene's trivalued logical system [31] is used to capture also missing values, i.e., the logical values are 1: »Present«, ½: »Don't know«, and 0: »Absent«.

The degrees of compatibility $\mu_{S_i}(u_k)$ are interpreted as binary fuzzy relationships $\mu_{PS}(P,S_i)$ between the patient P and the symptom S_i, that is,

$$\mu_{S_i}(u_k) = \mu_{PS}(P,S_i). \quad (1)$$

Example 1 (quantitative laboratory test):

$\mu_{\textit{serum glucose normal}}$ (130 mg/dl) = 0.17
$\mu_{\textit{serum glucose elevated}}$ (130 mg/dl) = 0.82
$\mu_{\textit{serum glucose highly elevated}}$ (130 mg/dl) = 0.00
$\mu_{\textit{serum glucose reduced}}$ (130 mg/dl) = 0.00
$\mu_{\textit{serum glucose highly reduced}}$ (130 mg/dl) = 0.00

Example 2 (sign aggregated from detailed information about the *radial, ulnar, dorsal,* and *volar motion of the hand joint*):

$\mu_{\textit{limited motion in the hand joint}}$ (»Yes«) = 1.00

2.1.2. *Medical relationships*

In CADIAG-2, medical knowledge is represented in form of rules. These rules contain relationships between antecedents and consequents. Rules with a single medical entity as antecedent express known associations between medical entities, whereas compound antecedents, which are combinations of symptoms, allow the definition of pathophysiological states and the incorporation of very specific, complex criteria for diagnosing diseases. The evaluation of compound antecedents is carried out by means of fuzzy logic[5].

In CADIAG-2, the overwhelming part of the stored medical knowledge consists of relationships between single antecedents and consequents. The present ratio between single and compound antecedent/consequent rules is about 23,000:120.

Two kinds of relationships – the necessity and sufficiency – define the associations between antecedents and consequents. Causal relationships are not explicitly represented. They are (whenever possible) expressed as associations and captured by associational relationships. In CADIAG-2, the employed relationships are:

- frequency of occurrence O of the antecedent with the consequent (necessity);
- strength of confirmation C of the antecedent for inferring the consequent (sufficiency).

These relationships are interpreted as binary fuzzy relationships between antecedents and consequents. They take their values μ_O and μ_C in $[0,1] \cup \{v\}$ with v: »No relationship«. Linguistic terms λ_O and λ_C such as *always, often, seldom, never, strong, weak,* etc., have been found semantically useful in order to characterize these relationships, although numerical values are stored in the knowledge base of CADIAG-2 (see [3, 32]). The general form of the CADIAG-2 rules is:

IF *antecedent* THEN *consequent* WITH (O, C). (2)

The relationship tupels (O, C) contain linguistic and numerical values λ_O and μ_O, and/or λ_C and μ_C. An antecedent is a sufficient criterion for concluding a consequent if the strength of confirmation is $\mu_C = 1.00$; it is a necessary criterion if the frequency of occurrence is $\mu_O = 1.00$; and it is an excluding criterion if the frequency of occurrence is $\mu_O = 0.00$ and the strength of confirmation is $\mu_C = 0.00$. Relationships with intermediate degrees are supportive criteria. Graded adverse criteria are not considered in CADIAG-2.

Example 3 (supportive):
IF *elevated pancreatic oncofetal antigen (POA) in serum*
THEN MAY BE *pancreatic cancer*
with (λ_O = *often* [μ_O = 0.80], λ_C = strong [μ_C = 0.70]).

Example 4 (necessary and sufficient):
IF (IF NOT) *rheumatoid arthritis,* and
splenomegaly, and
leukopenia under 4 giga/l
THEN (THEN NOT) *Felty's syndrome*
with (λ_O = *always* [μ_O = 1.00], λ_C = *always* [μ_C = 1.00]).

Example 5 (excluding):
IF *positive rheumatoid factor*
then not *seronegative rheumatoid arthritis*
with (λ_O = *never* [μ_O = 0.00], λ_C = *never* [μ_C = 0.00]).

The relationships frequency of occurrence $\mu_O = 1.00$ and strength of confirmation $\mu_C = 1.00$ are also applied to establish symptom and disease taxonomies. In these taxonomies, present sub-terms are sufficient criteria for inferring super-terms, e.g., *bacterial arthritis* implies *arthritis.* But vice versa, super-terms are necessary criteria for the presence of sub-terms, e.g., criteria excluding *infectious arthritis* also exclude all kinds of *infectious arthrites* such as *bacterial* and *viral arthrites, arthritis* by *fungi* or *rickettsiae.*

In CADIAG-2, the frequency of occurrence and strength of confirmation are expressible as proportions of the cardinalities of sets of patients. They can be numerically calculated from sample patient data with already diagnosed patients. Although the frequency of occurrence and the strength of confirmation suggest an interpretation as conditional probabilities – P(S/D) as frequency of occurrence and P(D/S) as strength of confirmation – the property of the symptoms of having attached degrees of compatibility $\mu_{PS}(P,S_i)$, unequal zero or unity, prevents a pure probabilistic calculation.

The formal apparatus allowing to calculate proportions of fuzzy sets is found in the concept of the relative sigma-count. For a general exposition on the cardinality of fuzzy sets see [33].

Provided that the patient data base contains N patients, we calculate for the frequency of occurrence

$$\mu_{SD^O}(S_i,D_j) = \Sigma\,\text{Count}(S_i/D_j) = \frac{\Sigma\,\text{Count}(S_i \cap D_j)}{\Sigma\,\text{Count}(D_j)} = \frac{\sum_{k=1}^{N} \text{Min}\,[\mu_{PS}(P_k,S_i); \mu_{PD}(P_k,D_j)]}{\sum_{k=1}^{N} (\mu_{PD}(P_k,D_j))} \quad (3)$$

and for the strength of confirmation

$$\mu_{SD^C}(S_i,D_j) = \Sigma\,\text{Count}(D_j/S_i) = \frac{\Sigma\,\text{Count}(S_i \cap D_j)}{\Sigma\,\text{Count}(S_i)} = \frac{\sum_{k=1}^{N} \text{Min}\,[\mu_{PS}(P_k,S_i); \mu_{PD}(P_k,D_j)]}{\sum_{k=1}^{N} (\mu_{PS}(P_k,S_i))}. \quad (4)$$

[5]) In fact, an extension of fuzzy logic allowing variables to take values in $[0,1] \cup \{v\}$ with v: »Don't know« is used (see [3]).

If there are only degress of compatibility $\mu_{PS}(P_k,S_i) \in \{0,1\}$ and clinical diagnoses $\mu_{PD}(P_k,D_j) \in \{0,1\}$ in the sample patient data base, i.e., if their is no uncertainty about the assignment of measured or observed findings into symptom categories and no uncertainty about the clinical diagnoses for these patients, then the above-mentioned calculations coincide with the calculations for P(S/D) and P(D/S).

The Σ Count interpretations for symptom combination/disease, symptom/symptom, and disease/disease relationships are conducted analogously to symptom/disease relationships.

Table 1 shows as an example the stored medical description of *pancreatic cancer,* a disease belonging to the differential diagnostic group of pancreatic diseases, a subgroup of gastroenterology in CADIAG-2 .

Table 1. The medical documentation of *pancreatic cancer* in CADIAG-2 with the relationships frequency of occurrence $\mu_{SD^O}(S_i,D_j)$ and strength of confirmation $\mu_{SD^C}(S_i,D_j)$. An empty place means that there is no known or a totally unspecific relationship between the symptom and the disease.

Pancreatic cancer	μ_{SD^O}	μ_{SD^C}
Anamnesis/known present disorder(s)		
age, 17–29 years	0.03	
age, 30–50 years	0.08	
age, 51–80 years	0.09	0.01
sex, male	0.59	
sex, female	0.41	
(heavy) smoker	0.35	
alcoholism	0.20	0.01
diabetes mellitus	0.20	0.02
latent diabetes	0.55	0.02
venous thrombosis	0.35	0.03
Present symptoms		
tiredness	0.52	
weakness	0.55	0.01
lack of concentration	0.23	
malaise	0.60	
decrease in physical/mental powers	0.50	
major debilitation	0.26	0.02
jaundice	0.54	0.02
nausea	0.44	
vomiting	0.34	
anorexia	0.50	0.01
weight loss	0.82	0.05
subfebrile temperature	0.25	0.01
bloating	0.18	
pain, abdominal	0.86	0.06
pain, abdominal, epigastric region, diffuse	0.08	
pain, abdominal, right epigastric region	0.54	
pain, abdominal, mid-epigastric region	0.20	
pain, abdominal, left epigastric region	0.04	
pain, abdominal, deterioration at hyperlordosis	0.35	0.03
pain, abdominal, radiation to the back	0.33	
pain, abdominal, radiation to the side(s)	0.05	
pain, abdominal, radiation to the hypogastrium	0.03	
pain, abdominal, persistent, mild	0.09	0.01
pain, abdominal, persistent, intensive	0.16	0.01
pain, abdominal, intermittent, mild	0.45	0.02
pain, abdominal, intermittent, colicky	0.16	0.01
pain, abdominal, acute onset	0.16	
pain, abdominal, insidious onset	0.70	0.02
pain, abdominal, unrelated to food intake	0.66	
pain, abdominal, postprandial	0.34	
defecation, constipation	0.18	0.01
defecation, diarrhea	0.23	0.01
defecation, steatorrhea	0.09	0.02
defecation, light stools	0.40	
defecation, tarry stools	0.07	0.03
miction, dark urine	0.38	
anemia-signs	0.45	0.02
depression	0.33	
Physical examination		
general condition, normal	0.15	
general condition, ill	0.85	0.01
nutritional status, normal	0.25	
nutritional status, reduced	0.75	0.02
strength, normal	0.30	
strength, reduced	0.70	0.01
skin, subicteric	0.06	0.02
skin, icteric	0.54	0.02
skin, subcutaneous fat necrosis	0.10	0.03
skin, thrombophlebitis, migrans	0.10	0.03
lymphnodes (LN), palpable LN metastasis	0.09	0.30
eyes, sclerae, icteric	0.60	
abdomen, ascites	0.28	0.01
abdomen, systolic murmur, left epigastrium	0.17	0.05
abdomen, palpable abdominal mass	0.30	0.05
abdomen, liver, hepatomegaly	0.50	0.01
abdomen, liver, palpation, suspicion (susp.) of liver metastasis	0.55	0.30
abdomen, gall bladder, Courvoisier's sign	0.35	0.07
abdomen, spleen, splenomegaly	0.30	0.01
Laboratory findings		
blood sedimentation rate (BRS), increased	0.75	0.01
serum, alkaline phosphatase (AP), elevated	0.72	
serum, glucose, elevated	0.48	0.02
serum, bilirubin, elevated	0.60	0.02
serum, glutamic-oxaloacetic transaminase (SGOT), elevated	0.65	0.02
serum, glutamic-pyruvic transaminase (SGPT), elevated	0.65	0.02
serum, lactate dehydrogenase (LDH), elevated	0.46	0.02
serum, gamma-GT (γGT), elevated	0.65	0.02
serum, amylase, elevated	0.21	0.02
serum, lipase, elevated	0.21	0.02
serum, electrophoresis, alpha-2 globulin, elevated	0.60	
serum, electrophoresis, beta-globulin, elevated	0.60	
serum, pancreatic oncofetal antigen (POA), elevated	0.80	0.65
serum, CA-19-9, elevated	0.80	0.65
serum, carcinoembryonic antigen (CEA), elevated	0.68	0.30
hemogram, leucocytosis	0.32	0.01
urine, amylase, elevated	0.10	
urine, glucose, elevated	0.20	
stool, hemoccult test, positive	0.45	0.03
stool, chymotrypsin, pathological finding	0.70	0.20
stool, fat content, increased	0.09	0.20
Clinical investigations		
exocrine pancreatic function test(s), pathological finding	0.80	0.50
X-ray, hypotonic duodenography, susp. of inflammatory/malignant pancreatic tumor	0.38	0.05
X-ray, endoscopic retrograde choledochopancreatography (ERCP), susp. of pancreatic tumor	0.85	0.80
X-ray, selective arteriography, susp. of pancreatic tumor	0.85	0.85
X-ray, computerized tomography (CT), susp. of pancreatic tumor	0.85	0.85
ultrasonography (US), susp. of pancreatic tumor	0.80	0.70
percutaneous aspiration biopsy (US/CT-guidance), positive cytology	0.65	1.00

2.2. *Inference engine*

2.2.1. *Inference and chaining*

The fuzzy inference mechanism applied in CADIAG-2 (for a formal, detailed description see [3]) allows inference under uncertainty. The basic rule on which the inference mechanism relies is the compositional rule of fuzzy inference advanced in [7]. It accepts patient's symptoms with attached degrees of compatibility and infers diagnoses with a certain degree of confidence expressing the degree to which the diagnosis can logically be concluded from given medical evidence. From the logical point of view, only the strength of confirmation can be applied for carrying out inferences from given symptoms (modus ponens). An exception is the case in which the frequency of occurrence is $\mu_O = 1.00$ and the respective criterion is definitely absent. It can then be inferred that the conclusion has to be rejected (Boolean modus tollens).

The formal representation is:

$$\begin{array}{l} P \text{ has } S_i \text{ with } \mu_{PS}(P,S_i), \text{ and} \\ \text{If } S_i \text{ implies } D_j \text{ with } \mu_{SD^c}(S_i,D_j), \\ \hline \text{Then P has } D_j \text{ with } \mu_{PD}(P,D_j). \end{array} \qquad (5)$$

This modus ponens syllogism is calculated using the Max-Min composition that provides very secure and reliable inference values and is considered to be »conservative«. For many symptoms S_i in patient P the formula is stated as:

$$\mu_{PD}(P,D_j) = \max_{S_i} \text{Min}\,[\mu_{PS}(P,S_i); \mu_{SD^c}(S_i,D_j)]. \qquad (6)$$

The compositional rule of inference is not only applied for drawing conclusions from given symptoms, but also for conclusions from given combinations of symptoms. Furthermore, inferences among symptoms and among diseases are based on the same principle.

Chaining is carried out by employing the same rule in subsequent steps, e.g.,

$$\begin{array}{l} P \text{ has } S_i \text{ with } \mu_{PS}(P,S_i), \text{ and} \\ \text{If } S_i \text{ implies } S_j \text{ with } \mu_{SS^c}(S_i,S_j), \text{ and} \\ \text{If } S_j \text{ implies } D_i \text{ with } \mu_{SD^c}(S_j,D_i), \text{ and} \\ \text{If } D_i \text{ implies } D_j \text{ with } \mu_{DD^c}(D_i,D_j), \\ \hline \text{Then P has } D_j \text{ with } \mu_{PD}(P,D_j). \end{array} \qquad (7)$$

Out of efficiency reasons, CADIAG-2 processes a fixed number of inference steps: (1) symptom/symptom inference to complete and extend patient's initial symptom pattern; (2) symptom/disease inferences; (3) symptom combination/disease inferences with symptoms as evidence only, at this point a primary set of diagnoses is obtained; (4) disease/disease inferences to complete and extend the primary set of diagnoses; (5) again symptom combination/disease inferences, but not with symptoms and already inferred diseases as evidence; and finally (6) again disease/disease inferences to complete and extend the final set of diagnostic results.

In this process, confirmed diagnoses are obtained from given evidence (symptoms, symptom combinations, and diseases) and confirming relationships. Excluded diagnoses are determined from given evidence with excluding relationships or from definitely absent evidence whose occurrence was indicated as absolutely necessary. Diagnostic hypotheses are generated if uncertain evidence is combined with confirming relationships and/or certain or uncertain evidence is combined with supportive, but not confirming relationships. In case of diagnostic hypotheses, it is tested if

$$\varepsilon \leq \mu_{PD}(P,D_j) \leq 0.99, \qquad (8)$$

where ε is a threshold precluding diagnoses with too little evidence. Usually, ε is assigned a small value, say, between 0.1 and 0.4, to allow a very »liberal« hypothesis generation.

2.2.2. *Combination of evidence*

Because the values $\mu_{PD}(P,D_j)$ for diagnostic hypotheses, inferred with the compositional rule of inference (Eqn. 6), are both independent of the number of pieces of evidence supporting diagnosis D_j and their frequency of occurrence with the consequent, a heuristic function (Eqn. 9) was introduced considering the number of present and partly present symptoms with supportive relationships to diseases D_j and calculating support scores SS_{D_j} according to which all diagnostic hypotheses are ranked in descending order. These scores reveal very clearly how strongly the various hypotheses are supported. The formula with which the support scores are calculated is an unnormalized function and the individual score becomes the higher the more symptoms support the hypothesis. It considers the degrees of presence (compatibility values) of the symptoms and weights the associated frequency of occurrence and strength of confirmation according to given weights α and β. At present, the value α is 0.09 and β is 0.91, thus $\alpha + \beta$ yields 1.00. The weights cause that the strength of confirmation determines the support score ten times stronger than the frequency of occurrence. The formula is:

$$SS_{D_j} = 100 \sum_{i=1}^{m^*} \Big\{ \alpha \text{MIN}[\mu_{PS}(P,S_i); \mu_{SD^o}(S_i,D_j)] + \beta \quad \text{MIN}[\mu_{PS}(P,S_i); \mu_{SD^c}(S_i,D_j)] \Big\}, \qquad (9)$$

where m^* is the number of present or partly present symptoms occurring in the medical documentation of diagnosis D_j.

Table 3 shows the diagnostic results inferred from the patient data given in Table 2. The above-described fuzzy inference engine was applied to obtain these results. The diagnostic process was carried out in the differential diagnostic group of pancreatic diseases, containing at present *acute* and *chronic pancreatitis, pancreatic cancer, cystic pancreatic fibrosis, pancreatic pseudocyst, annular pancreas, insulinoma, glucogonoma, Zollinger-Ellison syndrome* and *Verner-Morrison syndrome.*

2.3. *Explanation system, examination proposal generation, and unexplained symptom generation*

As explanation, CADIAG-2 presents every item of evidence that contributed to the respective diagnostic result along with its frequency of occurrence and its strength of confirmation. Usually, these are many single symptoms and only few symptom combinations. Diagnostic hypotheses are always presented with their calculated degrees of compatibility to the patient and their achieved support scores (cf. Table 3).

A natural component of an iteratively working, data-driven medical expert system is the generation of further useful examinations at the end of each iteration cycle. CADIAG-2 advances examination proposals for each of the diagnostic

EDV in Medizin und Biologie
EDP in Medicine and Biology

INHALTSVERZEICHNIS/CONTENTS BAND/VOLUME XVI (1985)

Computeranwendungen/Computer applications

Statistische Verfahren/Statistical Methods

hypotheses. All symptoms or symptom combinations that will improve the fuzzy inference results or the support scores but were not investigated or determinable yet are offered. They are categorized according to their degree of risk for the patient and their costliness (cf. Table 4).

Unexplained disorders or positive findings are those symptoms in the patient's symptom pattern that cannot be accounted for by any of the inferred diagnostic results. They make it necessary to continue the diagnostic process and search for other diagnoses the patient may suffer from. These symptoms

Table 2. Patient data of a clinical case from the 2nd Department for Gastroenterology and Hepatology (Director: Prof. Dr. G. Grabner) including data from patient's history, present symptoms, signs, laboratory test results, and clinical findings. The data were evaluated by the patient data fuzzy interpreter, which assigns to every symptom, sign, test result, and finding a degree of compatibility $\mu_{PS}(P,S_i)$ of the measured or observed concrete value with the semantic concept under consideration.

Patient's medical data

μ_{PS}		μ_{SD^O}	μ_{SD^C}
	Anamnesis/known present disorder(s)		
1.00	age, 30–50 years		
1.00	sex, female		
1.00	known previous disorder, peptic ulcer		
1.00	known present disorder, gastric ulcer		
1.00	known present disorder, duodenal ulcer		
1.00	known present disorder, multiple gastro-intestinal (GI)-ulcerations		
	Present symptoms		
1.00	malaise		
1.00	decrease in physical/mental powers		
1.00	nausea		
1.00	vomiting		
1.00	anorexia		
1.00	weight loss		
1.00	pain, abdominal		
1.00	defecation, diarrhea		
	Physical examination		
1.00	general condition, ill		
1.00	nutritional status, normal		
1.00	strength, normal		
1.00	abdomen, tenderness, diffuse		
1.00	liver, palpation, susp. of liver metastasis		
	Laboratory findings		
1.00	BSR, increased		
1.00	serum, AP, elevated		
0.80	serum, potassium, reduced		
1.00	serum, SGOT, elevated		
1.00	serum, SGPT, elevated		
1.00	serum, LDH, elevated		
0.90	serum, γGT, elevated		
1.00	serum, cholesterol, elevated		
0.33	serum, triglycerides, elevated		
	Clinical investigations		
1.00	gastric analysis, BAO > 15 mval hydrochl. acid/hour		
1.00	gastric analysis, BAO/PAO > 0.6		
0.40	gastric analysis, volume > 200 ml/hour		
1.00	X-ray, hypotonic duodenography, susp. of inflammatory/ malignant pancreatic tumor		
1.00	X-ray, ERCP, susp. of pancreatic tumor		
1.00	X-ray, CT, susp. of pancreatic tumor		

μ_{PS}		μ_{SD^O}	μ_{SD^C}
	Physical examination		
1.00	general condition, reduced	0.85	0.01
1.00	nutritional state, normal	0.25	
1.00	strength, normal	0.30	
1.00	liver, palpation, susp. of liver metastasis	0.55	0.30
	Laboratory findings		
1.00	BSR, increased	0.75	0.01
1.00	serum, AP, elevated	0.72	
1.00	serum, SGOT, elevated	0.65	0.02
1.00	serum, SGPT, elevated	0.65	0.02
1.00	serum, LDH, elevated	0.46	0.02
0.90	serum, γGT, elevated	0.65	0.02
	Clinical investigations		
1.00	X-ray, hypotonic duodenography, susp. of inflammatory/malignant pancreatic tumor	0.38	0.05
1.00	X-ray, ERCP, susp. of pancreatic tumor	0.85	0.80
1.00	X-ray, CT, susp. of pancreatic tumor	0.85	0.85

μ_{PD} **0.50** **Verner-Morrison Syndrome (10 symptoms: 171.4 points)**

μ_{PS}		μ_{SD^O}	μ_{SD^C}
	Present symptoms		
1.00	malaise	0.72	0.01
1.00	weight loss	0.65	0.05
1.00	defecation, diarrhea	0.30	0.11
	Physical examination		
1.00	general condition, reduced	0.78	0.02
1.00	nutritional state, normal	0.40	
1.00	strength, normal	0.30	
1.00	liver, palpations, susp. of liver metastasis	0.30	0.08
	Laboratory findings		
0.80	serum, potassium, reduced	0.90	0.02
	Clinical investigations		
1.00	X-ray, ERCP, susp. of pancreatic tumor	0.85	0.50
1.00	X-ray, CT, susp. of pancreatic tumor	0.85	0.50

μ_{PD} **0.50** **Glucagonoma (5 symptoms: 134.4 points)**

μ_{PS}		μ_{SD^O}	μ_{SD^C}
	Present symptoms		
1.00	weight loss	0.80	0.07
	Physical examination		
1.00	liver, palpation, susp. of liver metastasis	0.30	0.08
	Laboratory findings		
1.00	serum, cholesterol, elevated	0.38	0.01
	Clinical investigations		
1.00	X-ray, ERCP, susp. of pancreatic tumor	0.85	0.50
1.00	X-ray, CT, susp. of pancreatic tumor	0.85	0.50

μ_{PD} **0.50** **Insulinoma (5 symptoms: 122.3 points)**

μ_{PS}		μ_{SD^O}	μ_{SD^C}
	Present symptoms		
1.00	nausea	0.20	0.02
1.00	vomiting	0.15	0.01
	Physical examination		
1.00	liver, palpation, susp. of liver metastasis	0.30	0.08
	Clinicqal investigations		
1.00	X-ray, ERCP, susp. of pancreatic tumor	0.85	0.50
1.00	X-ray, CT, susp. of pancreatic tumor	0.85	0.50

Table 3. This Table shows diagnostic results of a CADIAG-2 diagnostic process with the data given in Table 2. The achieved inference values μ_{PD} (P,D_j), which are greater or equal the threshold ε ($\varepsilon = 0.10$ for pancreatic diseases) are printed out on the left hand side of this disease name. On the right hand side, one can find both the number of symptoms supporting the hypothesis and the calculated support score SS_{D_j}. The diagnostic hypotheses are ranked according to the support scores in descending order. The symptoms, which led to the generation of the hypotheses, are printed out together with their degrees of compatibility μ_{PS} (P,S_i) with the patient and their frequencies of occurrence μ_{SD^O} (S_i,D_j) and strengths of confirmation μ_{SD^C} (S_i,D_j) to the hypotheses. An empty place means that there is no known or a totally unspecific relationship between the symptom and the disease. The diagnosis with the highest point number *(Zollinger-Ellison Syndrome)* was also the clinical diagnosis for this patient.

Diagnostic hypotheses with explanations

μ_{PD} **0.50** **Zollinger-Ellison Syndrome (15 symptoms: 325.7 points)**

μ_{PS}		μ_{SD^O}	μ_{SD^C}
	Anamnesis/known present disorder(s)		
1.00	age, 30–50 years	0.70	0.02
1.00	sex, female	0.40	
1.00	known present disorder, duodenal ulcer	0.79	0.30
1.00	known present disorder, multiple GI-ulcerations	0.10	0.35
	Present symptoms		
1.00	nausea	0.20	0.01
1.00	vomiting	0.26	0.01
1.00	weight loss	0.65	0.05
1.00	defecation, diarrhea	0.36	0.05
	Physical examination		
1.00	liver, palpation, susp. of liver metastasis	0.30	0.08
	Laboratory findings		
0.80	serum, potassium, reduced	0.36	0.03
	Clinical investigations		
1.00	gastric analysis, BAO > 15 mval hydrochl. acid/hour	0.80	0.30
1.00	gastric analysis, BAO/PAO > 0.6	0.80	0.30
0.40	gastric analysis, volume > 200 ml/hour	0.80	0.30
1.00	X-ray, ERCP, susp. of pancreatic tumor	0.85	0.50
1.00	X-ray, CT, susp. of pancreatic tumor	0.85	0.50

μ_{PD} **0.85** **Pancreatic cancer (22 symptoms: 317.4 points)**

μ_{PS}		μ_{SD^O}	μ_{SD^C}
	Anamnesis/known present disorder(s)		
1.00	sex, female	0.41	
	Present symptoms		
1.00	malaise	0.60	
1.00	decrease of physical/mental powers	0.50	
1.00	nausea	0.44	
1.00	vomiting	0.34	
1.00	anorexia	0.50	0.01
1.00	weight loss	0.82	0.05
1.00	pain, abdominal	0.86	0.06
1.00	defecation, diarrhea	0.23	0.01

Table 4. Because the diagnoses presented in Table 3, which are the result of the CADIAG-2 diagnostic process with the data given in Table 2, could not be confirmed yet, proposals for further examinations are presented. The provided list contains the achieved inference values μ_{PD} (P,D_j), the calculated support scores, according to which the hypotheses are ranked, a list of symptoms not yet examined along with their frequencies of occurrence μ_{SD^O} (S_i,D_j) and strengths of confirmation μ_{SD^C} (S_i,D_j) to the hypotheses. The proposed examinations are categorized into three classes: Class-1-symptoms are those that can be examined very easily and that are cheap; Class-2-symptoms include more specific investigations; and Class-3-symptoms are invasive and/or expensive procedures.

Diagnostic hypotheses with examination proposals

μ_{PD} **0.50** **Zollinger-Ellison Syndrome (325.7 points)**

	μ_{SD^O}	μ_{SD^C}
Class-1-symptom		
abdominal pain, peptic ulcer-like	0.86	0.10
Class-2-symptoms		
serum, gastrin, elevated	0.10	0.60
serum, gastrin, significantly elevated	0.90	0.90
secretin-provocative-test, abnormal finding	0.93	0.90

μ_{PD} **0.85** **Pancreatic cancer (317.4 points)**

	μ_{SD^O}	μ_{SD^C}
Class-3-symptoms		
percutaneous aspiration biopsy (US/CT-guidance), positive cytology	0.65	1.00

μ_{PD} **0.50** **Verner-Morrison Syndrome (171.4 points)**

	μ_{SD^O}	μ_{SD^C}
Class-2-symptoms		
serum, magnesium, reduced	0.60	0.02
serum, VIP, abnormal finding	0.88	0.90

μ_{PD} **0.50** **Glucagonoma (134.4 points)**

	μ_{SD^O}	μ_{SD^C}
Class-2-symptoms		
serum, glucagon, elevated	0.98	0.88
arginine-stimulation test, specific abnormal finding	0.98	0.90
glucose-suppression test, specific abnormal finding	0.92	0.90

μ_{PD} **0.50** **Insulinoma (122.3 points)**

	μ_{SD^O}	μ_{SD^C}
Class-2-symptoms		
serum, insulin, elevated	0.95	0.80
serum, proinsulin, elevated	0.95	0.80
serum, C-peptide, elevated	0.95	0.80

are determined by the CADIAG-2 system and presented to the consulting physician. The physician can then switch to other diagnostic areas and try to confirm of hypothesize further diseases (cf. Table 5).

3. Conclusion

The practicability of CADIAG-2's concept of knowledge representation and fuzzy inference was confirmed by applying the expert system to about 500 clinical cases. The comparison of the CADIAG-2 diagnoses with clinical and pathological diagnoses yielded an accuracy of up to 93 % [3, 4].

The present goal is to advance with the clinical trials at the test clinics and gain further experience regarding the behavior of the system and the acceptance by the physicians, nurses, and medical technicians. These trials are also supposed to possess an educational function at the Vienna General Hospital in such a manner that the prospective medical users learn to apply and utilize these tools.

Furthermore, strong efforts are being made to improve and extend the medical documentation and incorporate further medical specialties.

Acknowledgement

The author gratefully acknowledges the contributions of G. Kolarz, M.D., W. Scheithauer, M.D., P. Peichl, M.D., H. Effenberger, M.D., and G. Grabner, M.D. Without their work, this project would not have been possible. The author is especially thankful to W. Scheithauer, M. D., who documented the group of pancreatic diseases and helped to prepare the example given in the tables. The author is also indebted to W. Bogard, M.Sc., A. Hatvan, M.Sc., F. Lipomersky, and P. Sachs for programming various parts of CADIAG-2. Furthermore, the author would like to thank I. Gröger and G. Sedivy for their help in managing the knowledge acquisition and the patient data documentation.

Table 5. Two findings out of patient's data (Table 2) could not be accounted for by the obtained diagnostic results (Table 2). There are now two possibilities: Either the consulting physician decides that these findings remain unexplained or further diagnoses in this or in other diagnostic areas have to be searched for.

Unexplained symptoms	
μ_{PS}	
	Present symptoms
1.00	abdomen, tenderness, diffuse
	Laboratory findings
0.90	serum, triglycerides, elevated

Literature

1. GRABNER, H., A. MARKSTEINER, W. DORDA, W. WOLF, G. GRABNER (1981): A Medical Information System. Conception and Clinical Usage. J. Clin. Comp. **10,** 154–169.
2. GRABNER, G. (Ed.) (1985): WAMIS – Wiener Allgemeines Medizinisches Informations System. Springer-Verlag, Berlin.
3. ADLASSNIG, K.-P., G. KOLARZ, W. SCHEITHAUER (1985): Present State of the Medical Expert System CADIAG-2. Meth. Inform. Med. **24,** 13–20.
4. ADLASSNIG, K.-P., G. KOLARZ, W. SCHEITHAUER, H. EFFENBERGER, G. GRABNER (1985): CADIAG: Approaches to Computer-Assisted Medical Diagnosis. Comput. Biol. Med. **15,** 315–335.
5. SPINDELBERGER, W., G. GRABNER (1968): Ein Computerverfahren zur Diagnostischen Hilfestellung. In: Fellinger, K. (Ed.) Computer in der Medizin – Probleme, Erfahrungen, Projekte. Verlag Brüder Hollinek, Wien, 189–221.
6. ZADEH, L. A. (1965): Fuzzy Sets. Information and Control **8,** 338–353.
7. ZADEH, L. A. (1973): Outline of a New Approach to the Analysis of Complex Systems and Decision Processes. IEEE Transactions on Systems, Man, and Cybernetics **SMC-3,** 28–44.
8. WEISS, S. M., C. A. KULIKOWSKI, S. AMAREL, A. SAFIR (1978): A Model-Based Method for Computer-Aided Medical Decision-Making. Artificial Intelligence **11,** 145–172.
9. WEISS, S. M., C. A. KULIKOWSKI, A. SAFIR (1978): Glaucoma Consultation by Computer. Comput. Biol. Med. **8,** 25–40.
10. SHORTLIFFE, E. H. (1976): Computer-Based Medical Consultation: MYCIN. Elsevier, New York–Oxford–Amsterdam.
11. BUCHANAN, B. G., E. H. SHORTLIFFE (Eds.) (1984): Rule-Based Expert Systems – The MYCIN Experiments of the Stanford Heuristic Programming Project. Addison-Wesley Publishing Company, Reading/Massachusetts.
12. MILLER, R. A., H. E. POPLE, J. D. MYERS (1982): INTERNIST-I, An Experimental Computer-Based Diagnostic Consultant for General Internal Medicine. New England Journal of Medicine **307,** 468–476.
13. MILLER, R. A. (1984): INTERNIST-1/CADUCEUS: Problems Facing Expert Consultant Programs. Meth. Inform. Med. **23,** 9–14.
14. WEISS, S. M., C. A. KULIKOWSKI (1979): EXPERT: A System for Developing Consultation Models. Proc. IJCAI, Tokyo, 942–947.
15. WEISS, S. M., C. A. KULIKOWSKI (1984): A Practical Guide to Designing Expert Systems. Rowman & Allenheld Publishers, New Jersey.
16. SANCHEZ, E. (1979): Compositions of Fuzzy Relations. In: Gupta, M. M., R. K. Ragade, R. R. Yager (Eds.) Advances in Fuzzy Set Theory and Applications. North-Holland Publishing Company, Amsterdam, 421–433.
17. SANCHEZ, E. (1979): Medical Diagnosis and Composite Fuzzy Relations. In: Gupta, M. M., R. K. Ragade, R. R. Yager (Eds.) Advances in Fuzzy Set Theory and Applications. North-Holland Publishing Company, Amsterdam, 437–444.
18. ADLASSNIG, K.-P. (1982): A Survey on Medical Diagnosis and Fuzzy Subsets. In: Gupta, M. M., E. Sanchez (Eds.) Approximate Reasoning in Decision Analysis. North-Holland Publishing Company, Amsterdam, 203–217.
19. BELFORTE, G., B. BONA, C. CRAVETTO, S. FREDIANI, M. MILANESE, M. MOLINO, L. SAITTA, R. TEMPO (1985): Selection and Assessment of Laboratory Tests for the Evaluation of Liver Functional Impairment. Meth. Inform. Med. **24,** 39–45.
20. BORTOLAN, B., R. DEGANI (1983): Comparative Evaluation of Three Models of Fuzzy Decision-Making in Electrocardiography. In: Van Bemmel, J. H., M. J. Ball, O. Wigertz (Eds.). MEDINFO 83. North-Holland Publishing Company, Amsterdam, 537–540.
21. KERRE, E. E. (1982): The Use of Fuzzy Set Theory in Electrocardiological Diagnostics. In: Gupta, M. M., E. Sanchez (Eds.). Approximate Reasoning in Decision Analysis. North-Holland Publishing Company, Amsterdam, 277–282.
22. FIESCHI, M., M. JOUBERT, D. FIESCHI, G. BOTTI, M. ROUX (1983): A Program for Expert Diagnosis and Therapeutic Decision. Med. Inform. **8,** 127–135.
23. LESMO, L., M. MARZUOLI, G. MOLINO, P. TORASSO (1984): An Expert System for the Evaluation of Liver Functional Assessment. Journal of Medical Systems **8,** 87–101.
24. LESMO, L., L. SAITTA, P. TORASSO (1982): Learning of Fuzzy Production Rules for Medical Diagnosis. In: Gupta, M. M., E. Sanchez (Eds.). Approximate Reasoning in Decision Analysis. North-Holland Publishing Company, Amsterdam, 249–260.
25. MILANESE, M., B. BONA, S. FREDIANI, L. SAITTA, C. CRAVETTO, G. MOLINO (1984): A Sequential Protocol for the Optimization of Diagnostic Procedures in Hepatology. Journal of Medical Systems **8,** 73–85.

26. SAITTA, L., P. TORASSO (1981): Fuzzy Characterization of Coronary Disease. Fuzzy Sets and Systems **5**, 245–258.
27. SANCHEZ, E., J. GOUVERNET, R. BARTOLIN, L. VOVAN (1982): Linguistic Approach in Fuzzy Logic of the W.H.O. Classification of Dyslipoproteinemias. In: Yager, R. R. (Ed.) Fuzzy Set and Possibility Theory – Recent Developments. Pergamon Press, New York, 582–588.
28. SMETS, P. (1983): Fuzzy Set Theory for Medical Decision Making. In: Beneken, J. E. W., S. M. Lavelle (Eds.) Objective Medical Decision-Making; Systems Approach in Acute Disease. Springer-Verlag, Berlin, 7–19.
29. SOULA, G., E. SANCHEZ (1982): Soft Deduction Rules in Medical Diagnostic Processes. In: Gupta, M. M., E. Sanchez (Eds.). Approximate Reasoning in Decision Analysis. North-Holland Publishing Company, Amsterdam, 77–88.
30. SOULA, G., B. VIALETTES, J.-L. SANMARCO (1983): Protis; A Fuzzy Deduction-Rule System: Application to the Treatment of Diabetes. In: Van Bemmel, J. H., M. J. Ball, O. Wigertz (Eds.). MEDINFO 83. North-Holland Publishing Company, Amsterdam, 533–563.
31. RESCHER, N. (1969): Many-Valued Logic. McGraw Hill-Book Company, New York.
32. ADLASSNIG, K.-P., G. KOLARZ (1982): CADIAG-2: Computer-Assisted Medical Diagnosis Using Fuzzy Subsets. In: Gupta, M. M., E. Sanchez (Eds.). Approximate Reasoning in Decision Analysis. North-Holland Publishing Company, Amsterdam, 219–247.
33. ZADEH, L. A. (1981): Test Score Semantics for Natural Languages and Meaning Representation via PRUF. Technical Note 247, AI Center, SRI International, Menlo Park CA. Also in: Rieger, B. B. (ed.). Empirical Semantics. Brockmayer, Bochum, 281–349.

Date of receipt: February 3rd, 1986.
The author's address: Dr. med. Klaus-Peter Adlassnig, Department of Medical Computer Sciences, University of Vienna, Garnisongasse 13, A-1090 Vienna, Austria.

EDV in Medizin und Biologie **17** (1/2), 20–27, ISSN 0300-8282

On the use of growth and decay functions for modelling stem profiles

C. Brink and K. von Gadow

Summary

One of the most useful tools for modelling the effects of environmental and stand treatment factors on stem form is a simple taper equation. This article presents four new taper functions for modelling stem profiles. These are not derived de novo, but from known growth and decay functions. The paper demonstrates that it is possible to modify any growth or decay function with certain structural properties to serve as a taper function.

Zusammenfassung

Wenn man die Auswirkungen von Umgebungs- und Behandlungsfaktoren auf die Form von Baumschäften untersuchen will, empfiehlt sich die Anwendung einer flexiblen Schaftgleichung mit begrenzter Parameterzahl. Eine Spline Approximation ist für diesen Zweck unbrauchbar. In diesem Beitrag werden vier neue Schaftgleichungen vorgestellt. Die Gleichungen werden nicht de novo, sondern von bekannten Wachstums- und Zerfallfunktionen abgeleitet. Es wird gezeigt, daß es möglich ist, jede Wachstums- oder Zerfallfunktion mit bestimmten strukturellen Eigenschaften so zu modifizieren, daß sie als Schaftgleichung verwendet werden kann.

1. Introduction

Foresters need to be able to estimate the stem form of trees and how it is affected by environment and stand treatment. One method of modelling the stem form of a tree involves the use of a taper, or stem profile equation which expresses radius r as a function of height h.

One of the first attempts to model the stem profile is BEHRE's (1923) hyperbola (see PRODAN (1965), p. 62):

$$q = \frac{x}{a+bx} \qquad (1.1)$$

where

x = relative tree height;
q = relative tree diameter;
a,b = parameters to be estimated.

More recent attempts to model stem profiles include ORMEROD's (1973) equation

$$d = D\left[\frac{H-h}{H-k}\right]^p \qquad (1.2)$$

where

d = diameter (cm) at tree height h (m);
D = diameter (cm) at breast height k (m);
H = total tree height (m) with $0 < h < H$;
p = parameter to be estimated ($p > 0$);
k = 1.35 (breast height).

Because of its simplicity, ORMEROD's equation appears to be rather popular (REED and BYRNE (1985)). Other noteworthy models of stem taper are, for example the dual equation system proposed by DEMAERSCHALK and KOZAK (1977), a system using 4th degree polynomials (MADSEN (1982)) and a recent approach involving the Chapman-Richards function (BINGING (1984)).

We find that a taper equation is a useful tool for modelling stem form. On the one hand, it is flexible enough to provide good estimate of radius at different heights, on the other hand it is simple enough to allow parameter smoothing, consequently permitting better insight into the relationship between stem form, tree dimensions and stand treatment.

In this article we present four new taper functions. Whereas the general tendency seems to be to derive such functions *de novo,* we take the methodologically simpler approach of modifying known growth and decay functions to serve as taper functions. The motivation for this approach is best explained by reference to an idealized stem form, as pictured in Fig. 1.1, which we call the *target shape.* In Fig. 1.1, h indicates tree height and r tree radius; b is breast height (1.35 m) and r_b is radius at breast height; H is the total tree height and r_0 the radius at the base of the stem (the initial radius). So if we view r as a function of h we have $r(0) = r_0$, $r(b) = r_b$ and $r(H) = 0$. The target shape is monotone decreasing, with the curvature at first anti-clockwise, then zero at a point of inflection, and then clockwise. In Fig. 1.1 the point of inflection occurs at the point (I,i), where height is I and the corresponding radius r(I) is i. We call that part of the target shape where $0 \leq h \leq I$ the *initial part,* and the part where $I \leq h \leq H$ the *terminal part* of the target shape. The point to note is that *change of curvature*

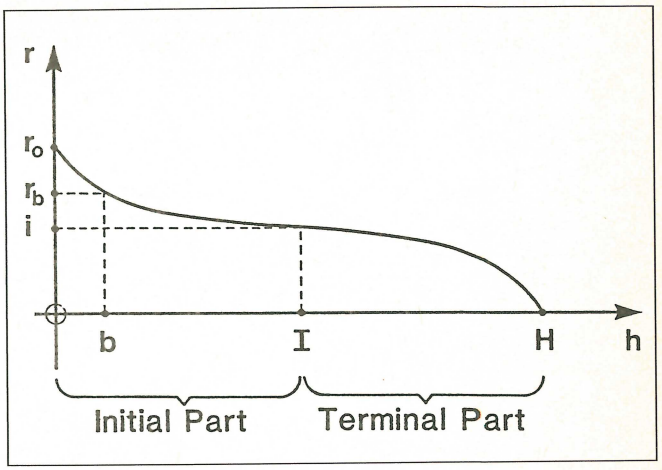


Fig. 1.1. Idealized stem profile or »target shape«.

takes place (mathematically: the second derivative of r changes sign). Whenever this happens, the function in question exhibits an S-shape around the point of inflection, no matter what the orientation of the function is with regard to the axes. In Figure 1.1 the S-shape is best seen by rotating the figure anti-clockwise through ninety degrees. It is well known that an S-shape may be modelled by a variety of functions, usually viewed as growth functions. In principle, then, any one of these functions can also be used as a taper function modelling stem forms. In practice, of course, the choice of such functions depends on considerations such as possible physical interpretation of parameters and techniques available for parameter estimation.

In this article we modify in §3 the logistic function and in §4 the Weibull growth function for use as taper functions. Beforehand, in §2, we obtain two taper functions by modifying a decay function.

We test each of the four taper functions presented here against ten *Eucalyptus cloeziana* trees, the measurements of which are listed in Table 1.1. The results are listed in tables, one for each model. In line with the minimum data usually available for a particular tree, we list in these tables *diameter at breast height over bark* (DBH O-B) and *total height* (H in Figure 1.1). However, for the testing of the models we do not

Table 1.1. Heights (h) and radii (r, under bark) of 10 *Eucalyptus cloeziana* trees.

Tree number																			
116		121		127		128		138		139		142		147		149		152	
h m	r cm	h m	r cm	h m	r cm	h m	r cm	h m	r cm	h m	r cm	h m	r cm	h m	r cm	h m	r cm	h m	r cm
0	9.78	0	11.43	0	8.38	0	7.75	0	8.51	0	7.37	0	9.52	0	6.86	0	9.91	0	9.78
0.6	8.89	0.6	10.29	0.6	7.37	0.6	6.73	0.6	7.62	0.6	6.73	0.6	8.89	0.6	6.35	0.6	8.76	0.6	9.02
1.2	8.38	1.2	9.40	1.2	6.86	1.2	6.10	1.2	6.60	1.2	6.10	1.2	8.13	1.2	5.46	1.2	8.26	1.2	8.51
1.35	8.36	1.35	9.34	1.35	6.81	1.35	6.02	1.35	6.14	1.35	6.06	1.35	7.92	1.35	5.26	1.35	8.21	1.35	8.45
2.4	8.00	2.4	8.89	2.4	6.73	2.4	5.97	2.4	5.59	2.4	5.84	2.4	7.62	2.4	5.08	2.4	7.75	2.4	8.13
4.9	7.37	4.9	7.87	4.9	6.22	4.9	5.08	4.9	5.08	4.9	5.08	4.9	6.86	4.9	4.83	4.9	7.37	4.9	7.24
7.3	6.60	7.3	7.24	7.3	5.59	7.3	4.32	7.3	4.32	7.3	4.57	7.3	6.22	7.3	3.94	7.3	6.35	7.3	6.60
9.8	5.97	9.8	6.35	9.8	4.95	9.8	3.56	9.8	3.68	9.8	3.81	9.8	5.46	9.8	3.18	9.8	5.72	9.8	5.97
12.2	5.08	12.2	5.59	12.2	4.19	12.2	2.92	12.2	2.92	12.2	3.05	12.2	4.57	12.2	1.52	12.2	4.95	12.2	5.08
14.6	4.32	14.6	4.44	14.6	3.30	15.2	1.40	14.9	1.40	15.2	1.52	14.6	3.81	13.1	0	14.6	3.94	14.6	4.19
17.1	3.30	17.1	3.43	17.1	1.65	18.6	0	18.0	0	18.3	0	17.1	1.90			17.7	1.90	17.1	2.79
20.1	1.65	19.2	1.65	19.8	0							19.8	0			21.0	0	18.9	1.40
21.6	0	21.3	0															20.7	0

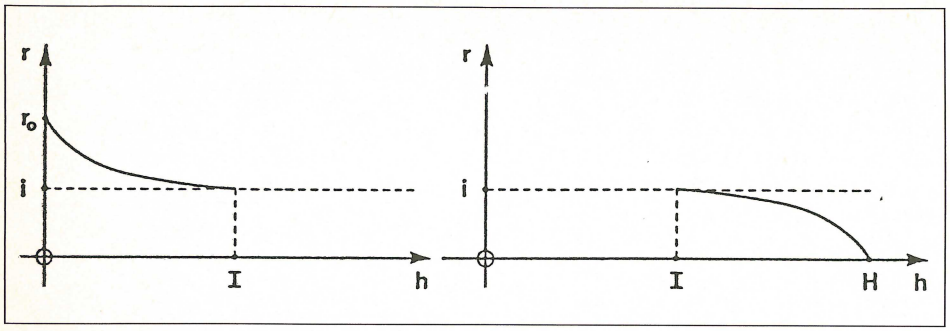


Fig. 2.1. Initial part and terminal part of target shape viewed separately. Both may be regarded as decay functions.

use DBH O-B but r_b – *radius at breast height under bark*. In fact, all the radii listed in Table 1.1 are taken under bark. Moreover, although nothing in the mathematical development depends on this, we try where possible to arrange matters in such a way that the taper function passes precisely through the points (b, r_b) and (H,0).

2. Application of the classic decay function

We start out by noting that if we view the initial part and the terminal part of the target shape separately, as in Figure 2.1, both of them may be considered as decay functions. Standard models of decay are therefore applicable in both cases. We choose to use what we call the *classic decay function,* namely the solution of the differential equation

$$\frac{dy}{dx} = k\,(y - B), \tag{2.1}$$

which says that the rate of change of y as a function of x is directly proportional to the difference between y and some constant (upper or lower) limit B. (Example: Newton's law of cooling says that the rate of cooling of a warm body is directly proportional to the difference between its own temperature and the (constant and lower) temperature of its environment.)

Using (2.1), we model the initial shape by a function α of h, using i as a lower limit and r_0 as an initial value. So (2.1) becomes

$$\frac{d\alpha}{dh} = -p\,(\alpha - i) \quad (p > 0) \tag{2.2}$$

and standard methods yield the solution

$$\alpha\,(h) = i + (r_o - i)\,e^{-ph}. \tag{2.3}$$

Similarly we model the terminal shape by a function β of h, using i as an upper limit and H as an initial *(sic!)* value. In this case the differential equation

$$\frac{d\beta}{dh} = -q\,(i - \beta) \quad (q > 0) \tag{2.4}$$

yields the solution

$$\beta\,(h) = i - ie^{q(h-H)}. \tag{2.5}$$

The models α and β are illustrated in Figure 2.2. Note that neither passes precisely through the point of inflection; this is of course a consequence of using i as a limiting value.

Note further that in each case, the proportionality constant is easily determined from one further data point. Thus if (X, x) is any data point in the initial part, and (Y, y) any data point in the terminal part, then (2.3) and (2.5) respectively show that

$$p = \frac{1}{X} \ln\left(\frac{r_0 - i}{x - i}\right) \tag{2.6}$$

and

$$q = \frac{1}{(Y - H)} \ln\left(\frac{i - y}{i}\right) \tag{2.7}$$

which can be used as initial values when estimating the parameters.

A first model of the target shape may now be constructed by the simple expedient of subtracting from α the difference between i and β. Call the resulting function r_1, then

$$r_1(h) = i + (r_0 - i)e^{-pH} - ie^{q(h-H)}. \tag{2.8}$$

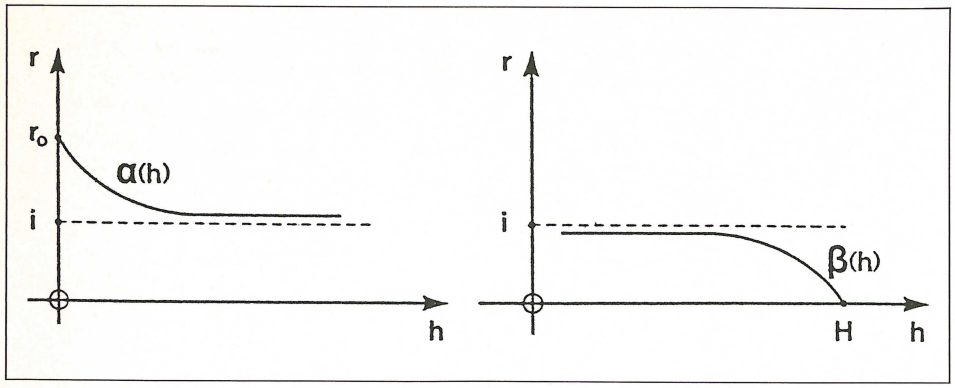


Fig. 2.2. Models for the initial part and terminal part, respectively, of the target shape. Both are obtained from the classic decay function.

Thus r_1 lies slightly below α in the initial part and slightly above β in the terminal part. In this function r_0 and H are empirically given, and the parameters i, q and p are calculated using standard nonlinear regression methods. The function is therefore fully determined by the data, is theoretically quite simple, and, as is shown in Table 2.1, is fairly accurate.

A second model may be constructed from the classic decay function by using in (2.2) a variable lower limit in place of the constant lower limit i; the idea being to bend down the graph of α in Figure 2.2 so as to cross the horizontal axis and thus model also the terminal shape. And in fact such a variable lower limit is at hand – we simply use β. Rechristening α as r_2, we get in this way the differential equation

$$\frac{dr_2}{dh} = -p\,(r_2 - \beta),$$

which may be rewritten as

$$\frac{dr_2}{dh} + pr_2 = p\beta \quad (p > 0) \tag{2.9}$$

to show that it is first-order linear. Applying the standard method of solution for this type of differential equation, and using the initial value r_0, we obtain

$$r_2(h) = i - \frac{pi}{p+q}\, e^{q(h-H)} + \left[r_0 - i + \frac{pi}{p+q}\, e^{-qH}\right] e^{-ph} \tag{2.10}$$

this being our second model. Again r_0 and H are given and the parameters are determined as in the first model. On the assumption that the initial part of r_2 closely follows the shape of α, and the terminal part the shape of β, the initial values for p and q found in (2.6) and (2.7) may be used as initial values for parameter estimation of r_2 as well.

To compare the two models r_1 and r_2, rewrite r_2 as

$$r_2(h) = i + (r_0 - i)\, e^{-ph} - \frac{pi}{p+q}\left[e^{q(h-H)} - e^{-ph-qH}\right], \tag{2.11}$$

then it is clear that, like r_1, r_2 is obtained by subtracting an auxiliary function from the classic decay function α of (2.3). Given that p, q, h and H are all positive, we observe that

$$\frac{pi}{p+q} < i$$

and

$$e^{q(h-H)} - e^{-ph-qH} < e^{q(h-H)}.$$

Consequently, $\frac{pi}{p+q}\left[e^{q(h-H)} - e^{-ph-pH}\right] < ie^{q(h-H)}$

and so, these being precisely the auxiliary functions subtracted from α in (2.11) and (2.8) to obtain r_2 and r_1 respectively, we conclude that r_2 lies between α and r_1. If, therefore, we take α and β as describing precisely the initial shape and the terminal shape respectively, then of r_1 and r_2, the latter will be the better model in the initial part, the former in the terminal part. (Since $\beta < r_1 < \alpha$, with $\alpha - r_1$ small in the initial part, and $r_1 - \beta$ small in the terminal part). But of course α and β are just models themselves, and so the relative accuracy of r_1 and r_2 can in practice not be settled in this way.

Table 2.1. Fitting equation (2.8) to the 10 *Eucalyptus cloeziana* trees in Table 1.1.

Tree No.	DBH O.B	Total Height	i	p	q	Error Mean Square
116	18.0	21.64	4.856	0.0720	0.1010	0.068
121	20.6	21.34	7.864	0.3109	0.1317	0.069
127	15.5	19.81	6.658	0.4059	0.1330	0.070
128	13.2	18.59	4.787	0.2376	0.1297	0.081
138	13.7	17.98	4.910	0.3976	0.1558	0.092
139	13.2	18.29	4.837	0.2292	0.1503	0.047
142	17.5	19.81	6.542	0.2507	0.1540	0.049
147	11.9	13.10	4.442	0.2561	0.1620	0.107
149	18.0	21.03	7.216	0.2446	0.1208	0.113
152	18.0	20.73	6.942	0.2170	0.1438	0.029

Table 2.2 Fitting equation (2.14) to the 10 *Eucalyptus cloeziana* trees listed in Table 1.1.

Tree No	DHB O.B.	Total Height	i	p	q	Error Mean Square
116	18.0	21.64	10.21	0.6894	0.0650	0.027
121	20.6	21.34	13.77	1.4654	0.0518	0.056
127	15.5	19.81	9.00	3.1447	0.0773	0.023
128	13.2	18.59	13.72	2.4542	0.0319	0.050
138	13.7	17.98	9.16	1.1155	0.0541	0.135
139	13.2	18.29	10.09	2.2126	0.0520	0.028
142	17.5	19.81	10.77	1.7110	0.0670	0.072
147	11.9	13.10	7.64	1.7711	0.0705	0.135
149	18.0	21.03	12.72	2.8174	0.0514	0.039
152	18.0	20.73	12.11	2.1786	0.0595	0.042

In some cases it may be desired of a function modelling stem profiles to pass *precisely* through selected points – e.g. radius at breast height. For the classic decay function α this is easily effected by using $r(b) = r_b$ as an initial value in (2.2) (instead of $r(0) = r_0$) to obtain

$$\alpha(h) = i + (r_b - i)\, e^{p(b-h)}). \tag{2.12}$$

For our first model r_1, this easy option is not available, since r_1 does not appear here as the solution of a differential equation. But r_2 does, so we may use $r\,(b) = r_b$ as an initial value in (2.9) to obtain

$$r_2(h) = i - \frac{pi}{p+q}\, e^{q(h-H)} + \left[r_b - i + \frac{pi}{p+q}\, e^{q(b-H)}\right] e^{p(b-h)} \tag{2.13}$$

instead of (2.10). Again, for purpose of comparison this may be rewritten as

$$r_2(h) = i + (r_b - i)\, e^{p(b-h)} - \frac{pi}{p+q}\left[e^{q(h-H)} - e^{q(b-H)+p(b-h)}\right] \tag{2.14}$$

where r_2 then again appears as the result of subtracting an auxiliary function from the classic decay function – (2.12) in this case. For comparison Table 2.2 does for r_2 what Table 2.1 did for r_1.

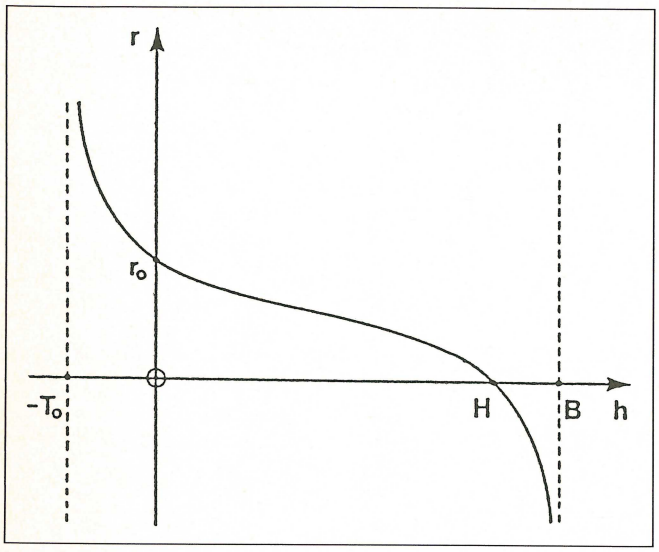


Fig. 3.1. The target shape extended into the second and fourth quadrants. It is assumed to be delimited by two vertical asymptotes.

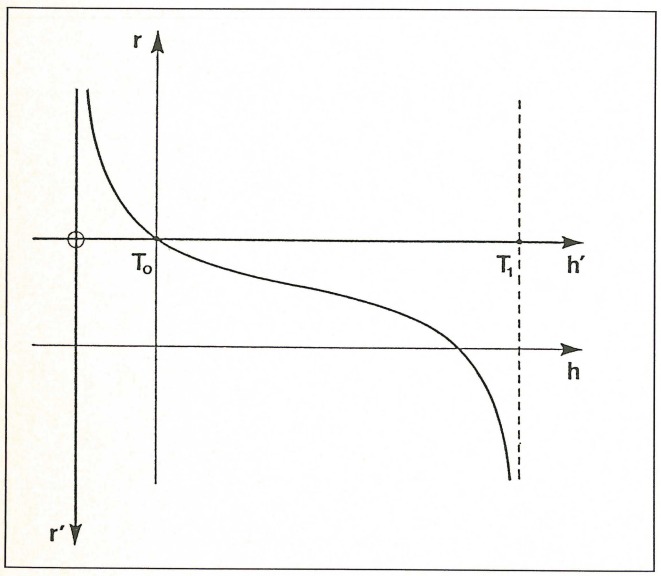


Fig. 3.2. The same graph as in Figure 3.1, but showing the new axes h′ and r′.

Table 3.1. Fitting equation (3.7) to the 10 *Eucalyptus cloeziana* trees listed in Table 1.1.

Tree No	DBH O.B.	Total Height	k	T_0	B	Error Mean Square
116	18.0	21.64	0.0086	11.779	27.54	0.118
121	20.6	21.34	0.0283	0.812	21.34	0.144
127	15.5	19.81	0.0090	20.992	22.27	0.186
128	13.2	18.59	0.0468	0.424	19.31	0.149
138	13.7	17.98	0.0518	0.174	17.98	0.169
139	13.2	18.29	0.0434	1.021	18.99	0.103
142	17.5	19.81	0.0341	1.046	20.30	0.199
147	11.9	13.10	0.0677	0.288	16.81	0.246
149	18.0	21.03	0.0157	5.962	22.86	0.228
152	18.0	20.73	0.0145	7.452	22.39	0.103

3. Application of the logistic function

Recall that both models of §1 were constructed by using a horizontal asymptote. This asymptote (the line r = i) was a purely theoretical entity which played no part in the eventual models, except insofar as it influenced the derivation of the equations constituting these models. In the same spirit we now construct a third model by using two vertical asymptotes. We assume, namely, that if the target shape extended also into the second and fourth quadrants it would naturally assume a shape as in Figure 3.1, with the lines $h = -T_0$ and $h = B$ acting as vertical asymptotes. In order to model this extended target shape, we first effect a translation and re-orientation of axes by the equations

$$h' = h + T_0$$
$$r' = r_0 - r, \qquad (3.1)$$

so that the target shape appears with respect to the new axes as in Figure 3.2; with $T_1 = B + T_0$. We then assume that

(a) the rate of change of r′ is inversely proportional to h′; and
(b) the rate of change of r′ is inversely proportional to $T_1 - h'$.

Each of these assumptions is modelled by a differential equation. Combining these in the standard way, and presenting the constant of proportionality as a reciprocal, we get the single differential equation

$$\frac{dr'}{dh'} = \frac{1}{kh'(T_1 - h')} \quad (k > 0). \qquad (3.2)$$

Using the initial value $h'(0) = T_0$ we obtain by standard methods the solution

$$r' = \frac{1}{kT_1} \ln\left[\frac{h'B}{T_0 (T_1 - h')}\right],$$

which, translated back to the original axes, becomes

$$r(h) = r_0 - \frac{1}{kT_1} \ln\left[\frac{(h + T_0) B}{(B - h) T_0}\right]. \qquad (3.3)$$

The constant of proportionality is easily determined from one further data point. E.g. from r(H) = 0 we obtain

$$k = \frac{1}{r_0 T_1} \ln\left[\frac{(H + T_0) B}{(B - H) T_0}\right] \qquad (3.4)$$

which can be used as initial value in the parameter estimation procedure.

As in the first model r_0 and H are empirically given. We do not at this stage attempt to offer an interpretation of the parameters T_0 and B, except to mention that B may be related to the concept of maximum possible height under optimum conditions.

The perceptive reader may have noticed that the present model is in fact a variation of the logistic model of a growth curve. This fact may be illustrated by rotating Figure 3.2 anti-clockwise through ninety degrees, in which case h′ appears as a function of r′ in the familiar S-shape of the logistic function. More precisely, the relationship is exhibited by the fact that

$$\frac{dh'}{dr'} = kh'(T_1 - h'), (k > 0) \qquad (3.5)$$

which follows directly from (3.2) by taking reciprocals, and is of course the differential equation of which the logistic curve is a general solution. In fact, solving (3.5) with initial valual $h'(0) = T_0$ yields the logistic function

$$h' = \frac{T_0 T_1}{(T_1 - T_0)\, e^{-T_1 k r'} + T_0}. \tag{3.6}$$

And the relationship between this function and our model (3.3) of a taper function is that each may be obtained from the other by using the translation equations (3.1).

Should it be required that the function (3.3) pass precisely through a selected point, this may easily be effected by using that point as an initial value in place of $h'(0) = T_0$. For example, if (3.3) is required to pass through (b, r_b), we obtain from (3.1) the information that $h'(r_0 - r_b) = b + T_0$, and using this as an initial value in solving (3.5) we get

$$r(h) = r_b - \frac{1}{kT_1} \ln \left[\frac{(h + T_0)(B - b)}{(B - h)(T_0 + b)} \right]. \tag{3.7}$$

Corresponding to (3.4) we obtain in this case

$$k = \frac{1}{r_b T_1} \ln \left[\frac{(H + T_0)(B - b)}{(B - H)(T_0 + b)} \right]. \tag{3.8}$$

An assessment of this model appears in Table 3.1.

4. Application of the Weibull function

The mathematical gymnastics of 3 (Application of the logistic function) notwithstanding, the model constructed there may be regarded as being obtained from the logistic function by a translation and re-orientation of axes. In effect, then, we obtained a taper function by the simple expedient of rotating a growth function. This is a general method which may be applied to any growth function, and with that a whole class of taper functions became available. For example, the known growth functions of WEIBULL, MITSCHERLICH, VON BERTALANFFY and GOMPERTZ all yield corresponding taper functions. In each case, the taper function obtained in this way, inherits all the structural characteristics of the original growth function.

Thus, for example, the taper function of 3 is symmetric around its point of inflection, since this is a property of the logistic function. This taper function is therefore not a satisfactory model of stem forms in which the curvature of the initial part differs from that of the terminal part. For such stem forms we obtain in this section a taper function from the Weibull growth function.

We use here the Weibull growth function as characterized by YANG, KOZAK and SMITH (1978) and illustrated in Figure 4.1:

$$h' = B - Be^{-(r'/\alpha)^{\beta}}, \tag{4.1}$$

where h' is a function of r', B is an upper limit for h', $\alpha > 0$ is a scale parameter and $\beta > 0$ is a shape parameter.

Adapting now to our present purposes the equations (3.1) we effect a clockwise rotation through ninety degrees by

$$\begin{aligned} h' &= h \\ r' &= r_0 - r, \end{aligned} \tag{4.2}$$

so that the curve of Figure 4.1 re-appears as in Figure 4.2, where we restrict it to the first quadrant. Substituting (4.2) into (4.1) leads to

$$\beta \ln (r_0 - r) - \beta \ln \alpha = \ln \left[\ln \left(\frac{B}{B - h} \right) \right], \tag{4.3}$$

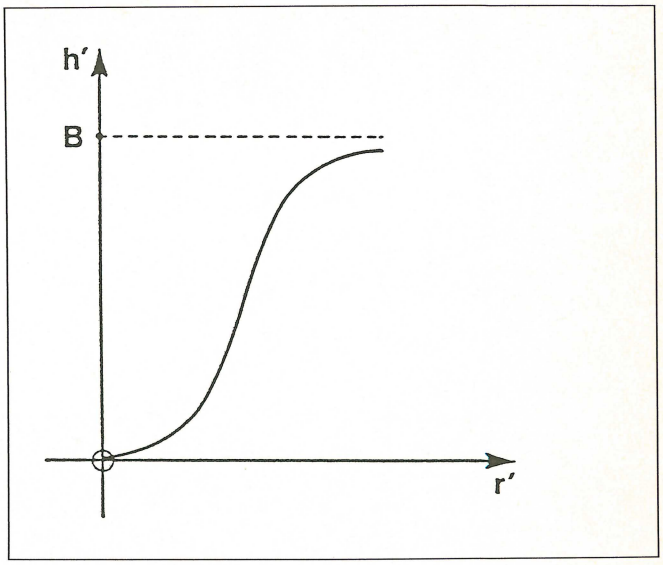


Fig. 4.1. Characteristic S-shaped growth function, here assumed to be modelled by the Weibull function (4.1).

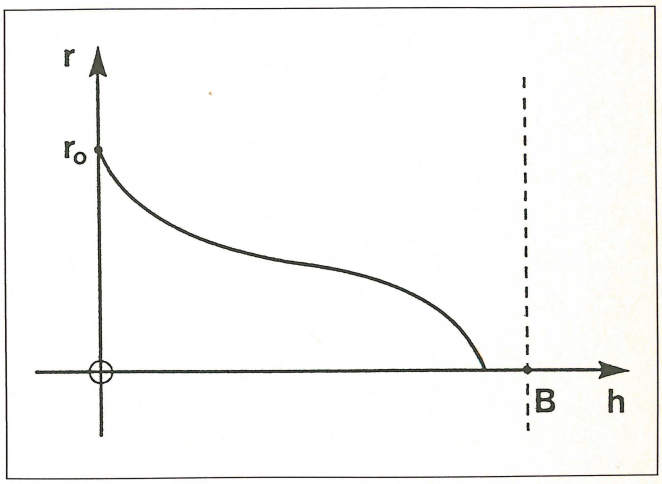


Fig. 4.2. The same graph as in Figure 4.1, but rotated clockwise, showing new axes and restricted to the first quadrant.

Table 4.1. Fitting equation (4.8) to the 10 *Eucalyptus cloeziana* trees listed in Table 1.1.

Tree No.	DBH O.B	Total Height	α	β	B	Mean Square Error
116	18.0	21.64	136.290	0.669	173.628	0.021
121	20.6	21.34	13.374	0.789	43.799	0.076
127	15.5	19.81	12.400	0.633	42.346	0.017
128	13.2	18.59	297.614	0.560	185.194	0.089
138	13.7	17.98	8.000	0.774	33.671	0.031
139	13.2	18.29	164.729	0.591	149.428	0.022
142	17.5	19.81	7.500	0.758	32.667	0.041
147	11.9	13.10	283.700	0.452	113.842	0.110
149	18.0	21.03	37.305	0.649	71.522	0.043
152	18.0	20.73	6.656	0.820	31.363	0.030

then solving for r we obtain

$$r(h) = r_0 - \alpha \exp.\left\{\frac{1}{\beta} \ln \left[\ln\left(\frac{B}{B-h}\right)\right]\right\}, \tag{4.4}$$

which is our fourth model.

(We briefly point out a mathematical subtlety in (4.4). The function r is in fact undefined at h = 0, nevertheless its graph will appear to originate at the point $(0,r_0)$. This is because r has the *limit-value* r_0 as h tends to 0. Formally:

$$\lim_{h \to 0} r(h) = r_0$$

We take this to be sufficient for the purpose of modelling stem forms. At any rate, (4.6) below could be used in place of (4.4), in which case r *is* defined at h = 0, and $r(0) = r_0$ as desired).

In this model, as in that of § 3, we suggest that B may be related to maximum possible height. With B and r_0 taken as known, the other parameters may be estimated by the same methods as is used for the Weibull growth curve. Equation (4.3), in fact, yields such a method. Let

$$\begin{aligned} Y &= \ln\left[\ln\left(\frac{B}{B-h}\right)\right] \\ X &= \ln(r_0 - r) \\ M &= \beta \\ C &= -\beta \ln \alpha \end{aligned} \tag{4.5}$$

then (4.3) appears in the form Y = MX + C. Use data points to plot X against Y, then fit a straight line through these points (e.g. by using least squares). This yields M and C, and then from (4.5) we obtain

$$\beta = M \text{ and } \alpha = e^{-C/\beta}.$$

It is possible to force the function (4.4) to pass precisely through a selected point. For example, if (4.4) is required to pass through (b, r_b), this may be effected by forcing the corresponding straight line Y = MX + C to pass through the point $\left(\ln(r_0 - r_b), \ln\left[\ln\left(\frac{B}{B-b}\right)\right]\right)$. Such a restriction on the straight line fitted through the (X, Y) points may of course have the effect that the overall fit of (4.4) to the original data is not the best possible with this method. This would be the price to pay for forcing r(h) through a selected point.

Another way of forcing (4.4) through, say, (b, r_b) would be to interpret the origin in Figure 4.1 not as the point $(0, r_0)$ but as the point (b, r_b). Instead of the equations (4.2) we then have

$$\begin{aligned} h' &= h - b \\ r' &= r_b - r \end{aligned} \tag{4.6}$$

which leads in the same way as before to

$$\beta \ln(r_b - r) - \beta \ln \alpha = \ln\left[\ln\left(\frac{B}{B-(h-b)}\right)\right] \tag{4.7}$$

and then to

$$r(h) = r_b - \alpha \exp\left\{\frac{1}{\beta} \ln\left[\ln\left(\frac{B}{B-(h-b)}\right)\right]\right\} \tag{4.8}$$

The equations corresponding to (4.5) are then

$$\begin{aligned} Y &= \ln\left[\ln\left(\frac{B}{B-(h-b)}\right)\right] \\ X &= \ln(r_b - r) \\ M &= \beta \\ C &= -\beta \ln \alpha \end{aligned} \tag{4.9}$$

An assessment of this model appears in Table 4.1. Note that (4.8) is undefined for $0 \leq h < b$, so that this model in effect ignores the stem profile beneath breast height. Measurements below breast height were therefore disregarded when calculating the mean square errors listed in Table 4.1. (Note further that in fact (4.8) is also undefined for h = b, but that the situation is entirely analogous to that which obtains for h = 0 in (4.4), and that the same comments concerning limits therefore apply.)

5. Conclusion

We briefly summarize here the main features of each of the four models presented in this article.

The first model, given by equation (2.8) is obtained from two classic decay functions by subtraction – simple enough, but not very elegant. The function has three parameters, two of which are shape parameters for the initial part and the terminal part respectively, and the third is intended to approximate the radius at the point of inflection. Given extra data points in the initial part and the terminal part, respectively, there are simple formulae giving initial values for estimating the shape parameters. The function does not pass precisely through total height (H;0), nor through radius at breast height $(b;r_b)$, and it cannot be forced to do so. Because of its simplicity, the square of this function is analytically integrable. If, therefore, we view the bole of a tree modelled by this function r_1 as being represented by the volume of revolution obtained by revolving r_1 around the horizontal axis, then

$$\text{volume} = \pi \int_0^H [r_1(h)]^2 \, dh. \tag{5.1}$$

Thus, given the integral of $r_1(h)^2$, finding volume is a simple calculation.

The second model, given by equation (2.14), arises as the solution of a differential equation. Like the first model it has three parameters, two for shape and one for size. Again the shape parameters are easily initialized given extra data points. In this case, too, the square of the function is analytically integrable. In addition, and here the second model has the advantage, the function passes precisely through the standard measurement point $(b;r_b)$, though not precisely through (H;0).

The third model, given by equation (3.7), may be viewed either as arising from a differential equation (which in turn arises from seemingly reasonable assumptions), or as arising directly from the logistic function by a translation and reorientation of axes. The function has three parameters, one for shape and two for size. One of the size parameters does not have an obvious physical interpretation. The shape parameter is easily initialized from an extra data point. The function passes precisely through $(b;r_b)$. If (3.8) is accepted as giving the precise value of k, then the function passes precisely through (H,0) as well.

The fourth model, given by equation (4.8), is obtained directly from the Weibull growth function by translation and re-orientation. It has three parameters, one for shape, one for scale and one for size. It does not arise from a differential equation, and there is no simple initialization procedure from a single data point. The function passes precisely through $(b;r_b)$, but this involves ignoring the stem profile beneath breast height. The function does not pass precisely through (H,0).

For a simultaneous comparison of all four models we present in Table 5.1 a summary of the four previous tables (MSE = mean square error). For this purpose we count a function as providing a »good fit« of a particular tree if it has MSE less than 0.1. On the basis of this table it would seem that the second and the fourth models are the most successful. But the sample of the trees used here is too small to give a definite verdict.

The main point of this article has been to show that taper functions modelling stem profiles may be obtained by using known growth and decay functions. We see in this a methodological simplification of the search for taper functions. There may not be a single »correct« or even »best« taper function – what is best may vary from one context to another. In our approach one may choose any growth or decay function having the structural properties required, and modify it to serve as a taper function.

We conclude with the remark that using growth or decay functions for obtaining taper functions may also lead to further taper functions. To illustrate this point, note that there is a structural similarity between the last two models (both obtained from growth functions). The third model, according now to equation (3.3), may be rewritten as

$$r(h) = r_0 - \frac{1}{KT_1}\left[\ln\left(\frac{B}{B-h}\right) + \ln\left(\frac{h+T_0}{T_0}\right)\right]. \qquad (5.2)$$

And the fourth model, according to equation (4.4), may be rewritten in the form

$$r(h) = r_0 - \alpha\left[\ln\left(\frac{B}{B-h}\right)\right]^{1/\beta}. \qquad (5.3)$$

Comparison of these two equations, (5.2) and (5.3), then shows that both models are of the form

$$r(h) = r_0 - C.F\left[\ln\left(\frac{B}{B-h}\right)\right],$$

Table 5.1. Summary of previous tables.

	model 1	model 2	model 3	model 4
Arithmetic mean of MSE's	0.0725	0.0607	0.164	0.048
Number of MSE's < 0.1	8	8	0	9
Smallest MSE	0.029	0.023	0.103	0.017
Largest MSE	0.113	0.135	0.246	0.110

where C is some constant and F is some function. That is, in both cases a taper function arises as some function F of $\ln\left(\frac{B}{B-h}\right)$. In (5.2) F is a linear function, in (5.3) F is a power function. So the models are similar in form, and differ in content only insofar as they modify

$$\ln\left(\frac{B}{B-h}\right)$$

in different ways, where B is some upper limit on h. Investigating other such modifications may be a fruitful line of further research.

Literature

BINGING, G. S., 1984: Taper equations for second-growth mixed conifers on Northern California. For. Sci. **30** (4), 1103–17.

DEMAERSCHALK, J. P., A. KOZAK, 1977: The whole-bole system – a conditioned dual-equation system for precise prediction of tree profiles. Can. J. For. Res. **7**, 488–497.

MADSEN, J. S., 1982: Determination of Stem Profiles in *Picea abies*. Faculty of Forestry, Royal Veterinary and Agricultural University, Copenhagen.

ORMEROD, D. W., 1973: A simple bole method. For. Chron. **49**, 136–138.

PRODAN, M., 1965: Holzmeßlehre. J. D. Sauerländer.

REED, D. D., J. C. BYRNE, 1985: A simple, variable form volume estimation system. For. Chron. **61** (2), 87–90.

YANG, R. C., A. KOZAK, J. H. G. SMITH, 1978: The potential of Weibull-type functions as flexible growth curves. Can. J. For. Res. **8** (4), 424–431.

Date of receipt: February 13th, 1986.

The authors' addresses: C. Brink, Department of Mathematics, University of Stellenbosch, Stellenbosch 7600, Republic of South Africa. K. von Gadow, z. Zt. Lehrstuhl für Waldwachstumskunde an der Universität München, Amalienstraße 52, D-8000 München 40.

EDV in Medizin und Biologie 17 (1/2), 28–34, ISSN 0300-8282

A model of the fetal thermodynamics during delivery and its clinical use

W. Dorda, B. Laminger, W. Wolf, R. Rudelstorfer, and G. Simbruner

Summary

A simple thermodynamic model of the fetus is designed and validated against measured data of heatflux from the scalp. In this model the fetal heatproduction, the fetal heatstorage and the heatflux from the fetus to the mother are taken into account. Heatflux is said to occur via fetal surface and via umbilical-placental circulation.

The aim of this study was to investigate the model by a simulation program under following dynamic conditions:

1. *unstressed fetus more than 1 hour before delivery. Here the heatproduction is assumed to be constant and the simulated scalp-heatflux is compared to the measured one to validate the model.*
2. *stressed fetus during the last ½ hour before delivery, where heatproduction can not be assumed to be constant. Here additionally scalp-heatflux and fetal heart rate are used as input to the model to calculate heatproduction.*
3. *abolished umbilical-placental bloodflow and its effect on the scalp-heatflux of the fetus.*
4. *various body sizes and its effect on the heatproduction of the fetus.*

This model allows to estimate the heatproduction of the fetus and to gain information about the fetal intrauterine conditions. First results showed that calculated heatproduction was related to indicators of metabolic status. Therefore, the thermodynamic model could provide a systematic basis for better fetal monitoring.

Key words:

Computersimulation, fetal thermodynamics, fetal heatproduction, fetal monitoring, delivery, mathematical model

Zusammenfassung

In dieser Arbeit wird ein thermodynamisches Modell vorgestellt, mit dessen Hilfe man das intrauterine Ergehen des Fetus abschätzen kann.

Das Modell berücksichtigt die fetale Wärmeproduktion und -speicherung sowie den Wärmefluß vom Fetus zur Mutter. Für letzteren wird angenommen, daß er sowohl über die Körperoberfläche als auch über den plazentaren Blutkreislauf erfolgt.

Folgende dynamische Vorgänge wurden mit Hilfe des Modells simuliert:

1. *Modellvalidierung:*
   Im Zeitraum, der mehr als 1 Stunde vor der Geburt liegt, kann die Wärmeproduktion als konstant angenommen werden. Der Vergleich zwischen gemessenem und simuliertem Wärmefluß ergab eine gute Übereinstimmung, wodurch das Modell zur Berechnung der Wärmeproduktion herangezogen werden kann.
2. *Simulation der Wärmeproduktion:*
   Unmittelbar vor der Geburt kann die Wärmeproduktion nicht mehr als konstant angesehen werden. Um diese berechnen zu können, wurden daher der gemessene Wärmefluß und die Herzfrequenz als zusätzliche Eingangsgrößen in dem Modell verwendet.
3. *Nabelschnurabklemmung:*
   Die Verminderung des plazentaren Blutflusses und seine Auswirkung auf den Wärmefluß über die Oberfläche des Fetus wurden mit Hilfe des Modells simuliert.
4. *Parameterempfindlichkeit:*
   Der Einfluß verschiedener Körpergrößen auf den berechneten Wärmefluß wurde untersucht.

Mit Hilfe dieses Modells ist es also möglich, die fetale Wärmeproduktion zu simulieren. Erste Ergebnisse zeigten, daß die errechnete Wärmeproduktion als ein Indikator für den metabolischen Status des Fetus angesehen werden kann. Aus diesem Grund könnte dieses einfache thermodynamische Modell eine systematische Grundlage für eine bessere Geburtsüberwachung darstellen.

1. Introduction

In utero the fetus is nourished by the mother via the placenta. Owing to this fact a large amount of the fetal blood flows to the placenta, where it comes in close contact with the maternal blood for the exchange of oxygen and food (Fig. 1). In case of placental insufficiency the intrauterine conditions of fetal life deteriorate. Particularly, during delivery when contractions additionally reduce the placental bloodflow, problems like hypoxic tissue damage with subsequent mental disability and even perinatal death may occur.

The knowledge about the physiology and pathophysiology of the human fetus before and during delivery is still very limited. Cardiotokogram, fetal blood gas analysis or transcutaneous pO_2 measurements have been employed to assess the fetal condition and to identify the risk of asphyxia (lack of oxygen). However as shown in recent publications (1, 2, 3) the information of fetal monitoring about the true status of the fetus is unsatisfactory. Rudelstorfer et al. (4) described a new and non-invasive method for heatflux measurements from the fetus during delivery. The fetus in utero is warmer than the surrounding maternal tissue. Due to this temperature gradient of about 0.3 °C heat flows from the fetus to the mother via the bodysurface and via the placenta, which functions like a heat exchanger by cooling the fetal blood. To maintain this continuous caloric loss, heat must be supplied by the fetal metabolism. Since 95 % of the energy in form of oxygen and organic substrate is transformed into heat, heatproduction is related to fetal metabolism. Based on these facts Simbruner (5) described a thermodynamic model in which the fetus is viewed as a warm nucleus and the maternal uterine environment as cooler surrounding shell.

Based on this model and intrauterine heatflux measurements it was our aim to estimate the heatproduction of the fetus to receive information about the intrauterine conditions.

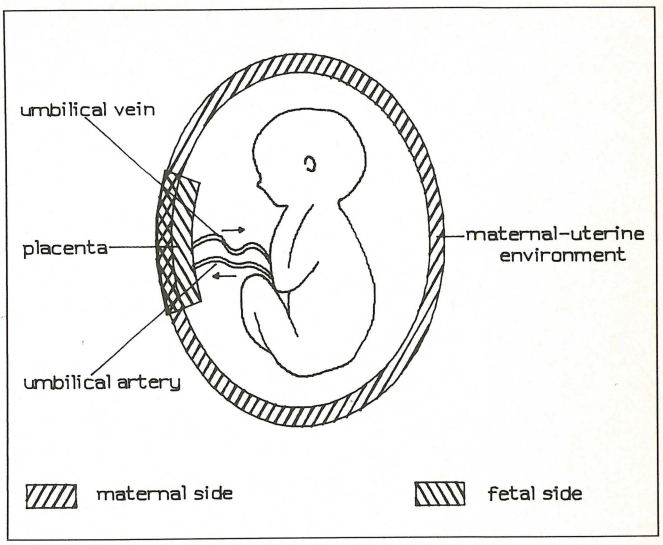


Fig. 1. Fetus with surrounding maternal uterine environment and umbilical placental circulation.

2. Method

The thermodynamic model of the fetus is based on an energy balance equation, where the change of fetal heat content (dQ/dt) is set to be equal to the metabolic heatproduction (M · W) minus the heatloss via the body surface (Hsurf) and via umbilical circulation (Hcirc):

$$dQ/dt = M \cdot W - Hsurf - Hcirc \quad (1)$$

(dimensions and abbreviations are all summarized in table 1).

Heat content of the fetus

The heat content Q is the product of fetal temperature (T-fet), body weight (W) and average heat capacity of fetal tissues (C). Average heat capacity was said to be C = 0.82 calorie per gramme and degree (6, 7)

$$Q = T\text{-}fet \cdot W \cdot 1000 \cdot 0.82$$

or

$$T\text{-}fet = Q/(W \cdot 1000 \cdot 0.82) \quad (2)$$

Metabolic heatproduction

Metabolic heatproduction (M) is related to oxygen consumption (VO_2). The proportionality factor is dependent on the constitutients of nutrition which it oxydized. Assuming that the fetus has a similar nutrition as the newborn it gains 4.83 cal out of one millilitre oxygen (8). Oxygen consumption of the unstressed human fetus was assumed to be 8.5 ml per kilogramme and minute, based on chronic animal experiments (9, 10) or direct measurements (11, 12). Accordingly the heatproduction per kilogramme of such a fetus is

$$M = 4.83 \cdot 8.5 \quad (3)$$

Heatloss via body surface

Heatloss via body surface (Hsurf) is calculated according to the law of Fourier

$$Hsurf = A \cdot k \cdot (T\text{-}fet - T\text{-}mat\text{-}envir) \quad (4)$$

Table 1. Abbreviations and dimensions of the variables.

A	cm^2	fetal surface area
C	cal/(g · °C)	average heat capacity of fetal tissues
FHR	beats/min	fetal heartrate
H	cm	fetal height
Hcirc	cal/min	heatloss via umbilical circulation
Hsurf	cal/min	heatloss via body surface
k	cal/(cm^2 · min · °C)	heattransfer coefficient
M	cal/(min · kg)	metabolic heatproduction per kilogramme
Q	cal	heat content
R	cal/ml · °C)	specific heat capacity of blood
t	min	time
T-art-umb	°C	temperature of arterial umbilical blood
T-fet	°C	fetal temperature
T-mat-envir	°C	temperature of maternal environment
T-ven-umb	°C	temperature of venous umbilical blood
uBF	ml/min	amount of umbilical bloodflow
V	ml/(kg · beats)	bloodvolume per heartbeat
VO 2	ml/(kg · min)	fetal oxygen consumption
W	kg	fetal bodyweight

Table 2. Height, weight and postpartal status of fetuses.

	height (cm)	weight (kg)	postpartal status
patient 1	49	3.5	»good«
patient 2	50	3.55	»good«
patient 3	51	3.3	»good«
patient 4	49	2.8	»bad«
patient 5	48	3.35	»bad«
patient 6	49	3.4	»bad«

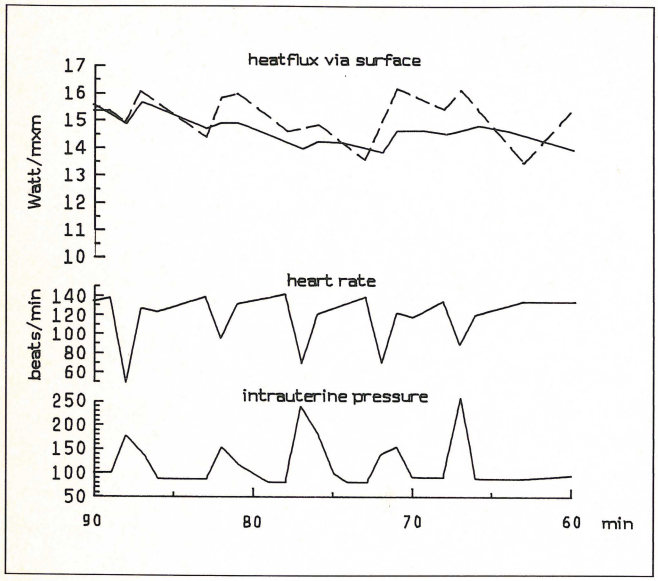


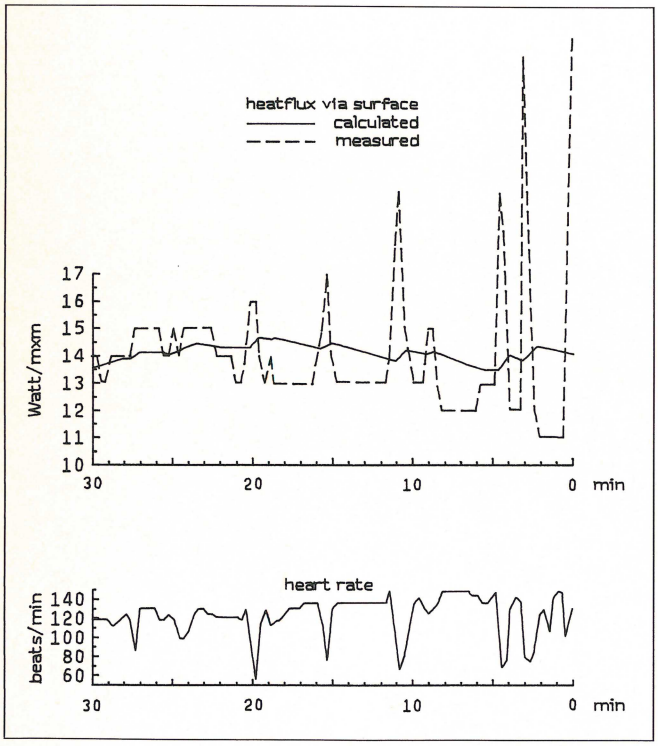
Fig. 2. Calculated (–) and measured (– –) heatflux via surface from 90 to 60 minutes before delivery.

where A is the surface area of the fetus, k is the heattransfer coefficient, T-fet is the temperature of the fetus and T-mat-envir the temperature of the maternal environment. The surface area A is estimated from the fetal weight (W) and height (H) according to the formula of Du Bois (13):

$$A = 6.29 \cdot (W \cdot 1000)^{0.55} \cdot H^{0.36} \quad (5)$$

As the fetus is floating in amniotic fluid, which is thermally similar to water, the heattransfer coefficient k of 0.066 cal/(cm^2 · min · °C) for man in water is used (14).

Fig. 3. Heatflux simulation of a fetus with good outcome (patient 1).



Heatloss via umbilical circulation and placenta

Heatloss via umbilical and placental circulation (Hcirc) is proportional to the amount of umbilical bloodflow (uBF), the temperature difference between arterial and venous umbilical blood (T-art-umb – T-ven-umb) and a constant for the heat capacity of blood (R)

$$Hcirc = uBF \cdot (T\text{-art-umb} - T\text{-ven-umb}) \cdot R \quad (6)$$

Umbilical blood flow (uBF) is calculated from the fetal heart rate and the bloodvolume which is pumped by each heartbeat through the umbilical circulation. The umbilical bloodvolume per heart beat and per kilogramme (V) in healthy human fetuses is derived from ultrasonic measurements (15): The umbilical blood flow is 110 ml/(kg · min) at a fetal heart rate (FHR) of 140/min, therefore $V = 110/140 = 0.786$ ml per heartbeat and kilogramme.

The value of the heat capacity of blood is taken from the literature (16) $R = 0.92$ cal/(ml · C). Rewriting the equation (6), one gets:

$$Hcirc = 0.786 \cdot FHR \cdot W \cdot (T\text{-art-umb} - T\text{-ven-umb}) \cdot 0.92 \quad (7)$$

Equations of the model

In this model several assumptions are made: Both the fetus and the maternal uterine environment are isothermal (i.e. they have the same temperature all over). Therefore the arterial blood leaving the fetus, flowing to the placenta, has the temperature of the fetus. The placenta functions as an efficient heat-exchanger. It cools the fetal blood down to the temperature of the maternal side of the placenta which is identical to the temperature of the maternal uterine environment. Thus the temperature of the venous blood leaving the placenta to return to the fetus has the temperature of the maternal environment which is assumed to be constant. Accepting these assumptions and substituting equation (3), (4), (5) and (7) into (1) one gets the equations of the model:

$$\begin{aligned} dQ/dt = {} & 41.055 \cdot W \\ & - 6.29 \cdot (W \cdot 1000)^{0.55} \cdot H^{0.36} \cdot 0.066 \cdot \\ & \quad (T\text{-fet} - T\text{-mat-envir}) \\ & - 0.786 \cdot FHR \cdot W \cdot (T\text{-fet} - T\text{-mat-envir}) \cdot 0.92 \end{aligned} \quad (8)$$

$$T\text{-fet} = Q/(W \cdot 1000 \cdot 0.82) \quad (2)$$

Using equation (8) and the measured heatflux of an actual delivery instead of the heatloss via body surface (equ. 4) you can calculate the heatproduction per kilogramme M, which is an indicator of the status of the fetus, as:

$$\begin{aligned} M = {} & (dQ/dt \\ & + 6.29 \cdot (W \cdot 1000)^{0.55} \cdot H^{0.36} \cdot 0.066 \cdot \\ & \quad (T\text{-fet} - T\text{-mat-envir}) \\ & + 0.876 \cdot FHR \cdot W \cdot (T\text{-fet} - T\text{-mat-envir}) \cdot 0.92)/W \end{aligned} \quad (9)$$

Numerical realisation

The calculations were done with the Continuous System Modeling Program APL/CSMP. Because of numerical reason the variable Q (fetal heat content) had to be transformed: the numerical values are very large compared with its changes dQ/dt. Therefore Q was transformed internally to $(Q - 30600 \cdot W)$.

Since heatflux of human fetuses is measured in Watt/m^2 the calculated values of the heatloss via surface per cm^2 must be transformed from cal/(min · cm^2) to Watt/m^2. Simulating the heatproduction (using equation 9) great peaks appear (see discussion below). Since only the trend of the heatproduction

is of clinical interest, a smoothing procedure had to be introduced: instead of the actual values of the heatflux via surface, the weightened mean of an adjacent interval of 5 minutes was used. The other summands of equation 9 were not changed.

In addition it is unavoidable that data during limited periods are not available due to technical difficulties or clinical circumstances like vaginal examination. Consequently linear interpolation from last to next available data was used.

3. Patients

In order to test the validity of the model we used

1. data from the literature in regard to weight, height (17), fetal heart rate (18), and scalp-heatflux (4, 5) of the healthy unstressed fetus during the first stage of labor to calculate the steady state;
2. data from 6 fetuses, three of which had a good and three of which had a bad outcome, characterized by the metabolic acidosis in the arterial umbilical blood. Lack of oxygen leads to an anaerobic, lactate and puruvatic acid producing metabolism. The important data of fetuses are listed in Table 2.

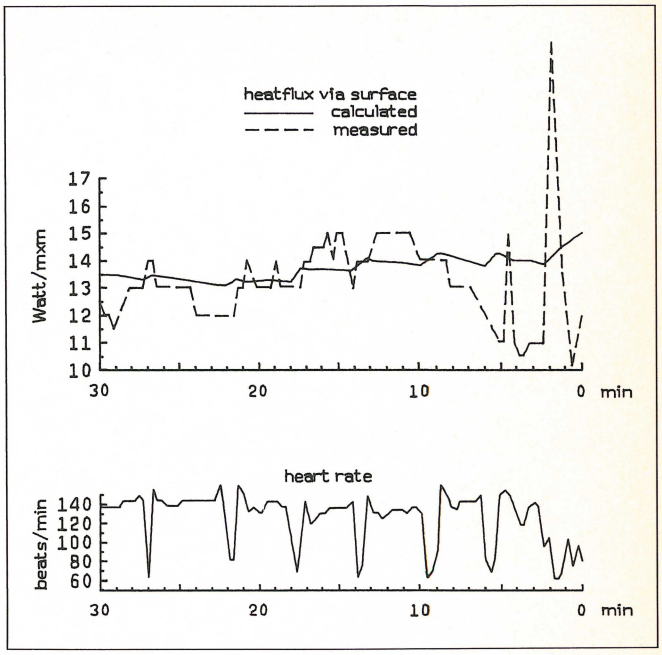


Fig. 4. Heatflux simulation of a fetus with bad outcome (patient 5).

4. Results

a) Steady state

First the steady state of the model was studied. An average fetus with a fetal heart rate FHR = 140 beats per minute, a fetal weight W = 3.25 kg, a fetal height H = 50 cm and a temperature of the maternal environment T-mat-envir = 37 °C is assumed. From equation (8) with dQ/dt = 0 follows T-fet = 37.28 °C. This result corresponds to measured data (19).

b) Heatflux simulation more than one hour before delivery – normal umbilical and placental circulation

In order to simulate the heatflux under dynamic conditions the actual course of the fetal heart rate during delivery was used instead of the constant FHR in equation (8). The results shown in Figure 2 were calculated for a fetus of weight of 3.25 kg and height of 49.3 cm. This example shows very well that the labor causes a decrease of the fetal heart rate. Therefore the heatloss via umbilical and placental circulation is diminished and consequently the heatflux via surface increases. After the disturbance the system returns slowly to the steady state.

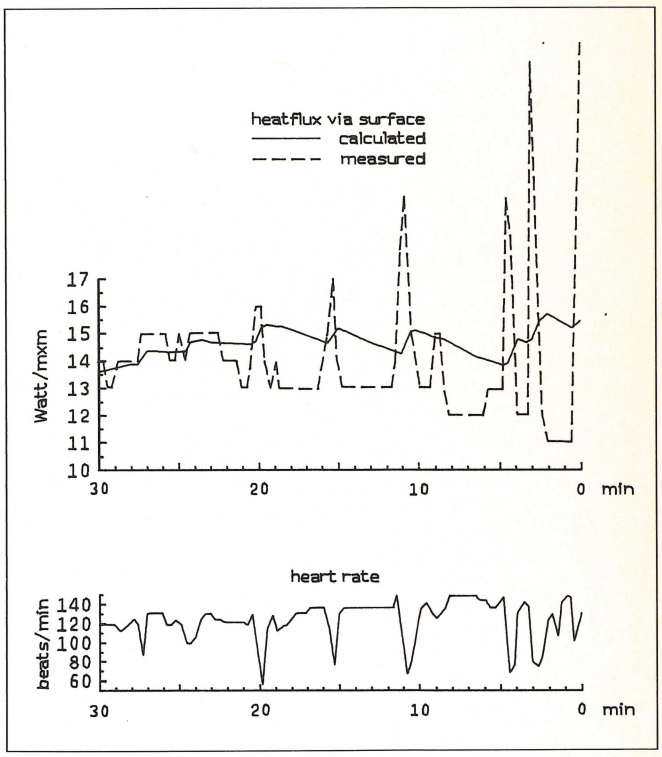
c) Heatflux simulation during the last half hour before delivery – normal umbilical and placental circulation

The results of the simulation of the fetal heatflux during the last 30 minutes before delivery are depicted in Figure 3 and 4. Figure 3 differs from Figure 4 in the outcome of the fetus. It should be pointed out, that the averaged, measured heatflux of the fetus with a bad outcome is lower than the calculated, ideal one. This difference could indicate impaired intrauterine conditions. In these examples the model does not describe properly the influence of the labor on the heatflux.

d) Heatflux simulations as in c), but without heatloss via umbilical and placental circulation during contractions

Since the pressure of the contractions may also hinder the umbilical and placental circulation the cooling function of the placenta could be still more diminished during labors. Abolished placental heatexchange during contractions results in a steeper rise of heatflux, but has little influence on the »decay« of heatflux during the contraction pauses (Fig. 5) when compared to standard circulatory conditions of our model (compare Fig. 3).

Fig. 5. Heatflux simulation assuming placental heat exchange during labors (patient 1).



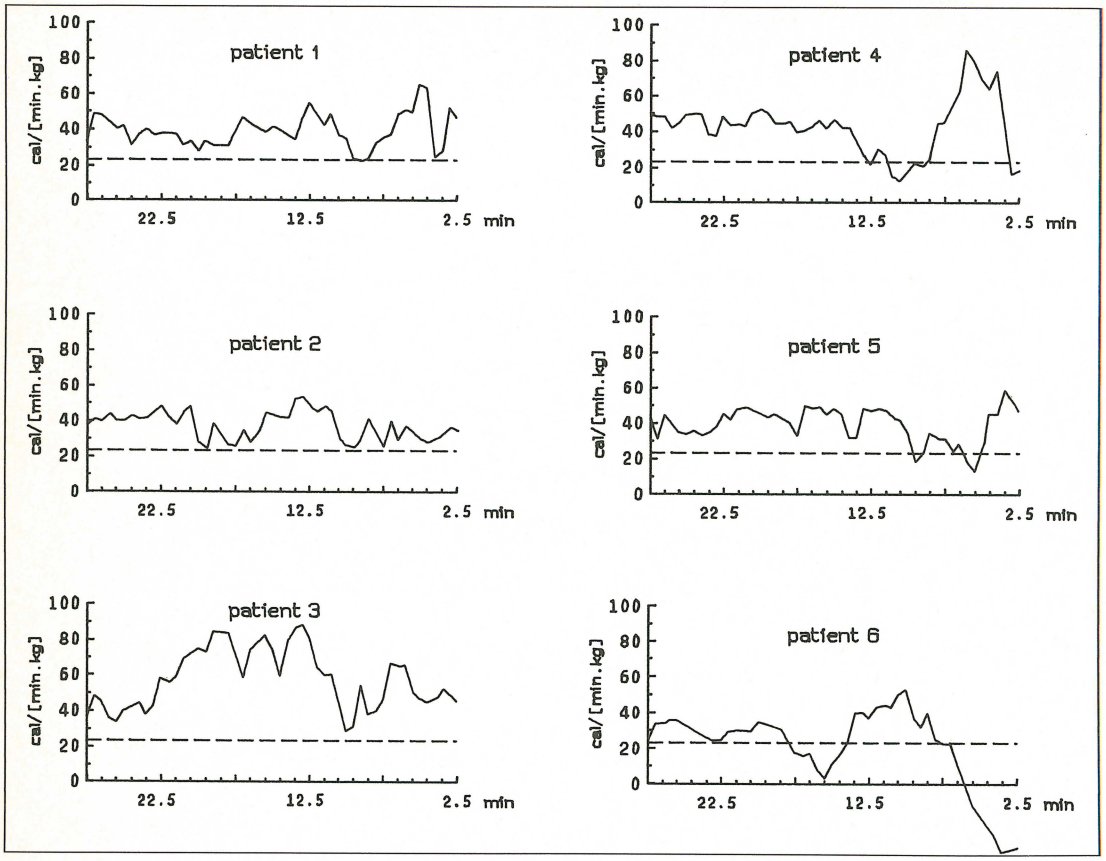


Fig. 6. Calculated heatproduction of fetuses with good (left side) and bad postpartale status (right side) (– – critical value).

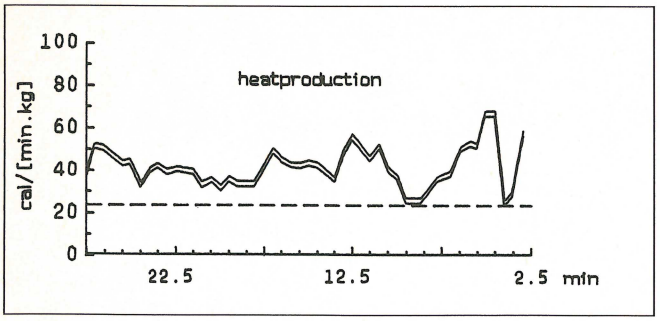
e) Clinical use of the model

As stated above the fetal heatproduction per kilogramme (M) may be an indicator of the status of the fetus. Therefore we used equation (9) to calculate this parameter of the last 30 minutes before birth. The results are represented in graphical form (Fig. 6) and as integrated values of M over 2½ minutes in Table 3. If 24 cal/(min · kg) is taken as critical value for a metabolism leading to acidosis, the thermomodel reveales that the calculated metabolic rate decreases in each of the fetus with postpartal acidosis, but never in the healthy fetuses below that critical value.

f) Influence of body sizes on the calculation of the heatproduction

The heatloss via surface depends on the size of the fetus, which can be estimated by ultrasonic measurements only. Since there is an error of up to 25 % (20), it was important to simulate the effect of body size, estimated with some error, on the outcome of the model predictions. We show in Figure 7 that the errors of ultrasonic size estimation should have little influence on the accuracy of the thermomodel.

Fig. 7. Influence if varying the body size on the calculation of heatproduction (H = 45 cm upper curve, H = 49.3 cm lower curve, – – critical value, patient 1).



5. Discussion

Steady state: All parameters or constants were either taken from the literature or directly measured, as all relations between variables and constants were derived from known laws in physics. Because of this it is really remarkable that the calculated temperature difference between fetus and mother of 0.28 °C is very similar to the one measured in the human pregnant mother. This difference was found to range between 0.1 and 0.2 °C by Walker and Wood (19), between 0.2 and 0.25 °C by Peltonen et al. (21) and to be about 0.3 °C by Adamsons (22). We conclude that the model, especially for the unstressed fetus with a normal umbilical circulation and placental function, yields a fair image of reality.

Dynamic conditions: The principle of the model is that any change in fetal temperature and thus metabolism is associated with a change in surface and circulatory heatloss. The dynamic version of the model was used in two manners:
First the heatflux was simulated assuming a constant metabolic rate i.e. heatproduction. This can be reasonably assumed for all conditions where the fetus is unstressed, because he is not subjected to the effects of labor i.e. intrauterine contractions and changes in placental blood flow (Fig. 2).

The simulated heatflux half an hour before delivery was close to the measured one between the laborpains, but differed considerably during labor (Fig. 3, 4). Possible reasons for this difference will be discussed below.

Secondly the heatproduction was calculated from measured heatflux and fetal heart rate for fetuses subjected to the stress of labor. As pointed out the measured heatflux (now an input to the model!) fluctuated more than the calculated one (Figures 3, 4). Therefore in our model the calculated heatproduction could not be constant during labor, but showed the right dimension between the laborpains. To eliminate the peaks in order to get only the trend of the heatproduction the smoothing procedure was used. During the second stage of labor, usually about half hour before birth, metabolic rate and surface heatflux are known to differ in newborns with good and bad outcome (1, 2, 3, 4, 5). Outcome is characterized by the acid base status and metabolic acidosis indicates a low aerobic and raised anaerobic metabolism respectively.

When the heatproduction was calculated from heatflux and fetal heart rate, the result was quite reasonable and in accordance with pathopysiological concepts: Each baby, who developed a metabolic acidosis had a calculated heatproduction (cal/min · kg), which dropped below a critical level. In contrast, babies with a good outcome (no acidosis) always had a heatproduction (cal/min · kg) above that critical level. We conclude from these findings, that the model yields valuable information about the condition of the fetus. More clinical data are needed to confirm the relationship between calculated heatproduction and indicators of metabolic rate.

Assumptions and possible improvements of the model: The validity of the assumptions like T-fet is equal to T-art-umb etc. are discussed in detail elsewhere (5). Apparently the errors introduced by these assumptions are small and consistent. Consequently the thermodynamic model will provide a systematic diagnostic basis. Deviation from the normal will be judged as pathologic.

That the model would yield a better simulation of heatflux during labor, a rapidly increasing and decreasing heatproduction during uterine contractions has to be postulated. However from the pysiological point of view this seems very unlikely, because such changes could not take place within half a minute and oxygen transport across the placenta and in the umbilical circulation decreases during the initial phase of contractions. Therefore the rise of heatflux observed during contractions cannot be explained by metabolic changes.

The decreased umbilical and placental flow during labors does not only effect the oxygen exchange, but also the heat exchange. A more diminished heatloss via placenta caused by the pressure of the contractions could be responsible for the peaks of the fetal temperature and heatflux from the surface. The simulation of heatflux when there is no umbilical or placental circulation during labors showed a somewhat faster rise in heatflux (compare Fig. 3 with 5), but still not fast and high enough to match experimental values. Furthermore the decay in heatflux is slower than it should be. We believe it would be useful to improve the model by taking into account a compensatory high placental-umbilical bloodflow after the intrauterine contraction. This would lead to a more rapid fall in heatflux.

Two other factors could explain the discrepancy between the simulated and measured heatflux. Both are related to local differences of metabolic or circulatory heattransport in the fetus. In the model an even distribution of heatproduction is assumed. However it is well known that the head contains the metabolically most active brain. The fetal brain might have a metabolic rate which is 5 to 10 times higher than in the rest of the body (23, 24). Furthermore the head is connected to the body via the cerebral circulation, which might be decreased during the pressure rise caused by contractions. Especially the venous circulation with its low pressure system and consequently heattransport out of the brain might be reduced. We suggest, in analogy to the change of umbilical-placental bloodflow, to make the circulatory heatloss from the brain a function of contractions. The increased metabolic rate of the brain and the diminished heatloss from it during contractions then might yield a better agreement between simulated and measured heatflux values. Still, another problem arises from the fact that the head itself is not a homogenous tissue. The »hot« brain is contained in an isolating shell. Preliminary results in rabbits, subjected to hypoxia, indicate that the temperature change at the brain surface is much less than in the core (25). These problems of the inhomogenity of heatproduction and vascular circuits in series encountered in the fetus probably can only be solved by enlarging the model into a two (or multi-)compartmental system. Whether such a multicompartmental model will be really advantageous to the one investigated is an open question. It will no longer be accessible for experimental validation and it is doubtful whether it will be more useful to the clinician than the model presented.

Table 3. Integrated values of simulated heatproduction.

min	pat. 1	pat. 2	pat. 3	pat. 4	pat. 5	pat. 6
0.0– 2.5	44.48	40.92	40.56	45.64	40.00	37.76
2.5– 5.0	39.08	43.12	42.16	44.32	35.88	27.12
5.0– 7.5	34.80	43.08	64.60	47.92	44.36	29.68
7.5–10.0	32.40	29.92	80.12	46.80	41.68	29.72
10.0–12.5	42.24	33.84	74.56	41.44	47.04	>12.44
12.5–15.0	40.64	46.68	77.12	35.32	42.12	27.84
15.0–17.5	46.44	41.16	57.28	>20.48	45.04	45.68
17.5–20.0	28.44	32.16	42.40	28.48	28.84	34.08
20.0–22.5	47.08	32.48	57.20	67.52	>23.68	> 5.88

Literature

1. LOW, J. A., S. R. PANCHAM, D. N. WORTHINGTON: Intra partum fetal heart rate profiles with and without fetal asphyxia. Am. J. Obstet. Gynecol **127,** 729 (1973).
2. TEJANI, N., L. MANN, A. BHAKTHAVATHSALAN: Fetal heart rate patterns and fetal pH. Obstet. Gynecol. **48,** 460 (1976).
3. SYKES, G. S. et al.: Fetal distress and the condition of newborn infants. Brit Med J **287,** 943 (1983).
4. RUDELSTORFER, R., G. SIMBRUNER, G. BERNASCHECK, A. M. ROGAN, ST. SZALAY, H. JANISCH: Heatflux from the fetal scalp during labor and fetal outcome. Arch. Gynecol. **233,** 85–91 (1983).
5. SIMBRUNER, G.: Thermodynamic models for diagnostic purposes in the newborn and fetus. Facultas Verlag GmbH Wien, pp. 170, 1983, ISBN 3-85076-133-9.
6. ADAMSONS, K. J., G. M. GANDY, L. S. JAMES: The influence of thermal factors upon oxygen consumption of the new-born human infant. J. Pediatr. **66,** 495 (1965).
7. RYSER, G., E. JEQUIER: Study by direct calorimetrie of thermal balance on the first day of life. Eur. J. Clin. Invest **2,** 176 (1972).
8. KARLBERG, P.: Determination of standard energy metabolism (basal metabolism) in normal infants. Acta Paediatric Scand. (suppl) 1952, 41 Suppl. **89,** 73.
9. LORIJN, R. H., L. D. LONGO: Clinical and physologic implications of increased fetal oxygen consumption. Am J. Obstet. Gynecol. **136,** 451 (1980).
10. BATTAGLIA, F. C., G. MESCHIA: Principale substrates of fetal metabolism. Physiol. Rev. **58,** 499 (1978).
11. HIRVONEN, L., J. LIND, T. PELTONEN: Kreislauf und Atmung im Fetal- und Neonatalstadium. Acta unversitatis ouluensis series D Medica Nr. 80, Oulu (1982).

12. ROMNEY, S. L., D. E. REID, J. METCALF, C. S. BURWELL: Oxygen utilization by the human fetus in utero. Am J. Obstet. Gynec. **70,** 791 (1955).
13. DU BOIS, D., E. F. DU BOIS: A formula to estimate the approximate surface area if height and weight be known. Arch. int. med. (1916).
14. GOLDMAN, R. F., J. R. BECKENRIDGE, E. REEVES, Kl. BECKMANN: Wet versus dry suit approaches to water immersion protective clothing. Aerospace Med. **37,** 485 (1966).
15. EIK-NES, S. H., A. D. BRUBAKK, M. K. ULSTEIN: Measurement of human fetal blood flow. Br. Med. J. **1,** 283 (1980).
16. ATZLER, E., F. RICHTER: Die spezifische Wärme des Blutes und des Serums unter besonderer Berücksichtigung der Methodik. Biochem. Z. **100,** 193 (1919).
17. USHER, R., F. MCLEAN: Intrauterine Growth of Liveborn Caucasians Infants at Sea Level: Standards Obtained from Measurements in 7 Dimensions of Infants. Born between 25 and 44 Weeks of Gestation. J. Pediat. **74,** 901 (1969).
18. HON, E.: An Atlas of Fetal Heart Rate Patterns. Harthy Press, New Haven, Connecticut (1968).
19. WALKER, D. W., C. WOOD: Temperature relationship of the mother and fetus during labor. Am. J. Obstet Gynec. **107,** 83 (1970).
20. CAMPBELL, S., D. WILKIN: Ultrasonic measurement of fetal abdo men circumference in the estimation of fetal weight. Brit. J Obstet. Gynec. **82,** 689 (1975).
21. PELTONEN, R., G. ROOTH, A. HUCH, R. HUCH: The differenc between the fetal and maternal temperature during delivery. Eur Congress Perinatology (Uppsala) Abstract p 188 (1976).
22. ADAMSONS, K. Jr., M. E. TOWELL: Thermal Homeostasis in th Fetus and Newborn. Anesthesiology **26,** 531 (1965).
23. CROSS, K. W., D. STRATTON: Aural temperture of the newbor infant. Lancet 1974 ii: 1179.
24. STRATTON, D.: Aural temperature of the newborn infant. Arch Dis. Childh. **52,** 865 (1977).
25. personal communication.

Date of receipt: April 4th, 1986.

The authors' addresses: W. Dorda, B. Laminger and W. Wolf, Department of Medica Computer Sciences (Head: Prof. Dr. G. Grabner), University of Vienna, Garnisongasse 13 A-1090 Wien, Austria. R. Rudelstorfer, 2nd Department of Obstetrics and Gynecology (Head: Prof. Dr. H. Janisch), University of Vienna, Spitalgasse 23, A-1090 Wien, Austria. G Simbruner, Clinic of Pediatrics (Head: Prof. Dr. E. Zweymüller), Department o Neonatology (Director: Prof. Dr. O. Thalhammer), University of Vienna, Währinger Gürte 74, A-1090 Wien, Austria.

EDV in Medizin und Biologie **17** (1/2), 34–42, ISSN 0300-8282

Risk factor identification for cerebrovascular disease and myocardial infarction by stepwise discriminant analysis

K. P. Pfeiffer*), G. Zenker**) and G. Jürgens***)

Summary

In the last years some biochemical variables have been identified as possible risk factors for diseases of the cardiovascular system. In this study the discriminatory power of lipoprotein (a) (Lp(a)) together with other biochemical variables, which are known to be possible risk factors, is analysed. Lp(a) is not correlated to this variables and therefore an additional improvement for discrimination between control and risk groups is expected. In addition to the absolute values of the biochemical variables some ratios of this variables have been computed to consider the aspects of a multifactorial risk concept. Because some variables deviate from a normal distribution also after transformations, only nonparametric methods, and especially a stepwise kernel discriminant analysis procedure, have been used for the examination of differences.

The comparison of a group of survivors of myocardial infarction and a group of patients with cerebrovascular insult with control groups shows that Lp(a) as a single variable is a good discriminator. In a multivariate analysis using a kernel discriminant function together with some other biochemical variables selected before the additional discriminatory power of Lp(a) is rather small.

Zusammenfassung

In den letzten Jahren wurden einige biochemische Variable als mögliche Risikofaktoren für Erkrankungen des Herz-Kreislauf-Systems untersucht. In dieser Studie wird das Trennvermögen von Lipoprotein (a) (Lp(a)) zusammen mit anderen biochemischen Variablen, die als mögliche Risikofaktoren bekannt

*) Physiologisches Inst., Harrachgasse 21/5
**) II. Med. Abt. d. Landeskrankenhauses Graz
***) Inst. f. Med. Biochemie, Harrachgasse 21/3
alle: Karl-Franzens-Universität Graz, A-8010 Graz, Austria

sind, untersucht. Lp(a) ist mit diesen Variablen nicht korreliert und könnte daher eine Verbesserung der Unterscheidung zwischen Kontroll- und Risikofällen bringen. Zusätzlich zu den Absolutwerten der biochemischen Variablen wurden im Sinne des multifaktoriellen Risikokonzeptes auch Quotienten von Variablen untersucht. Da einige Variable auch nach Transformationen von einer Normalverteilung abweichen, wurden zur Untersuchung der Unterschiede nur parameterfreie Verfahren, und speziell eine stufenweise Kernfunktionsdiskriminanzanalyse, verwendet.

Die Vergleiche einer Gruppe von Patienten nach einem Herzinfarkt und einer Patientengruppe mit cerebrovasculärem Insult mit Kontrollgruppen zeigten, daß Lp(a) als einzelne Variable betrachtet eine gute Trennung der Gruppen ermöglicht. Bei einer multivariaten Untersuchung mittels Kernfunktionsdiskriminanzanalyse ist die zusätzliche Trennkraft von Lp(a) relativ gering, wenn andere biochemische Variable schon als Diskriminatoren ausgewählt wurden.

1. Introduction

In the last years there was a strong interest in the identification of risk factors for cerebrovascular ischemic events and myocardial infarction [10, 14, 15, 17, 19]. Especially lipoproteins, like high or low density lipoproteins, have gained much attention and have been already accepted in clinical preventive care. Most of these biochemical variables are highly correlated, whereas lipoproteins (a) Lp(a) seems to be not correlated to the other lipoproteins and therefore additional information may be expected.

Taking into consideration the important role of plasma lipids and lipoproteins for the development of ischemic heart disease, their role is not as well defined in the development of ischemic cerebrovascular disease [8, 15, 16]. A Japanese cerebral infarction study [12] did not show elevated total lipoprotein cholesterol levels. Even no significantly lowered serum levels for high density lipoprotein and apolipoprotein A have been found in Italians suffering from cerebrovascular disease [1]. In both Italian and Japanese patients with transient ischemic attack, high density lipoprotein cholesterol (HDL-C) was the only lipid parameter significantly lowered, however only in male subjects [6]. In Japanese cerebral infarction patients HDL-C and the ratio HDL-C/LDL-C (low density lipoproteine cholesterol) was found to be significantly lowered in men and women [11, 18].

In this study we want to evaluate the role of Lp(a) simultaneously with other biochemical variables and ratios of variables as possible risk factors for cerebrovascular insult and/or myocardial infarction. In the multifactorial risk concept it is assumed that not only the absolute amount of one factor is important. For some factors also the relation of two factors should be considered and therefore some ratios have also been computed as additional »independent« variables. Furthermore some studies on e.g. atherosclerosis have shown that ratios are good discriminators [14]. A further reason for the analysis of ratios is that the acceptance of a factor for clinical applications depends on its simplicity.

The requirement of simplicity of the results seems to be also one of the reasons for the application of only univariate procedures or the application of simple multivariate models, also when the conditions for the application of these procedures are not fulfilled. To avoid misleading conclusions we started our analysis with nonparametric comparisons of characteristic values of the groups and after this we proceeded to a nonparametric kernel discriminant analysis with the aim to identify some good discriminators between control groups and risk groups. A simultaneous comparison of four groups has been performed to get some information about the possibility to differentiate between different diseases. This difference may be reflected in different values of the factors. The classification results have been weighted by different cost matrices according to the medical consequences of misclassification.

2. Case and control groups

The criteria for an exclusion of one of the groups are inflammatory, liver, thyroid, endocrine or renal diseases. None of the subjects was on a therapy known to cause changes of serum lipids or lipoprotein levels. Characteristic variables of the groups are summarized in table 1.

2. 1. Case Groups

The case group of this study represents all patients with cerebrovascular ischemic events (CI) or myocardial infarction (MI) of the Department of Internal Medicine II and the University Clinic of Neurology of the University Hospital in Graz during 1984.

Within the case groups we distinguished between three groups. The first one is group d(31) with n31 = 94 patients with CI, but without coronary artery disease (CAD) proven by ECG and exercise testing. The second case group d(32) consists of n32 = 54 patients with CI and with CAD. The sono

Table 1. Description of the symptoms and age and sex structure of the case control groups.

	sono score	CAD	CI	MI	n	<50	50–75	>75	50–75	
									m	w
Control group d(1)	0	–	–	–	38	8	29	1	6	23
Control group d(2)	>0	–	–	–	38	3	28	7	16	12
Case group d(31)	?	–	+	–	94	10	62	22	37	25
Case group d(32)	>0	+	+	–	54	0	32	22	18	14
Case group d(4)	?	+	–	+	42	8	33	1	25	8

CAD coronary artery disease
CI cerebrovascular ischemic attack
MI myocardial infarction
? sono score not determined

score of these groups is between 0 and 5. After an examination of possible differences in biochemical variables we merged these groups together to group d(3) for further analysis. A third case group d(4) consists of n4 = 42 survivors of myocardial infarction (MI) with CAD.

2. 2. Control Groups

The criteria for the inclusion of a person into the control groups are: −age > 40 years
- exclusion of a CAD by electro cardiographic examination during and after exercise
- exclusion of peripheral arterial disease
- no signs or history of cerebrovascular disease.

All control persons have been selected randomly from in- and outpatients of the University Clinic of Neurology and of the Department of Internal Medicine II in Graz during the interval of the sampling of the case group.

To eliminate possible influences of also small injures, a continuous wave Doppler examination and duplex scanning of the extra-cranial carotid arteries was performed. We differentiated between a control group d(1) (n1 = 38) with sone score 0 an a second control group d(2) (n2 = 38) with sono score greater than 0. In this group n = 10 people have sono score 1, n = 17 have sono score 2, and n = 11 people have sono score 3.

Because of the different age-distribution in different groups the main part of this analysis was restricted to subjects between 50 and 75 years old.

3. Characterization of the data

After the subject has been fasting about 12–14 hours the following biochemical variables have been determined: Lp(a) using the Laurell technique [9] and an own monospecific antiserum [7]. Total cholesterol (TC), high density lipoprotein cholesterol (HDL-C) and triglycerides (TG) using routine laboratory methods. Low density cholesterol (LDL-C) was calculated according to the Friedewald formula. Assuming that the diseases under consideration are caused by a cluster of risk factors, where also the relations of the variables are important factors, the following ratios of variables have been computed for each subject: LDL-C/HDL-C, HDL-C/TC, LPa(a)/HDL-C. For all subjects of group d(2) and d(3) a sonoscore has been determined by Doppler and duplex scanning to measure the extent of vascular restrictions. As possible confounding variables age and sex have been included in the study. Further confounding factors, like nicotin consumption, diabetes or hypertension have been documentend, but they have not been included as possible discriminators in this study.

4. Statistical methods

4. 1. Univariate comparisons

The first step of our analysis of possible risk factors for cerebrovascular diseases and myocardial infarction is a description of the location, dispersion and distribution of single variables in different groups expressed by quartiles and the results of the Kolmogoroff-Smirnow-test for normal distribution (Table 2). To fulfill the necessary conditions for the application of parametric statistical procedures, like t-test, ANOVA and linear discriminant analysis, some variables have been transformed by a logarithmic or square root function. But also these transformations did not lead to good approximations to a normal distribution for some variables as can be shown by the Kolmogoroff-Smirnow-test.

Furthermore for some variables the variance is different for different groups. Considering these prepositions the application of parametric procedures is not indicated and may not lead to optimal classification results. Therefore all comparisons have been done by nonparametric Wilcoxon-U-tests or Kruskal-Wallis-H-tests [3]. For the evaluation of the correlation between different variables, the Spearman-rank-correlation coefficient was computed [3].

A reliability analysis of the data before computation of any characteristic parameter did not show any outlier from a medical point of view. Therefore we decided to eliminate no observation.

As can be seen from table 2, there is a big overlap of most variables for the case and control groups and therefore a high error rate has to be expected, if single variables are used for classification.

For univariate classification usually a normal range or cut-off point is determined according to the criterion of the minimal error rate or minimal cost function. The apparent error rate (aER) is determined by reclassification of the data used for the determination of the cut-off point. The apparent error rate is an over-optimistic, biased estimation of the true error rate (tER), which would be expected if test data, which are different from the learning data, are classified. Also if we use clinically accepted risk factors, like HDL-C or the ratio LDL-C/HDL-C for classification, the apparent error rate (aER) between control group d(1) and case group d(4) is more than 41 % in this study. Studying Lp(a), the variable whose discriminatory power was of special interest for this study, the apparent error rate aER = 32.5 % for the classes d(1) and d(4). The error rate would not decrease if the ratio Lp(a)/HDL-C is used (aER = 32.5 %) for discrimination of d(1) and d(4).

4.2. Kernel discriminant analysis

Kernel functions K(x, d(i), H) can be used for the estimation of the uni- and multivariate probability density function [4, 5, 13]:

$$\hat{f}(x, d(i), H) = \frac{1}{n(i)} \sum_{j=1}^{n(i)} K(x, x(ij), H) \qquad [1]$$

of each class d(i), i = 1, . . ., m from the observations x(ij), j = 1, . . ., n(i). In this study only continuously distributed variables have been used for classification and therefore multinormal kernels of dimension q:

$$K(x, x(ij), H) = \frac{1}{(2\pi)^{q/2} |H|^{1/2}} \exp(-1/2(x - x(ij))\, H^{-1}\, (x - x(ij))) \qquad [2]$$

have been used. The estimation of the smoothing factors is performed according to a maximum likelihood criterion:

$$L(H) = \prod_{i=1}^{m} \prod_{j=1}^{n(i)} \hat{f}(x(ij), d(i), H) \longrightarrow \max. \qquad [3]$$

For simplification the elements h(ij), i = 1, . . ., p, j = 1, . . ., p for a p-dimensional probability density function with i ≠ j have been set to zero:

$$h(ij) = \emptyset,\ i \neq j$$

Table 2. Characterization of the variables of the case and control groups d(i), i = 1, . . ., 4 (50–75 years old). Results of the Kolmogoroff-Smirnow-Test (A) for normal distribution and the Kruskal-Wallis-Test (H*) for the comparison of four groups. tER is the estimated true error rate in %, which results from a simultaneous classification of four classes using the kernel-estimated probability density function of one variable.

	|i|	min	Q25	Q50	Q75	max	A	H*	tER
Age	|1|	50.0	55.5	60.0	67.0	74.0	0.079	3.20	62.1
	|2|	50.0	59.0	63.0	68.5	75.0	0.086		25.0
	|3|	50.0	58.0	63.0	70.0	75.0	0.091		89.4
	|4|	50.0	54.0	62.0	67.0	75.0	0.112		100.0
Lp(a)	|1|	0,0	3.0	7.0	13.5	41.0	0.156	9.00*	100.0
	|2|	1.0	4.0	9.0	27.5	138.0	0.270*		25.0
	|3|	0.0	6.0	10.0	32.0	129.0	0.223**		100.0
	|4|	0.0	6.8	15.0	23.8	55.0	0.150		54.5
TC	|1|	109.0	183.3	219.0	247.8	327.0	0.095	2.22	100.0
	|2|	103.0	182.5	220.0	243.5	305.0	0.536**		75.0
	|3|	117.0	178.0	207.0	244.0	348.0	0.085		93.6
	|4|	113.0	164.5	203.0	229.8	377.0	0.576**		54.5
HDL-C	|1|	31.0	38.8	46.0	57.0	86.0	0.690**	9.81*	41.4
	|2|	24.0	36.0	45.5	56.5	78.0	0.102		100.0
	|3|	16.0	34.0	41.0	48.0	89.0	0.089		81.9
	|4|	18.0	28.0	38.0	48.8	52.0	0.129		45.5
TG	|1|	40.0	101.5	138.0	173.5	498.0	0.379**	4.05	27.6
	|2|	88.0	128.5	168.5	284.0	476.0	0.202		78.6
	|3|	53.0	108.0	148.5	231.0	428.0	0.383**		91.5
	|4|	60.0	96.5	121.0	205.5	424.0	0.178		81.8
LDL-C	|1|	20.4	119.6	143.6	154.1	259.4	0.168	1.89	41.4
	|2|	35.8	90.5	124.4	157.3	231.8	0.131		60.7
	|3|	60.4	102.8	128.1	158.6	259.6	0.102		70.2
	|4|	56.4	99.6	128.4	160.5	268.4	0.100		100.0
LDL-C/HDL-C	|1|	0.15	0.26	0.35	0.43	2.35	0.300**	6.01	17.2
	|2|	0.14	0.24	0.43	0.57	1.01	0.119		100.0
	|3|	0.08	0.26	0.32	0.39	0.71	0.426**		25.5
	|4|	0.13	0.20	0.30	0.40	0.81	0.164		75.6
HDL-C/TC	|1|	0.11	0.18	0.23	0.27	0.44	0.112	4.95	93.1
	|2|	0.10	0.16	0.24	0.29	0.42	0.101		100.0
	|3|	0.06	0.17	0.20	0.24	0.39	0.479**		44.7
	|4|	0.10	0.14	0.19	0.25	0.38	0.128		42.4
Lp(a)/HDL-C	|1|	0.00	0.05	0.15	0.32	0.70	0.164	12.91**	100.0
	|2|	0.02	0.09	0.18	0.51	2.26	0.293*		50.0
	|3|	0.00	0.14	0.27	0.77	4.20	0.214**		90.4
	|4|	0.00	0.17	0.42	0.69	1.61	0.424**		57.6

Table 3. Comparison of some biochemical variables of two age classes (50–75 years, >75) by a Wilcoxon-U-Test (U) only for classes d(2) and d(3).

	50–75			>75			
	$\bar{x}$	s	n	$\bar{x}$	s	n	$\hat{z}$
control group d(2)							
Lp(a)	18.11	26.70	28	22.43	35.00	7	0.309
TC	216.50	50.58	28	251.71	76.34	7	0.990
HDL-C	47.75	13.72	28	42.00	12.21	7	0.969
LDL-C	127.31	50.20	28	155.00	38.68	7	1.608
LDL-C/HDL-C	0.449	0.236	28	0.282	0.095	7	1.815
case group d(3)							
Lp(a)	22.37	25.06	94	13.75	11.14	44	0.991
TC	212.57	52.23	94	194.09	52.38	44	1.953
HDL-C	42.14	13.16	94	43.23	10.11	44	0.544
LDL-C	135.49	42.01	94	118.92	43.66	44	2.198
LDL-C/HDL-C	0.335	0.126	94	0.420	0.236	44	1.960

$\hat{z}$ Transformation of the rank-sum of the Wilcoxon-U-Test to normal distribution

Table 4. Comparison of some biochemical variables of the male and female groups by a Wilcoxon-U-Test.

	male			female			
	$\bar{x}$	s	n	$\bar{x}$	s	n	$\hat{z}$
control group d(2), 50–75 years							
Lp(a)	14.13	11.54	16	23.42	38.91	12	0.604
TC	198.69	47.04	16	240.25	46.73	12	2.112
HDL-C	48.19	13.34	16	47.17	14.78	12	0.209
case group d(3), 50–75 years							
Lp(a)	23.86	28.57	55	20.28	19.23	39	0.035
TC	210.07	54.24	55	216.10	49.74	39	0.725
HDL-C	39.69	12.86	55	45.59	12.96	39	2.130

Table 5. Smoothing factors $\hat{h}(ii)$ of the kernel functions and some results of the stepwise kernel discriminant analysis for two groups.

5.a: d(1) – d(2), n1 = 29, n2 = 28

	p = 1			p= 2		p = 3
Lp(a)	0.557	–	–	0.556	–	0.590
TC	–	0.698	–	–	–	–
LDL-C/HDL-C	–	–	0.573	0.573	0.573	0.573
HDL-C/TC	–	–	–	–	0.587	0.587
aER	42.1	47.4	24.6	22.8	3.5	1.8
tER	47.4	73.7	26.3	29.8	8.8	14.0
Q1	0.362	0.244	0.377	0.132	0.138	0.045
Q2	–0.6	–0.1	6.3	5.3	6.1	6.4

5.b: d(1) – d(4), n1 = 29, n2 = 44

	p = 1			p = 2	p = 3	p = 4
Lp(a)	0.689	–	–	0.791	–	–	0.799
HDL-C	–	0.597	–	0.597	0.597	0.597	0.597
LDL-C/HDL-C	–	–	0.545	–	0.479	0.479	0.479
HDL-C/TC	–	–	–	–	–	0.564	0.564
aER	43.5	25.8	25.8	24.2	4.8	1.6	1.6
tER	45.2	27.4	29.0	33.9	8.1	3.2	6.5
Q1	0.269	0.268	0.402	0.071	0.136	0.063	0.016
Q2	–0.3	1.3	6.9	1.1	18.0	21.1	21.5

5.c: d(1) – d(3), n1 = 29, n3 = 94

	p = 1			p = 2		p = 3	p = 4	p = 5
Lp(a)	0.540	–	–	0.553	–	–	–	0.689
HDL-C	–	0.589	–	–	–	–	0.562	0.562
LDL-C	–	–	–	–	–	0.577	0.577	0.577
LDL-C/HDL-C	–	–	0.570	0.570	0.570	0.570	0.570	0.570
HDL-C/TC	–	–	–	–	0.548	0.548	0.548	0.548
aER	40.7	47.2	13.8	13.8	4.1	0.8	0.0	0.0
tER	46.3	50.4	15.4	16.3	6.5	4.1	1.6	0.8
Q1	0.344	0.272	0.324	0.118	0.141	0.047	0.021	0.007
Q1	–0.8	0.1	19.2	21.9	33.0	39.7	58.3	59.7

5.d: d(2) – d(4), n2 = 28, n4 = 33

	p = 1			p = 2		p = 3	
Lp(a)	–	–	–	0.583	–	0.583	0.583
TC	0.699	–	–	–	0.796	0.700	–
HDL-C	–	0.598	–	0.598	0.598	0.598	0.598
HDL-C/TC	–	–	0.591	–	–	–	–
Lp(a)/HDL-C	–	–	–	–	–	–	0.545
aER	37.7	37.7	41.0	21.3	29.5	14.8	16.4
tER	37.7	37.7	45.9	29.5	36.1	26.2	26.2
Q1	0.251	0.257	0.258	0.085	0.061	0.021	0.045
Q2	0.1	0.3	0.0	2.0	1.3	4.7	6.0

and the diagonal elements h(ii), i = 1, . . ., p have been estimated parallel to the variable selection process [13]. In table 5 the estimated smoothing factors h(ii), i = 1, . . ., p of different variables for different model orders p are listed.

An observation x(ij), i = 1, . . ., m, j = 1, . . ., n(i) was classified according to the following criterion of maximal probability:

$$\text{x(ij) is classified into class d(k), if } \hat{f}(x(ij), d(k), H) = \max_{l=1,\ldots,m} (\hat{f}(x(ij), d(l), H) \quad [4]$$

For the estimation of the true error rate the leaving-one-out or jackknife probability density:

$$\hat{f}_J(x(ij), d(i), H) = \frac{1}{n(i) - 1} \sum_{\substack{i=1 \\ j \neq i}}^{n(i)} K(x(ij), d(i), H) \quad [5]$$

has been computed. $\hat{f}_J$ has also been used in the maximum likelihood criterion [eq. 3].

4. 3. Model evaluation

There is no unique or best criterion for model evaluation in discriminant analysis [6, 13]. The disadvantage of the error rate is that the continuous information contained in the probability function for each observation x(ij) is reduced to binary states: right or wrong classification. For a two group classification sensitivity, specificity and positive and negative predictive value may be used for characterization of the discriminatory power of a model.

For many medical decision problems a misclassification of a control person into the case group and vice versa may be of different importance. Therefore a weighting of the classification matrix W:

		Classified into			
		d(1)	d(2)	d(m)	
real state	d(1)	w(11)	w(12)	w(1m)	[6]
	d(2)	w(21)	. . . . .		
	.				
	.				
	d(m)	w(m1)	. . . . .		

by a cost matrix C:

$$C = \begin{matrix} c(11)\ c(12) \ldots\ldots\ldots\ c(1m) \\ c(21) \ldots\ldots\ldots\ldots\ \cdot \\ \cdot \qquad\qquad\qquad\qquad\ \cdot \\ \cdot \qquad\qquad\qquad\qquad\ \cdot \\ c(m1) \ldots\ldots\ldots\ldots\ c(mm) \end{matrix} \quad [7]$$

and the evaluation of the model quality by a cost function:

$$COST = \sum_{i=1}^{m} \sum_{j=1}^{m} w(ij) * c(ij) \quad [8]$$

can be performed. The determination of the elements of the cost matrix C depends on the problem under consideration and is arbitrary. We will show in this study that a weighting of misclassification with c(ij) ≠ 1 for i ≠ j has often a very practical, medical background.

Another criterion for model evaluation is [6]:

$$Q1 = \frac{1}{n} \sum_{i=1}^{m} \sum_{j=1}^{n(i)} \hat{f}_J(x(ij), d(i), H). \quad [9]$$

This criterion evaluates the mean probability of the observations x(ij) for classification into class d(i). To use this criterion Q1 for model evaluation in stepwise variable selection the decrease with increasing model order has to be considered. One simple possibility to do this, is to compare Q1(p) for model order p with the product of Q1 for model order 1 for the p variables, which have been selected. If Q1(p) − Q1[1]. . .Q1[p] > e, than continue the variable selection, where Q1[j] is Q1 of model order p = 1 for variable j.

In this study another model evaluation criterion Q2:

$$Q2 = \sum_{i=1}^{m} \sum_{j=1}^{n(i)} v(ij). \quad [10]$$

$$r(ij) = \frac{\hat{f}_J(x(ij), d(i), H)}{\max_{k \neq i} \hat{f}(x(ij), d(k), H)}$$

$$v(ij) = \begin{cases} r(ij) - 1 & \text{if } r(ij) > 1 \\ 1/r(ij) - 1 & \text{else} \end{cases}$$

was used to indicate the power of the decisions for classification.

5. Univariate and multivariate comparisons

5. 1. Age-dependent analysis

In class d(3) 30 % (n = 44) of all subjects are elder than 75 years, whereas in class d(1) only 3 % (n = 1), in class d(2) only 18 % (n = 7) and in class d(4) only 2 % (n = 1) are elder than 75 years. A comparison of the location of the biochemical variables of class d(3) for the age group between 50–70 years (n = 94) an the group > 75 years (n = 44), shows significant differences ($p < 0.05$) for TC, TG, LDL-C, LDL-C/HDL-C and HDL-C/TC. There is also a significant difference between this age classes in group d(2) for LDL-C and LDL-C/HDL-C (table 3).

Within all the age-groups of the 4 classes we could not find any substantial correlation between age and biochemical variables. The substantial differences of single biochemical variables between age groups have been the reason for a restriction of the further analysis to groups between 50 and 75 years old.

5. 2. Sex-dependent analysis

There are some results in the literature [2] that there are differences especially in Lp(a) in women before and after menopause. To identify possible differences in the biochemical variables under consideration we compared men and women between 50 and 75 years old inside the classes by Wilcoxon-U-test. Only HDL-C in groups d(1) and d(3) and TC in d(2) shows significant ($p < 0.05$) differences between men and women (Tab. 4). All other variables have not found to be substantially different for men and women within the four classes.

Only within control group d(2) we found a significant correlation ($p < 0.05$) between age and Lp(a) for men and women. All other variables are not correlated to the age.

From the inhomogenity between the results in different groups and the low significance level ($p < 0.05$) of all differences and correlations we concluded that sex should not be treated as a confounding variable.

5. 3. Relations to the sono score

From a clinical viewpoint, there is a difference between control group d(1) (n1 = 29) and control group d(2) (n2 = 28) because of the sono score >0 in group d(2). But there is no substantial difference in the location of any single biochemical variable or ratio in the population between 50 and 75 years. The minimal true error rate (tER) is 26.3 % for LDL-C/HDL-C (the transformed rank-sum of the Wicoxon-U-test z = 0.85), whereas aER for all other single variables is tER > 38 %. The true error rate was computed from the univariate distribution estimated by normal kernels according to the criterion of maximal probability described above.

The correlation of the sono score in groups d(2) and d(32) to the biochemical variables is not significant ($p < 0.05$). Only in group d(32) a significant correlation between sono score and Lp(a) (rS = 0.34, n = 94) respectively Lp(a)/HDL-C (rS = 0.31, n = 94) has been found.

By stepwise KDA LDL-C/HDL-C and HDL-C/TC have been identified to be substantial for classification. Using these variables the estimated true error rate was reduced to 8.8 %. For this model the mean relation of classification Q2 = 6.1 and the mean probability Q1 = .138. Lp(a) as an additional variable seems not to be important (table 5.a).

Considering the »clinical significance« and the multivariate classification results, then there is no reason to merge these groups into one control group.

5. 4. The role of CAD

Within case group d(3) with cerebrovascular ischemic events we distinguished between a group d(31) (n31 = 62) without CAD and a group d(32) (n32 = 32) with CAD. The aim of a comparison between these groups is to study the possible additional influence of CAD. There is no significant difference ($p < 0.05$) between these groups for any variable. The estimated true error rate for single variables is always >42.5 %. By stepwise KDA this result could not be improved to an error rate <40.0 %.

From these results we conclude that CAD has no additional effect on the variables under consideration if CI is present. Therefore these groups were merged to one case group d(3) (n3 = 94) for the further analysis.

A comparison of subjects of the case group d(31) without CAD and the subjects of the case groups d (32) and d(4) – all of them are with CAD – shows LDL-C/HDL-C as the best single discriminator (aER = 37 %). Adding further variables, like Lp(a) or HDL-C would not improve the classification results substantially.

5. 5. Comparison of one control group with one case group

The main interest of this study from a medical viewpoint is the examination of the difference between the control group d(1) (n1 = 38) and the patient group with myocardial infarction (MI) and CAD d(4) (n4 = 42) and the identification of the best discriminators for these groups (table 5.b). A comparison of the location of these groups by a Wilcoxon-U-test shows a highly significant difference only for Lp(a) respectively Lp(a)/HDL-C. The difference of all other variables are not significant. If the analysis is restricted to the age groups of 50 to 75 years (n1 = 29, n4 = 33), then the locations of Lp(a), Lp(a)/HDL-C and also HDL-C are significantly different ($p < 0.01$). The estimated true error rate is 27.4 % for HDL-C (z = 2.35). Lp(a) is not correlated to any variable or ratio without Lp(a)/HDL-C in all groups.

By stepwise variable selection for KDA the next variable selected was LDL-C/HDL-C. The apparent error rate decreased to 8.1 % (aER = 4.8 %). If HDL-C is considered simultaneously with Lp(a), than tER = 33.9 %. The best classification results for these groups using KDA have been achieved for model order p = 3 using the following variables: HDL-C, LDL-C/HDL-C, HDL-C/TC: tER = 3.2 %, aER = 1.6 %, Q2 = 21.06. But there are also some other models of order p = 3 out of 9 possible variables which lead to very good discrimination and classification results. The additional importance of Lp(a) for multivariate discrimination of these groups is rather poor if HDL-C or LDL-C/HDL-C is present in the model (table 5.b).

If the location of the variables of the CI group d(3) (n3 = 94) is compared with the control group d(1) (n1 = 29) the similar variables, namely Lp(a), HDL-C and LDL-C/HDL-C, show significant differences ($p < 0.01$) (table 5.c). For LDL-C/HDL-C the true error rate becomes tER = 15.4 %. But the sensitivity and specificity for this variable are rather different, which may be due to the big difference of the sample size.

The location of the variables HDL-C, LDL-C and Lp(a)/HDL-C of groups d(3) and d(4) are also significant different ($p < 0.05$) from those of control group d(2).

Stepwise variable selection for KDA leads to a model of order p = 5 including the variables LDL-C/HDL-C, HDL-C/TC, TG, HDL-C and Lp(a). This model classifies only 2 of 123 subjects wrong (aER = 1.6 %). This seems to be an over-optimistic result, but on the other hand it shows the enormeous discriminatory power of biochemical variables in KDA. For the discrimination of control group d(2) from case group d(3) HDL-C, TC and also Lp(a) have been identified. But the error rates are rather high compared to the previous results (table 5.d).

5. 6. More than two groups DA

The examples above have shown that it is possible to achieve very good discrimination between two groups. But it is of special interest to which extent it is possible to differentiate simultaneously between more than two groups. For linear DA the main problem of more than two group classification is the determination of best cut-off points. In KDA, where the estimated multivariate probability of an observation in each class d(i) is used for classification, the determination of cut-off points is not important. An element is classified in this class d(i), for which its probability becomes a maximum.

A more important problem in the evaluation is the determination of an appropriate cost function. The misclassification of an observation of class d(i) into class d(j), $i \neq j$ has to be evaluated by a cost matrix C, whose elements must be determined subjectively. The values of the matrix C represent the cost of misclassification from a medical point of view in this study. If the criterion for variable selection is a cost function, then for different cost matrices C different »best« variable-combinations have to be expected.

There are many medical aspects which indicate a separate treatment of each group and therefore a simultaneous classification of four groups d(i), i = 1, . . . , 4 has been performed.

The costs of misclassification have been determined from from a medical point of view. For comparison, three different

cost matrices C_1 (unweighted), C_2 (symmetric) and C_3 (asymmetric) have been examined:

$$C_1 = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix} \quad C_2 = \begin{pmatrix} 0.0 & 0.5 & 2.0 & 2.5 \\ 0.5 & 0.0 & 1.5 & 2.0 \\ 2.0 & 1.5 & 0.0 & 1.0 \\ 2.5 & 2.0 & 1.0 & 0.0 \end{pmatrix} \quad C_3 = \begin{pmatrix} 0.0 & 0.5 & 1.5 & 1.5 \\ 0.5 & 0.0 & 1.0 & 1.0 \\ 2.0 & 1.5 & 0.0 & 0.5 \\ 2.5 & 2.0 & 0.5 & 0.0 \end{pmatrix}$$

In cost matrix C_3 the misclassification of a person of the case group into one of the control groups is weighted heavier than vice versa. A misclassification within the control or case groups seems to be not so important. A comparison of the stepwise selected variables with a symmetric cost matrix C_2 and an »unweighted« cost matrix C_1 (Tab. 6) resulted approximately in the same variables.

Using only single variables for classification, the true error rate is rather high (Tab. 2) even for variables like Lp(a) or HDL-C, which are significantly different. For more than two variables generally the discriminatory power of this kernel discriminant function is very high (tER = 25.0 % for p = 3). Only the subjects of class d(2) – the control group with sono score >0 – have been classified very frequently into one of the case groups. For each class d(j), j = 1, 2, 3, 4 the portion of the cost for cost matrix C_k:

$$co(k,j) = \sum_{j=1}^{4} w(i,j) * c_k(i,j) \qquad [11]$$

was computed. From table 6 we can see that the portion of the cost for group d(2), co(k,2) is always higher than the portion of the sample size (n2 = 28, 15.2 %) in this study.

6. Medical aspects and discussion

Univariate comparisons of possible risk factors for cerebrovascular ischemic events and myocardial infarction show that Lp(a) and the relation Lp(a)/HDL-C are very good discriminators. From the results of stepwise variable selection for kernel discriminant analysis we conclude that combinations of HDL-C, LDL-C, TC and especially ratios of these variables like LDL-C/HDL-C are important discrimators and the role of Lp(a) decreases, when some of these variables are present.

From the viewpoint of experimental design the good discrimination results between the two control groups have been surprising and show the importance of the selection of »healthy« control subjects. But considering the results of the four group DA, we can see a high error rate for class d(2) (table 6) than for any other class. Many of the subjects have been classified into one of the case groups, which may indicate that in group d(2) a high risk becoming a case exists. The good classification results for the other groups d(1), d(3) and d(4) indicate a high sensitivity of the selected biochemical variables.

Generally it was surprising for us that for most comparisons only three variables have been found to be substantial and the addition of further variables would not improve the model quality. This relatively low model order is of course to some extent due to the low sample size and for an application of the results for preventive risk determination the analysis of a larger sample is necessary. Furthermore, most of the biochemical variables except Lp(a) are highly correlated and the addition of this variables would not contain much information.

From tables 5 and 6 the bias of the over-optimistic estimation of the error rate by reclassification becomes obvious if aER and tER are compared. The model with the lowest error rate for one model order p must not be the best model for an other evaluation criterion. Model evaluation by other criteria than the error rate may lead to different suboptimal variable sets. Therefore the simultaneous computation of more than one model evaluation criterion for one model is recommended.

Acknowledgment

We thank Dr. G. BONE and Dr. K. NIEDERKORN from the University Clinic of Neurology in Graz and Dr. P. KÖLTRINGER from the hospital Barmherzige Brüder in Graz for sampling of the data.

Part of this work was supported by the Austrian Nation Bank, Project Nr. 2631 and by the Austrian Research Council, Project Nr. 5158.

Literature

1. AIRO, R., C. M. FERRARI, 1985: High Density Lipoproteins and Apolipoprotein A in Cerebrovascular Disease. Atherosclerosis, **54**, 343–346.
2. ALBERS, J. J., P. WAHL, W. R. HAZZARD, 1984: Quantitative Genetic Studies of the Human Plasma Lp(a) Lipoprotein. Biochem. Genet., **11**, 475–486.
3. BÜNNING, H., G. TRENKLER, 1978: Nichtparametrische Statistische Methoden. Walter de Gruyter, Berlin/New York.
4. HABBEMA, J. D. F., J. HERMANS, K. VAN DEN BROEK, 1974: A Stepwise Discriminant Analysis Program using Density Estimation. Compstat 1974, Proceedings in Computational Statistics (G. Bruckmann, Ed.) 101–110, Physica-Verlag, Vienna.
5. HAND, D. J., 1981: Discrimination and Classification. Wiley, Chichester.

Table 6. Comparison of four groups d(1) (n1 = 29, 16 %), d(2) (n2 = 28, 15 %), d(3) (n3 = 94, 51 %) and d(4) (n4 = 33, 18 %) by stepwise kernel discriminant analysis, estimated smoothing factors $\hat{h}(ii)$ of the kernel functions and results of the model evaluation for different cost matrices C_1, C_2, C_3 (for symbols see text).

	p = 1		p = 2		p = 3
Lp(a)	0.548	–	–	–	0.586
HDL-C	–	–	–	0.592	–
LDL-C	–	–	–	0.573	–
LDL-C/HDL-C	–	0.578	0.578	0.578	0.578
HDL-C/TC	–	–	0.557	–	0.557
aER	78.3	39.1	32.1	12.0	25.5
tER	80.4	44.6	35.3	25.0	37.0
Q1	0.341	0.310	0.135	0.037	0.047
Q2	–0.3	–0.1	0.3	1.3	0.2
co(1,1)	29.0	5.0	4.0	5.0	5.0
co(1,2)	7.0	28.0	26.0	17.0	17.0
co(1,3)	94.0	24.0	23.0	14.0	33.0
co(1,4)	18.0	25.0	12.0	10.0	13.0
$COST_1$	148.0	82.0	65.0	46.0	68.0
co(2,1)	34.5	11.5	9.5	4.0	10.0
co(2,2)	12.0	33.0	45.5	20.5	27.5
co(2,3)	126.5	33.5	24.0	14.0	35.5
co(2,4)	35.0	48.0	12.0	20.5	19.0
$COST_2$	208.0	126.0	91.0	59.0	92.0
co(3,1)	24.5	40.5	6.0	3.5	6.5
co(3,2)	6.5	26.0	26.0	12.5	16.5
co(3,3)	112.0	54.0	13.0	7.0	21.0
co(3,4)	34.5	17.0	6.0	17.5	15.5
$COST_3$	177.5	137.5	51.0	40.5	59.5

6. HILDEN, J., J. D. F. HABBEMA, B. BJERREGARD, 1978: The measurement of performance in probabilistic diagnosis II. Meth. Inform. Med., **17.**
7. JÜRGENS, G., G. M. KOSTNER, 1975: Studies on the Structure of the Lp(a) Lipoprotein. Isolation and Partial Characterisation of the Lp(a) Specific Antigen. Immunogenetics, **1,** 560–569.
8. KANNEL, W. B., T. GORDON, T. R. DAWBER, 1974: Role of Lipids in the Development of Brain Infarction: the Framingham Study. Stroke, **5,** 679–685.
9. LAURELL, C. B., 1972: Electroimmuno Assay. Scand. J. Clin. Lab. Invest. Suppl., **29,** 21–37.
10. MURAI, A., T. MIYAHARA, N. FUJIMOTO, M. MATSUDA, M. KAMEYAMA, 1986: Lp(a) Lipoprotein as a Risk Factor for Coronary Heart Disease and Cerebral Infarction. Atherosclerosis, **59,** 199–204.
11. MURAI, A., T. TANAKA, T. MIYAHARA, M. KAMEYAMA, 1981: Lipoprotein Abnormalities in the Pathogenesis of Cerebral Infarction and Transient Ischemic Attack. Stroke, **12,** 167–172.
12. NOMA, A., S. MATSUSHITA, T. KOMORI, K. ABE, H. OKABE, K. KURAMOTO, M. MURAKAMI, 1979: High and Low Density Lipoprotein Cholesterol in Myocardial and Cerebral Infarction. Atherosclerosis, **32,** 327–331.
13. PFEIFFER, K. P., 1985: Stepwise Variable Selection and Maximum Likelihood Estimation of Smoothing Factors of Kernel Functions for Nonparametric Discriminant Functions Evaluated by Different Criteria. Computers and Biomedical Research, **18,** 46–61.
14. PILGER, E., H. PRISTAUTZ, K. P. PFEIFFER, G. M. KOSTNER, 1983: Risk Factors for Peripheral Atherosclerosis. Retrospective Evaluation by Stepwise Discriminant Analysis. Arteriosclerosis, **3,** 57–63.
15. RÖSSNER, S., K. G. KJELLIN, K. L. METTINGER, A. SIDEN, C. E. SÖDERSTRÖM, 1978: Dyslipoproteinemia in Patients with Ischemic Cerebro-Vascular Disease. A Study of Stroke before the age of 55. Atherosclerosis, **30,** 199–209.
16. RÖSSNER, S., 1982: Serum Lipoproteins and Ischemic Vascular Disease. On the Interpretation of Serum Lipid versus Serum Lipoprotein Concentrations. J. Cardiovasc. Pharmacol., **4,** 201–205.
17. SCHRIEWER, H., G. ASSMANN, M. SANDKAMP, 1984: The Relationship of Lipoprotein (a) (Lp(a)) to Risk Factors of Coronary Heart Disease. J. Clin. Chem. Clin. Biochem., **22,** 591–596.
18. SIRTORI, R. C., G. GIANFRANCHESCHI, I. GRITTI, G. NAPPI, G. BRAMBILLA, P. PAOLETTI, 1979: Decreased High Density Lipoprotein-Cholesterol Levels in Male Patients with Transient Ischemic Attacks. Atherosclerosis, **32,** 205–211.
19. G. ZENKER, P. KÖLTRINGER, G. BONE, K. NIEDERKORN, K. P. PFEIFFER, G. JÜRGENS, 1986: Lipoprotein (a) as a Strong Indicator for Cerebrovascular Disease. Stroke, (in press).

Date of receipt: April 24th, 1986.
The authors' address: Univ.-Doz. Dipl.-Ing. Dr. Karl P. Pfeiffer, Physiologisches Institut, Karl-Franzens-Universität Graz, Harrachgasse 21/5, A-8010 Graz, Austria.

NACHRICHTEN/NEWS

1. Conference of the International Federation of Classification Societies

June 29 to July 1, 1987, Aachen, Federal Republic of Germany

The International Federation of Classification Societies (IFCS) is organizing its First Conference from June 29 to July 1, 1987 at the Technical University Aachen (RWTH Aachen), Aachen, Federal Republic of Germany. This Conference is devoted to the presentation of theoretical and applied papers on »Classifications« and related methods of data analysi in the broad sense. This includes mathematical and statistical as well as practical investigations, real case studies, papers on numerical and algorithmic aspects, applications in special fields of knowledge, and the interface between classification and the Information Sciences.

Papers are invited for this meeting. Suitable main topics include, e.g.:

1. Discrimination, classification, aggregation and clustering methods
2. Pattern recognition methods
3. Linear, nonlinear, and algebraic methods of data analysis
4. Similarity and distance measures, measurement theory, qualita of data
5. Multidimensional scaling and seriation
6. Probabilistic models for data analysis and classification
7. Graphical representation of structures and classifications
8. Biological taxonomy, systematics, microbiological classification (molecules, strings etc.)
9. Comparison of classifications, consensus methods
10. Analysis of tree- and graph-like patterns
11. Classification for information retrieval systems
12. Expert systems, data bases, and classification methods
13. Algorithmic aspects and software for computers and micro-computers
14. Applications in specific fields: Archeology, biology and medicine, chemistry, computer science, documentation and information sciences, economics and marketing, engineering and technology, psychology and social sciences etc.

If you plan to attend the conference please write to the conference chairman: Prof. Dr. H. H. Bock, IFCS 87, Institut für Statistik und Wirtschaftsmathematik, Technical University Aachen, Wuellnerstraße 3, D-5100 Aachen, FRG, who will provide all further information concerning the program and your registration.

If you intend to present a paper, please send an English abstract of at most one page, classified into one of the 14 subjects given above. A selection of papers will be published in a proceedings volume. Deadline for submitting papers is **January 15, 1987.**

4. Konferenz über die wissenschaftliche Anwendung von Statistik-Software

23. bis 26. März 1987, Mannheim, Bundesrepublik Deutschland

Die 4. Konferenz über die wissenschaftliche Anwendung von Statistik-Software wird diesmal wieder vom Zentrum für Umfragen, Methoden und Analysen (ZUMA) organisiert und durchgeführt. Tagungsort ist die Universität Mannheim.

Themenschwerpunkt sind Neuerungen bei SPSS, BMDP, SAS etc., Programmvergleiche und -bewertungen, Grafik-Einsatzmöglichkeiten, exploratorische Datenanalyse, Statistik-Software im PC-Bereich, Statistik-Expertensysteme, Statistik-Software in Forschung und Ausbildung.

Es wird gebeten, die Kurzfassungen der Beiträge bis zum 15. September 1986 an die folgende Tagungsadresse zu schicken:
Software-Konferenz
ZUMA e. V.
B 2,1
Postfach 59 69
6800 Mannheim 1.

Die Kurzbeiträge sollten so abgefaßt sein, daß sich der Programmausschuß ein hinreichend differenziertes Bild vom Inhalt des Beitrags machen kann. Eine Rückmeldung über die Annahme der Beiträge erfolgt bis zum 1. November 1986. Geplant ist auch diesmal wieder die Herausgabe eines Tagungsbandes, in dem qualifizierte Beiträge veröffentlicht werden sollen.

Konferenzsprachen sind Deutsch und Englisch. Der Tagungsbeitrag beläuft sich auf 75,– DM.

BUCHBESPRECHUNGEN/BOOK REVIEWS

KELTER, U.
Parallele Transaktionen in Datenbanksystemen
Reihe Informatik, Bd. 51
1985, 207 S., DM 28,–
Bibliographisches Institut, Mannheim–Wien–Zürich

Thema des Buches ist das sogenannte Concurrency-Control-Problem für zentrale Datenbanken: Wie müssen parallele Transaktionen in einer zentralen Datenbank gesteuert werden, ohne daß gewisse Anomalien, z. B. Verlust von Änderungen oder Inkonsistenzen, eintreten? Ein sehr ähnliches Problem tritt bei parallel genutzten Datenstrukturen in parallelen Programmen auf, z. B. in der Betriebssystemprogrammierung. Zur besseren Übertragbarkeit der Ergebnisse wird in diesem Buch möglichst wenig auf spezielle Strukturen von Datenbanken Bezug genommen, statt dessen wird das Hauptgewicht auf allgemeine Prinzipien, Methoden und Techniken gelegt. Behandelt werden sowohl die analytisch/theoretischen Aspekte des Problems als auch alle wesentlichen praktischen Verfahren, nämlich Sperr-, Zeitstempel- und optimistische Verfahren für zentrale Datenbanken. Der Bereich Recovery in Datenbanken, der in enger Wechselwirkung mit dem Bereich Concurrency-Control steht, wird in einer ausführlichen Übersicht vorgestellt.

MAURER, H. (Hrsg.)
Überblicke Informationsverarbeitung 1985
1985, 235 S., DM 44,–
Bibliographisches Institut, Mannheim–Wien–Zürich

Dieser Band umfaßt wieder eine Reihe von Beiträgen zu aktuellen Themen, insbesondere zu Bildschirmtext. Behandelt werden u. a. der Einfluß der Hochintegration (VLSI) auf die Rechnerarchitektur, neuere Entwicklungen des rechnerunterstützten Lernens, die Verwendung von formalen Potenzreihen im Bereich der Automaten und der formalen Sprachen, Telesoftware, ein Beitrag über Logo sowie der Entwicklungsstand der anwendungsorientierten Softwaretechnologie. Den Abschluß bildet ein historischer Beitrag über Ideen für die Speicherung von Daten in Computern. Ge.

WRICKE, G. und WEBER, W. E.:
Quantitative Genetics and Selection in Plant Breeding
1986, 406 S., DM 198,–
Walter de Gruyter, Berlin–New York

Das Buch setzt sich zum Ziel, die Prinzipien der Selektion in der Pflanzenzüchtung für Studenten, Pflanzenzüchter und Angewandte Genetiker darzustellen. Es gliedert sich in 13 Kapitel: In den ersten vier Kapiteln werden auf 140 Seiten die Grundlagen der Populationsgenetik, der Quantitativen Genetik und des Schätzens von Variabilitätsparametern beschrieben. Die Konzepte der Selektion und ihre Anwendung bei Fremd- und Selbstbefruchtern beanspruchen die nächsten 80 Seiten. Spezielle Kapitel, je mit 15 bis 35 Seiten, sind der Selektion bei synthetischen Sorten, bei Hybridsorten, in frühen Generationen von Selbstbefruchtern, bei Autotetraploiden und der Selektion auf mehrere Merkmale sowie den Selektionsgrenzen gewidmet. In einem statistischen Anhang (25 S.) werden die wichtigsten Verteilungen, Tabellen und Formeln zur Matrixalgebra sowie zu linearen Modellen bei fixierten und zufälligen Effekten zusammengestellt.

Schon aus dem Umfang des Werkes wird deutlich, daß die Autoren sich nicht auf die grundlegenden Elemente des Stoffes beschränkt, sondern eine mehr enzyklopädische Darstellung bevorzugt haben. Mit über 320 Referenzen haben sie eine Fülle von Material verarbeitet, wobei vorwiegend neuere Literatur zitiert wird. Dieses Auswahlkriterium kommt dem heutigen Leser eines Fachbuches entgegen, der sich ja meist auf einem bestimmten Sektor einen Überblick über den neuesten Stand verschaffen möchte.

Für den Leser dieser Zeitschrift mag von Interesse sein, daß die wichtigsten hierarchischen und faktoriellen Pläne der Pflanzenzüchtung angesprochen sind und naturgemäß die Anwendung von Varianzkomponenten einen breiten Raum einnimmt; doch kann das Werk ein entsprechendes Statistikbuch auf diesem Gebiet nicht ersetzen – und will es sicherlich auch nicht.

Der Druck ist exzellent. Der Text ist durch viele anschauliche Tabellen, Grafiken oder numerische Beispiele aufgelockert, so daß man das Buch gerne zur Hand nimmt. Gelegentliche Druckfehler, soweit überhaupt festzustellen, behindern das Lesen nicht. So stellt das Werk ein empfehlenswertes Handbuch für jeden wissenschaftlich arbeitenden Pflanzenzüchter, insbesondere für jeden Diplomanden und Doktoranden, dar. Es dürfte schnell zur Standardreferenz auf dem Gebiet werden. Nur der Preis mag einer weiten Verbreitung etwas im Wege stehen. H. F. Utz

LORENZ, R. J.
Grundbegriffe der Biometrie
1984, 241 S., DM 48,–
G. Fischer Verlag, Stuttgart

Die Anforderungen an ein Lehrbuch der Biometrie haben sich in den letzten Jahrzehnten mit der zunehmenden Verbreitung der Computer wesentlich verändert. Standen früher Rechenanweisungen im Mittelpunkt, so sind es heute Hilfen für eine zweckentsprechende Wahl eines Verfahrens und Hilfen für die Interpretation der Ergebnisse.

In diesem ersten Band einer Buchreihe „Biometrie" hat es der Autor verstanden, sein in vielen Kursen gezeigtes pädagogisches Geschick in eine lesbare Form zu bringen.

Das Ergebnis ist ein wirklich empfehlenswertes Lehrbuch neuen Stils der Biometrie. Ge.

HARTUNG, J. und ELPELT, B.
Multivariate Statistik
Lehr- und Handbuch der angewandten Statistik
1984, 820 S., DM 98,–
R. Oldenbourg Verlag, München – Wien

Zunächst muß die mühsame Arbeit der Zusammenstellung der Vielzahl von Methoden und Verfahren vorbehaltlos anerkannt werden. Es zeigt sich dabei aber auch, daß ein umfassendes Buch über statistische Verfahren einen entsprechenden Umfang haben muß.

Das vorliegende Buch ist zwar ein Lehrbuch, aber vielmehr ein Handbuch für jeden, der mit der Auswertung und Interpretation komplexen Datenmaterials beschäftigt ist. Erfreulich ist dabei, daß die dargestellten Verfahren durch instruktive Beispiele und viele Abbildungen erläutert werden. Insgesamt ein empfehlenswertes Werk, das in keinem biometrischen Labor fehlen sollte. Ge.

FILLBRANDT, H.
Verteilungsfreie Methoden in der Biostatistik
EDV-Programmband
Mit einem Geleitwort von G. A. Lienert
1986, 416 S., DM 68,–
Verlag A. Hain Meisenheim GmbH, Königstein/Ts.

Das Fehlen von Computerprogrammen, speziell für verteilungsfreie Methoden, bedingt, daß vielfach nicht vertretbare Verfahren benutzt werden. Diese Lücke versucht der vorliegende Programmband zu schließen.

Nun muß man aber bedenken, daß heute als Arbeitshilfsmittel PCs oder Bildschirmterminals an Großcomputern eingesetzt werden. Die »Anwender« erwarten dabei eine leichte »Menuesteuerung«. Die dazu erforderlichen Programme sollten auf einem Datenträger (Disketten oder Magnetband) in einer »modernen« Programmiersprache verfügbar sein. Diesen Erfordernissen entsprechen die Listings von ALGOL-60-Programmen, die zum Teil schon über 10 Jahre alt sind, leider nicht. Es ist zu hoffen, daß die erstellten Programme die Grundlage für ein wirklich brauchbares Programmpaket bilden, das dann sicher auch zur vermehrten Nutzung der verteilungsfreien Methoden beitragen wird. Ge.