

2 Methoden

Diese Arbeit beruht auf den Daten von 1289 AVM-Patienten aus drei unterschiedlichen Datenbanken. Der größte der drei Datensätze ist eine multizentrische AVM-Datenbank. Sie enthält Daten verschiedener Zentren: dem Hôpital de Bicêtre, Le Kremlin Bicêtre / Frankreich (n=393, im weiteren als „Paris“ bezeichnet), dem Universitätsklinikum Benjamin Franklin, Berlin / Deutschland (n=116, „Berlin“), sowie von weiteren Orten in verschiedenen Ländern Asiens (n=153), Riad / Saudi-Arabien (n=30), Bangkok / Thailand (n=61) und Singapur(n=62) [50,51]. Die asiatischen Patientenkohorten sollen aus Fallzahlengründen im weiteren Verlauf dieser Arbeit zusammen betrachtet werden („Asien“). Die Columbia AVM Databank (n=337, „New York“) [52] und die AVM-Datenbank der University of Toronto (n=290, „Toronto“) [44] enthalten Patientendaten, die in dem jeweiligen Zentrum erhoben worden sind.

Analysiert und dargestellt werden dementsprechend fünf verschiedene Datensätze, die den fünf geographisch abgegrenzten Zentren bzw. Orten der Behandlung entsprechen (Berlin, Paris, Asien, New York und Toronto). Die Daten dieser Zentren werden bezüglich der im folgenden dargestellten demographischen, morphologischen und klinischen Charakteristiken verglichen.

2.1 Untersuchte Variablen

Untersucht wurden demographische Charakteristiken, morphologische Charakteristiken, die insbesondere zur Klärung des Risikos eines invasiven neurochirurgischen Eingriffs herangezogen werden, und klinische Charakteristiken, die durch die Art der initialen Krankheitspräsentation definiert wurden.

2.1.1 Demographische Charakteristika

Untersucht wurden

(a) das Patientenalter zum Zeitpunkt der klinischen Erstmanifestation der AVM und

(b) das Geschlecht.

Als klinische Erstmanifestation gilt das Indexereignis (z.B. intrazerebrale Blutung, Krampfanfall, Kopfschmerz, etc.), dessen weitere Abklärung zur Diagnose der AVM mittels zerebraler Bildgebung (CT, MRT, und/oder Angiographie) führte. Beim Alter wurden Lage und Streuungsmaße für das Lebensalter in Jahren verglichen. Zum Vergleich der empirischen Altersverteilung des Patientengutes der jeweiligen Zentren wurde das erreichte Lebensjahrzehnt der Patienten herangezogen.

2.1.2 Morphologische Charakteristika

Die untersuchten morphologische Merkmale entsprechen den drei Elementen der "Spetzler-Martin Skala": Maximale AVM-Größe, Art der venösen Drainage, sowie die „Eloquenz“ der betroffenen Hirnregion:

(a) Die Größe der AVM wird über den maximalen Durchmesser des AVM-Nidus bestimmt. Entsprechend der Spetzler-Martin Skala wurden AVMs unter 3 cm als klein, AVMs zwischen 3 und 6 cm als mittelgroß und AVMs über 6 cm als groß eingestuft.

(b) Das Merkmal "Art der venösen Drainage" wurde dahingehend dichotomisiert, ob die venöse Drainage ausschließlich über oberflächliche Venen erfolgte, oder ob auch eine Komponente mit tiefer venöser Drainage vorlag. AVMs mit ausschließlich oberflächlicher Drainage entsprechen weitgehend der Untergruppe der rein kortikalen und der kortiko-subkortikalen AVMs in der eingangs dargestellten topographischen Einteilung nach Berenstein und Lasjaunias. Die zweite Gruppe dieses Merkmals umfasst kortiko-ventrikuläre AVMs, tief lokalisierte AVMs und AVMs des Plexus chorioidei.

(c) Als „eloquente“ Hirnregionen werden die von Spetzler definierten funktionell bedeutenden Hirnareale bezeichnet (Siehe Tabelle 1: Spetzler-Martin Skala). Entsprechend der Aufteilung der Spetzler-Martin Skala wurde hier zwischen "AVM in eloquenter Hirnregion" und "AVM außerhalb eloquenter Hirnregionen" unterschieden [58].

2.1.3 Klinische Charakteristika

Die untersuchten klinischen Merkmale beschreiben die Art der klinischen Erstmanifestation der jeweiligen AVM. Im Einzelnen wurden folgende Merkmale erhoben:

- (a) AVM-assoziierte intrakranielle Blutung,
- (b) epileptischer Krampfanfall,
- (c) Kopfschmerzen und
- (d) fokale neurologische Defizite.

Als intrakranielle Blutung wurden Blutungen in das Hirnparenchym, in die Ventrikel und/oder den Subarachnoidalraum gewertet. Bei Anfällen wurde nach generalisierten und fokalen Anfällen unterschieden. Bei den Kopfschmerzen wurden nur chronische Beschwerden erfasst. Die neurologischen Defizite wurden als reversibel, persistierend oder progredient eingestuft.

2.2 Datenerhebung, Datenvorbereitung und -Datenprüfung

Bei dem in dieser Arbeit angestrebten Vergleich der demographischen, morphologischen und klinischen Charakteristiken handelt es sich um eine Sekundäranalyse auf der Ebene nicht aggregierter Daten. Die Daten liegen somit als Patientendatensätze von drei verschiedenen, voneinander unabhängig geplanten und konzipierten AVM-Datenbanken vor. Die Phase der Datenerhebung bestand aus der Übernahme von Datenbankauszügen aus den drei Datenbanken.

Für die Datenvorbereitung implizierte dies, dass sie auf vorhandenes, gegebenenfalls heterogenes Datenmaterial aufsetzen musste. Das Ziel dieser Phase war es deshalb, die Datenbankauszüge zu vereinheitlichen und sie in ein gemeinsames Datenformat zu überführen.

Die Datenbankbetreiber verwendeten verschiedene Datenbankmanagementsysteme zur Verwaltung der Daten. Da die Organisation des Datenerhebungsprozesses und die dabei eingesetzten Datenverarbeitungstechniken auch Rückschlüsse auf die Qualität und Vereinbarkeit der Daten erlaubt, sollen die hierzu vorliegenden Erkenntnisse kurz dargestellt werden.

2.2.1 Datenerhebung in Paris, Berlin und Asien

Als Datenbankmanagementsystem (DBMS) der multizentrischen Datenbank wurde das unter dem Betriebssystem Macintosh laufende Programm FileMaker Pro 3.0 der Firma Claris Corporation eingesetzt. Dabei handelt es sich um eine einfach zu bedienende, desktopzentrierte und ursprünglich nichtrelationale Anwendung, die es erlaubte, ohne Programmierkenntnisse Eingabemasken zur Unterstützung der Dateneingabe durch eine Benutzerführung, Auswahllisten und Eingabeüberprüfungen zu definieren.

Über die genaue Arbeitsorganisation des Erhebungs- und Eingabeprozesses und den Zeitraum, über den sich dieser hingezogen hat, ist nichts bekannt. Die Daten wurden als Excelfile übergeben.

2.2.2 Datenerhebung in New York

Die New Yorker Datenbank setzte zur Speicherung der AVM-Daten ab 1990 das unter dem Betriebssystem Macintosh laufende Produkt 4th Dimension 1.0 (später 4D Server 1.0) des Herstellers ACIUS ein. Bei der verwendeten Variante handelte es sich um ein vollrelationales, client-server-basiertes, netzwerkfähiges, jedoch nicht SQL-kompatibles DBMS. Ab 1998 wurde die Datenbankanwendung dann auf ein webbasiertes Frontend mit Zugriff auf SQL-Datenbanken (Microsoft Access und Oracle) über standardisierte Datenbankschnittstellen portiert. Für die Eingabe der Daten standen in beiden Anwendungen graphische Eingabemasken mit Benutzerführung, Auswahllisten für vorgegebene Werte und Eingabeüberprüfungen zur Verfügung [95].

Die Eingabe der Daten erfolgte dezentral durch das Fachpersonal der Abteilungen Neurologie, Neuroradiologie und Neurochirurgie, wobei bereits die Eingabemasken so angelegt waren, dass den Abteilungen nur die Eingabe derjenigen Daten möglich war, die in die eigene Zuständigkeit fielen [95]. Die Daten wurden als CSV-Datenfiles (ASCII-Datei) aus den verschiedenen Tabellen der 4th Dimension Anwendung exportiert und unter Auflösung der relationalen Bezüge in das Statistikpaket SPSS importiert.

2.2.3 Datenerhebung in Toronto

Informationen über die verwendete Software und die Organisation des Erhebungs- und Eingabeprozesses sowie den Zeitraum, über den die Erhebung der Daten erfolgte, liegen für Toronto nicht vor. Der Datensatz aus Toronto enthielt zudem deutlich weniger Merkmale als Datensätze der beiden anderen Zentren. So fehlten mit Ausnahme der intrakraniellen Blutung alle Angaben aus dem Bereich der klinischen Merkmale. In der Gruppe der morphologischen Merkmale fehlten Angaben zur Drainage der AVM. Ebenso lagen bezüglich der demographischen Merkmale keine Angaben zum Alter der Patienten bei Einschluss in die Studie vor. Die Daten wurden manuell von Papierauszügen in eine Exceldatei übertragen.

2.2.4 Zusammenführung der Daten

Die Zusammenführung der unterschiedlichen Dateien zu einem Datensatz erfolgte mit dem Statistikpaket SPSS 8.0. Im Rahmen der Übernahme der Werte in SPSS wurden vor allem die Codes, die den nominalskalierten Merkmalen in den Datensätzen zugeordnet waren, vereinheitlicht. Zudem wurden fehlende Werte einheitlich mit dem Wert -2 kodiert. SPSS kennzeichnet wie die meisten Statistikpakete fehlende Werte explizit und berücksichtigt diese bei der Analyse.

Zusätzlich wurden die Variablen auf das Vorliegen einheitlicher Definition hin überprüft und gegebenenfalls angepasst. Im Fall der Drainage ergaben sich dadurch Abweichungen, dass in New York das Merkmal Drainage mit den Kategorien "oberflächliche Drainage", "oberflächliche und tiefe Drainagekomponenten" und "ausschließlich tiefe Drainage" erfasst worden war, während in der multizentrischen Datenbank nur die Kategorien "oberflächliche Drainage" sowie und "oberflächliche und tiefe Drainagekomponenten" vorlagen. Die Kategorien wurden durch Zusammenlegen der Kategorien "oberflächliche und tiefe Drainagekomponenten" und "ausschließlich tiefe Drainage" vereinheitlicht.

Im Rahmen der Zusammenführung der Datensätze war es außerdem erforderlich, die klinischen Merkmale neurologisches Defizit und Krampfanfall in dichotome Indikatorvariablen (reversibles, persistierendes oder progredientes neurologisches Defizit bzw. generalisierter oder fokaler Krampfanfall, jeweils mit den Ausprägungen "trifft zu", "trifft nicht zu" und "fehlender Wert") umzuwandeln. Im New Yorker Originaldatensatz waren beide Merkmale jeweils durch zwei Variablen kodiert worden, wobei eine Variable das Vorliegen oder Nichtvorliegen des Merkmals und die zweite Variable den Typ des Merkmals beschrieb. Im Originaldatensatz der multizentrischen Studie enthielten beide Merkmale eine zusätzliche Kategorie ("Normal" im Fall des neurologischen Defizits und "Abwesend" bei dem Merkmal Krampfanfall), um explizit festhalten zu können, dass das Merkmal nicht vorlag.

2.2.5 Vorbereitung der Analyse fehlender Werte

Für die Analyse der Muster an Auslassungen und unvollständiger Eingaben wurde aus den zusammengeführten Daten aller Zentren mit Ausnahme von Toronto ein zweiter Datensatz generiert, der unvollständige Eingaben in den drei Merkmalsgruppen demographische, morphologische und klinische Daten erfasste. Das Zentrum Toronto wurde von der Analyse ausgeschlossen, da hier von Vorneherein nur für einige der untersuchten Merkmale Werte vorlagen.

Durch ein automatisiertes Verfahren wurde pro Patient und Merkmalsgruppe eine Datenzeile mit den Variablen "Merkmalsgruppe", "fehlender Wert" und "Zentrum" angelegt. Dies führte zu einer Gesamtzahl von $n=2997$ insgesamt auf vollständige Eingabe zu überprüfenden Merkmalsgruppen. "Fehlender Wert" erfasste dabei pro beobachteter Gruppe, ob mindestens eines der Merkmale ausgelassen worden war (Tabelle 17).

Tabelle 17: Aufbau des Datensatzes zur Analyse der fehlenden Werte

Patient Nr.	Merkmalsgruppe (Merkmal)	Fehlender Wert ¹	Zentrum
1	Demographisches Merkmal ²	Nein	Berlin
1	Morphologisches Merkmal ³	Nein	Berlin
1	Klinisches Merkmal ⁴	Nein	Berlin
2	Demographisches Merkmal ²	Nein	Berlin
2	Morphologisches Merkmal ³	Ja	Berlin
2	Klinisches Merkmal ⁴	Nein	Berlin
...
220	Demographisches Merkmal ²	Nein	Paris
220	Morphologisches Merkmal ³	Nein	Paris
220	Klinisches Merkmal ⁴	Nein	Paris
...

¹Mindestens ein Wert der Merkmalsgruppe wurde ausgelassen.
²Geschlecht, Alter.
³Eloquenz der Lokation der AVM, Drainage der AVM, Größe der AVM.
⁴Intrakranielle Blutung, chronische Kopfschmerzen, Generalisierter oder fokaler Krampfanfall, persistierendes, progredientes oder reversibles neurologisches Defizit.

2.3 Statistische Analysen

2.3.1 Bivariate Verfahren

Die Mittelwerte intervallskalierteter stetiger Merkmale wurden durch eine univariate Varianzanalyse auf Unterschiede zwischen den Zentren getestet [96,97]. Bei signifikanten Unterschieden wurde angesichts der relativ großen Fallzahlen die Varianzaufklärung des untersuchten Zusammenhanges bestimmt. War die Varianzaufklärung geringer als ein Prozent, wurde von einem unbedeutenden Resultat ausgegangen.

Bei dichotomen oder diskreten, nominalskalierten Merkmalen wurde der χ^2 -Test auf Unabhängigkeit eingesetzt, um bezüglich der Häufigkeit dieser Merkmale Unterschiede zwischen den Zentren festzustellen.

Da die Signifikanz eines χ^2 -Test auf Unabhängigkeit nur ausdrückt, dass die beobachtete Zellbelegungen insgesamt von den aufgrund der Randverteilungen erwarteten Häufigkeiten abweichen, wurde für die darüber hinausgehende Identifikation derjenigen Zellen, die besonders von den Abweichungen betroffen waren, die Methode der Residualanalyse mittels adjustierter Residuen im Sinne einer Post-Hoc Analyse eingesetzt [98,99].

Der Grundgedanke der Residualanalyse besteht in der Interpretation der Differenz zwischen der beobachteten Zellhäufigkeit n_i (i ist die i -te Zelle einer Kontingenztafel) und der bei Zutreffen der Annahme erwarteten Zellhäufigkeit m_i . Die Größe dieses Residuums ist auch abhängig vom Umfang der Stichprobe. Die nach der Formel

$$e_i = \frac{n_i - m_i}{\sqrt{\frac{m_i}{n}}}$$

standardisierten Residuen e_i sind dagegen unabhängig von dem Stichprobenumfang und eignen sich deshalb bereits für einen Vergleich zwischen verschiedenen Kontingenztafeln. Unter der Annahme der Unabhängigkeit sind sie zwar mit einem Mittelwert von 0 asymptotisch normalverteilt, ihre Standardabweichung ist jedoch kleiner als 1. Der als adjustiertes Residuum bezeichnete Quotient r aus dem standardisierten Residuum und dem Standardfehler der Zelle ist dagegen standardnormalverteilt. Somit kann das adjustierte Residuum r_i als Indikator für das Vorliegen eines signifikanten Unterschieds zwischen der beobachteten und der erwarteten Häufigkeit einer einzelnen Zelle i einer Kontingenztafel interpretiert werden [100,101,102].

Da dieser Ansatz der Situation multiplen Testens entspricht, wurden das Signifikanzniveau α , anhand dessen die adjustierten Residuen auf Signifikanz zu testen waren, entsprechend dem Bonferroni-Kriterium korrigiert. Wegen des geringen Mehraufwandes wurde auf die häufig verwendete Approximation über $\alpha' = \alpha/n$, (n = Anzahl der Felder der Kontingenztafel) verzichtet und stattdessen das korrigierte Signifikanzniveau α' über $\alpha' = 1 - (1 - \alpha)^{1/n}$ berechnet [99,100,103]. Im Fall einer Kontingenztafel mit 8 Zellen sind beispielsweise 8 adjustierte Residuen auf Signifikanz zu testen. Die Bonferroni-Korrektur besagt, dass die auf einem

nominellen Signifikanzniveau von $\alpha = 0,05$ zu testenden Residuen in diesem Fall erst ab einem Wert von $\alpha' = 0,00625$ (da $1-(1-0,00625)^8=0,05$) bzw. $z = 2,50$ als signifikant bewerten werden können [96,103].

Sowohl für stetige als auch für dichotome (kategoriale) Merkmale wurden Konfidenzintervalle berechnet. Im Fall dichotomer Merkmale wurden die Konfidenzintervalle über die Normalverteilungsapproximation der Binomialverteilung bestimmt, wenn eine hinreichende Approximation gegeben war. Davon wurde ausgegangen, wenn die Bedingung $n \cdot p \cdot q \geq 9$ erfüllt war (n = Stichprobenumfang, p = relative Häufigkeit und $q = 1 - P$) [99].

2.3.2 Multivariate Verfahren

Die in der vorliegenden Arbeit eingesetzten multivariaten Verfahren sind sogenannte loglineare Modelle. Diese Verfahren, die zur Analyse der Zusammenhänge zwischen mehreren kategorialen und nominalskalierten Variablen eingesetzt werden können, stellen in gewisser Weise eine multivariate Erweiterung des χ^2 -Tests auf Unabhängigkeit dar [102].

Loglineare Modelle können als einerseits Erweiterung des χ^2 -Tests auf Unabhängigkeit auf mehr als zwei Variablen angesehen werden. Andererseits stellt der χ^2 -Tests auf Unabhängigkeit, nur eines von mehreren möglichen loglinearen Modellen dar [102].

Bei mehr als zwei Variablen steigt die Anzahl dieser Modelle rasch an. Es stellt sich deshalb die Frage, nach welchen Kriterien ein geeignetes Modell ausgewählt werden kann. Ein Gütekriterium für loglineare Modelle ist die Sparsamkeit, mit der Informationen zur Berechnung von Randverteilungen aus der Stichprobe entnommen werden. Dies wird über die Anzahl der Freiheitsgrade quantifiziert. Je größer die Anzahl der Freiheitsgrade ist, desto sparsamer ist das Modell. Ein weiteres Gütekriterium ist der "Fit" des Modells. Wie bei dem χ^2 -Test auf Unabhängigkeit kann eine χ^2 -verteilte Prüfgröße gebildet und auf Signifikanz geprüft werden. Für einen guten Fit sollte das Modell auch bei einer hohen Schranke nicht signifikant werden. [102,104,105,106]

Loglineare Modelle wurden in der vorliegenden Arbeit angepasst, um den Effekt dritter Variablen auf einen zu untersuchenden Zusammenhang zu kontrollieren. Durch die Integration intervenierender Variablen in das Modell soll zum einen der Einfluss dieser Variablen kontrolliert und die χ^2 -verteilte Prüfgröße entsprechend bereinigt werden. Der dazu verwendete Ansatz ist der modellbasierte Test auf bedingte Unabhängigkeit. Zum anderen soll die Verteilung fehlender Werte in Abhängigkeit von Zentren und Merkmalsgruppen untersucht werden, um zu einer aussagekräftigen Modellierung dieser Zusammenhänge zu gelangen. Dazu werden die Methoden der Residualanalyse und der Zerlegung von Interaktionseffekten in relevante orthogonale Kontraste eingesetzt.

Der modellbasierte Test auf bedingte Unabhängigkeit

Tests auf bedingte Unabhängigkeit untersuchen, ob beobachtete Zusammenhänge bestehen bleiben, wenn mögliche intervenierende Variablen kontrolliert werden [100]. Zur Gruppe der Tests auf bedingte Unabhängigkeit gehört zum Beispiel der in der klinischen Forschung häufig eingesetzte Mantel-Haenszel-Test. Er wird typischerweise verwendet, um zwei oder mehr nach der Kaplan-Meier-Methode geschätzte kumulierte Überlebensraten zu vergleichen [107,108,109]. Allgemein können damit $k \times 2 \times 2$ Kontingenztabelle auf Unabhängigkeit getestet werden, wobei der Einfluss der ordinalen bzw. metrischen Variable kontrolliert wird, die als Schichtungsmerkmal der $k \times 2 \times 2$ Kontingenztabelle bzw. k Straten dient. Neben der Zeit kann zum Beispiel auch die Dosishöhe als Schichtungsmerkmal eingesetzt werden [96,100]. Verallgemeinerungen des Mantel-Haenszel-Tests auf den Vergleich von mehr als zwei Überlebensraten oder auf mehr als ein Schichtungsmerkmal sind möglich [107].

Alternativ kann der Test auf bedingte Unabhängigkeit auch über ein loglineares Modell erfolgen. Das loglineare Modell (XZ, YZ) ist eine Erweiterung des bivariaten χ^2 -Tests auf Unabhängigkeit (X, Y) . X und Y stehen dabei für den zu untersuchenden Zusammenhang, Z stellt die zu kontrollierende intervenierende Variable dar. Die Hypothese der bedingten Unabhängigkeit kann direkt über den Anpassungstest des Modells $G^2[(XZ, YZ)]$ geprüft werden. Durch den direkten Tests wird jedoch im Unterschied zum Mantel-Haenszel-Test auch die drei-Faktoren-

Interaktion (XYZ) mitgetestet. In die Begrifflichkeit der Überlebensanalyse übersetzt entspricht eine signifikante drei-Faktoren-Interaktion dem (mehrfachen) Überkreuzen von zwei (oder mehr) Überlebenskurven [100,107]. Ein loglineares Modell, das genau dem Mantel-Haenszel-Test entspricht, erhält man über den Test von

$$G^2[(XZ, YZ)|(XY, XZ, YZ)] = G^2(XZ, YZ) - G^2(XY, XZ, YZ).$$

Die Differenz beider loglinearen Modelle ist eine ebenfalls χ^2 -verteilte Prüfgröße, die den Zusammenhang zwischen X und Y nach Korrektur des Einflusses von Z testet. Wie der Mantel-Haenszel-Test sollte dieser modellbasierte Test auf bedingte Unabhängigkeit nicht eingesetzt werden, wenn eine signifikante drei-Faktoren-Interaktion vorliegt [100,107]. In der vorliegenden Arbeit wurde der modellbasierte Test auf bedingte Unabhängigkeit eingesetzt, um den Einfluss intervenierender Variablen festzustellen und zu korrigieren.

Loglineare Modellierung unter Verwendung von Kontrasten

Die Prüfstatistik G^2 eines loglinearen Modells kann in mehrere ebenfalls χ^2 -verteilte Komponenten zerlegt werden, die voneinander unabhängig sind. Diese Technik der Dekomposition lässt sich zur Durchführung von Einzelvergleichen nutzen, mit denen herausgefunden werden kann, welche Ausprägungen und Kombinationen der beteiligten Merkmale für die Modellierung der Zusammenhänge relevant sind. Einzelvergleiche erlauben somit wie die Analyse der adjustierten Residuen einen detaillierteren Blick auf die zugrundeliegenden Zusammenhänge [100]

Im Rahmen der Varianzanalyse und der Modellierung von Zusammenhängen auf der Grundlage des Allgemeinen Linearen Modells werden häufig sogenannte Einzelvergleiche oder Kontraste eingesetzt, um voneinander linear unabhängige bzw. orthogonale Teilaspekte eines Zusammenhanges zu schätzen [99]. Will man beispielsweise nicht nur testen, ob ein Zusammenhang zwischen der Höhe der Dosis eines Medikamentes und seiner biologischen Wirksamkeit besteht, sondern ob der Zusammenhang (nach entsprechender Transformation der Ausgangsdaten) linear ist - d.h. ob bei zunehmender Dosishöhe auch eine Zunahme des

biologischen Effektes erfolgt - können orthogonale Polynome als Kontraste verwendet werden. Zusammen mit dem linearen Trend kann bei v Ausprägungen des Faktors Dosishöhe auf insgesamt $v-1$ Trends getestet werden, wobei jeder Trend einem Polynom einer bestimmten Ordnung entspricht. Da der lineare Trend ein Polynom erster Ordnung ist und da bei v Ausprägungen $v-1$ Trends existieren, ist das höchste Polynom, auf das getestet werden kann ein Polynom $v-1$ -ter Ordnung. Bei vier Ausprägungen kann so auf einen linearen, quadratischen und kubischen Trend getestet werden [99].

Andere orthogonale Kontraste testen auf Unterschiede zwischen den Ausprägungen eines Faktors. Die als "umgekehrte Helmert-Kontraste" bezeichneten Einzelvergleiche ermöglichen die Zusammenfassung der Ausprägungen nach folgendem Schema: 1 gegen 2, 1+2 gegen 3, 1+2+3 gegen 4, ..., 1+2+3+...+ $k-1$ gegen k . Bei k Ausprägungen eines Faktors ergeben sich analog zu den orthogonalen Polynomen $k-1$ voneinander unabhängige Einzelvergleiche, wobei der erste Einzelvergleich die zweite mit der ersten Ausprägung und der letzte Einzelvergleich die letzte mit allen vorherigen Ausprägungen vergleicht [99,110].

In der vorliegenden Arbeit wurden umgekehrte Helmert-Kontraste verwendet, um loglineare Modelle zu spezifizieren, die über die Analyse von Haupt- und Interaktionseffekten hinausgehen und eine differenziertere Betrachtung der Zusammenhänge zwischen Herkunft und Qualität der Daten erlauben.

2.3.3 Durchführung der Analysen

Die Datenanalyse wurde mit dem Statistikpaket SPSS 8.0 und dem Tabellenkalkulationsprogramm Microsoft Excel 2000 durchgeführt.

Eingesetzt wurden die SPSS-Module anova, crosstabs, genlog und loglinear sowie einige Funktionen zur Rekodierung und Transposition der Daten. SPSS gibt bei den Modulen crosstabs, genlog und loglinear sowohl Pearsons χ^2 -Test also auch den Likelihood Ratio χ^2 -Test aus. Aus Gründen der Konsistenz und Vergleichbarkeit wurde auch bei den bivariaten Tests der Likelihood Ratio χ^2 -Test zur Beurteilung der Signifikanz der untersuchten Zusammenhänge herangezogen.

Für eine hinreichende Approximation der diskreten Häufigkeiten durch die stetige χ^2 -Verteilung muss eine ausreichende Belegung der Zellen mit erwarteten Häufigkeiten gewährleistet werden. Mit erwarteten Häufigkeiten unter 5 sollten nicht mehr als 20% der Zellen belegt sein. [99,111]. Die Zellbelegungen der erwarteten Häufigkeiten wurden deshalb anhand der Ergebnisoutputs der entsprechenden SPSS-Module überprüft. Die Grenzwerte sind in jedem Fall – auch bei dem mit stark asymmetrischen Randverteilungen ausgestatteten Test der fehlenden Werte – eingehalten worden.

Mit Microsoft Excel wurden die Bonferroni-Korrekturen der adjustierten Residuen durchgeführt und die Differenzen von χ^2 -verteilter Prüfgrößen auf Signifikanz geprüft.