



R-COD-Browser

(Mortality **R**isks for all **C**auses of **D**eath)

**Webbasierte Software zur Ermittlung von Mortalitätsraten in Deutschland
für beliebige Todesursachen nach ICD-10 im Zeitraum 1998-2003**

- Arbeitspapier -

Version 1.0

Stand: 07.06.2006

Thema: R-COD-Browser (Mortality Risks for all Causes of Death):

Webbasierte Software zur Ermittlung von Mortalitätsraten in Deutschland für beliebige Todesursachen nach ICD-10 im Zeitraum 1998-2003

Kontakt:

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG)

Dillenburger Str. 27

51105 Köln

Internet: www.iqwig.de

Tel: 0221/35685-0

Fax: 0221/35685-1

Bemerkung:

Im folgenden Text wurde bei der Angabe von Personenbezeichnungen jeweils die männliche Form angewandt. Dies erfolgte ausschließlich zur Verbesserung der Lesbarkeit.

Inhaltsverzeichnis

	Seite
Inhaltsverzeichnis	2
Tabellenverzeichnis	3
Abbildungsverzeichnis.....	3
Abkürzungsverzeichnis	3
1. Ziel des Projekts.....	4
2. Hintergrund.....	4
3. Methoden	5
3.1 Arbeitsgruppe.....	5
3.2 Datengrundlagen.....	6
3.3 Programmierung	6
3.4 Berechnungen	7
3.5 Ergebnisüberprüfungen.....	7
4. Beschreibung des R-COD-Browsers	9
5. Diskussion und Ausblick	12
6. Zusammenfassung	13
7. Literatur	13
Anhang: Weiterer Zeitplan.....	14

Tabellenverzeichnis

Tabelle 1	Ausgewählte Klassen zur Ergebnisüberprüfung des R-COD-Browsers
Tabelle 2	Sterbefälle je 100000 Einwohner nach ICD-Kapiteln ermittelt aus der GBE-BUND-Seite [4]

Abbildungsverzeichnis

Abbildung 1	Startbildschirm des R-COD-Browsers
Abbildung 2	Anzeige nach Wahl der Gruppe I60 der zerebrovaskulären Krankheiten
Abbildung 3	Mortalitätsraten für die Todesursache der zerebrovaskulären Krankheiten
Abbildung 4	Mortalitätsraten für Todesursache der zerebrovaskulären Krankheiten getrennt nach Altersgruppen für das Jahr 2003

Abkürzungsverzeichnis

Abkürzung	Bedeutung
ASP	Active Server Pages
DIMDI	Deutsches Institut für Medizinische Dokumentation und Information
GBE-BUND	Gesundheitsberichterstattung des Bundes
G-BA	Gemeinsamer Bundesausschuss
HTML	Hypertext Markup Language
ICD	International Statistical Classification of Diseases and Related Health Problems
ICD-9	International Statistical Classification of Diseases and Related Health Problems – Ninth Revision
ICD-10	International Statistical Classification of Diseases and Related Health Problems – Tenth Revision
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
R-COD	Mortality Risks for all Causes of Death
RKI	Robert Koch Institut
SQL	Structured Query Language
WHO	World Health Organization
XMS	Extensible Markup Language

1. Ziel des Projekts

Ziel des Projekts ist die Entwicklung einer webbasierten Software, mit deren Hilfe Mortalitätsraten pro Kalenderjahr für beliebige Todesursachengruppen auf der Basis von ICD-Codes für Deutschland ermittelt und übersichtlich dargestellt werden können (R-COD-Browser). Die Mortalitätsraten sollen getrennt nach Kalenderjahr und wahlweise gemeinsam oder getrennt nach Geschlechtern sowie nach Altersgruppen dargestellt werden. Die Software soll webbasiert entwickelt werden, damit das Tool im Internet von interessierten Wissenschaftlern aus den Bereichen Epidemiologie und Public Health verwendet werden kann.

2. Hintergrund

Das Statistische Bundesamt sammelt im Rahmen der amtlichen Todesursachenstatistik Informationen über alle in Deutschland auftretenden Todesfälle. Insbesondere werden auch die auf den Totenscheinen dokumentierten primären Todesursachen in der jeweils gültigen Kodierung nach der Internationalen Statistischen Klassifikation der Krankheiten und verwandter Gesundheitsprobleme (ICD) gesammelt. Seit 1998 ist hierbei in Deutschland die 10. Version der Internationalen Statistischen Klassifikation der Krankheiten und verwandter Gesundheitsprobleme (ICD-10) gültig [1]. Die zusammengefassten Daten der deutschen amtlichen Todesursachenstatistik werden von der World Health Organization (WHO) in die internationale WHO Mortalitätsdatenbank eingepflegt. Diese Daten sind im Internet verfügbar [2] und können somit für wissenschaftliche Zwecke verwendet werden. Diese von der WHO zum Download bereitgestellten Datenbanken sind allerdings sehr umfangreich und müssen je nach Verwendungszweck mehr oder weniger stark aufbereitet werden.

Für viele epidemiologische Fragestellungen ist es hilfreich für interessierende Todesursachen nationale Mortalitätsraten zu kennen. Die entsprechende Information ist im Prinzip in den WHO Datenbanken verfügbar, aber nicht ohne weiteres abrufbar. Um einen leichteren Zugriff auf diese Daten zu ermöglichen,

wird im Rahmen dieses Projekts eine webbasierte Software – der R-COD-Browser – entwickelt, mit deren Hilfe Mortalitätsraten pro Kalenderjahr für beliebige Todesursachengruppen auf der Basis von ICD-Codes für Deutschland ermittelt und übersichtlich dargestellt werden können.

Speziell für maligne Erkrankungen gibt es bereits ein solches Tool, das vom Robert Koch Institut (RKI) im Internet zur Verfügung gestellt wird [3]. Dieses RKI-Tool ist aber auf bösartige Neubildungen beschränkt und basiert noch auf ICD-9. Jährliche Mortalitätsraten für Deutschland in Abhängigkeit nach Geschlecht, Alter und Kalenderjahr nach ICD-10 lassen sich über die Webseite der Gesundheitsberichterstattung des Bundes (GBE-BUND) berechnen [4]. Über die GBE-BUND-Seite lassen sich die Sterberisiken aber nur für anhand der ICD-10-Systematik vorgegebene Todesursachengruppen berechnen und nicht für einzelne Todesursachen oder für Gruppierungen über verschiedene Kapitel hinweg. Der R-COD-Browser dagegen deckt sämtliche Todesursachen nach ICD-10 ab, wobei eine freie Wahl der Todesursachengruppierung möglich ist.

3. Methoden

3.1 Arbeitsgruppe

Die Programmierung des R-COD-Browsers wurde von Frau Arzu Dursun im Jahr 2005 im Rahmen einer Diplomarbeit [5] im Studiengang "*Angewandte Mathematik*" der Fachhochschule Koblenz durchgeführt (Erstgutachter: Prof. Dr. Manfred Berres, Zweitgutachter: Prof. Dr. Jürgen Krämer, externer Gutachter: PD Dr. Ralf Bender, Abgabe der Arbeit: November 2005). Die Arbeit wurde betreut durch PD Dr. Ralf Bender (IQWiG, Köln) für den Bereich Biostatistik sowie durch Dr. Guido Noelle (damals IQWiG, Köln) für den Bereich Medizinische Informatik. Ferner war Dr. Jürgen Wellmann (Institut für Epidemiologie und Sozialmedizin, Universität Münster) für den Bereich Epidemiologie am Projekt beteiligt.

3.2 Datengrundlagen

Die erste Datenbasis des R-COD-Browsers bezüglich der Todesursachen stellen die Codes nach ICD-10 dar. Die deutsche Version von ICD-10 wird vom Deutschen Institut für Medizinische Dokumentation und Information (DIMDI) zur Verfügung gestellt [6]. Die deutsche Version von ICD-10 (Version 2004) wurde von der Webseite des DIMDI in Form von Metadaten entnommen und für die weitere Softwareentwicklung in den R-COD-Browser eingepflegt.

Die zweite Datenbasis des R-COD-Browsers bezüglich der Mortalitätsdaten stellen die Anzahlen der Verstorbenen sowie die der Bevölkerung von Deutschland aus der WHO Mortalitätsdatenbank [2] dar. Anfang 2005 standen hier für ICD-10 Daten der Jahre 1998-2001 zur Verfügung. Diese Daten wurden vom Institut für Epidemiologie und Sozialmedizin der Universität Münster in Form von SAS-Dateien aufbereitet, um z.B. für epidemiologische Auswertungen verwendet werden zu können. Diese Dateien wurden nach Microsoft Access überführt und in den R-COD-Browser integriert.

Die dritte Datenbasis des R-COD-Browsers bezüglich der Mortalitätsdaten stellen die Anzahlen der Verstorbenen sowie die der Bevölkerung von Deutschland für die Jahre 2002 und 2003 dar. Diese wurden der amtlichen Todesursachenstatistik des Statistischen Bundesamtes entnommen, die als Excel-Files zur Verfügung stehen [7,8]. Diese Daten wurden nach Microsoft Access überführt, mit Hilfe von SQL-Abfragen in die Struktur der Daten aus den Jahren 1998-2001 gebracht und in den R-COD-Browser integriert.

3.3 Programmierung

Im R-COD-Browser stellt Microsoft Access die grundlegende Datenbank dar. Zur Softwareentwicklung wurde die Hypertext Markup Language (HTML) mit JavaScript verwendet. Dynamische Webseiten wurden hierbei mit Hilfe von Active Server Pages (ASP) erzeugt. Für den Transfer und die Speicherung der benötigten Daten wurde die Extensible Markup Language (XML) eingesetzt. Eine ausführliche Beschreibung der technischen Grundlagen gibt Dursun [5].

3.4 Berechnungen

Die wesentlichen Berechnungen im R-COD-Browser bestehen – in Abhängigkeit von Todesursachengruppe, Kalenderjahr, Geschlecht und Altersgruppe – aus der durch Summierung gebildeten Anzahl der Gestorbenen (G) sowie der Personenzahl (P) der entsprechenden Bevölkerung. Hieraus wird die rohe Mortalitätsrate (R) pro 100000 berechnet durch

$$R = \frac{G}{P} \times 100000$$

3.5 Ergebnisüberprüfungen

Während der Programmierung des R-COD-Browsers wurden laufend Überprüfungen der Berechnungen durchgeführt. Dabei wurde von allen Jahren, allen Kapiteln und allen Gruppen mindestens eine Krankheitsklasse als Stichprobe entnommen und die Ausgaben mit den entsprechenden Daten in der Datenbank verglichen. Auftretende Fehler wurden in den Datenbankabfragen korrigiert. In der letzten Version des Browsers (Version 1.0) stimmten in allen unsystematisch durchgeführten Stichproben die ausgegebenen Ergebnisse mit den Daten der grundlegenden Datenbank überein.

Um das Einlesen der Daten aus der Datenbank an den richtigen Stellen im Browser bei der endgültigen Version 1.0 des R-COD-Browsers systematisch zu überprüfen, wurden nochmals gezielte Stichprobenvergleiche durchgeführt. Es wurden aus vier willkürlichen Kapiteln jeweils 2-3 Klassen aus verschiedenen Gruppen eines Kapitels für die Jahre 1998 und 2003 ausgewählt. Tabelle 1 zeigt die ausgewählten Kapitel und Klassen, die zur Überprüfung herangezogen wurden. Alle Raten für diese Klassen stimmten mit den Daten in der Datenbank überein.

Zum Schluss wurde auf Kapitelebene für alle Jahre ein Vergleich mit den Ergebnissen der GBE-BUND-Seite [4] vorgenommen. Sämtliche Mortalitätsraten der GBE-BUND-Seite (Tabelle 2) konnten auch mit dem R-COD-Browser erzeugt werden mit Ausnahme von Kapitel 19 (S00-T89), da diese ICD-Codes in

der WHO Mortalitätsdatenbank in den Jahren 1998-2001 nicht enthalten sind. Es ist somit insgesamt davon auszugehen, dass sowohl die Datenbanken als auch die Berechnungen des R-COD-Browsers bei der Version 1.0 korrekt sind.

Tab. 1: Ausgewählte Klassen zur Ergebnisüberprüfung des R-COD-Browsers

Nr.	Kapitel und Klasse	ICD-Code
1. Kapitel: Bestimmte infektiöse u. parasitäre Krankheiten (A00-B99)		
1	Infektiöse Darmkrankheiten	A02.0
2	HIV-Krankheit	B20.0
3	Sonstige Infektionskrankheiten	B99
2. Kapitel: Neubildungen (C00-DD48)		
4	Bösartige Neubildungen der Verdauungsorgane	C15.1
5	Bösartige Neubildungen der Brustdrüse	C50.0
10. Kapitel: Krankheiten des Atmungssystems (J00-J99)		
6	Akute Infektionen der oberen Atemwege	J03.9
7	Lungenkrankheiten durch exogene Substanzen	J61
18. Kapitel: Symptome und abnorme klinische und Laborbefunde (R00-R99)		
8	Symptome, die das Kreislaufsystem und das Atmungssystem betreffen	R04.0
9	Allgemeinsymptome	R54

Tab. 2: Sterbefälle je 100000 Einwohner nach ICD-Kapiteln ermittelt aus der GBE-BUND-Seite [4]

Kapitel mit Codes nach ICD-10	Kalenderjahr					
	1998	1999	2000	2001	2002	2003
1) A00 - B99	9.5	11.1	12.3	12.6	12.7	13.2
2) C00 - D48	266.3	263.5	263.3	258.8	261.2	260.3
3) D50 - D90	2.0	2.0	2.2	2.3	2.3	2.5
4) E00 - E90	28.3	28.4	28.8	29.6	32.0	33.0
5) F00 - F99	11.8	11.6	10.5	9.5	9.6	10.3
6) G00 - G99	19.4	19.1	19.6	20.2	21.5	22.4
7) H00 - H59	0.0	0.0	0.0	0.0	0.0	0.0
8) H60 - H95	0.0	0.0	0.0	0.0	0.0	0.0
9) I00 - I99	501.5	494.7	480.7	475.7	477.4	480.6
10) J00 - J99	59.8	62.7	63.0	58.9	65.0	70.3
11) K00 - K93	49.4	48.9	49.5	49.7	50.7	51.2
12) L00 - L99	0.6	0.5	0.5	0.5	0.5	0.6
13) M00 - M99	2.3	2.2	2.2	2.3	2.3	2.4
14) N00 - N99	13.1	13.1	13.0	13.6	14.9	16.0
15) O00 - O99	0.1	0.0	0.1	0.0	0.0	0.0
16) P00 - P96	2.0	2.0	1.8	1.8	1.8	1.7
17) Q00 - Q99	2.2	2.0	2.0	2.0	1.9	2.0
18) R00 - R99	28.8	27.5	28.9	27.1	24.9	26.3
19) S00 - T98	42.2	41.5	42.0	41.5	41.6	41.9
20) V01 - Y98	42.2	41.5	42.0	41.5	41.6	41.9

4. Beschreibung des R-COD-Browsers

Nach Aufruf des R-COD-Browsers erscheint der Startbildschirm, der aus zwei nebeneinander gesetzten Frames besteht (siehe Abbildung 1). Auf der linken Seite wird die Systematik von ICD-10 in einer anklickbaren Baumstruktur angezeigt. Die Kapiteltitle bilden nach Kapitelnummer geordnet die äußeren Punkte der Baumstruktur und sind durch davorstehende Plus-Bullet-Zeichen gekennzeichnet. Durch Anklicken erscheinen die ICD-Gruppen innerhalb der Kapitel.

Im Hauptteil des Startbildschirms wird eine Suchfunktion angeboten, mit dem sich bestimmte Krankheitsbilder suchen lassen. In einem weiteren Fenster werden direkt alle einzelnen ICD-Codes zur Auswahl angeboten, so dass Benutzer, die mit den ICD-10-Codes vertraut sind, direkt die gewünschte Todesursachengruppe zusammenstellen können.

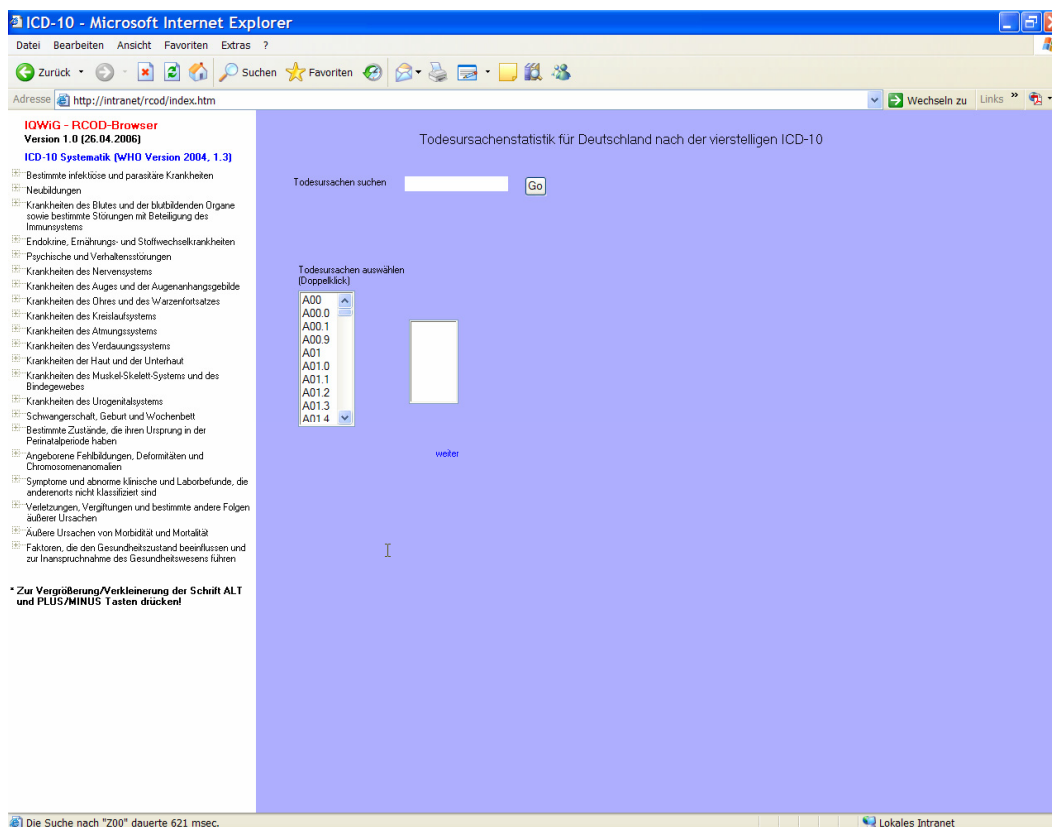


Abb. 1: Startbildschirm des R-COD-Browsers

Bei Verwendung des linken Teils erscheinen beim Anklicken einer Gruppe im Hauptteil des Bildschirms alle verfügbaren ICD-Codes dieser Gruppe. In Abbildung 2 wird diese Funktion am Beispiel der Gruppe I60 der zerebrovaskulären Krankheiten illustriert. Die einzelnen ICD-Codes lassen sich nun durch Anklicken der Kästchen frei kombinieren. Alternativ kann auch das gesamte Kapitel oder die gesamte Gruppe durch Anklicken ausgewählt werden.

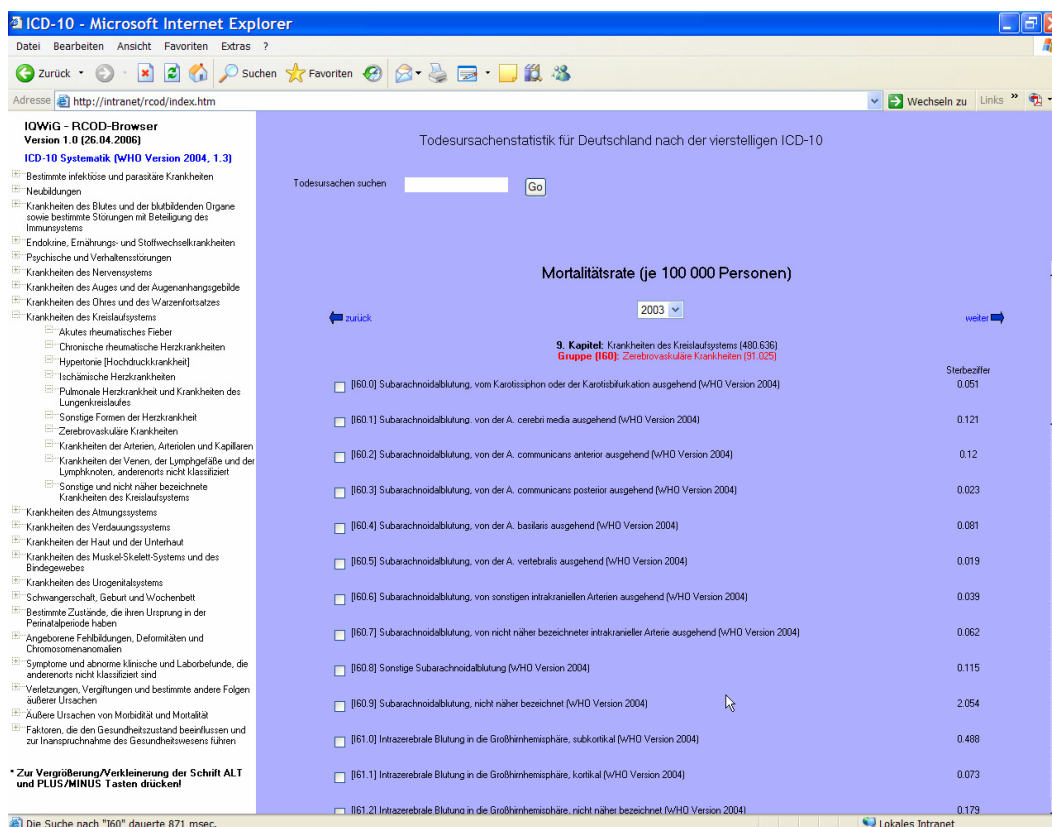


Abb. 2: Anzeige nach Wahl der Gruppe I60 der zerebrovaskulären Krankheiten

Wird z.B. die gesamte Gruppe I60 der zerebrovaskulären Krankheiten ausgewählt, so öffnet sich ein neues Fenster, in dem die Mortalitätsraten in Deutschland für die Todesursache der zerebrovaskulären Krankheiten für die Jahre 1998-2003 für beide Geschlechter getrennt und gemeinsam angezeigt werden (Abbildung 3). Die Jahre sind hierbei wiederum anklickbar. Nach Klick auf ein gewünschtes Jahr werden die Mortalitätsraten getrennt nach Altersgruppen angezeigt (Abbildung 4).

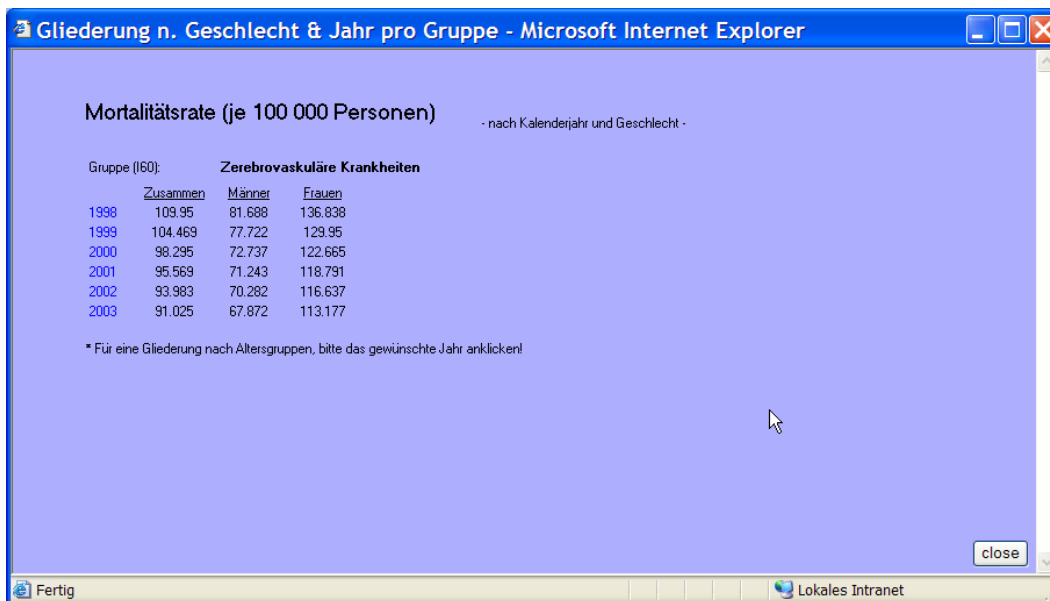


Abb. 3: Mortalitätsraten für die Todesursache der zerebrovaskulären Krankheiten

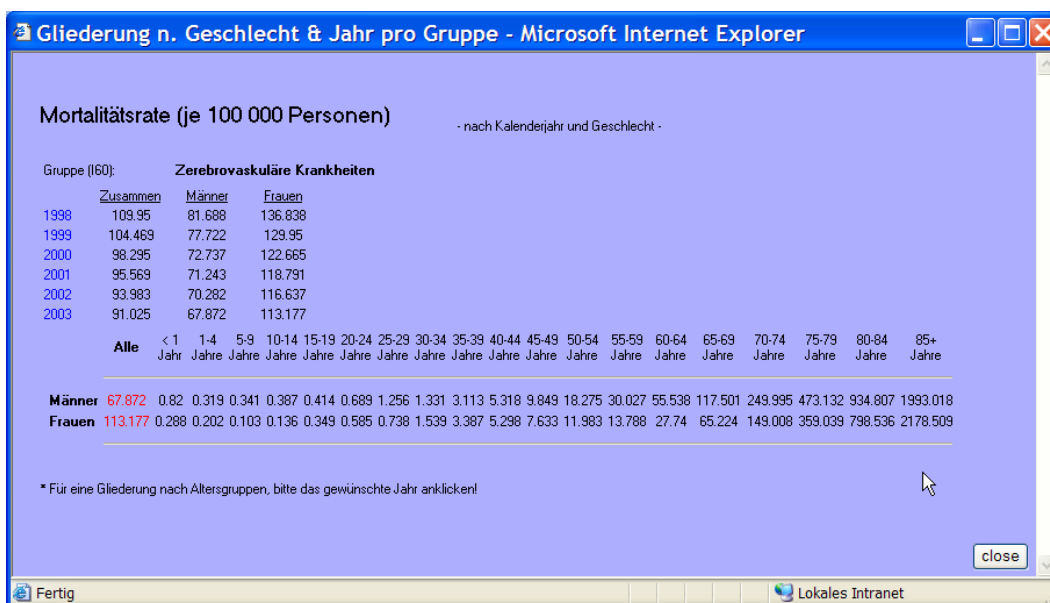


Abb. 4: Mortalitätsraten für Todesursache der zerebrovaskulären Krankheiten getrennt nach Altersgruppen für das Jahr 2003

5. Diskussion und Ausblick

Der R-COD-Browser ist in der Lage, die in Deutschland in den Jahren 1998-2003 beobachteten Mortalitätsraten für beliebig kombinierbare Todesursachengruppen nach ICD-10 in Abhängigkeit von Kalenderjahr, Geschlecht und Alter darzustellen. In der vorliegenden Version 1.0 ist es aber nur möglich, die jährlichen Mortalitätsraten zu berechnen und auszugeben.

Eine interessante Erweiterung der Methodik stellt die Möglichkeit zur Berechnung von Mortalitätsraten über mehrere Jahre dar (z.B. 5-Jahres-Risiken). Solche Berechnungen sind methodisch deutlich schwieriger, da hierbei das Problem der konkurrierenden Risiken berücksichtigt werden muss [9]. Aufgrund der vorliegenden Datengrundlage, bei der in Abhängigkeit von Kalenderjahr, Geschlecht und Alter nur die jährlichen Todesfälle und die dazugehörige Bevölkerungsgröße bekannt sind, können die üblichen Methoden der Überlebenszeitanalyse für konkurrierende Risiken [10] nicht angewendet werden, da diese voraussetzen, dass personenbezogene Informationen über eine klar definierte Kohorte für mehrere Jahre zur Verfügung stehen.

Es gibt spezielle methodische Ansätze, um Mortalitätsraten über mehrere Jahre mit Berücksichtigung konkurrierender Risiken auch auf der Basis amtlicher Mortalitätsdaten zu berechnen [9]. Zur Schätzung von Krebsmortalitäten wurden in den letzten Jahren neuere Verfahren entwickelt [11,12]. Diese Methoden werden z.B. bei der Software DevCan des National Cancer Instituts angewendet [13]. Eine Anpassung und Implementierung dieser Methoden in den R-COD-Browser war im Rahmen der Diplomarbeit von Frau Dursun, die primär die technische Umsetzung des R-COD-Browsers mit modernen Verfahren der Medizinischen Informatik zum Ziel hatte, nicht möglich. In einer für die Zukunft geplanten Version 2.0 des R-COD-Browsers sollen Methoden zur Berechnung von Mortalitätsraten über mehrere Jahre jedoch implementiert werden.

Zudem sollen in regelmäßigen Abständen die Daten der amtlichen Todesursachenstatistik für die Jahre ab 2004 in den R-COD-Browser eingepflegt werden.

6. Zusammenfassung

Mit Hilfe des R-COD-Browsers können jährliche Mortalitätsraten für Deutschland im Zeitraum 1998-2003 für beliebig kombinierbare Todesursachengruppen nach ICD-10 in Abhängigkeit von Kalenderjahr, Geschlecht und Alter berechnet und übersichtlich dargestellt werden. Die Möglichkeit der freien Kombinierbarkeit sämtlicher Todesursachen nach ICD-10 bietet sonst kein anderes der zurzeit verfügbaren Tools. Im Moment kann das Tool im Intranet des IQWiG verwendet werden (<http://intranet/rcod/>). Nach Einstellung in die Internet-Seite des IQWiG steht das Tool weltweit allen wissenschaftlichen Instituten sowie interessierten Wissenschaftlern für epidemiologische Zwecke zur Verfügung (siehe Anhang).

7. Literatur

- [1] World Health Organization. *International Statistical Classification of Diseases and Related Health Problems. Tenth Revision (ICD-10). Volume 1: Tabular List*. Geneva: World Health Organization, 1992.
- [2] World Health Organization. *Mortality Data Base* [online].
Available from: <http://www3.who.int/whosis/menu.cfm?path=whosis,mort&language=english>.
Zugriff am 24.05.2006.
- [3] Robert Koch Institut. *Alterspezifische Mortalität von Krebskrankheiten* [online].
Available from: http://www.rki.de/cln_006/nn_226994/DE/Content/GBE/DachdokKrebs/Datenbankabfragen/Altersverteilung/altersverteilung__node.html__nnn=true. Zugriff am 24.05.2006.
- [4] Gesundheitsberichterstattung des Bundes. *Sterbefälle je 100.000 Einwohner (ab 1998, Region, Alter, Geschlecht, ICD-10)* [online].
Available from: http://www.gbe-bund.de/oowa921-install/servlet/oowa/aw92/WS0100/_XWD_PROC?_XWD_2/2/xs_standard_neu/F.IND326. Zugriff am 30.05.2006.
- [5] Dursun A. *Entwicklung einer webbasierten Software zur Ermittlung von Sterberisiken für beliebige Todesursachen*. Diplomarbeit im Studiengang Mathematik, RheinAhrCampus Remagen, Fachhochschule Koblenz, 2005.
- [6] Deutsches Institut für Medizinische Dokumentation und Information. *Internationale Statistische Klassifikation der Krankheiten und verwandter Gesundheitsprobleme, 10. Revision, Version 2006. Vollständige amtliche Ausgabe, vierstellige ausführliche Systematik* [online], 2006.
Available from: <http://www.dimdi.de/dynamic/de/klassi/diagnosen/icd10/htmlamt12006/fr-icd.htm>.
Zugriff am 23.05.2006.

- [7] Statistisches Bundesamt. *Todesursachenstatistik 2002, Stand: 31.12.2002* [CD-ROM]. Wiesbaden: Statistisches Bundesamt, 2004.
- [8] Statistisches Bundesamt. *Todesursachen in Deutschland, Fachserie 12, Reihe 4, 2003* [Excel-File]. Wiesbaden: Statistisches Bundesamt, 2005.
- [9] Chiang CL. Competing risks in mortality analysis. *Annu. Rev. Public Health* 1991: 12, 281-307.
- [10] Tsiatis AA. Competing risks. In: Armitage P, Colton T (eds): *Encyclopedia of Biostatistics*, pp. 1025-1035. Chichester: Wiley, 2005.
- [11] Fay MP, Pfeiffer R, Cronin KA, Le C, Feuer EJ. Age-conditional probabilities of developing cancer. *Stat. Med.* 2003: 22, 1837-1848.
- [12] Fay MP. Estimating age conditional probability of developing disease from surveillance data. *Popul. Health Metr.* 2004: 2, 6.
- [13] National Cancer Institute. *DevCan - Probability of Developing or Dying of Cancer* [online].
Available from: <http://srab.cancer.gov/devcan/>. Zugriff am 24.05.2006.

Anhang: Weiterer Zeitplan

Arbeitsschritt	Termin (geplant)
Weitergabe des Arbeitspapiers an den G-BA	Juni 2006
Veröffentlichung des Arbeitspapiers und des R-COD-Browsers im Internet	Vier Wochen nach Weitergabe an den G-BA