

Evangelische  
Fachhochschule  
Darmstadt

## **Eine Untersuchung zum Zusammenhang von Patientenzuständen und Pflegeaufwand – Vorschläge für empirisch abgesicherte (Patienten-) Fallgruppen**

Fachbereich  
Pflege- und Gesundheitswissenschaft  
Diplomarbeit, 07. Mai 2007

Betreut durch: Herrn Prof. Nielsen und Frau Prof. Dr. Höhmann

Madlen Fiebig  
Matrikelnummer: 923935



**“If it matters, measure it!”**

(Singleton-Green)

## **Zusammenfassung**

Thema: In einer im Rahmen der Diplomarbeit durchgeführten Untersuchung wird der Zusammenhang zwischen Patientenzuständen und Pflegeaufwand analysiert. Darüber hinaus werden Vorschläge für empirisch abgesicherte Patientenfallgruppen auf der Grundlage bereits früherer Konstruktionsversuche aus mannigfachen nationalen und internationalen Studien skizziert. Des Weiteren werden bereits umgesetzte Modelle der adäquaten Abbildung der Pflege im DRG-System aufgezeigt und diskutiert.

Fragestellung: Mit welcher Wahrscheinlichkeit können Patienten im akutstationären Setting, in Abhängigkeit vom ePA-CasemanagementScore<sup>®</sup>, in die korrekte LEP<sup>®</sup>-Minutengruppe eingeordnet werden?

Design und Methode: Die Zusammenhangsuntersuchung basiert auf einer quantitativen, nicht experimentellen, deskriptiven Querschnittsstudie. Im Rahmen der Querschnittsstudie werden Daten zu einem konkreten Zeitpunkt an einer Stichprobe erhoben.

In dieser Untersuchung wird der Zustand eines Patienten mittels des ergebnisorientierten PflegeAssessment-AcuteCare<sup>®</sup> (ePA-AC<sup>®</sup>) erhoben. Der Pflegeaufwand wird mit Hilfe des Instruments „Leistungserfassung in der Pflege“ (LEP<sup>®</sup>) gemessen.

Analysemethode: Zur Beantwortung der Fragestellung wurde eine multinominal logistische Regression durchgeführt. Ziel dieser Analysemethode war es, die Abhängigkeitsstruktur zwischen abhängiger und unabhängiger Variable zu eruieren.

Ergebnisse: Die Ergebnisauswertung zeigt eine sehr gute Modellanpassung und Trennschärfe auf Variablenebene bei Modellen aus Gruppen, deren Interventionen sich in Fähigkeitsbeeinträchtigungen von Patienten begründen. Weniger gute Resultate kann für Modelle aus den Gruppen der medizinisch induzierten Interventionen und erbrachten Pflegeminuten ermittelt werden.

Empfehlungen: Die nächsten Schritte müssen eine Optimierung und Verfeinerung der Modelle zur Rechnung der logistischen Regression beinhalten.

Schlüsselwörter: Patientenzustand; Pflegeaufwand; Pflegekosten; pflegerrelevante Fallgruppen; Krankenhausfinanzierung; Kostenhomogenität

## **Abstract**

Theme: In a study carried out as part of a thesis the interrelationship is analyzed between patient condition and the level of nursing care. In addition, recommendations for empirically based patient case groups are outlined on the basis of earlier design tests from a broad range of national and international studies. Moreover, already implemented models of adequate representation of nursing care in the DRG system are described and discussed.

Research question: Using the ePA-CasemanagementScore<sup>®</sup>, what is the probability of patients in acute care settings being placed in the correct time allotment group of the nursing workload measurement system LEP<sup>®</sup>?

Design and Methodology: The study of the interrelationship is based on a quantitative, non-experimental, descriptive cross-sectional study. In the framework of the cross-sectional study, data are collected from a sampling taken at a specific point in time.

In this study, the condition of a patient is measured by means of the results-oriented PflegeAssessment-AcuteCare<sup>®</sup> (ePA-AC<sup>®</sup>). The nursing care work load is measured by LEP<sup>®</sup>.

Analytic method: The question was answered by conducting a multi-nominal logistic regression analysis. The objective of this analytic method is to establish the dependency structure between dependent and independent valuable.

Results: The assessment of the results show a very good model adaptation and discriminatory power on the variable level for models from groups whose interventions are based on impaired patient abilities. Less good results were established for models from the groups of medically induced interventions and delivered nursing care minutes.

Recommendations: The next steps must include an optimization and refinement of the models for calculating the logistic regression.

Keywords: Patient condition; nursing workload; nursing cost; nursing related groups; hospital charges; cost homogeneity

# Inhaltsverzeichnis

<b>EINLEITUNG .....</b>	<b>6</b>
<b>1 HINTERGRUND DES PROJEKTS .....</b>	<b>7</b>
1.1 PROJEKTINTENTION.....	7
1.2 FORSCHUNGSFRAGE .....	9
<b>2 BEGRIFFSBESTIMMUNG UND THEORETISCHE GRUNDLAGEN .....</b>	<b>10</b>
2.1 BEGRIFFSKLÄRUNG .....	10
2.2 STAND DER AKTUELLEN PFLEGETHEORETISCHEN FACHDISKUSSION .....	15
2.3 ZUR BEDEUTUNG DER PFLEGEKLASSIFIKATIONSSYSTEME IM DRG-SYSTEM.....	19
2.4 ZUR BEDEUTUNG DER PATIENTENKLASSIFIKATION IM DRG-SYSTEM.....	20
<b>3 KONZEPTUALISIERUNG DER UNTERSUCHUNG .....</b>	<b>22</b>
3.1 QUERSCHNITTSDESIGN .....	22
3.2 METHODE .....	23
3.2.1 <i>Auswahl der Erhebungsinstrumente</i> .....	24
3.2.2 <i>Stichprobenbeschreibung</i> .....	28
3.2.3 <i>Hintergrund zum Datenmaterial</i> .....	33
<b>4 LOGISTISCHE REGRESSIONSANALYSE .....</b>	<b>34</b>
4.1 DATENAUFBEREITUNG .....	35
4.2 MODELLFORMULIERUNG .....	38
4.3 SCHÄTZUNG DER LOGISTISCHEN REGRESSIONSFUNKTION.....	39
4.4 PRÜFUNG DES MODELLS .....	40
<b>5 ERGEBNISSE .....</b>	<b>42</b>
<b>6 INTERPRETATION UND DISKUSSION DER ERGEBNISSE .....</b>	<b>54</b>
<b>7 KONKLUSION UND DIE RELEVANZ FÜR DIE PFLEGE .....</b>	<b>61</b>
<b>8 EMPFEHLUNGEN .....</b>	<b>65</b>
<b>LITERATURVERZEICHNIS.....</b>	<b>67</b>
<b>INTERNETVERZEICHNIS .....</b>	<b>72</b>
<b>ABKÜRZUNGSVERZEICHNIS .....</b>	<b>73</b>
<b>ABBILDUNGSVERZEICHNIS.....</b>	<b>74</b>
<b>TABELLENVERZEICHNIS.....</b>	<b>75</b>
<b>ANHANG.....</b>	<b>76</b>
<b>PERSÖNLICHE ERKLÄRUNG .....</b>	<b>77</b>

## Einleitung

Einer der größten Kostenblöcke in der Krankenhausfinanzierung entfällt auf die Pflege. Im Jahr 2004 wurden 18,3 Milliarden Euro von insgesamt 60,4 Milliarden Euro für pflegerische Leistungen entrichtet (vgl.: Statistisches Bundesamt, 2007).

Doch die adäquate Berücksichtigung der Pflege im DRG-System fehlt bis heute. Zahlreiche Studien, die in dieser Arbeit im nachfolgenden zum Teil erläutert werden, geben deutliche Hinweise, dass Pflege zur Verbesserung der Homogenität in einzelnen DRGs beitragen kann. Weiter zeigen Studien, dass pflegerische Informationen zur Kostengewichtung sich nicht, wie lange angenommen, in Gänze aus den medizinischen Diagnosen speisen lassen.

Es ist außer Frage, dass Pflege ihren Beitrag in der DRG Diskussion zur Stabilisierung der Kostenhomogenität sowie der Kostengewichtung leisten kann. Die sich stellende Frage ist: Wie Pflege sich an betriebswirtschaftlichen Konfigurationen beteiligen kann. Die Arbeit greift diese Fragestellung auf und arbeitet Vorschläge für eine opportune Einbindung der Pflege im Fallpauschalensystem heraus. Hierbei soll einer der Schwerpunkte die Entwicklung von pflegerelevanten Fallgruppen bilden. Die Arbeit hat realistischer Weise nicht den Anspruch diese Fallgruppen zu konzipieren, sondern es sollen Wege aufgezeigt werden, wie diese Konzeption ermöglicht werden könnte.

Im ersten Teil dieser Arbeit werden Beweggründe für die Auseinandersetzung mit diesem Thema im Rahmen der Diplomarbeit erläutert. Ebenso wird in diesem Abschnitt eine theoretische Übersicht zur aktuellen Fachdiskussion im Zusammenhang mit Pflege im DRG-System gegeben. Im darauf folgenden Kapitel 3 werden das zur Beantwortung der Forschungsfrage zu wählende Studiendesign und die sich daraus ableitende Methodenauswahl vorgestellt.

Ausgehend von dem hier zu untersuchenden Thema des Zusammenhangs zwischen Patientenzustand und Pflegeaufwand wird eine logistische Regression zur Analyse der erhobenen Daten verwendet. Kapitel 4 gibt einen Einblick in das regressionsanalytische Modell und skizziert grundlegende Zusammenhänge der logistischen Regression.

Die aus der Regressionsanalyse gewonnenen Ergebnisse werden im Kapitel 5 aufgezeigt und im nachfolgenden Kapitel 6 kritisch diskutiert. Im Rahmen des 7. Kapitels wird der zweite Teil des Diplomthemas, die Erarbeitung von Vorschlägen für mögliche Patientenfallgruppen, aufbereitet und diskutiert.

## 1 Hintergrund des Projekts

Die Ausgaben im Gesundheitssektor steigen seit Jahren stetig an. Deutschland verzeichnete im Jahr 2004 Gesundheitsausgaben von insgesamt 234,0 Milliarden Euro. Das Statistische Bundesamt errechnet einen durchschnittlichen Anstieg der Ausgaben für Gesundheitsversorgung und –leistung von 2,6% pro Jahr seit 1995. Das entspricht einem Anstieg der Ausgaben seit 1995 bis 2004 von etwa 47,5 Milliarden Euro (vgl.: Statistisches Bundesamt, 2007).

Der größte Anteil für Gesundheitsausgaben wird für die ambulante und stationäre sowie teilstationäre Versorgung entrichtet. Im Ganzen wurden 2004 für diese Bereiche des Gesundheitswesens 198,2 Milliarden Euro der insgesamt 234,0 Milliarden Euro ausgegeben. Von diesem Anteil wiederum entfielen 85,4 Milliarden Euro auf stationäre Institutionen. 60,4 Milliarden Euro werden hierbei für Krankenhäuser aufgewendet.

Mit diesen Zahlen liegt Deutschland im Vergleich aller OECD-Staaten mit einem Anteil der Gesundheitsausgaben am Bruttoinlandsprodukt von 10,6% im oberen Drittel. Ebenso befindet sich Deutschland im internationalen Vergleich der „Pro-Kopf-Gesundheitsausgaben“<sup>1</sup> mit Platz drei von acht auf den vorderen Rängen (vgl.: Statistisches Bundesamt, 2006:15-18).

Die oben aufgeführten Gesundheitsausgaben im Milliardenbereich geben einen ersten Hinweis auf die Bedeutsamkeit der Thematik „Krankenhausfinanzierung“ im deutschen Gesundheitssystem. Aber nicht nur die hohen Ausgaben für Gesundheitsversorgung geben Anlass, sich verstärkt mit einer bedarfsgerechten Finanzierung der Gesundheitsleistungen auseinander zu setzen, sondern auch um eine gesundheitsökonomisch adäquate Finanzierung über alle Berufsgruppen des Gesundheitssektors im Krankenhaus zu erreichen. Die Wirtschaftlichkeit der erbrachten Leistungen von Kliniken muss angestrebt und forciert werden. Wobei die Effizienz von Interventionen nachweislich dargestellt werden sollte, d.h. Kennzahlen für einen Wirtschaftlichkeitsnachweis müssen generiert und offen gelegt werden. Der Aspekt der Effizienzprüfung muss sich selbstverständlich auch hier auf die gesamten Gesundheitsdienstleistungsberufe in einem Krankenhaus beziehen.

### 1.1 Projektintention

Die Berufsgruppe der Pflegenden nimmt in den Handlungsdimensionen des Kliniksektors einen unverzichtbaren Part ein. Das pflegerische Personal, die „front-line“-Mitarbeiter, wie sie von Blaudszun (2000:57) charakterisiert werden, besitzen ein bis

---

<sup>1</sup> Der Vergleich der Gesundheitsausgaben wird in US-Dollar je Einwohner angegeben. Umgerechnet wird mit Hilfe von Bruttoinlandsprodukt-Kaufkraftparitäten.

dato noch nicht voll ausgeschöpftes Innovationspotential für die Krankenhäuser, in denen sie agieren. Ihre Potentiale erstrecken sich über die Bereiche Marktbeobachtung, Kundenorientierung sowie Betriebswirtschaft und Management. Das ungenügende Ausschöpfen der Möglichkeiten dieser Berufsgruppe ist unter anderem der Berufsgruppe selbst geschuldet, da das Auflösen der Intransparenz ihrer Arbeit und Leistungen nur sehr verhalten erfolgt.

Somit bleibt pflegerische Arbeit zu einem Großteil für andere Berufsgruppen im Krankenhaus in der Unsichtbarkeit verhaftet (vgl.: Bamert, 2003:142). Diese Gegebenheit kann über das Nutzen standardisierter Instrumente aufgefangen werden. Das Handeln aus scheinbar subjektiven Beweggründen sollte über diese Instrumente transformiert werden, hin zu nachvollziehbaren und begründeten Aktivitäten. Dienlich dafür sind gleichermaßen leistungserfassungs- und zustandsbezogene Instrumente. Neben der Abbildung pflegerischer Leistungen und deren Begründungen können diese Instrumente für eine leistungsorientierte Kostenkalkulation genutzt werden. Diese ökonomischen Überlegungen setzen häufig bei Pflegenden Assoziationen einer Verschlechterung ihrer Arbeitssituation frei. Doch vor dem Hintergrund realistischer Existenzgefährdungen von Kliniken muss sich zukunftsorientiert von diesen antiquierten Denkweisen gelöst werden. Auch Pflege kann den stetig steigenden Kostendruck auf die Leistungserbringer im Gesundheitssektor nicht ignorieren.

Pflege als eine der größten Kostenblöcke in der Krankenhauslandschaft sollte sich eigeninitiativ in die Diskussion der Krankenhausfinanzierung einbringen. Pflege muss ihren Stellenwert verdeutlichen und ihr Know-how in der Kostenkalkulation ausspielen. Mittels innovativer Impulse kann Pflege sich im Zeitalter der DRGs aus seiner oft unterschätzten Bedeutung im Gesamtgefüge Krankenhaus heraus lösen. Des Weiteren muss die Berufsgruppe der Pflege ihre Möglichkeiten in der Kennzahlengewinnung mittels standardisierter Instrumente heraus arbeiten, um hierüber auch eine Kostentransparenz ihres Handelns bzw. ihrer Outputs zu produzieren. „Die Kostentransparenz ist sowohl für die interne Mittelverteilung in der Klinik als auch für die Verhandlung mit externen Stellen von großer Bedeutung“ (Bamert, 2003:143).

Die bisher existierenden Systeme zur Fallkostenkalkulation in der Pflege haben gegenüber dem Prinzip des DRG-Systems den entscheidenden Nachteil der retrospektiven Herangehensweise. Daher ist es, so auch in dieser vorliegenden Arbeit, aus pflegewissenschaftlicher und praxisorientierter Sicht von Interesse Wege und Möglichkeiten aufzuzeigen, wie sich Pflege im DRG-System bzw. über die Konzeption von pflegerelevanten Fallgruppen optimal abbilden lässt.

## 1.2 Forschungsfrage

In einschlägiger Fachliteratur wird verstärkt darauf hingewiesen, dass die investierte Pflegezeit über Patienteninformationen zum aktuellen und potentiellen Patientenzustand gut und objektiv erklärt werden kann (vgl.: Deutscher Pflegerat, 2007:7). Damit diese Zustandsdaten der Patienten auch in der Tat einen objektiven und nachvollziehbaren Charakter erhalten, ist eine Voraussetzung die Nutzung standardisierter Instrumente. Welche Instrumente für diese Aufgabe zum Einsatz kommen sollten, wird in der pflegewissenschaftlichen Fachwelt derzeit diskutiert. Der Deutsche Pflegerat (2007) zeigt zwei mögliche Überlegungen zur Generierung von Daten zum Patientenzustand bzw. zum Ausprägungsgrad von Pflegebedürftigkeit. Auf der einen Seite stehen diverse Assessmentinstrumente (z.B.: RAI, FIM, ePA-AC<sup>®</sup>) und dem gegenüber stehen Klassifikationssysteme wie bspw. NANDA und ICF. Eine Entscheidung, welche Instrumente bzw. Pflegeklassifikationssysteme letztendlich genutzt werden können, wird sich über die jeweilige Erklärungskraft der Zeitvarianzen im DRG-System ergeben (vgl.: Deutscher Pflegerat, 2007:7).

Vor dem Hintergrund, dass Aussagen zum Patientenzustand den Pflegezeitverbrauch äußerst gut erklären, soll in dieser Arbeit der Zusammenhang beider Aspekte untersucht werden. Diese Erkenntnisse können wiederum für die Konstruktion von pflegerelevanten Fallgruppen dienlich sein.

Für die Untersuchung des Zusammenhangs wird das standardisierte ergebnisorientierte Assessmentinstrument - AcutCare<sup>®</sup> (ePA-AC<sup>®</sup>) verwendet, welches klinische Indikatoren von Patienten erfasst (vgl.: Hunstein et al., 2005:397). Über dieses Assessmentinstrument sollen Informationen zum Patientenzustand ermittelt werden. Die verbrauchte Pflegezeit eines Patienten wird mit Hilfe des Leistungserfassungsinstrumentes LEP<sup>2</sup> generiert.

Folgende Forschungsfrage liegt dieser Untersuchung zugrunde: **„Mit welcher Wahrscheinlichkeit können Patienten im akutstationären Setting, in Abhängigkeit vom ePA-CasemanagementScore<sup>®</sup>, in die korrekte LEP-Minutengruppe eingeordnet werden?“**

Eine detaillierte Beschreibung des Untersuchungsplans sowie der eingesetzten Instrumente erfolgt im Kapitel 3.

---

<sup>2</sup> LEP = Leistungserfassung in der Pflege

## 2 Begriffsbestimmung und theoretische Grundlagen

In diesem Kapitel sollen die hier verwendeten Begriffe, G-DRG, Patientenzustand, Pflegebedürftigkeit, Pflegebedarf und Pflegeaufwand, aus pflegetheoretischer Sicht definiert werden. Die Bestimmung der Wortbedeutungen von obigen Begriffen ist für das bessere Verständnis von Bedeutung, da sie die Grundpfeiler der fortlaufenden Ausführungen darstellen.

In einem weiteren Schritt wird ein Bild zur aktuellen Diskussion der Pflegewissenschaft in Bezug auf die Thematik Pflege und DRGs skizziert.

### 2.1 Begriffsklärung

#### G-DRG-System

Die grundsätzliche Einführung einer neuen Krankenhausfinanzierung ist erstmals im Gesundheitsreformgesetz 2000 im §17b des Krankenhausfinanzierungsgesetzes erwähnt. In diesem Absatz des Paragraphen 17b ist die Implementierung einer leistungsgerechten Vergütung durch ein pauschalisiertes, diagnose-orientiertes Entgeltsystem geregelt (vgl.: Peters-Alt, 2005:21).

Das in Deutschland seit dem Jahre 2003 zur Anwendung kommende DRG-basierte Krankenhausfinanzierungssystem wurde aus einer Reihe von weltweit existierenden Diagnose-Related-Group-Systemen ausgewählt. Hierbei standen in einem der letzten Entscheidungsschritte vier Systeme zur Auswahl (vgl.: Peters-Alt, 2005:22). Letztendlich entschied sich Deutschland für die „Australian Refined Diagnosis Related Groups (AR-DRG)“ (Australian Government Department of Health and Ageing, 2007). AR-DRGs werden seit 1999 in einzelnen Staaten in diversen Formen in Australien eingesetzt (vgl.: Fischer, 2005:12). Das in Australien publizierte Entgeltsystem, AR-DRG-System Version 4.1, diente Deutschland als Basis für die Implementierung eines eigenen Entgeltsystems (vgl.: InEK, 2006). Das AR-DRG-System besitzt im Gegensatz zu den drei anderen favorisierten Systemen einen „höheren Differenzierungsgrad bei der Berücksichtigung von Begleiterkrankungen und Komplikationen“ (Peters-Alt, 2005:22-23).

Die AR-DRG Version 4.1 wurde im Jahre 2003 für die deutsche Krankenhausfinanzierung adaptiert. Das German-Diagnosis-Related-Group-System (G-DRG) wurde seitdem stark modifiziert und weiterentwickelt (vgl.: Fischer, 2005:12).

Die DRGs werden nach König definiert als Fallpauschalen, welche das oberste Ziel, die Homogenität der Behandlungsfallgruppen, haben (vgl.: König, 2003:62). Die Homogenität berücksichtigt dabei zwei Blickwinkel. Zum einen die Sicht der klinischen Ähnlichkeit, d.h.: „(...) dass die Patienten der gleichen DRG bezüglich

des Krankheitsbildes und der Problemstellung einander“ homogen sein müssen (Fischer, 2002:57). Und zum anderen müssen die Fallgruppen auch aus einer ökonomischen Perspektive eine Einheitlichkeit vorweisen. Eine übermäßige Streuung der Behandlungsfallkosten in der gleichen DRG darf nicht die Regel sein (vgl.: Fischer, 2002:57). Das DRG-Patientenklassifikationssystem hat somit das Ziel, das Output eines Krankenhauses zu beschreiben und deren Kosten transparent und vergleichbar werden zu lassen, um hierüber die Vergütung der einzelnen klinischen Produkte zu verhandeln (vgl.: Fischer, 2007:1).

Wie die Kosten sich im DRG-System im Einzelnen errechnen, wird an dieser Stelle der Arbeit nicht erläutert. In Teilen wird die Thematik der Kostenkalkulation im Krankenhaus in den Kapiteln 2.2 und 7 dargestellt.

### Patientenzustand

Die Erhebung eines Patientenzustands hat das Ziel, die Situation der Pflegebedürftigkeitsausprägung zu beschreiben (vgl.: Bartholomeyczik et al., 2000:107). Diese Zustandsbeschreibung sollte idealerweise in einer standardisierten Form erfolgen, da so die Vergleichbarkeit zwischen den Einschätzungen eines Patienten, aber auch zwischen den Patienten untereinander sichergestellt werden kann.

Zur Identifikation eines Zustandes von Patienten können unterschiedliche standardisierte Informationssysteme genutzt werden. Nur einige Möglichkeiten seien an dieser Stelle genannt, so zum Beispiel Screeninginstrumente, Assessmentinstrumente und Pflegeklassifikationssysteme. All diese Instrumente haben gemeinsam, dass sie das Ausmaß einer Beeinträchtigung und gegebenenfalls auch das Ressourcenpotential von Patienten erfassen möchten. Der Umfang der Erfassung kann sich jedoch auf unterschiedliche Dimensionen von Pflegebedürftigkeit erstrecken sowie eine differierende Detaillierungstiefe einnehmen.

### Pflegebedürftigkeit

Derzeit kursiert eine Unzahl an Darstellungen von Pflegebedürftigkeit. Doch Bartholomeyczik (2004) merkt kritisch an, dass gerade in Untersuchungen zur Pflegebedürftigkeit implizit vorausgesetzt wird, dass ein allgemeingültiges Verständnis darüber, was Pflegebedürftigkeit ist, vorausgesetzt wird. Häufig werden zur Messung der Pflegebedürftigkeit Assessmentinstrumente eingesetzt. In der Regel, so stellt Bartholomeyczik fest, sind dies Instrumente, die auf ADL (activities of daily living) aus den 60er Jahren basieren. Die ADL fokussieren jedoch nur eine sehr eingeschränkte Perspektive auf mögliche Pflegebedürftigkeitsbereiche von Patienten.

Bartholomeycziks konstruktive Kritik gegenüber mangelnden detailliert dargelegten Definitionen von Pflegebedürftigkeit richtet sich ebenso an pflegetheoretische Modelle. Die Beanstandung wird begründet mit einer scheinbar willkürlichen Zuordnung bzw. Auswahl der jeweiligen Kategorien in den Modellen. Jedoch können diese Modelle als Konstrukte zur Orientierung und Strukturierung verstanden werden. Die theoretischen Modelle lassen sich des Weiteren für Klassifikationszwecke einsetzen (vgl.: Bartholomeyczik, 2004:389, 391). Die Kategorien sollten nicht unreflektiert und zweckentfremdet für den Einsatz von Items in Assessmentinstrumenten genutzt werden. Doch bedauerlicherweise ist es zurzeit eine häufig praktizierte Art und Weise um Pflegebedürftigkeit abbilden zu wollen (vgl.: Dassen et al., 2001; Bartholomeyczik, 2004:391).

Momentan existiert eine rechtliche Definition von Pflegebedürftigkeit im §14 SGB XI. Als Pflegebedürftig werden hier Personen eingestuft, welche „der Hilfe bei gewöhnlichen und regelmäßig wiederkehrenden Verrichtungen bedürfen“ (SGB XI, <http://db03.bmgs.de/gesetze/sgb11x014.htm>; Stand: 02.05.07). Ein Kritikpunkt an der Begriffsbestimmung im SGB XI richtet sich auf den nur kleinen erfassten Ausschnitt der Bereiche, in denen eine Person pflegebedürftig sein kann. Unbestreitbar ist jedoch, dass eine betroffene Person auch über die hier erfassten Bereiche, wie Körperpflege, Ausscheidung, Ernährung und Teilaspekte von Mobilität, eingeschränkt sein kann. So finden existenzielle Themenbereiche wie Angst, Schmerz, Kommunikation und soziale Partizipation keine Berücksichtigung (vgl.: Bartholomeyczik, 2004:390). Das Defizit der Deklaration von Pflegebedürftigkeit in den Bereichen der demenziell Erkrankten und psychisch beeinträchtigten Personen wird in der Politik allmählich realisiert. Eine Umgestaltung der Definition wird zunehmend im Zuge des Reformgedankens der Pflegeversicherung diskutiert. Eine sinnvollere und umfassendere Begriffsklärung bietet das interdisziplinäre System der International Classification of Functioning, Disability and Health (ICF). Die ICF beschreibt Pflegebedürftigkeit über den Begriff „Funktionale Gesundheit“ (DIMDI, 2005:4), wobei die Formulierung „Funktional“ über das rein körperliche einer Person hinausgeht (vgl.: Bartholomeyczik, 2004:392). Drei Hauptaspekte stehen dabei im Vordergrund. Diese sind zum einen Körperfunktionen und –strukturen sowie Aktivitäten und Partizipation am gesellschaftlichen Leben. Darüber hinaus spielen Umwelt- und persönliche Faktoren eine nicht unerhebliche Rolle (vgl.: DIMDI, 2005:4-5). All diese aufgeführten Aspekte stehen in Beziehung zu einander, sind aber nicht zwingend abhängig von einander (vgl.: Bartholomeyczik, 2004:392).

Über diese umfangreiche Begriffsbestimmung, im Gegensatz zur Definition von Pflegebedürftigkeit nach §14 SGB XI, wird die Komplexität der Bereiche in denen eine Person Pflegebedürftigkeit erfahren kann deutlich.

In allen bestehenden Definitionen herrscht Konsens darüber, dass „Pflegebedürftigkeit (...) vorrangig als beeinträchtigte Autonomie bei der Lebensgestaltung und die dadurch bedingte Abhängigkeit von Hilfe bei der Gestaltung des Alltags und krankheitsbedingter Erfordernisse bezeichnet werden“ kann (Bartholomeyczik, 2004:392).

Soll Pflegebedürftigkeit erhoben werden, müssen auch mögliche Gefährdungen für eine Beeinträchtigung erfasst werden. Des Weiteren ist es wichtig die Kennzeichen für die Gründe einer Beeinträchtigung zu eruieren sowie zu dokumentieren. Die Erfassung der Gründe von Pflegebedürftigkeit bzw. die Messung der Störungen erfolgt in der Regel in Zusammenhang mit standardisierten Assessmentinstrumenten. Bartholomeyczik (2004) nennt beispielhaft den Barthel-Index, Functional Independence Measure (FIM), Pflegeabhängigkeitsskala (PAS), EasyCare und Resident Assessment Instrument (RAI) (vgl.: Bartholomeyczik, 2004:392).

In der vorliegenden Arbeit werden Daten die Auskunft über den Patientenzustand, welcher gleichzeitig auch den Grad der Beeinträchtigung eines Patienten angeben soll, mittels des ergebnisorientierten Pflegeassessments-AcuteCare<sup>®</sup> (ePA-AC<sup>®</sup>) erhoben. Das ePA-AC<sup>®</sup> ermöglicht, neben der Defiziteruierung eines Zustands, Rückschlüsse auf die Selbstfähigkeitspotentiale eines Patienten. Eine ausführliche Beschreibung zu diesem Instrument kann dem Kapitel 3.2.1 entnommen werden.

### Pflegebedarf

Im Pflegeprozess prospektiv verwendete Patientenklassifikationen erfassen über diese Ausrichtung der Perspektive den Pflegebedarf. Retrospektiv wird der Pflegeaufwand erhoben (vgl.: Baumberger, 2001:5). Mittels der Angaben zum Pflegebedarf kann unterstützend die Personalplanung in stationären Einrichtungen gestaltet werden (vgl.: Bartholomeyczik et al., 2000:106).

Der Pflegebedarf definiert das Erfordernis eines Patienten an pflegerischer Leistung. Dieses Erfordernis wird durch den Gesundheitszustand des Patienten begründet (vgl.: Fischer, 1999:11). Vor diesem Hintergrund ist es bedeutsam, dass die Leistungserfassung sich nicht nur auf erbrachte Pflegeinterventionen stützt, sondern auch Bezug nimmt auf den Zustand eines Patienten. Gleichwohl ist es wichtig strikt zwischen der notwendigen und der tatsächlich erbrachten Leistung zu differenzieren (vgl.: Bartholomeyczik et al., 2000:106).

Gerade aus ökonomischer Sicht ist diese Unterteilung von großem Interesse. Über die Differenz von erforderlicher und erbrachter Leistung können Aussagen zur Wirtschaftlichkeit eines Unternehmens bzw. Krankenhauses getroffen werden. Die Wirtschaftlichkeit sinkt mit Zunahme an erbrachten Tätigkeiten über die als notwendig identifizierten Leistungen. Ökonomisch optimal wäre ein Gleichgewicht zwischen obligatorischer und erbrachter Leistung (vgl.: Fischer, 1999:11). Über das Verhältnis von Bedarf und Aufwand kann des Weiteren eine Transparenz hergestellt werden, um zu identifizieren, welche Interventionen gemäß des Patientenzustands geplant wurden, jedoch aufgrund beschränkter personeller Mittel nicht durchgeführt werden konnten.

Es stellt sich nun die Frage: Wie kann der Bedarf an Leistungen identifiziert werden? In einigen Patientenklassifikationen wird der Interventionsbedarf gleichbedeutend mit geplanten Leistungen verwendet. Dies folgt der Annahme, dass das Planen von Tätigkeiten schon die Erklärung für die Notwendigkeit selbiger liefert. Doch eine sachlogische Begründung bleiben diese so konstruierten Systeme schuldig. Zu nennen sind hier Instrumente, wie bspw. PRN und PLAISIR.

Besonders schwerwiegend ist diese Gleichsetzung, wenn die planmäßigen Interventionen als Teil einer Tarifgrundlage in ein Vergütungssystem einfließen. Solche Instrumente spornen indirekt die Leistungserbringer an, nicht mehr Tätigkeiten als die geplant dokumentierten Leistungen auszuführen. Ein solcher Ansatz in der Patientenklassifikation steigert nicht bzw. erzeugt gar nicht erst die Motivation der Leistungserbringer, die Quantität der geplanten Leistungen auf ein Optimum an Notwendigkeit zu minimieren. Aus diesen genannten Gründen wird in der Literatur davon abgeraten dieser Idee der Bestimmung des Leistungsbedarfs zu folgen.

Ein anderer diskutierter Ansatz identifiziert den Leistungsbedarf unmittelbar aus der Patientenzustandsbeurteilung (vgl.: Fischer, 1999:11-12). Nach Bartholomeyczik et al. lässt sich der Interventionsbedarf von Patienten direkt aus der Person selbst und aus deren Umfeld ableiten (Bartholomeyczik et al., 2001:185).

Die vorliegende Arbeit folgt in ihren weiteren Ausführungen diesem Grundgedanken.

### Pflegeaufwand

Der Aufwand allgemein wird nach Dörrie und Preißler (2002) als ein „(...) nach Handels- und Steuerrecht bewerteter Wertverzehr“ eines Betriebes während einer abgegrenzten Abrechnungsperiode definiert. Weiterhin verstehen sie unter Aufwand den betrieblichen Input (z.B.: Sach- und Personalmittel) „unabhängig davon, ob der Wertverzehr dem Betriebszweck dient oder nicht“ (Dörrie, Preißler, 2002:40).

Der Begriff Pflegeaufwand meint jedoch in dieser Arbeit nicht den Wert der für die Durchführung einer Leistung zu erbringen ist, sondern die bereits erfolgten Interventionen. Diese Begriffsbestimmung des Aufwands folgt der Definition von Baumberger (2001:5). Ebenso definiert Isfort (2001:45) den pflegerischen Aufwand „als die Gesamtheit der tatsächlich erbrachten Pflegeleistungen“.

Über den Pflegeaufwand werden nicht ausschließlich nur die verrichteten Pflegeleistungen abgebildet, sondern auch Pflegekosten lassen sich auf diese Weise ermitteln (vgl.: Baumberger, 2001:6). „Aufgrund von Daten aus Pflegeaufwandsmessungen können Pflegekosten in DRG-bezogenen Fallpauschalen auf der Ebene der Kostengewichtung berücksichtigt werden“ (Fischer, 2002; Baumberger, 2001:6). In der hier vorliegenden Zusammenhangsstudie wird der Pflegeaufwand mit dem Erhebungsinstrument LEP-Nursing 3 Version 3.0.0 erhoben. Es existieren unterschiedliche Arten von Instrumenten die den Pflegeaufwand messen. Zum einen gibt es Prototypenmodelle, wie bspw. PPR und zum anderen bestehen Instrumente aus dem Genre des Faktorenmodells, wie zum Beispiel LEP. Die Thematik des Unterschieds zwischen beiden Modellen ist bei tiefer gehendem Interesse in Fischer (2002) nachzulesen. In dieser Arbeit wird lediglich im Kapitel 3.2.1 auf das Faktorenmodell am Beispiel von LEP eingegangen.

## **2.2 Stand der aktuellen pflegetheoretischen Fachdiskussion**

Die Einführung eines pauschalisierten Entgeltsystems bedingt die Ablösung der vorherigen Selbstkostenmethode durch die Beitragsmethode. Die Beitragsmethode steht für das Prinzip der Leistungsdeckung und nicht wie bisher für die Kostendeckung. Zielintention ist das Erreichen von Einsparungen und wirtschaftlichen Optimierungen in Krankenhäusern sowie ein gesundheitsökonomisches Denken zu fördern. Durch den Ansatz eines leistungsbezogenen Vergütungssystems wird das Denken in Grenzwerten gefordert. Das bedeutet, eine Gewinnmaximierung in Kliniken wird erreicht, wenn der Grenzumsatz gleich den Grenzkosten ist. Mit dieser neuen Perspektive in der Krankenhausfinanzierung soll das Denken in Durchschnittswerten eliminiert werden.

Das DRG-System ermöglicht eine bessere Beschreibung der intermediären Produkte eines Krankenhauses und mittels DRGs können Patientenfälle klinisch sinnvoll kategorisiert werden.

Die Einteilung der Patienten erfolgt hierbei über medizinische Diagnosen und Prozeduren sowie den Nebendiagnosen (vgl.: Baumberger, 2001:1-2). Diese Kategorien basieren jedoch so gut wie ausschließlich auf ärztlichen Informationen.

Ziel ist es Kategorien mit möglichst homogenen Fallgruppen zu konzipieren, d.h. die Behandlungsfälle in den jeweiligen Einordnungen sollten mit einer relativ ähnlichen klinischen Problematik behaftet sein. Der dahinter liegende Gedanke ist, dass diese Fälle auch einen ähnlichen Kostenverbrauch haben werden und somit kostenhomogene Gruppen erreicht werden können (vgl.: Fischer, 2002:8). In einem nächsten Schritt soll über diese Patientenklassifikation die Leistung eines Krankenhauses ermittelt und gemessen werden. Baumberger (2001:2) argumentiert, dass so die „DRGs als Bezugseinheit für Fallpauschalen“ dienen können, im Falle einer ausreichenden Homogenität der Behandlungskosten (vgl.: Baumberger, 2001:1-2). Jedoch weisen die Behandlungskosten nicht die gewünschte Kongruenz in den einzelnen DRGs auf.

Sollte eine Einheitlichkeit innerhalb der einzelnen DRG vorliegen, so müsste sich dies auch auf den Pflegeaufwand in den Kategorien beziehen. Doch diverse Studien zeigen eine gegensätzliche Tendenz. Eine hohe Streuung von Fallkosten und Aufenthaltstagen sowie dem Pflegeaufwand zeichnen sich verstärkt ab. Detaillierte Ausführungen zu einzelnen Studien die sich mit der Inhomogenität von DRGs auseinandersetzen können bei Fischer (2002:59ff.) nachgelesen werden. Allgemeiner Konsens in all diesen von Fischer (2002) vorgestellten Studien ist, dass die DRGs nur zu einem geringen Teil den Pflegeaufwand erklären können.

Dieses Problem ist aber keineswegs eine neue Erkenntnis, denn schon 1985 hat Halloran im Rahmen einer multiplen Regressionsanalyse Ergebnisse gewonnen die belegen, dass „nursing condition explained twice the variation in daily nursing workload (52,4%) than medical condition (26,3%)“ (Halloran, 1985:433).

Auch wenn Halloran Bezug nimmt auf ein US-amerikanisches DRG-System, hätte Deutschland möglicherweise diese Phänomene frühzeitiger berücksichtigen müssen.

Fischer (2002) bezieht sich in seinen Ausführungen auf einige Studien die das Problem der Kosteninhomogenität innerhalb der einzelnen DRGs abhandeln. Zur Erklärung dieser Uneinheitlichkeit im DRG-System werden unterschiedliche Ursachen diskutiert. Zum einen könnte die nicht genügende Effizienz des Betriebes ein Moment der ursächlichen Homogenitätsschwankungen darstellen. Zum anderen besteht auch die Möglichkeit, dass die verwendete Patientenklassifikation nicht differenziert genug ist. Auch eine mangelnde Codierqualität und eine ungeeignete Kalkulationsmethode werden als Anlass für Ungleichheiten im System gesehen. Fischer (2002) sieht darüber hinaus eindeutige Gründe für diese Problematik in der Gestaltung und Anwendung des DRG-Systems. Ferner zählen „unpräzise Falldefinitionen“, „Missachtung weiterer kostenrelevanter Patientenmerkmale“ und „Mög-

lichkeiten zur Manipulation der Gruppenzuteilung“ (Fischer, 2002:67) bedingt durch unzulängliche Begriffsbestimmungen der Hauptdiagnosen zu möglichen Erklärungsgründen (vgl.: Fischer, 2002:67). Einige dieser angesprochenen Probleme haben nicht unerhebliche Auswirkungen auf den pflegerischen Sektor in einem Krankenhaus.

Vor diesem Hintergrund wird verstärkt über die Existenz von independenten kostenbedeutenden Kriterien konferiert (vgl.: Fischer, 2002:8). In diese Überlegungen sollte nun auch die Pflege integriert werden. Fischer (2002:8) verfolgt die Idee, dass gerade in der Pflege Leistungen verrichtet werden, die nicht unmittelbar im Zusammenhang mit der medizinischen Diagnose stehen müssen. Ebenso vertritt Baumberger (2001:1-2) die Position der Integration pflegerischer Kriterien im DRG-System, zur Stabilisierung der Kostenhomogenität. Als bedeutsame pflegerische Kennzeichen werden Aspekte der Pflegediagnosen und des Pflegeaufwands, zum Beispiel in Form eines minimalen Datensatzes, angesehen (vgl.: Baumberger, 2001:1-2).

Doch diese Ansätze sind zurzeit noch Visionen. Momentane Ansätze zur Berücksichtigung der Pflege im DRG-System hängen immer zwei Arten an: Einerseits können über Klassifikationen und andererseits kann über Kostengewichtung sowie ergänzende Kriterien die pflegerischen Leistungen und Kosten abgebildet werden. Die drei von Fischer (2002) aufgeführten Möglichkeiten sollen kurz skizziert werden.

#### 1. Abbildung der Pflegekosten im DRG-Kostengewicht

Hierbei werden Pflegeleistungen in die Gesamt-Fallkostenberechnung mit eingerechnet (vgl.: Fischer, 2002:91). Derzeit erfolgt dies über den Einsatz der Pflege-Personal-Regelung (PPR). Doch die Kritik an der Verwendung der PPR ist nicht mehr zu überhören. Der Deutsche Pflegerat (DPR) äußerte sich kürzlich ebenfalls kritisch zur Nutzung der PPR im Rahmen der Kostengewicht-Berechnung (vgl.: DPR, 2007:5).

#### 2. Abbildung der Pflege durch den Einsatz der CC-Kategorien

Bei dieser Art der Darstellung findet ein Mapping zwischen Pflegediagnosen und ICD-10-Codes statt. Diese ICD-10-Codes beeinflussen wiederum die CC-Kategorien. CC steht für Komplikations- und Co-Mobilitätsstufen im DRG-System. Sie werden auch gewöhnlich als Nebendiagnosen deklariert.

Nach Fischer (2002) erscheint es sachlich sinnvoller und zweckdienlicher das Mapping zwischen Pflegediagnosen und ICD-10-Codes einer anders verlaufenden Systematik zu zuordnen. Nicht die ICD-10 sollte, Fischers Ansicht nach,

zu Pflegediagnosen klassifiziert, sondern Pflegediagnosen nach ICD-10 transcodiert werden (vgl.: Fischer, 2002:97).

Das in dieser Untersuchung zur Anwendung kommende Assessmentinstrument ePA-AC<sup>®</sup> folgte der von Fischer (2002:97) vorgeschlagenen Mapping-Systematik. Hier erfolgte eine Transcodierung der ePA<sup>®</sup>-Items nach ICD-10. Über diese Herangehensweise wird deutlich, welche pflegerischen Diagnosen bzw. in diesem Fall Items nicht in der ICD-10-Codierung abgebildet werden können.

Fischer versuchte 2002 NANDA<sup>3</sup>-Pflegediagnosen nach ICD-10 zu übersetzen. Dabei zeigten die Ergebnisse, dass gerade einmal ein Fünftel der NANDA-Diagnosen sich in der ICD-10 wieder gefunden haben (vgl.: ebd.:98). Vor dem Hintergrund dieser Ergebnisse kann die in der pflegewissenschaftlichen Fachwelt weit verbreitete Annahme gefestigt werden, dass Pflegenden Patienten im Gegensatz zu Medizinern aus einem anderen Kontext heraus beurteilen.

Für eine optimale Darstellung der Pflege mittels CC-Kategorien werden erweiterte CC-Listen angedacht. Die Erweiterung sieht vor ergänzende Elemente anderer Klassifikationssysteme ins System zu integrieren. In diesem Zusammenhang seien Klassifikationssysteme wie bspw. die ICF oder auch Pflegediagnosen erwähnt. Solche Bestrebungen wurden bereits schon 1996 von der Firma 3M versucht umzusetzen (vgl.: Fischer, 2002:95-97). Doch bisher haben sich diese Bemühungen noch nicht nachhaltig durchsetzen können.

### 3. Abbildung von Pflege im Zusammenhang mit pflegebezogenen Gruppen

Die relative Unabhängigkeit einiger von der Pflege eruiertes Patientenprobleme von medizinischen Diagnosen konnte in diversen Studien nachgewiesen werden. Beispielhaft sei an dieser Stelle die Studie von Welton und Halloran (2005) angeführt. In einer retrospektiven Analyse kombinierten Welton und Halloran Patientendaten des Pflegeklassifikationssystems „nursing diagnosis terminology (NDX)“ mit den „payer refined DRG (APR-DRG)“. Die so generierten Ergebnisse zeigten, dass „(...) the contention that nursing care is an independent predictor of patient hospital outcomes. These nursing data are not redundant with the medical diagnosis, in particular, the DRG“ (Welton et al., 2005:541-549)<sup>4</sup>. Die somit gewonnenen Erkenntnisse bilden die Basis für Überlegungen in Richtung Konzeption von gesonderten pflegerelevanten Fallgruppen. Dieser Ansatz wird momentan vom DPR (2007) verstärkt forciert. Der DPR hat kürzlich einen Projektauftrag zur Entwicklung von Nursing Related

---

<sup>3</sup> North American Nursing Diagnosis (Nordamerikanischer Verbund für Pflegediagnosen)

<sup>4</sup> Siehe ergänzend Fischer 2002, DPR 2007

Groups an Pflegewissenschaftler vergeben. Vor diesem Hintergrund wird unter anderem die aktuelle Brisanz dieser Thematik deutlich.

Auf die sich ergebende Chance Pflege über diesen Weg ins DRG-System als integralen Bestandteil zu implementieren bzw. die Begegnung beider Systeme auf gleicher „Augenhöhe“ zu fördern wird in Kapitel 7 detailliert erläutert. In dem sich anschließenden Kapitel soll die Bedeutsamkeit und Notwendigkeit der favorisierten Methoden zur Generierung pflegerischer Kennzeichen erörtert werden.

### **2.3 Zur Bedeutung der Pflegeklassifikationssysteme im DRG-System**

An dieser Stelle sei nochmals explizit darauf hingewiesen, dass Pflegeklassifikationssysteme bzw. Pflegediagnosen und Diagnosen zur DRG-Klassifikation auf keinen Fall synonym verwendet werden dürfen. Pflegeklassifikationssysteme werden in dieser Arbeit in Anlehnung an Etzel (2003) als Pflegediagnosen verstanden.

Pflegediagnosen können aus zwei Perspektiven diskutiert werden. Auf der einen Seite steht die Bedeutung der Diagnostik in der Pflege und auf der anderen Seite stellt dieses Klassifizierungssystem eine Informationsgrundlage zur Systematik von Begrifflichkeiten da (Bartholomeyczik, 2003:11).

Pflegediagnosen gewinnen bezüglich der Finanzierungsdebatte in Krankenhäusern an Bedeutung. König (2003:61) sieht den Einsatz von Pflegediagnosen im Zusammenhang mit der Vervollständigung der medizinisch relevanten Dokumentation, durch die Ergänzung pflegerischer Kennzeichen gegeben. Somit fungieren Pflegediagnosen im DRG-System als Instrumente, die eine Mitgestaltungskraft bzgl. der Leistungs- und Kostenfaktoren im Krankenhaus besitzen.

Nach Auffassungen von Fischer (1999:10) stehen Pflegediagnosen im Kontext pflegerischer Aussagen, welche den Gesundheitszustand eines Patienten beschreiben, die Ziele der pflegerischen Behandlung definieren und evaluieren. Im Vordergrund der Beschreibungen des Gesundheitszustandes stehen jene Aspekte, die Auslöser für ein mögliches pflegerisches Intervenieren darstellen.

Folgt man der Definition von Randi Mortensen (1996) können „Pflegediagnosen (...) als Phänomene bezeichnet werden, welche Gesundheitsprobleme und Gesundheitszustände umfassen, auf welche die Pflege einwirkt: vorbeugend, beeinflussend oder fördernd“.

Im Unterschied zu den Pflegediagnosen sind Angaben zu Pflegeleistungen Aussagen über Aktivitäten, welche die Pflegenden (als Leistungserbringer) ausführen. Das sind demnach zwei unterschiedliche Perspektiven, die auseinander

zuhalten sind. Pflegeklassifikationssysteme systematisieren Interventionen bzw. Aktivitäten durch die Kategorisierung. Interventionen sind hierbei ein Bündel aus Aktivitäten. Patientenklassifikationen bewegen sich hingegen eher auf der Durchführungsebene und gehen der Frage nach: Welche Maßnahmen müssen durchgeführt werden bzw. sind bereits erfolgt?

#### **2.4 Zur Bedeutung der Patientenklassifikation im DRG-System**

In der Pflege wird versucht den Aufwand erbrachter Interventionen mit Hilfe von Leistungserfassungsinstrumenten abzubilden. Diese Erhebungsinstrumente stellen gleichzeitig die Grundlage zur Pflegepersonalkostenkalkulation dar (Peters-Alt, 2005:69). Zur Ermittlung der Pflegepersonalkosten im DRG-System wird derzeit immer noch die PPR verwendet. Diese ist jedoch, so schreibt der Deutsche Pflegerat (2007), „keine dauerhaft geeignete Lösung“. Ihre Kritik an diesem Instrumentarium zielt auf die ungenügende Erfassung bzw. Abbildung des pflegerischen Behandlungsaufwands (vgl.: DPR, 2007:4). In der Fachliteratur wird eine Ablösung der PPR durch weitaus differenziertere Leistungserfassungsinstrumente diskutiert. An vorderster Stelle steht in dieser Diskussion das Instrument „Leistungserfassung in der Pflege (LEP)“. Auf dieses Instrument und dessen Bedeutung in der hier durchgeführten Untersuchung wird im Kapitel 3.2.1 Bezug genommen.

Der Weg zur Kostenkalkulation von Pflegepersonalkosten ausschließlich über Informationen der bereits erfolgten Leistungsdurchführung, ungeachtet des Leistungsbedarfs, kann nach Hunstein und Bartholomeyczik (2001:25) nicht als ein valides Verfahren zur Abbildung des Kostenanteils angesehen werden. Darüber hinaus besteht mittlerweile in der Pflegewissenschaft allgemeiner Konsens darüber, dass es nicht genügt nur den Aufwand einer pflegerischen Versorgung abzubilden, sondern dass eine Verknüpfung zu den jeweiligen Zuständen des Patienten hergestellt werden muss (vgl.: DPR, 2007:5; Etzel, 2003:63, Fischer, 1999:11).

Diese Forderungen werden nicht nur aus Gründen einer optimierten Kostenkalkulation laut, sondern auch aus den Überlegungen heraus hierüber Prozessanalysen auf inhaltlicher Ebene durchführen zu können (vgl.: DPR, 2007:5). Aus diesen generierten Daten können Aussagen zum Verhältnis von erbrachter und notwendiger Leistung ermöglicht werden. Über die Verbindung beider Aspekte lassen sich Ausführungen zur Wirtschaftlichkeit treffen (vgl.: Fischer, 1999:11).

Baumberger (2001:4) gibt jedoch zu bedenken, dass es zur Verwendung von Pflegediagnose- und Pflegeaufwandsdaten in Beziehung zum DRG-System derzeit nur wenige Untersuchungen existieren. Auch 2007 hat sich die Studienlage in

Deutschland nicht nennenswert verändert. Jedoch existieren zahlreiche Studien und Pilotprojekte im Ausland, welche Versuche unternommen haben, die Beziehung zwischen Pflegeklassifikationssystemen und dem örtlichen DRG-System zu analysieren. Einige wurden bereits im vorherigen Kapitel angesprochen. Des Weiteren kann sich an ausländischen Beiträgen zur möglichen Integration pflegerelevanter Indikatoren im DRG-System orientiert werden. Im Kapitel 7 werden diverse Umsetzungsmodelle skizziert.

### **3 Konzeptualisierung der Untersuchung**

In diesem Abschnitt sollen die einzelnen Arbeitsschritte der Untersuchung erläutert werden. Im Rahmen der Konzipierung des Studiendesigns wird eine Strukturierung des Vorgehens vorgenommen, die Variablen werden eruiert und definiert, die Hypothese formuliert und der theoretische Bezugsrahmen für die Analysemethode hergeleitet und skizziert.

#### **3.1 Querschnittsdesign**

Diese Zusammenhangsuntersuchung basiert auf einer quantitativen, nicht-experimentellen, deskriptiven Querschnittsstudie.

Im Rahmen einer Querschnittsstudie werden Daten zu einem konkreten Zeitpunkt an einer Stichprobe erhoben. Rückschlüsse auf die Richtung der Kausalität können mit dieser Art von Studiendesign nicht getätigt werden (vgl.: Polit et al., 1999:162).

In dieser vorliegenden Arbeit werden retrospektiv Daten analysiert, die bereits in der Pflegeprozessdokumentation im regulären Stationsablauf erhoben wurden. Das Datenmaterial, welche die unabhängige und die abhängige Variable bilden, wurde zeitnah erfasst. Eine Definition der Variablen befindet sich im Kapitel 3.2.

Der verwendete Datenbestand zur Analyse des Zusammenhangs von Patientenzustand und Pflegeaufwand wurde vom Kantonsspital Uri in Altdorf (CH) zur Verfügung gestellt. Die Daten wurden im Zeitraum vom 01. Oktober 2006 bis 31. Januar 2007 erhoben.

Die im Rahmen der Konzeptualisierung des Studiendesigns aufkommenden forschungsethischen Überlegungen orientieren sich an der Einhaltung des pflegerischen Ethik-Kodex. Der Ethik-Kodex „Human Rights Guidelines for Nurses in Clinical and other Research“ wurde von der American Nurses Association entworfen. Im Vordergrund dieser ethischen Richtlinie steht die Berücksichtigung und Würdigung der ethischen Grundprinzipien (vgl.: Polit et al, 1999:97-105).

Zur Datengewinnung wurden den in die Stichprobe aufgenommenen Patienten keine zusätzlichen Untersuchungen zu teil, da die Erhebungsdaten im routinemäßigen Ablauf des Pflegeprozesses erhoben wurden. Somit kann auch eine Zunahme der Arbeitsintensität bedingt durch diese Studie ausgeschlossen werden.

Alle im Rahmen der Untersuchung verwendeten Daten werden streng vertraulich behandelt und die Regeln des Datenschutzes, entsprechend dem Bundesdatenschutzgesetz (vom 20.12.1990), werden eingehalten. Die Anonymität der Daten wird dadurch gewährleistet, dass weder Patientennamen noch real existierende Patientenfallnummern für die Analyse gespeichert werden. Jeder Patient erhält vor der

Ausleitung der Daten eine neu zugewiesene fiktive Fallnummer. Somit kann ein späteres Zurückführen auf den tatsächlichen Patienten ausgeschlossen werden. Die Daten zur Analyse sind passwortgeschützt aufbewahrt und somit für Dritte nicht zugänglich.

Das Nichtveröffentlichen von einzelnen Patientenfalldaten, sondern lediglich in kumulierter Form, wurde zugesichert.

### 3.2 Methode

Mit dem Assessmentinstrument ePA-AC<sup>®</sup> werden die Daten erhoben, die für den Grad der Pflegebedürftigkeit bzw. dem Zustand eines Patienten stehen. Diese Daten besitzen ein „gemischtes“ Skalenniveau (ordinal- und binär skalierte Daten). Das Instrument LEP (Leistungserfassung in der Pflege) wird verwendet, um den Pflegeaufwand abzubilden. Diese Daten sind in ihrem Ursprung intervallskaliert. Sie werden jedoch im Rahmen der Gruppierung der abhängigen Variablen in kategoriale Daten umcodiert (siehe unterstehende Abbildung 1).

Fragestellung	Abhängige Variable (Pflegeaufwand)	Unabhängige Variable (Patientenzustände)
Mit welcher Wahrscheinlichkeit können Patienten im akutstationären Setting, in Abhängigkeit vom ePA-CasemanagementScore (definiert als prädiktiver Faktor), in die korrekte LEP-Minutengruppe eingeordnet werden?	<p><i>Gruppen von LEP-Minuten</i></p> <p>Es werden unterschiedliche Modelle gerechnet mit mannigfachen Zeitkorridoren und drei LEP-Interventionsgruppen (siehe Kap. 4.1):</p> <ul style="list-style-type: none"> <li>▪ 100 Minuten</li> <li>▪ 40 Minuten</li> <li>▪ 20 Minuten</li> </ul>	<p>10 Variablen (=10 CMS-Items aus dem ePA-AC<sup>®</sup>)</p> <p>u.a.: SPF Fortbewegung, SPF Mobilität, etc. (siehe Seite 26).</p> <p>In der Analyse wird mit dem niedrigsten CMS-Wert eines Patienten gerechnet.<sup>5</sup></p>

**Abbildung 1: Variablendarstellung**

Ein Fall wird in dieser Untersuchung als eine Einschätzung an einem Aufenthaltstag eines Patienten definiert. Anders ausgedrückt, eine ePA-Zwischeneinschätzung bezogen auf den erstmalig aufgetretenen jeweiligen niedrigsten CMS mit den

<sup>5</sup> Die Wahl für den niedrigsten CMS<sup>®</sup> erfolgte aus dem Grund, dass nicht zu jedem Aufenthaltstag eines Patienten auch ein CMS<sup>®</sup>-Wert vorlag. Somit konnte der Mittelwert nicht errechnet werden. Und eine Interpolierung der Werte ist in diesem Fall nicht zulässig. Eine detailliertere Begründung kann der Seite 30 entnommen werden.

dazugehörigen LEP-Minutenwerten und LEP-Interventionen an einem Tag eines Patienten wird als ein Fall betrachtet.

### 3.2.1 Auswahl der Erhebungsinstrumente

Damit vergleichbare Daten in dieser Untersuchung zur Verwendung kommen können, werden die Potentiale standardisierter Assessmentinstrumente genutzt. In der Regel können diese Daten über den routinemäßigen Versorgungsprozess gewonnen werden. Mit Nutzung der standardisierten Assessmentinstrumente kann eine Transparenz über den Zustand des Patienten mit den dazugehörigen pflegerischen Interventionen ermöglicht werden. Jedoch kann sich die Qualität dieser Transparenz auf einem unterschiedlichen Ausprägungsniveau befinden.

Vor dem Hintergrund der Bestrebungen der vorliegenden Arbeit, den Zusammenhang von Patientenzustand und Pflegeaufwand zu erörtern, um einen Anteil zur Sichtbarkeit sowie Budgetierung der pflegerischen Leistungen in der Krankenhausfinanzierung beizutragen, werden zwei ausgewählte standardisierte Instrumente vorgestellt. Diese beiden Instrumente bilden die Grundlage für die Operationalisierung der abhängigen und unabhängigen Variablen der Analyse-methode.

#### Ergebnisorientiertes PflegeAssessment-AcuteCare<sup>®</sup> (ePA-AC<sup>®</sup>)

Im Jahr 2002 begann die Entwicklung des ergebnisorientierten PflegeAssessments-AcuteCare<sup>®</sup> (ePA-AC<sup>®</sup>). Die Entwicklung vollzog sich schrittweise über einen Zeitraum von vier Jahren. Bewusst wurden bei der Entwicklung dieses Instruments Theorie- sowie Praxisphasen kombiniert. Nach anfänglich theoretischer Entwicklung des Assessments wurde eine fortwährende Anpassung bzw. Modifizierung in Zusammenarbeit mit den Pflegefachkräften der Stationen vorgenommen. Somit sollte die Einbindung der Bedürfnisse und Erfordernisse der Pflegenden als Anwender des Instruments sichergestellt werden. Zielsetzung dieser abduktiven Vorgehensweise war es, ein wissenschaftlich begründetes und gleichzeitig praktikables Instrument für die Praxis zu entwickeln.

Die Entwicklung des Kategoriensystems und dessen Items erfolgte über den Einsatz vielfältiger Methoden, wie zum Beispiel Literaturanalyse, Analyse bestehender Instrumente, Expertendiskussion, ENP<sup>6</sup>-Mapping und Rückmeldungen aus der Pflegepraxis (vgl.: Hunstein, Dintelmann et al., 2005:400, in: Oud et al., 2005). Das ePA-AC<sup>®</sup> ist als Screeninginstrument für den akutstationären Bereich konzipiert und umfasst 50 Items in 10 Kategorien. Die Items wurden in einem der Entwicklungsschritte der Internationalen Klassifikation der Funktionsfähigkeit,

---

<sup>6</sup> „European Nursing care Pathways“ (Wieteck, 2004)

Behinderung und Gesundheit (ICF) zugeordnet. Die Items des ePA-AC<sup>®</sup> erstrecken sich über die Komponenten „Aktivität und Partizipation“, „Körperfunktion“, „Körperstruktur“ und „Kontextfaktoren“ der ICF (vgl.: Hunstein et al., 2007, in: Oud et al., 2007).

Die Items sind je nach Inhalt dichotom oder 4-skaliert, wobei jeweils der Wert 4 die vollkommene Selbstständigkeit und der Wert 1 die vollständige Unselbstständigkeit beschreibt.

Des Weiteren beinhaltet das Assessment den CaseManagementScore<sup>®</sup> (CMS<sup>®</sup>), welcher sich aus 10 Selbstpflegefähigkeitsitems des Gesamtinstruments zusammensetzt. Über den CMS<sup>®</sup> soll mittels eines Gesamtwertes der 10 Items das post-stationäre Versorgungsdefizit eines Patienten vorausgesagt werden. Darüber hinaus soll der Schweregrad der Pflegeabhängigkeit bestimmt werden. Der Score kann auch als Steuerungsinstrument für pflegerische Prozessabläufe genutzt werden, wie z.B. für Übergaben am Bett, zur Bestimmung der Durchführung einer Pflegevisite, etc. Zur Veranschaulichung des Instruments kann ein ePA-AC<sup>®</sup>-Screenshot (DV-Version) im Anhang (Seite 82) eingesehen werden.

Eine Intention des ePA-AC<sup>®</sup> ist das mehrmalige Einschätzen relevanter Kennzeichen und Symptome von Pflegebedürftigkeit im akutklinischen Setting. Auf der Basis von Fremdeinschätzungen<sup>7</sup> werden wesentliche Kennzeichen und Merkmale von Patientenfähigkeiten und deren Beeinträchtigungen zu unterschiedlichen Messzeitpunkten ermittelt. Über das in einer stationären Aufenthaltsepisode wiederholte Erfassen der Patientenfähigkeiten wird die Möglichkeit zur Evaluation des Erfolges pflegerischer Interventionen geboten. Die Einschätzungen der Patienten werden mindestens zu zwei Messzeitpunkten erhoben. Eine Einschätzung muss zwingend bei Aufnahme und bei Entlassung eines Patienten erfolgen. Je nach stationsintern definierten Zeitintervallen sowie bei Zustandsveränderungen des Patienten erfolgen während eines Patientenaufenthalts Zwischeneinschätzungen.

Das Screeninginstrument wird momentan an der HSK (Dr. Horst Schmidt Klinik) auf 19 Stationen in der Papierversion routinemäßig angewendet. Auf vier Stationen gibt es eine Datenverarbeitungslösung (DV-Lösung). Des Weiteren wird das ePA-AC<sup>®</sup> im Kantonsspital Uri in Altdorf (CH) eingesetzt.

An der HSK wurde im Zeitraum von März bis Juni 2006 das Instrument in seiner bis dahin bestehenden Form (ePA-AC<sup>®</sup>-Beta) einer wissenschaftlichen Testung wesentlicher Gütekriterien unterzogen. Im Fokus der Testung stand die Analyse von Interrater-Reliabilität, interner Konsistenz sowie Konstruktvalidität und prognostischer Validität. Die statistische Testung der Daten für die Interrater-Reliabilität und

---

<sup>7</sup> Eine Ausnahme hierbei bildet das Item zur Beurteilung der Schmerzintensität, welches auf dem Prinzip der Selbsteinschätzung beruht.

der Konstruktvalidität sowie Änderungssensitivität befindet sich derzeit in der Abschlussphase. Publikationsreife Ergebnisse liegen zum gegenwärtigen Zeitpunkt noch nicht vor.

In dieser Untersuchung des Zusammenhangs von Patientenzuständen und Pflegeaufwand bildet das Assessmentinstrument ePA-AC<sup>®</sup> die Grundlage aus der die notwendigen Informationen zum Zustand bzw. zu den Fähigkeiten des Patienten ausgeleitet werden. Die Entscheidung für die Wahl des ePA-AC<sup>®</sup> bedingten sachliche Überlegungen. Dieses Instrument stellt ein auf seine Güte getestetes Assessmentinstrument dar, welches den Anspruch hat Patientenfähigkeiten zu erfassen. Und die Daten des ePA-AC<sup>®</sup> in der Ergänzung mit Leistungserfassungsdaten konnten vom Kantonsspital Uri bereitgestellt werden. Somit ist die Wahl das ePA-AC<sup>®</sup> als zustandsbezogenes Instrument in dieser Studie zu verwenden nahe liegend.

Die zur statistischen Zusammenhanganalyse verwendeten Items werden auf die 10 CMS<sup>®</sup>-Items reduziert. „Der CMS<sup>®</sup> weist laut den Entwicklern auf das Risiko hin, dass nach dem stationärem Aufenthalt im Krankenhaus auf Grund von Pflegebedürftigkeit weitere Versorgungs- und Unterstützungsleistungen erforderlich sein können“ (Schlarmann, 2007:6-7). Schlarmann (2007:24) konnte im Rahmen seiner Analysen für den CMS<sup>®</sup> eine Sensitivität von 80,65% und eine Spezifität von 93,73% bei einem Cut-off-point  $\leq 32$  ermitteln. Vor dem Hintergrund dieser Zahlen und gemäß der kongruenten Definition des CMS<sup>®</sup> mit der Bestimmung der unabhängigen Variable wird dieser Scorewert in die hier durchzuführende logistische Regressionsanalyse einbezogen.

In der bereits im Vorfeld durchgeführten Untersuchung mit der hier zum Einsatz kommenden Stichprobe zeichnete sich eine Erklärungskraft der Varianz von Pflegezeiten mittels der Verwendung des CMS<sup>®</sup> von 60% ab (Hunstein et al., 2007b).

Folgende Items bilden in ihrer Summe den CaseManagementScore<sup>®</sup> im ePA<sup>®</sup>:

<b>CMS<sup>®</sup>-Item</b>
Selbstpflegefähigkeit Aktivität/Fortbewegung
Selbstpflegefähigkeit Nahrungsaufnahme Essen
Selbstpflegefähigkeit Nahrungsaufnahme Trinken
Selbstpflegefähigkeit Urinausscheidung durchführen
Selbstpflegefähigkeit Stuhlausscheidung durchführen
Selbstpflegefähigkeit Körperpflege Oberkörper
Selbstpflegefähigkeit Körperpflege Unterkörper
Selbstpflegefähigkeit An- und Auskleiden Oberkörper
Selbstpflegefähigkeit An- und Auskleiden Unterkörper
Fähigkeiten Kenntnisse zu erwerben

(vgl.: Hunstein et al., 2006)

Die höchste erreichbare Summe des CMS<sup>®</sup> liegt bei einem Punktwert von 40, d.h. es liegt keine Beeinträchtigung der Fähigkeiten des Patienten vor. Der niedrigste mögliche Punktwert befindet sich bei 10. Das bedeutet, dass bei diesem Patient eine vollständige Beeinträchtigung der Selbstpflegefähigkeiten<sup>8</sup> vorliegt.

Der CMS<sup>®</sup> bildet vor diesem Hintergrund den prädiktiven Faktor für die Bestimmung des Pflegeaufwands während eines Patientenaufenthalts im klinischen Setting.

#### Leistungserfassung in der Pflege (LEP<sup>®</sup>)

Das Instrument „Leistungserfassung in der Pflege“ ist eine in der Schweiz entwickelte und dort weit verbreitete Patientenklassifikation. LEP gehört der Gruppe der Faktorenmodelle an. Faktorenmodelle unterscheiden sich von Prototypenmodelle dahingehend, dass sie einzelne Zeitwerte entsprechenden Interventionen zuordnen und keine Zeitkorridore, wie z.B. das Instrument PPR, verwenden. Bei Faktorenmodellen steht die Gewichtung der Pflegeleistungen, und nicht die Gewichtung der Pflage tage im Vordergrund (siehe ergänzend Fischer, 2002:140-145).

Die LEP-Methode enthält 981 Variablen in ihrem Katalog der Version Nursing 3, davon sind 72 Pflegevariablen auf der vierten Gliederungsebene<sup>9</sup>.

Eine Variable setzt sich generell aus den Komponenten Variablennennung, Definition und einem vorbelegtem Zeitwert zusammen (vgl.: Brosziewski, 2001:60). Aus diesen Tätigkeitsvariablen wird der unmittelbare, patientenbezogene Pflegeaufwand errechnet. Dies ergibt in der Summe allen patientenbezogenen pflegerischen Aufwands einer Station den Gesamtpflegeaufwand. In einer Gegenüberstellung wird der unmittelbare Pflegeaufwand der pflegerischen Personalarbeitszeit zugeordnet. Die unterschiedlichen Kompetenzniveaus der Pflegenden (z.B. Pflegepersonen in Einarbeitungszeit, Auszubildende, etc.) können hierbei entsprechend prozentual berücksichtigt werden.

Leistungen die von Pflegenden erbracht werden, sich jedoch nicht unmittelbar einem Patienten zuordnen lassen, werden bei LEP gesondert errechnet. Genannt seien hier Tätigkeiten, die sich im Bereich der Administration, Ausbildung, Managementaufgaben, Service und Übergaben etc. bewegen. Diese nicht patientenbezogen zuordenbaren Leistungen werden an der Personalarbeitszeit gemessen und prozentual in Form des C-Werts ausgedrückt (vgl.: Bamert, 2003:140-141). LEP wird von Pflegefachkräften angewendet und ist unter anderem im akutklinischen Bereich einsetzbar.

---

<sup>8</sup> Selbstpflegefähigkeiten werden im ePA-AC<sup>®</sup> nicht kongruent zu der Definition von Dorothea Orem verstanden, sondern beziehen sich auf den Umfang der Fähigkeiten, Aktivitäten selbstständig durchzuführen (Hunstein et al., 2007b).

<sup>9</sup> LEP splittet seine Tätigkeitsvariablen in insgesamt fünf Gliederungsebenen.

LEP versteht sich als ein in der Sozialwissenschaft entwickeltes System (vgl.: Brosziewski, 2001:60). Die Einsatzmöglichkeiten der LEP-Datennutzung erstrecken sich vom Controlling und Management über die klinikinterne Planung bis hin zur externen Pflegeleistungsdokumentation. Bamert erwägt des Weiteren LEP im Zusammenhang mit der Durchführung von Benchmarking und als Basis epidemiologischer Fragestellungen einzusetzen (vgl.: Bamert, 2003:140).

Im Rahmen der Untersuchung zum Zusammenhang von Patientenzuständen und Pflegeaufwand wurde sich für das aufwandsbezogene Instrument LEP entschieden, da es zweckgebunden aussagekräftige und standardisierte Daten über die aufgewendeten patientenbezogenen Leistungen einer Klinik generiert.

### ePA<sup>®</sup> und LEP

Aus der Kooperation zwischen der LEP AG und dem ePA<sup>®</sup>-Entwicklerteam ist eine Verknüpfung beider Instrumente hervorgegangen. Die Verknüpfung beinhaltet eine theoriebasierte Kategorisierung der LEP-Interventionen zu den fachlich entsprechenden ePA<sup>®</sup>-Items.

Derzeit wird die Verbindung mit ePA<sup>®</sup> und LEP im Kantonsspital Uri in Altdorf getestet. Hierfür steht den Pflegenden eine Softwarelösung zur Verfügung. Das Softwareprogramm wurde so entwickelt, dass es bei erhobenen Kennzeichen und Symptomen von Pflegebedürftigkeit mittels des ePA-AC<sup>®</sup> automatisiert Vorschläge für mögliche Interventionen auf LEP-Basis unterbreitet. Diese Vorschläge sollen auch als Empfehlungen verwendet werden und nicht als dogmatische Gegebenheit hingenommen werden.

## **3.2.2 Stichprobenbeschreibung**

### Einschlusskriterien

In die Stichprobe eingeschlossen wurden Patienten die im Zeitraum Oktober 2006 bis Januar 2007 auf einer der in die Studie einbezogenen Stationen im Kantonsspital Uri in Altdorf aufgenommen waren. Hierzu zählten fünf interdisziplinäre Stationen.

Das Mindestalter der in die Studie eingeschlossenen Patienten wurde auf größer 12 Jahre festgesetzt. Die Begründung der Mindestaltersgrenze liegt in der originären Konzeption des ePA-AC<sup>®</sup>. Das ePA-AC<sup>®</sup> wurde nicht entwickelt für den Bereich der Kinderklinik. Dieses Instrument eruiert Fähigkeiten bzw. die Einschränkungen dieser Fähigkeiten. Bei Kindern kann ein natürlicher Hilfebedarf vorliegen der nicht über das ePA-AC<sup>®</sup> erfasst wird. Die Gewichtung der Fähigkeitseinschränkungen berücksichtigt den Aspekt des natürlichen Hilfebedarfs nicht in ihrer Operationalisierung. Hunstein et al. orientieren sich bei ihrer Begründung der Festlegung des Mindestalters an der Definition des Hilfebedarfs bei Kinder nach § 14 SGB XI

(vgl.: Hunstein et al., 2006:6, Medizinischer Dienst der Spitzenverbände der Krankenkassen, 1997:45). Mit der Festsetzung des Mindestalters von Patienten kann somit ausgeschlossen werden, dass das Assessmentinstrument bei Patienten, die nicht der Zielgruppe entsprechen, eingesetzt wurde.

### Ausschlusskriterien

Im Rahmen der Datenbereinigung wurden ePA-AC<sup>®</sup>-Einschätzungen der Patientenfälle bei folgenden Kriterien ausgeschlossen:

- Generell aus der Stichprobe eliminiert wurden Patientenfälle die kürzer als 24h auf einer Station verweilten, d.h. Erst- und Abschlusseinschätzung eines Patienten erfolgten am selben Tag.
- Bei Patientenfällen die mehr als eine Zwischeneinschätzung an demselben Tag aufwiesen, wurden diese Einschätzungen auf ihre Plausibilität überprüft. Konnte die Begründung einer zweiten Einschätzung am gleichen Tag auf der Datengrundlage rekonstruiert werden (z.B.: Patient hatte einen operativen Eingriff oder eine signifikante Zustandsveränderung), wurde die letzte gültige Einschätzung diesen Tages in die Stichprobe aufgenommen. Die erste Einschätzung wurde ausgeschlossen. Stellte sich kein Unterschied zwischen den mehrfach erfolgten Einschätzungen da, wurde ebenfalls die zuletzt erhobene Einschätzung als gültig in die Stichprobe aufgenommen und die anderen Einschätzungen des gleichen Tages wurden ausgeschlossen.
- Hatte ein Patient am selben Tag eine Zwischen- und eine Abschlusseinschätzung wurde die abschließende Einschätzung in die Stichprobe eingeschlossen.

Bei den erhobenen LEP-Interventionen eines Patienten wurden Fälle (entspricht einer ePA-Einschätzung) ausgeschlossen bei denen:

- Nur die geplanten Leistungen dokumentiert waren, jedoch die tatsächlich erfolgte Leistung nicht durchgeführt wurde.
- Die LEP-Interventionen mit hohen Minutenwerten (>100), wie z.B. Kleiden 900 Minuten oder Infusion 160 Minuten, sind über eine Kreuztabelle auf ihre sachlogische Verständlichkeit überprüft worden. Die Einschätzungen zu den hohen unplausiblen Minutenwerten wurden aus der Stichprobe eliminiert.
- In einem letzten Schritt wurden Patienten mit einem CMS von 40 (entspricht = keine Beeinträchtigung in den Selbstpflegefähigkeiten) auf unplausible hohe LEP-Minutenwerte überprüft. Hohe Minutenwerte sind definiert als Werte >11 Minuten. Dieser Grenzwert wurde gebildet unter der Überlegung, dass Patienten

die keine Einschränkungen in ihren Selbstpflegefähigkeiten haben, auch keine sachlogische Erklärung für Leistungen, die sich aus der Pflegebedürftigkeit eines Patienten heraus begründen, in Minutenbereichen über 10 Minuten vorweisen können. Diese betrachteten Minutenwerte beziehen sich auf Leistungen, die sich auf die CMS-Items beziehen lassen. Das betrifft zum Beispiel die LEP-Interventionen Mobilisation, Lagern, Kleiden, Ausscheidung etc. Auf diesen Aspekt der Kategorisierung der LEP-Interventionen wird im Kapitel 4.1 eingegangen.

### Endgültige Stichprobe

Nach Berücksichtigung der Einschlusskriterien entsprachen den Studienanforderungen 1115 Einschätzungen von 436 Patienten mit insgesamt 21648 einzelnen LEP-Interventionen.

Nach Setzung der Filterfunktion „niedrigster CMS“ wurde die Anzahl der Einschätzungen<sup>10</sup> auf 436 gesetzt. Die Entscheidung für den niedrigsten CMS-Wert ist unter der Überlegung getroffen, dass ein Tag gewählt werden sollte, der bei jedem Patienten vorhanden ist. Dieser Tag sollte jedoch weder der Tag der Aufnahme, noch den Entlassungstag darstellen. Eine höhere Gewichtung mehrfach eingeschätzter Patienten konnte über diesen Weg ebenso ausgeschlossen werden. Der Maximum-CMS-Wert wurde in der Betrachtung nicht berücksichtigt, da aufgrund der nicht normalverteilten Stichprobe ein Deckeneffekt im Bereich der 40iger CMS-Werte gegeben ist.

Ein Selektionseffekt kann nach Ausschluss von 679 Einschätzungen nicht ausgeschlossen werden. Es ist vorstellbar, dass Patienten die weniger als 24h auf einer Station verweilen, dennoch einen äußerst hohen Pflegezeitverbrauch aufweisen könnten. Ebenfalls kann die Entscheidung für den ersten jemals aufgetretenen Minimum-CMS<sup>®</sup> und Maximum-CMS<sup>®</sup> einen Selektionseffekt bewirken. Kritisch zu bemerken ist bei dieser Entscheidung, dass möglicherweise dies nicht die Tage mit den höchsten Interventionszeiten bzw. größten Interventionsanzahlen sind.

Das Durchschnittsalter der Patienten in der Stichprobe (N=436) beträgt 69,16 Jahre. Die Gesamtstichprobe hat einen Frauenanteil von 50,9% und einen Männeranteil von 49,1%.

---

<sup>10</sup> Einschätzungen können in diesem Fall mit Patienten gleichgesetzt werden, da bei jedem Patienten nur der erste jemals aufgetretene niedrigste CMS (entspricht dem CMS-Minimum) berücksichtigt wurde.

Zu den drei häufigsten Hauptdiagnosen in dieser Stichprobe zählten „Eingriffe an unteren Extremitäten + Humerus, ausgenommen Hüfte, Fuß und Femur“, „Eingriffe an großen Gelenken“ und „Schlaganfall“.

Die durchschnittliche Verweildauer in dieser Stichprobe (N=436) liegt bei einer Spanne zwischen 3 und 135 Verweildauertagen bei 15,6 Tagen. Der Median-Wert liegt bei 12 Tagen. In der Angabe der durchschnittlichen Verweildauer sind die Ausreißer<sup>11</sup> nicht rausgerechnet. In nachfolgender Abbildung werden die Ausreißer über ein Box-Plot verdeutlicht.

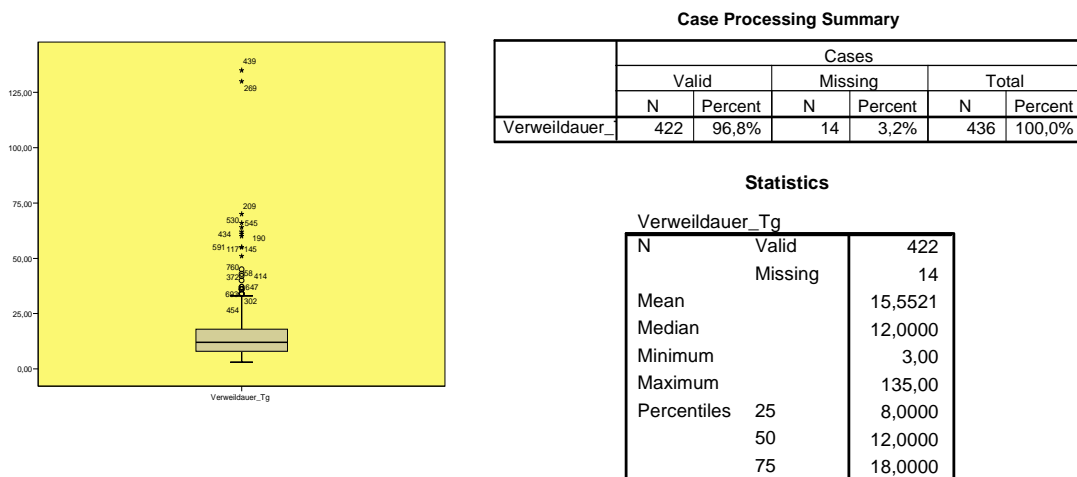


Abbildung 2: Darstellung der Ausreißer bei der Verweildauer (N=436 Einschätzungen)

In der Abbildung werden die Ausreißer in Form von „Kreisen“ und extreme Ausreißer in Form von „Sternen“ symbolisiert.

Angaben zu dem Verhältnis von Unter- und Überliegern<sup>12</sup> in dieser Stichprobe können in der Tabelle 1 eingesehen werden.

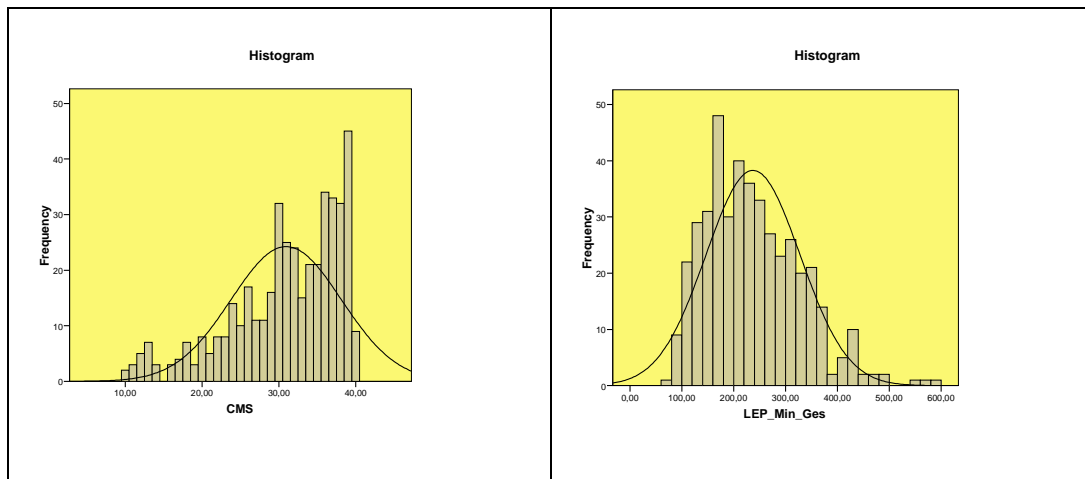
	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 1	52	11,9	11,9	11,9
1=Low Outliers	3	,7	,7	12,6
2=Inliers	327	75,0	75,0	87,6
3=High Outliers 1	40	9,2	9,2	96,8
4=High Outliers 2	14	3,2	3,2	100,0
Total	436	100,0	100,0	

Tabelle 1: Unter- und Überlieger in der Stichprobe (N=436)

<sup>11</sup> Als Ausreißer bezeichnet man enorm große und kleine Werte, die von den restlichen Werten stark abweichen (vgl.: Brosius, 2006:391). Zur Berechnung der Ausreißer gibt es unterschiedliche Methoden (siehe Fischer, 2002:51).

<sup>12</sup> Unter- und Überlieger sind jene Fälle die nach dem AP-DRG-System als solche definiert und identifiziert sind.

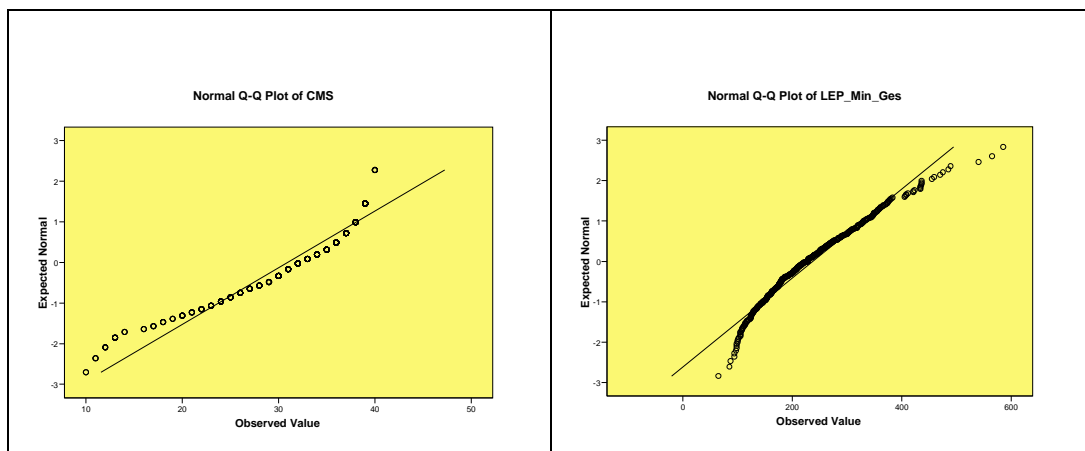
Mit einer Prozentzahl von 11,4 nehmen die Überlieger einen Anteil von ca. einem Siebentel an der Gesamtstichprobe ein. In der regulären Verweildauer wurden 327 Fälle (Patienteneinschätzungen) behandelt. Die Unterlieger nehmen mit 0,7% bezogen auf die Gesamtstichprobe einen verhältnismäßig geringen Anteil ein. Eine Verteilung der CMS<sup>®</sup>-Werte und der einzelnen LEP-Minutenwerte (bei N=436) kann der graphischen Abbildung 3 entnommen werden.



**Abbildung 3: Verteilung CMS und LEP-Minuten in der Stichprobe (N=436)**

Über diese Darstellung wird ersichtlich, dass bei der Variable „CMS<sup>®</sup>“ (linke Abbildung) eine linksschiefe Verteilung vorliegt. Für die Variable „LEP-Minuten-Gesamt“ wird eine steile, leicht rechtsschiefe Verteilungsform ersichtlich.

In einem nächsten Auswertungsschritt wurde die Normalverteilung des CMS<sup>®</sup> sowie der LEP-Minuten in der Gesamtstichprobe getestet. Mit Hilfe eines Q-Q-Diagramms soll die Normalverteilung dargestellt werden.



**Abbildung 4: Normalverteilungsdiagramm für die Variablen CMS und LEP-Minuten\_Gesamt**

Im Falle einer Normalverteilung müssten die einzelnen Punkte direkt der geraden Linie, welche die theoretische Normalverteilung abbildet, folgen. Sowohl bei der

Variable CMS als auch bei der Variable LEP-Minuten ist dies nicht eindeutig der Fall. Beide Grafiken zeigen eine Abweichung von der Normalverteilung in den Extrembereichen. Zur Prüfung der statistischen Bedeutsamkeit dieser Abweichung wurde der Lilliefors-Test<sup>13</sup> durchgeführt. In beiden Variablen stellte sich ein signifikantes Ergebnis dar. Das bedeutet, dass eine signifikante Abweichung von einer Normalverteilung vorliegt. Detaillierte Darstellungen und Tabellen können dem Anhang (S.85-88) entnommen werden.

Vor dem Hintergrund einer nicht gegebenen Normalverteilung und anderer nicht erfüllbarer Anforderungen an eine lineare Regressionsanalyse wurde sich in dieser Untersuchung zur Beantwortung der Forschungsfrage für eine logistische Regression entschieden. Im Kapitel 4 wird diese Analyseverfahren im Einzelnen dargestellt.

### 3.2.3 Hintergrund zum Datenmaterial

Die durchgeführte Untersuchung basiert auf Daten aus dem Kantonsspital Uri in Altdorf (CH). Die Daten sind aus dem dortigen System in anonymisierter Form der Abteilung Pflegeforschung/ -entwicklung der HSK, Horst Schmidt Klinik Wiesbaden, zur Verfügung gestellt.

Das Kantonsspital Uri ist ein Akutspital mit einem Beleg- und Konsiliarärztesystem. Im Jahr 2006 verzeichnet das Spital 3.710 stationäre Fälle (vgl.: <http://www.ksuri.ch/15.html>; Stand: 28.04.2007). Das Spital setzt sich zusammen aus den Kliniken Medizin, Chirurgie, Geburtshilfe und Gynäkologie sowie der interdisziplinären Intensivstation und der Notfallstation. Die mittlere Aufenthaltsdauer beträgt 11,5 Tage. Insgesamt sind in der Klinik in Altdorf 370 Mitarbeiter beschäftigt (vgl.: Urner Kantonalbank, .2005:27, zitiert nach: Jahresbericht Kantonsspital Uri, 2003).

---

<sup>13</sup> Lilliefors-Test ist geeignet für Datensätze deren Populationsparameter geschätzt werden müssen. Bei diesem statistischen Test bedeutet ein signifikantes Ergebnis eine signifikante Abweichung von der Normalität (vgl.: Brosius, 2006:401).

## 4 Logistische Regressionsanalyse

In der Praxis der Sozialforschung steht man häufig vor dem Problem lediglich auf nominal- oder ordinalskalierte Variablen zugreifen zu können. Hier sei ein Beispiel für die Thematik Pflegebedürftigkeit genannt:

- Liegt eine Pflegebedürftigkeit bei einem Patienten vor: ja/nein
- Welche Ausprägung der Pflegebedürftigkeit liegt vor: Stufe 1/ Stufe2/ Stufe3 (vgl.: SGB XI)
- Selbstpflegefähigkeiten: unselbständig/ umfangreiche Unterstützung/ geringe Unterstützung/ selbständig (vgl.: Hunstein et al., 2006)

Diese Ausprägung der Skalenqualität verhindert das Anwenden der linearen Regressionsanalyse, da fundamentale Voraussetzungen dieser Analyseverfahren nicht erfüllt werden. Das ist zum einen die Bedingung des Vorliegens der „varianzhomogene(n) Residuen“ (Baltes-Götz, 2006:6) und einer Normalverteilung. Ein anderer zu wählender Ansatz ist die Diskriminanzanalyse. Diese Analyseverfahren verfolgen die Grundidee der geradlinigen Gruppenzugehörigkeitsprognose. Jedoch auch dieser methodische Ausgangspunkt lässt sich nicht bei oben aufgeführten Fragestellungen verwenden. Auch diese Analyseverfahren fordern die Bedingungen, dass:

- die Daten Intervallskalierungen aufweisen
- häufig eine Normalverteilung in einer Stichprobe zu den Variablenausprägungen vorhanden ist
- eine „Homogenität der Kovarianzmatrixen“ existiert (Baltes-Götz, 2006:6)

Die in dieser Arbeit verwendete Stichprobe kann den eingeforderten Grundvoraussetzungen dieser beiden Analyseverfahren nicht genügen. Eine Normalverteilung in der Stichprobe liegt nicht vor (siehe Kap. 3.2.2) und ebenso wenig werden intervallskalierte Variablen verwendet. Daher können diese Methoden für die Analyse der erhobenen Daten zur Beantwortung der Forschungsfrage ausgeschlossen werden.

Einen alternativen Ansatz für Erklärungsmodelle mit kategorialen Kriteriumsvariablen bilden log-lineare bzw. logit-lineare Methoden. Diese Methoden werden jedoch in der regressionsanalytischen Fachliteratur als sehr komplexe und „verwirrende“ Modelle beschrieben und „als relativ unhandlich“ (Baltes-Götz, 2006:6) diskutiert.

Baltes-Götz (2006) empfiehlt ein Analysemodell, welches eng verwandt ist mit log-linearen Methoden, aber für die Untersuchung kategorialer Variablen geeigneter erscheint. In diesem Zusammenhang wird die *logistische Regressionsanalyse* erwähnt (vgl.: Baltes-Götz, 2006:6).

Die logistische Regression ist wie auch die Diskriminanzanalyse und Regressionsanalyse ein strukturen-prüfendes Verfahren. Die logistische Regression hat die Fragestellung nach der Wahrscheinlichkeit für das Eintreten eines bestimmten Ereignisses zu Grunde gelegt.

Das heißt, dass dieses Verfahren aus empirischen Ereignisbeobachtungen eine Wahrscheinlichkeit für das Eintreten dieser Ereignisse ableitet (vgl.: Backhaus et al., 2003:418-420).

Übertragen auf die durchzuführende Untersuchung dieser Arbeit wird daher die Wahrscheinlichkeit des Pflegeaufwands einer Person in Abhängigkeit von der Einflussgröße, Pflegebedürftigkeit bzw. Patientenzustand gesucht. Eine Kongruenz zur aufgestellten Forschungsfrage wird ersichtlich.

Viele verwendete Test- und Schätzverfahren die bei einer logistischen Regression Anwendung finden, werden gleichermaßen bei der linearen Regressionsanalyse herangezogen.

- „(...) der Test zur globalen Nullhypothese, dass alle Parameter außer dem konstanten Term gleich 0 sind,
- die Tests zu den Nullhypothesen für die einzelnen Parameter
- Bestimmtheitsmaße zur Beurteilung der Modellrelevanz“ (Baltès-Götz, 2006:7)

Durch die Wahrscheinlichkeitsmodellierung in einer logistischen Regressionsanalyse kommen einige Eigenheiten zum Tragen. Auf die zugrunde gelegte Modellgleichung wird zu einem späteren Zeitpunkt Stellung bezogen.

Im Rahmen einer logistischen Regressionsanalyse können binäre Kriterien (d.h., dichotome Ausprägungen) sowie multinominale Kriterien (d.b., mehr als zwei Ausprägungen) verwendet werden (vgl.: Backhaus et al., 2003:420).

Für eine logistische Regression bietet das Programm „Statistical Package for Social Sciences“ (SPSS) drei mögliche Prozeduren an: Logistic Regression (für binäre Variablen), NOMREG (für multinominale Variablen) und PLUM (für ordinale Variablen). Logistic Regression und NOMREG verwenden zwar einen differierenden Algorithmus, welcher jedoch in den Ergebnissen nicht von besonderer Bedeutung ist (vgl.: Baltès-Götz, 2006:7). Zu dieser Erkenntnis kommen auch Backhaus et al. (2003:420).

#### **4.1 Datenaufbereitung**

Nach Abschluss der Datenbereinigung unter Beachtung der Ein- und Ausschlusskriterien der Stichprobe (siehe Kapitel 3.2.2) wurde im Rahmen der Datenaufbereitung für die statistische Analyse eruiert, ob sich die LEP-Interventionen aus

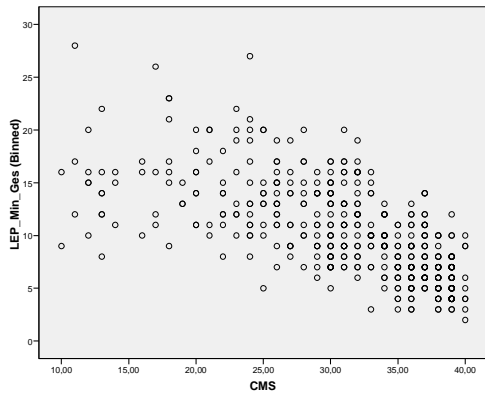
den ePA-AC<sup>®</sup>-Items herleiten lassen. In einem zweiten Schritt wurde ein Zuordnung zwischen LEP-Interventionen und dem Internationalen Klassifikationssystem ICF durchgeführt. Hierbei stand im Vordergrund, nach Kategorisierung der Interventionen in die ICF-Komponenten<sup>14</sup>, zu ermitteln, welche dieser Interventionen sich in die Kategorie „Aktivität und Partizipation“ einordnen lassen. Dies erfolgte unter der Überlegung, dass Pflegebedürftigkeit in weitestgehend allen Definitionen in Verbindung mit einer „beeinträchtigten Autonomie bei der Lebensgestaltung und die dadurch bedingte Abhängigkeit von Hilfe bei der Gestaltung des Alltags und krankheitsbedingter Erfordernisse“ (Bartholomeyczik, 2004:392) gebracht wird.

Der Ursprung der Kategorisierung der LEP-Interventionen liegt im Gedanken, die Betrachtung von Pflegeaufwand für die zu untersuchende Fragestellung zu reduzieren. Die Reduktion richtet sich dabei auf den Teil der Leistungen die in enger Verbindung mit Pflegebedürftigkeit von Patienten zu bringen sind. Und Pflegebedürftigkeit wird hierbei als eine Form des Ausdrucks der Selbstpflegefähigkeiten angesehen und genau dieser Aspekt soll mittels CMS-Items erfasst werden. Daher lag es nahe, eine Reduzierung der LEP-Interventionen hinsichtlich dieser Gesichtspunkte vorzunehmen. Die Leistungen wurden schließlich für die Analyse in die drei folgenden Gruppen kategorisiert: *Erstens* die Gruppe der so genannten pflegerisch induzierten Leistungen. Hier sind Interventionen zusammengefasst, die durch beeinträchtigte Selbstpflegefähigkeiten eines Patienten ausgelöst werden. *Zweitens* die Gruppe der medizinisch induzierten Leistungen, welche überwiegend aus diagnostischen und therapeutischen Gründen der Medizin ausgelöst werden. Und *drittens* die Gruppe in der keine Differenzierung der LEP-Interventionen vorgenommen wurde.

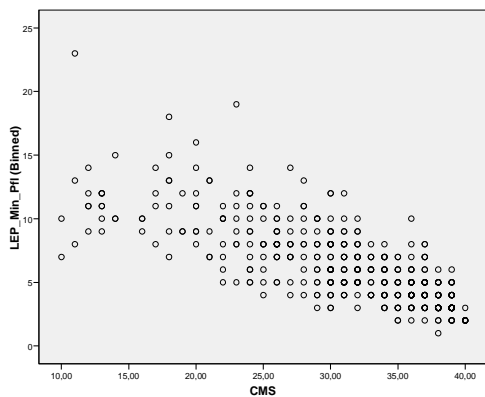
Die Einheit der pflegerischen Leistungen beinhaltet nachstehende LEP-Interventionen (hier sind nur LEP-Variablenobergruppen genannt): Mobilisation, Lagern, Umbetten, Körperpflege, Kleiden, Bett-/Liegeplatz, Essen/Trinken, Ausscheidung. Eine differenzierte Kategorisierung kann im Anhang (S. 89-91) eingesehen werden. Eine Gegenüberstellung der Gruppen, mittels eines Scatterplot, in Hinblick auf einen Zusammenhang zwischen dem CMS<sup>®</sup>-Minimum eines Patienten und den jeweiligen gruppierten LEP-Minuten bzw. Interventionen ergeben erste Hinweise auf eine Zusammenhangstendenz in den Einheiten. Beispielhaft wird hier der Vergleich der drei Gruppen in dem LEP-Minutenintervall von 20Minuten dargestellt. Die ausführlichen Abbildungen befinden sich im Anhang.

---

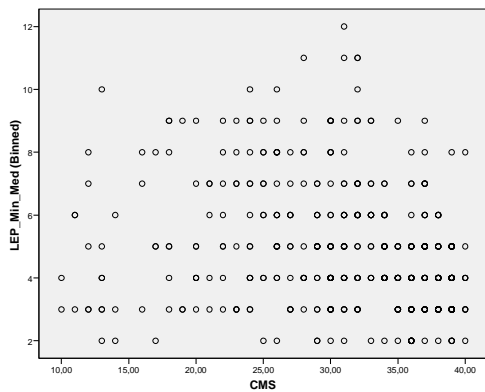
<sup>14</sup> ICF beinhaltet die Komponenten „Körperfunktion und –struktur, Aktivität und Partizipation“ sowie die als Kontextfaktoren definierte Komponenten „Umweltfaktoren, personenbezogene Faktoren“ (ICF, 2005:17).



Scatterplot: LEP\_Minuten der Gruppe 3 im 20Min. Intervall



Scatterplot: LEP\_Minuten der Gruppe 1 im 20Min. Intervall



Scatterplot: LEP\_Minuten der Gruppe 2 im 20Min. Intervall

### Abbildung 5: Gegenüberstellung der LEP-Interventionsgruppen

Objektiv muss jedoch angemerkt werden, dass dieser relativ eindeutige Unterschied sich mit Zunahme der Minutenintervalle abschwächt (siehe Anhang).

Im nachfolgenden Kapitel wird das Modell der in dieser Untersuchung zur Anwendung kommenden logistischen Regression formuliert und erläutert.

Es sei ausdrücklich darauf hingewiesen, dass die Darstellungen lediglich als Hinweise auf mögliche Zusammenhänge betrachtet werden können. Eine statistische Signifikanz wird über diese Abbildungen nicht gegeben. Doch es gibt eine Tendenz an, dass die Entscheidung der Gruppierung der Interventionen in die drei Gruppen, insbesondere der Teilung zwischen medizinischen und pflegerischen Leistungen gerechtfertigt ist. In dem Scatterplot der Gruppe 1 und 2 wird dieser Gedanke verdeutlicht. Hier ist zu ersehen, dass die Gruppe der pflegerischen Interventionen (Gr. 1, Mitte) einen ersten Hinweis auf einen recht eindeutigen linearen Zusammenhang gibt, wenn auch im unteren CMS-Bereich eine beginnende Zerstreung auffällig ist. Wohingegen im Scatterplot der medizinischen Leistungen (Gr. 2, unten) eine Linearität zu keinem Zeitpunkt ersichtlich ist. Es lässt die Vermutung zu, dass Leistungen die im medizinischen Kontext stehen (z.B. Infusionen anhängen, ärztliche Assistenz, Medikamente verabreichen, etc.) den pflegerischen Aufwand nur latent erklären.

## 4.2 Modellformulierung

Die logistische Regression versucht nicht, anders als die lineare Einfachregression, Beobachtungsschätzungen der abhängigen Variablen vorzunehmen. Ihr besonderes Interesse gilt der Ableitung der Eintrittswahrscheinlichkeit dieser Beobachtungswerte (vgl.: Backhaus et al., 2003:422). Vor diesem Hintergrund werden folgende Determinanten zugrunde gelegt: Ausprägung  $y=1$  entspricht der Annahme, dass das Ereignis „ein erhöhter Pflegeaufwand ist zu verzeichnen“ eintritt und die Ausprägung  $y=0$  ist mit dem Ereignis „ein erhöhter Pflegeaufwand ist *nicht* zu verzeichnen“ gleichzusetzen. Dementsprechend gilt in dieser Analyse, dass  $[P(y=1)]$  als Pflegeaufwandswahrscheinlichkeit eines Patienten zu interpretieren ist. Damit die Wahrscheinlichkeit des Eintritts eines definierten Ereignisses bestimmt werden kann, so schreiben Backhaus et al. (2003:423), „wird unterstellt, dass eine nicht empirisch beobachtete latente Variable „Z“ existiert, die die (...) Ausprägungen der abhängigen Variablen (Y) in Abhängigkeit der Ausprägungen der unabhängigen Variablen  $X_j$  erzeugen kann“. Eine Verknüpfung zwischen dem Response (abhängige Variable) und den beobachteten Prädiktoren (unabhängige Variable) wird über diese besagte latente Variable Z hergestellt.

Das Resultat „erhöhter Pflegeaufwand“ wird über die Variable Z als aggregierte Einflussstärke der unabhängigen Variablen gedeutet. Ferner wird die Annahme zu Grunde gelegt, dass „die verschiedenen Einflußgrößen  $X_j$  durch eine Linearkombination die latente Variable Z erzeugen“ (Backhaus et al., 2003:423).

Damit letztendlich Aussagen zur Wahrscheinlichkeit getroffen werden können muss eine Wahrscheinlichkeitsfunktion verwendet werden. Mittels der Wahrscheinlichkeitsfunktion kann das Ereignis  $y=1$  oder  $y=0$  über die Bedingungen der aggregierten Einflussstärke Z gebildet werden. Vor diesem Hintergrund verwendet die logistische Regression eine logistische Funktion (p) die sich wie folgt darstellt:  $p = e^z / (1 + e^z)$ <sup>15</sup>

Die Wahrscheinlichkeit für das Eintreffen der Ereignisses  $y=1$  unter Einsatz der logistischen Funktion wird anhand des logistischen Regressionssatzes errechnet. „Dabei spiegeln der Parameter  $\beta_0$  und die Regressionskoeffizienten  $\beta_j$  (...) die Einflussstärke der jeweils betrachteten unabhängigen Variablen  $X_j$  auf die Höhe der Eintrittswahrscheinlichkeit  $P(y=1)$  wider“ (Backhaus et al., 2003:423). Daraus ergibt sich, dass eine Wahrscheinlichkeitsbeziehung zwischen den unabhängigen Variablen und dem Ereignis  $y=1$  über die logistische Funktion hergeleitet wird. Häufig ist sie auch als Linking-Funktion betitelt (vgl.: Backhaus et al., 2003:423).

---

<sup>15</sup> „e“ steht in der logistischen Funktion für die Eulerische Zahl (2,71828183)

Für das hier zu konzipierende Analysemodell sollen nachstehend grundlegende Annahmen definiert werden:

Es gilt die Hypothese zu untersuchen: „Je niedriger die Selbstpflegefähigkeiten (Grad der Pflegebedürftigkeit) gemessen mit den CMS-Items des ePA-AC, desto höher ist der Pflegeaufwand gemessen mit dem LEP-Instrument“. In der Graphik können die grundsätzlichen Zusammenhänge der logistischen Regression entnommen werden.

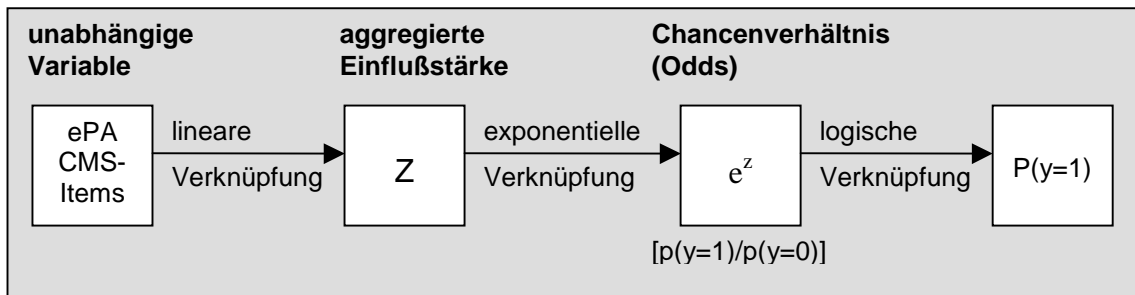


Abbildung 6: Zusammenhänge der logistischen Regression (nach Backhaus et al., 2003:426)

Geprüft werden soll die Abhängigkeitsstruktur zwischen den Variablen, d.h. wie die Modulation der abhängigen Variable sich über die Modulation von unabhängigen Variablen (Einflussgrößen) deuten lässt (vgl.: Tutz, 2000:4).

### 4.3 Schätzung der logistischen Regressionsfunktion

Im Rahmen der logistischen Funktion erfolgt gewöhnlich die Parameterschätzung über die Maximum-Likelihood-Methode. Durch die Maximum-Likelihood-Schätzung soll eine Maximierung der Wahrscheinlichkeit für das Erhalten der beobachteten Erhebungsdaten erzielt werden. Dies geschieht über die Bestimmung der Parameter, welche gleichzeitig auch für die einflussnehmenden Gewichte der unabhängigen Variable stehen. Im Zusammenhang mit empirischen Erhebungen ergeben sich die Ereignisse  $y=1$  oder  $y=0$  für jede Person. Vor diesem Hintergrund sollte die Schätzung der Parameter für die jeweils personenbezogene Betrachtung eine der zwei möglichen Wahrscheinlichkeiten,  $p(y=1)$  oder  $p(y=0)$ , ermitteln. Die statistische Maximierung der Likelihoodfunktion erfolgt in der hier verwendeten SPSS Version 15 über den Newton-Raphson-Algorithmus (vgl.: Backhaus et al., 2003:428). Der Maximum-Likelihood-Schätzung muss demnach mittels des Prozesses der Iteration errechnet werden, da er sich originär nicht genau angeben lässt (vgl.: Tutz, 2000:74). Mit Hilfe des Algorithmus werden in den einzelnen Iterationsschritten die Parameterschätzungen dahingehend verändert, dass die gewichteten Beobachtungen der unabhängigen Variable letzten Endes in dem Ergebnis der Maximierung der Wahrscheinlichkeit für das Eintreten des Ergebnisses  $y=1$  münden (vgl.: Backhaus et al., 2003:430).

Nach Durchführung der multinominal logistischen Regressionsanalyse kann aus dem Iterationsprotokoll bzw. aus dessen Ergebnis die Bestimmung der Regressionsgleichung hergeleitet werden. In dieser Abbildung 6 wird jedoch lediglich ein Teil exemplarisch dargestellt. Das Iterationsprotokoll bezieht sich auf die Gruppe mit den 20 Minuten Intervallen und auch hier nur auf die Gruppe der pflegerischen Leistungen, d.h. Leistungen die aufgrund einer Beeinträchtigung der Selbstpflegefähigkeiten (=Pflegebedürftigkeit) durchgeführt werden.

Für die durchgeführte multinominal logistische Regression ergibt sich nach 16 Iterationsschritten unter der Verwendung der Maximum Likelihood-Methode folgendes Schätzergebnis:

$$\text{Logits } (Z_k) = -113,065 \text{ (Konstante)} + 30,35 \text{ (Gruppe der LEP\_Min\_Pfl } \leq 6) + 17,4 \text{ (Gruppe der LEP\_Min\_Pfl 6,01 bis 26,00)} + 16,74 \text{ (Gruppe der LEP\_Min\_Pfl 26,01 bis 46,00)} \dots +/\text{-}N_{\text{Gruppen}}$$

**Abbildung 7: Regressionsgleichung**

Aus dieser Regressionsgleichung kann in einem weiteren Schritt der Verlauf der logistischen Funktion ermittelt werden. „Die mit Hilfe der logistischen Funktion erzeugte Wahrscheinlichkeitsverteilung (...) weist einen s-förmigen Verlauf auf“ (Backhaus et al., 2003:424).

Nach der Formulierung des Regressionsmodells und der Schätzung der logistischen Regressionsfunktion muss das Modell auf seine Güte geprüft sowie die Variablen auf ihre Trennkraft untersucht werden. Das nachstehende Kapitel 4.4 beschreibt die notwendige Vorgehensweise und verweist auf statistische Testverfahren zur Modellprüfung.

#### **4.4 Prüfung des Modells**

Das Gesamtmodell muss abschließend einer Prüfung unterzogen werden. Dabei sind zwei Fragen von besonderer Bedeutung. Zum einen: „Wie gut können die Parameterschätzungen in ihrer Gesamtheit das definierte Regressionsmodell abbilden? (Backhaus et al., 2003:437)“. Und zum anderen: „Liegen extreme Beobachtungsfälle vor, die als Ausreißer anzusehen sind und eine Eliminierung oder aber aufgrund ihres gehäuften Auftretens ggf. eine Modellveränderung erfordern?“ (ebd.:437).

Bei der Beurteilung der Güte des Modells geht es primär um die Frage der Trennschärfe. Die Beantwortung dieser Frage kann über drei Möglichkeiten bestimmt werden.

Folgende Gütekriterien bieten sich hier für an:

1. Prüfung der Güte auf der Grundlage der LogLikelihood (LL-Funktion)
  - Analyse des abweichenden Verhaltens (Devianz)
  - Likelihood Ratio-Test (Chi-Quadrat-Test)
2. Pseudo-R-Quadrat-Statistik
  - Cox und Snell
  - Nagelkerke
  - McFadden
3. Beurteilung der Ergebnisse aus der Klassifikation

Die hier erwähnten Gütekriterien werden in Kapitel 5, im Rahmen der Ergebnisdarstellung der durchgeführten Untersuchung, näher erläutert.

## 5 Ergebnisse

Die gesamte Auswertung erfolgt jeweils unter der Verwendung eines 95%-Konfidenzintervalls. Für alle im Rahmen der Untersuchung gerechneten Modelle wird im ersten Schritt lediglich das Modell mit ausschließlichem Einbezug der Haupteffekte berechnet. Zu jedem Modell werden in einem weiteren Schritt ergänzend die Wechselwirkungseffekte mit einbezogen über die Funktion „gesättigte Modelle“. Die Ergebnisdarstellung folgt der Systematik, dass zu Beginn der Auswertung exemplarisch die Gruppe der 20iger LEP-Minutenintervalle detailliert dargelegt wird. Zu jeder dieser Gruppe werden die drei LEP-Interventionsgruppen (1. Gesamtinterventionen, 2. pflegerisch induzierte- und 3. medizinisch induzierte Interventionen) untereinander abgebildet. Sich an diese Auswertung anschließend werden die Gruppen der 40iger- und 100er- LEP-Minutenintervalle in einer weniger ausführlichen dennoch inhaltlich vollständigen Form abgebildet.

Diese mannigfache Berechnung unterschiedlicher Modelle folgt der Überlegung, dass sich so eine Annäherung an ein Modell mit sehr guter Trennschärfe und akzeptabler Erfolgsquote ergeben kann.

Für die im Rahmen der Beantwortung der aufgeführten Forschungsfrage erfolgten multinominalen Regressionsanalyse mittels SPSS Version 15.0 wurden die folgenden Ergebnisse ermittelt.

Für die Gruppe mit den 20iger Minutenintervallen in der Kategorie „Gesamtinterventionen/-minuten“ sind die nachstehenden Fälle verarbeitet:

Die CMS-Items wurden im Rahmen der Analyse in sechs Gruppen eingeteilt. Über diesen Schritt sollte die Fallzahl in den einzelnen LEP-Minutengruppen erhöht werden. Jede Gruppe enthält fünf Items.

Aus dieser Tabelle kann heraus gelesen werden, dass alle Möglichkeiten der CMS-Gruppen in diesem Rechenmodell vorhanden sind. In der Gruppe der LEP-Minutenintervalle befindet sich kein Fall (bezogen auf einen Tag) unter 60 Minuten. Der höchste in dieser

CMS (Binned)					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	<= 15,00	20	4,6	4,6	4,6
	16,00 - 20,00	25	5,7	5,7	10,3
	21,00 - 25,00	45	10,3	10,3	20,6
	26,00 - 30,00	87	20,0	20,0	40,6
	31,00 - 35,00	106	24,3	24,3	64,9
	36,00+	153	35,1	35,1	100,0
	Total	436	100,0	100,0	

**Tabelle 2: Verteilung in den CMS-Gruppen**

Stichprobe aufgetretene Minutenwert pro Tag liegt bei 580 Minuten.

Die Gesamtstichprobe für diese Berechnung liegt bei N=436 Fälle (erste Zwischen-Einschätzung mit niedrigsten CMS-Wert bei einem Patienten während einer Aufenthaltsperiode).

Mittels des Histogramms wird eine linksschiefe (rechtssteile) Verteilung in der Häufigkeitsverteilung der CMS-Werte in der Stichprobe deutlich.

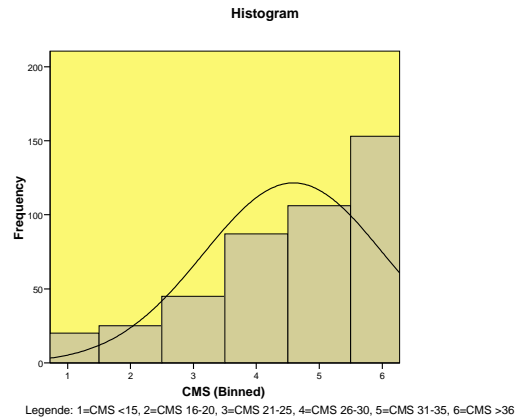


Abbildung 8: Histogramm zur Tabelle 2

LEP_Min_Ges (Binned)					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	60,01 - 80,00	1	,2	,2	,2
	80,01 - 100,00	9	2,1	2,1	2,3
	100,01 - 120,00	22	5,0	5,0	7,3
	120,01 - 140,00	29	6,7	6,7	14,0
	140,01 - 160,00	31	7,1	7,1	21,1
	160,01 - 180,00	48	11,0	11,0	32,1
	180,01 - 200,00	30	6,9	6,9	39,0
	200,01 - 220,00	40	9,2	9,2	48,2
	220,01 - 240,00	36	8,3	8,3	56,4
	240,01 - 260,00	33	7,6	7,6	64,0
	260,01 - 280,00	27	6,2	6,2	70,2
	280,01 - 300,00	23	5,3	5,3	75,5
	300,01 - 320,00	26	6,0	6,0	81,4
	320,01 - 340,00	20	4,6	4,6	86,0
	340,01 - 360,00	21	4,8	4,8	90,8
	360,01 - 380,00	14	3,2	3,2	94,0
	380,01 - 400,00	2	,5	,5	94,5
	400,01 - 420,00	5	1,1	1,1	95,6
	420,01 - 440,00	10	2,3	2,3	97,9
	440,01 - 460,00	2	,5	,5	98,4
	460,01 - 480,00	2	,5	,5	98,9
	480,01 - 500,00	2	,5	,5	99,3
	540,01 - 560,00	1	,2	,2	99,5
	560,01 - 580,00	1	,2	,2	99,8
	580,01+	1	,2	,2	100,0
Total		436	100,0	100,0	

Tabelle 3: Verteilung in den LEP-Minuten-Gruppen

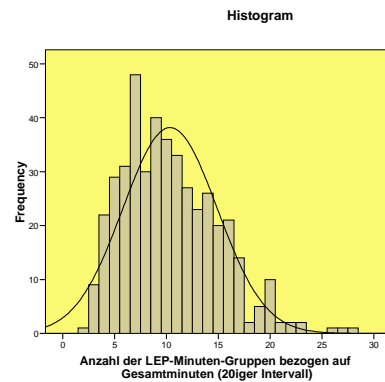


Abbildung 9: Histogramm zur Tabelle 3 (N=436)

Im Kontrast zur CMS-Verteilung zeigt sich bei den LEP-Minuten eine leicht rechtsschiefe Verteilung.

Informationen zur Güte der Modellanpassung sind der Tabelle 4 zu entnehmen. Der Chi-Quadrat-Wert von 257,654 ergibt sich aus der Differenz des Log-Likelihood-Wertes (LL-Wert) des Null-Modells (578,329) und des LL-Wertes des vollständigen Modells (320,675).

Model Fitting Information				
Model	Model Fitting Criteria	Likelihood Ratio Tests		
	-2 Log Likelihood	Chi-Square	df	Sig.
Intercept Only	578,329			
Final	320,675	257,654	24	,000

Tabelle 4: Information zur Modellanpassung

Aus dem Signifikanzniveau von 0,000 kann die Schlussfolgerung gezogen werden, dass die Nullhypothese des Likelihood Ratio Tests abgelehnt werden muss. Das

bedeutet, dass das Modell eine insgesamt gute Trennkraft für die Differenzen der Gruppen aufzeigt.

Angaben zur Güte der Anpassung werden durch die Pearson  $\chi^2$ -Statistik und der Devianz angegeben. Beide Werte ermitteln die „Abweichung vom Idealfall“ (Backhaus et al., 2003:447).

Diese Werte können aus Tabelle 5 heraus gelesen werden.

Die Pearson-Statistik<sup>16</sup> zeigt eine Signifikanz von 0,505 und veranschaulicht somit, nach Backhaus et al. (2003:462-463), ein mäßig gutes Ergebnis im Rahmen dieses Anpassungstests. Die Devianz mit 0,612 deutet auf eine mäßige Anpassung des Modells hin.

Goodness-of-Fit

	Chi-Square	df	Sig.
Pearson	95,163	96	,505
Deviance	91,458	96	,612

Tabelle 5: Güte der Anpassung

Pseudo R-Square

Cox and Snell	,446
Nagelkerke	,448
McFadden	,104

Abbildung 10: Pseudo R-Square

Mittels der Pseudo R-Quadrat-Statistik wird das  $R^2_{\text{logit}}$  bestimmt. Die Cox und Snell Statistik weist mit einem Wert von 0,446 ein gutes Ergebnis vor.

Dieser Test hat jedoch den Nachteil nur Werte  $<1$  zu erreichen. Den Wert einer perfekten Übereinstimmung von Modell und Daten kann der Nagelkerke-Test erreichen (vgl.: Backhaus et al., 2003:441). Der Nagelkerke-Test mit 0,448 zeigt ebenso ein gutes Ergebnis. Aus dem Nagelkerke-Test lassen sich, aus dem hier gerechneten Fall, 44,8% der Varianz bzgl. der Gruppenzugehörigkeit erklären. Der McFadden-Test hingegen weist einen nicht akzeptablen Wert auf, d.h. die Trennkraft der unabhängigen Variablen ist mit 0,104 in einem nicht annehmbaren Bereich.

Dieser Test hat jedoch den Nachteil nur Werte  $<1$  zu erreichen. Den Wert einer perfekten Übereinstimmung von Modell und Daten kann der Nagelkerke-Test erreichen (vgl.: Backhaus et al., 2003:441).

Likelihood Ratio Tests

Effect	Model Fitting Criteria	Likelihood Ratio Tests		
	-2 Log Likelihood of Reduced Model	Chi-Square	df	Sig.
Intercept	523,326	202,651	24	,000
CMS_Groupen_5	578,329	257,654	24	,000

The chi-square statistic is the difference in -2 log-likelihoods between the final model and a reduced model. The reduced model is formed by omitting an effect from the final model. The null hypothesis is that all parameters of that effect are 0.

Abbildung 11: Likelihood-Quotienten-Test

Schlussfolgerung zu, dass die Nullhypothese, die keinen Einfluss der CMS-Variable auf die Gruppentrennung unterstellt, abgelehnt werden muss.

Die Tabelle „Parameterschätzung“ setzt sich zusammen aus dem Regressionskoeffizienten (B), die dazu gehörigen Standardfehler, der Wald-Statistik

Mit Hilfe des Likelihood-Quotienten-Test kann die Güte des Modells auf Variablenebene beurteilt werden (vgl.: Backhaus et al., 2003:464). In dem hier vorliegendem Modell zeigt sich ein hoher Chi-Quadrat-Wert und ein Signifikanzniveau  $<5\%$ . Dies lässt die

<sup>16</sup> Pearson in der Nähe von Eins deutet auf ein hohes Maß an Abhängigkeit der Merkmale hin (vgl.: Benninghaus, 2001:116).

sowie Angaben zur Signifikanz und zur odd ratio (Exp(B)). Die außen links in der Tabelle (siehe Anhang) dargestellten Blöcke bilden die Segmente (repräsentiert 20iger LEP-Minuten) mit den jeweiligen CMS-Gruppen. Die Referenzkategorie bildet in allen hier aufgezeigten Rechenmodellen immer die höchste LEP-Minuten-Gruppe. „Aus dieser Tabelle können die Wirkungsrichtung und die Wirkungsstärke der Variablen abgelesen werden“ (Backhaus et al., 2003:465). Demnach kann dieser Tabelle entnommen werden, ob die Beobachtungen eher der Referenzgruppe angehören oder bereits richtig ihrer Gruppe zugeordnet sind (vgl.: Backhaus et al., 2003:465). Für die Gruppe der 20iger Minutenintervalle in der Gruppe der LEP-Gesamtminuten bedeutet ein positiver Wert, wie z.B. 29,154 der CMS-Gruppe 1 aus der Minutengruppe 60,01-80, dass die Beobachtung nicht der Referenzgruppe angehört, sondern richtig eingeordnet wurde. Umgekehrt weist ein negativer Regressionskoeffizient darauf hin, dass die CMS-Variable eher der Referenzgruppe angehört (vgl.: ebd:465).

Zur Überprüfung, ob die Schätzergebnisse in der Tat auf die Trennschärfe schließen lassen ist in Abhängigkeit zum Wald-Test zu beurteilen. Hierbei wird die Streuung der geschätzten Größen beurteilt (vgl.: ebd:465). Ein niedriger Wert in der Wald-Statistik und ein entsprechend hohes Signifikanzniveau bedeuten, dass die Null-Hypothese angenommen werden muss. Die Variablen haben damit keinen Einfluss auf die Trennung zwischen der Referenzgruppe und der jeweils betrachteten relevanten Gruppe. In dem aktuell betrachteten Rechenmodell zeigen sich im Wald-Test weitestgehend Werte von 0,000 und ein Signifikanzniveau nahe Eins. Die sich daraus ableitende Erkenntnis ist, dass die Variablen eine geringe Trennkraft haben. Die Wirkungsstärke und –richtung dieser Variablen kann ebenfalls der Tabelle aus der Spalte Exp (B) entnommen werden. Exp (B) bezeichnet das odd ratio bzw. synonym verwendet den Effektkoeffizienten (vgl.: ebd., 2003:466). So gibt dieser Wert an um das Wievielfache sich das Chancenverhältnis<sup>17</sup> bei Anstieg des X-Wertes um eine Einheit erhöht. Mit anderen Worten, bei einer Erhöhung des Skalenwertes um einen Faktor, verändert sich das Chancenverhältnis eben um das X-fache zum Vorteil der entsprechenden Minuten-Gruppe. Die Werte deuten auf eine Veränderung des Chancenverhältnisses in Abhängigkeit von der Variablenausprägung hin. Die Ergebnisse des Konfidenzintervalls und die sich daraus ergebene Vertrauenswahrscheinlichkeit des jeweiligen Exp (B) – Wertes können aus diesem Rechenmodell nicht ermittelt werden, da die Obergrenzen als Systemfehler angegeben werden (Erläuterung siehe Ergebnisdiskussion).

---

<sup>17</sup> Rechenweg für Chancenverhältnis:  $p(\text{Minuten-Gruppe X})/p(\text{Referenzgruppe})$

Zur Beurteilung der Gesamtgüte des Modells muss die Klassifikationstabelle beachtet werden. In dieser Kreuztabelle wird in der Zeile die beobachtete und in der Spalte die vorhergesagte bzw. geschätzte Gruppenzugehörigkeit jeder einzelnen Gruppe abgebildet. Hierbei wird der beobachtete Wert immer der Gruppe für die sich die größte Wahrscheinlichkeit ermitteln lässt zugeordnet (vgl. Backhaus et al., 2003:467).

Der Klassifikationstabelle (siehe Anhang, S.101) kann nun die Erfolgsquote für das korrekte Einordnen eines Falls in die richtige Gruppe entnommen werden. Für dieses Rechenmodell ergibt sich ein Wert von **14,4%**, dass ein Patient mit seinem entsprechenden Pflegebedürftigkeitsgrad (CMS) in die richtige Pflegeaufwandsgruppe (LEP-Minuten-Gruppe) zugeordnet wird.

Für die Gruppe mit den 20iger Minutenintervallen in der Kategorie „pflegerisch induzierte Interventionen/-minuten“ sind die nachstehenden Fälle verarbeitet:

Im Rahmen der Darstellung der Fallverteilung wird eine linksschiefe Verteilung erkennbar. Die häufigsten Fälle befinden sich in den Gruppen 3-6. Demzufolge ist die höchste Fallzahl im niedrigen Minutenwertebereich. Die Gesamtstichprobe umfasst N=436 Fälle mit einer Gesamtgruppenanzahl von 19.

Häufigkeitsverteilung der Fälle in die LEP-Minuten-Gruppen bezogen auf Pflegeminuten

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	<= 6,00	1	,2	,2	,2
	6,01 - 26,00	18	4,1	4,1	4,4
	26,01 - 46,00	66	15,1	15,1	19,5
	46,01 - 66,00	68	15,6	15,6	35,1
	66,01 - 86,00	67	15,4	15,4	50,5
	86,01 - 106,00	40	9,2	9,2	59,6
	106,01 - 126,00	39	8,9	8,9	68,6
	126,01 - 146,00	42	9,6	9,6	78,2
	146,01 - 166,00	25	5,7	5,7	83,9
	166,01 - 186,00	23	5,3	5,3	89,2
	186,01 - 206,00	16	3,7	3,7	92,9
	206,01 - 226,00	14	3,2	3,2	96,1
	226,01 - 246,00	6	1,4	1,4	97,5
	246,01 - 266,00	5	1,1	1,1	98,6
	266,01 - 286,00	2	,5	,5	99,1
	286,01 - 306,00	1	,2	,2	99,3
	326,01 - 346,00	1	,2	,2	99,5
	346,01 - 366,00	1	,2	,2	99,8
	426,01+	1	,2	,2	100,0
	Total	436	100,0	100,0	

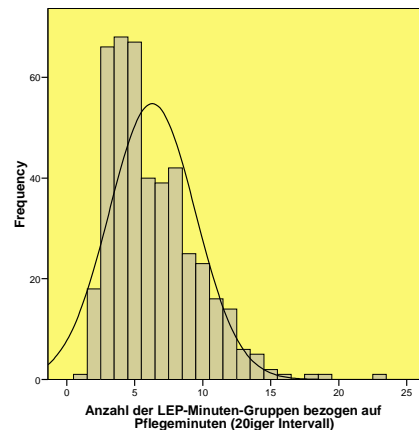


Abbildung 12: Histogramm zur Tabelle 6 (N=436)

Tabelle 6: Häufigkeitsverteilung

Model Fitting Information

Model	Model Fitting Criteria	Likelihood Ratio Tests		
	-2 Log Likelihood	Chi-Square	df	Sig.
Intercept Only	566,850			
Final	204,056	362,794	18	,000

Tabelle 7: Information zur Modellanpassung

Ein hoher Chi-Quadrat Wert und ein Signifikanzniveau <0,5% expliziert eine gute Trennschärfe zwischen den Gruppen des Modells. Die Nullhypothese muss verworfen werden.

Goodness-of-Fit

	Chi-Square	df	Sig.
Pearson	34,260	72	1,000
Deviance	35,241	72	1,000

Tabelle 8: Güte der Anpassung

Pseudo R-Square

Cox and Snell	,565
Nagelkerke	,569
McFadden	,173

Tabelle 9: Pseudo R-Quadrat

Anhand des Nagelkerke-Tests zeigt sich eine Varianz von 56,9% bzgl. der Gruppenzugehörigkeit der CMS-Variablen. Die Erklärungskraft des Modells liegt somit bei 56,9%. Die McFadden-Statistik hingegen weist mit 0,173 ein nicht genügendes Ergebnis vor. Die Trennkraft der Variablen ist demnach nicht ausreichend.

Likelihood Ratio Tests

Effect	Model Fitting Criteria	Likelihood Ratio Tests		
	-2 Log Likelihood of Reduced Model	Chi-Square	df	Sig.
Intercept	459,366	255,310	18	,000
CMS_Groupen_5	566,850	362,794	18	,000

The chi-square statistic is the difference in -2 log-likelihoods between the final model and a reduced model. The reduced model is formed by omitting an effect from the final model. The null hypothesis is that all parameters of that effect are 0.

Tabelle 10: Likelihood-Quotienten-Test

Die Regressionskoeffizienten sind mit Werten von 11,992 bis 30,353 durchweg im positiven Bereich und deuten auf eine korrekte Zuordnung in die entsprechenden LEP-Minuten-Gruppen hin. Ungeachtet dessen zeigt die Wald-Statistik niedrige Werte zwischen 0,000 und 0,003 und ein Signifikanzniveau nahe Eins. Vor diesem Hintergrund muss die Null-Hypothese akzeptiert werden. Die unabhängigen Variablen haben nach diesen Ergebnissen nur einen unbedeutenden Einfluss auf die Trennschärfe der LEP-Minuten-Gruppen.

Dieses Modell zeigt diskrepante Ergebnisse auf, da die Güte des Gesamtmodells sowie die Güte auf Variablenebene sehr gute Resultate zeigen, jedoch die Trennschärfe nicht gegeben scheint.

Der Effektkoeffizient (Exp (B)) repräsentiert durchgehend Werte im positiven Bereich. Angaben zur Vertrauenswahrscheinlichkeit können nicht gegeben werden, da das Konfidenzintervall mit seiner Obergrenze in diesem Modell als Systemfehler definiert wird. Mit einem aus der Klassifikationstabelle entnommenem Ergebnis zeigt sich eine Erfolgsquote für die korrekte Einordnung eines Falls in die richtige Gruppe von **26,8%**.

In der Statistik der Güte der Anpassung verdeutlichen die Ergebnisse eine sehr gute Güte der Modellanpassung.

Die Resultate der Pseudo R-Quadrat-Statistik veranschaulichen über den Cox- und Snell-Test, nach Backhaus et al. (2003:448), ein gutes Ergebnis mit einem Wert von 0,565.

Die Höhe des Chi-Quadrats und das entsprechend niedrige Signifikanzniveau bestätigen jedoch die Annahme einer sehr guten Modellanpassung bzw. Trennschärfe zwischen den einzelnen Gruppen.

Für die Gruppe mit den 20iger Minutenintervallen in der Kategorie „medizinisch induzierte Interventionen/-minuten“ sind die nachstehenden Fälle verarbeitet:

Häufigkeitsverteilung der Fälle in die LEP-Minuten-Gruppen bezogen auf medizinische Minuten

	N	Marginal Percentage
LEP_Min_Med (Binned)		
2,01 - 22,00	25	5,7%
22,01 - 42,00	90	20,6%
42,01 - 62,00	108	24,8%
62,01 - 82,00	79	18,1%
82,01 - 102,00	45	10,3%
102,01 - 122,00	34	7,8%
122,01 - 142,00	24	5,5%
142,01 - 162,00	22	5,0%
162,01 - 182,00	4	,9%
182,01 - 202,00	4	,9%
202,01 - 222,00	1	,2%
Valid	436	100,0%
Missing	0	
Total	436	
Subpopulation	6	

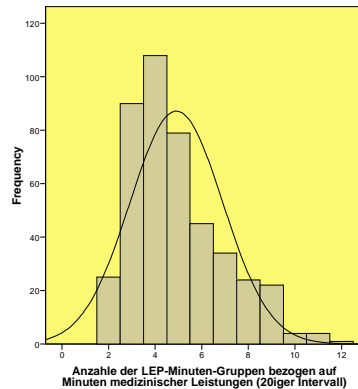


Tabelle 11: Häufigkeitsverteilung

Abbildung 13: Histogramm zur Tabelle 11 (N=436)

Bei der Betrachtung der Fallverteilung wird eine linksschiefe Verteilung ersichtlich.

Die häufigsten Fälle befinden sich in den Gruppen mit niedrigen Minutenwerten. Die Gesamtstichprobe umfasst N=436 Fälle und einer Gesamtgruppenanzahl von 11.

Im Vergleich zu den zwei voran gestellten Modellen wird hier ein eher niedriges Chi-Quadrat ersichtlich. Trotz des Signifikanzniveaus von 0,002 sollten diese Ergebnisse gedeutet werden als ein Modell mit geringer Trennschärfe.

Die Ergebnisse aus der Pearson-Statistik sowie der Devianz (Tabelle 13) weisen auf eine ungenügende Güte des Modells hin.

Diese unbefriedigenden Werte in der Modellanpassung und Güte werden durch die Pseudo R-Quadrat Statistik, welche die Güte des Gesamtmodells ausdrückt, bestätigt. Alle drei enthaltenen Tests weisen mangelnde Resultate aus. Die Schlüsse einer

Model Fitting Information

Model	Model Fitting Criteria	Likelihood Ratio Tests		
	-2 Log Likelihood	Chi-Square	df	Sig.
Intercept Only	249,777			
Final	222,643	27,135	10	,002

Tabelle 12: Information zur Modellanpassung

Goodness-of-Fit

	Chi-Square	df	Sig.
Pearson	70,258	40	,002
Deviance	76,091	40	,001

Tabelle 13: Güte der Anpassung

Pseudo R-Square

Cox and Snell	,060
Nagelkerke	,061
McFadden	,016

Tabelle 14: Pseudo R-Quadrat

Likelihood Ratio Tests

Effect	Model Fitting Criteria	Likelihood Ratio Tests		
	-2 Log Likelihood of Reduced Model	Chi-Square	df	Sig.
Intercept	238,348	15,706	10	,108
CMS_Groupen_5	249,777	27,135	10	,002

The chi-square statistic is the difference in -2 log-likelihoods between the final model and a reduced model. The reduced model is formed by omitting an effect from the final model. The null hypothesis is that all parameters of that effect are 0.

Tabelle 15: Likelihood-Quotienten-Test

schwachen Trennschärfe in diesem Modell werden desgleichen im Likelihood-Quotienten-Test widerspiegelt. Die Bestätigung der Hypothese wird legitimiert durch ein relativ niedriges Chi-Quadrat im Vergleich zu dem Chi-Quadrat der Pearson-Statistik.

Die Regressionskoeffizienten (B-Werte) befinden sich durchgehend in allen LEP-Minuten-Gruppen, bezogen auf dieses Modell, im negativen Bereich. Das signalisiert, dass die CMS-Variablen nicht korrekt eingeordnet sind und vielmehr der Referenzgruppe angehören. Folglich zeigt die Wald-Statistik Werte zwischen 0,008 bis 0,772 und ein hohes Signifikanzniveau mit 0,380 bis 0,903 in den diversen LEP-Minuten-Gruppen. Die Konsequenz aus diesen Resultaten ist die Akzeptanz der Null-Hypothese. Die CMS-Variablen haben demnach keinen Einfluss auf die Trennung zwischen der LEP-Referenzgruppe (202,01-222,00 Minuten) und der jeweils betrachteten relevanten LEP-Minuten-Gruppe.

Der Effektkoeffizient (Exp (B)) weist Werte zwischen 0,456 und 0,928 auf. Die Werte deuten auf eine Veränderung des Chancenverhältnisses in Abhängigkeit von der Variablenausprägung hin. Die Ergebnisse des Konfidenzintervalls und die sich daraus ergebende Vertrauenswahrscheinlichkeit des jeweiligen Exp (B) – Wertes liegt mit seiner Untergrenze unter Eins und mit der Obergrenze über Eins. Daraus lässt sich ableiten, dass ein tatsächliches Eintreten eines angezeigten positiven Effekts nicht mit einer 95%-Vertrauenswahrscheinlichkeit bestätigt werden kann. Ableitend muss festgehalten werden, dass in diesem Zusammenhang die CMS-Variablen eine geringe Bedeutung haben.

Die Erfolgsquote für das korrekte Einordnen eines Falls in die richtige Gruppe wird mit **24,5%** angegeben. Ungeachtet dessen ist zu vergegenwärtigen, dass die Güte des Modells sowie die Güte auf Variablenebene als ungenügend identifiziert wurde.

Für die Gruppe mit den 40iger Minutenintervallen in der Kategorie „Gesamtinterventionen/-minuten“ ergeben sich folgende Ergebnisse:

Die Information zur Modellanpassung mittels des Chi-Quadrats zeigt einen Wert von 244,416 und ein Signifikanzniveau von 0,000. Diese Ergebnisse deuten tendenziell auf eine **gute Modellanpassung** hin.

Die Güte der Anpassung ausgedrückt über die Pearson-Statistik repräsentiert einen Wert von 0,047 und die Abweichung ein Ergebnis von 0,248. Sie suggerieren hierüber eine unzureichende Güte des Modells. Die Resultate deuten eine Unabhängigkeit der Variablen an. Jedoch muss bei der Interpretation der Werte beachtet werden, dass es zu Testverfälschungen kommen kann, da dieser Anpassungstest als eine Grundbedingung die gleiche Zahl an Kovariatenmuster wie Beobachtungen einfordert (vgl.: Backhaus et al., 2003:463).

Der Cox- und Snell-Test mit einem Wert von 0,429 sowie die Nagelkerke-Statistik repräsentieren gute Ergebnisse. Hieraus ergibt sich für dieses Modell eine Erklärungskraft von 43,5%. Ungeachtet dessen weisen die Ergebnisse des

McFadden-Tests weniger gute Werte auf. Mit 0,129 befindet sich die Trennkraft der unabhängigen Variablen in einem nicht befriedigenden Bereich.

Über den Likelihood-Quotienten-Test und einem dazugehörigem Signifikanzniveau von 0,000 bestätigen sich die zuvor ermittelten guten Ergebnissen für die Trennschärfe des Modells. Die Wald-Statistik repräsentiert Werte zwischen 436,912 und 967,330 mit einem in alle CMS-Gruppen identischen Signifikanzniveau von 0,000. Abschließend lässt sich für dieses Regressionsmodell ein Klassifikationsergebnis von **24,8%** ermitteln.

Für die Gruppe mit den 40iger Minutenintervallen in der Kategorie „*pflegerisch induzierte Interventionen/-minuten*“ ergeben sich folgende Ergebnisse:

Bei der Überprüfung der Modellanpassung in diesem Regressionsmodell zeigt sich ein Chi-Quadrat Wert von 338,716 sowie ein Signifikanzniveau von 0,000. Die Resultate explizieren eine **sehr gute Modellanpassung**.

Gleichermaßen weisen die Pearson-Statistik mit 0,981 und die Devianz von 0,980 sehr gute Ergebnisse aus. Beide Werte befinden sich in der Nähe der Eins und deuten demnach auf ein hohes Maß an Abhängigkeit der Merkmale hin.

Desgleichen stellt das Testergebnis der Cox- und Snell-Statistik mit einem Wert von 0,540 ein gutes Resultat dar. Nach dem Nagelkerke-Test kennzeichnet das Modell eine Erklärungskraft von 55,6% und zeigt somit ein gutes Ergebnis. Auch die Statistik des McFadden-Tests stellt mit 0,220 ein gutes Ergebnis dar. Hierüber wird die gute Trennkraft der CMS<sup>®</sup>-Variablen bestätigt.

Der Likelihood-Quotienten-Test kann die Güte des Modells auf Variablenebene beurteilen (vgl.: Backhaus et al., 2003:464). In dem hier vorliegendem Modell zeigt sich ein Chi-Quadrat-Wert von 338,716 und ein Signifikanzniveau <5%. Das bedeutet, dass das restrigierte Modell nicht dem unrestrigierten Modell vorzuziehen ist, da das unrestrigierte im Gegensatz zum restrigierten Modell tatsächlich signifikant besser ist (vgl.: [http://www.lrz-muenchen.de/~wlm/ilm\\_I7.htm](http://www.lrz-muenchen.de/~wlm/ilm_I7.htm), Stand: 18.04.2007). Die Nullhypothese, die keinen Einfluss der CMS<sup>®</sup>-Variablen auf die Gruppentrennung unterstellt, muss folglich abgelehnt werden.

Weniger gute Resultate werden jedoch über die Wald-Statistik deutlich. Hier liegen die Werte bei 0,000 bis 0,002 und das Signifikanzniveau weist Werte nahe Eins aus. Doch folgt man den Anmerkungen von Hauck und Donner (1977:851-853), welche schlechtere Testeigenschaften für die Wald-Statistik prognostizieren im Vergleich zum Likelihood-Quotienten-Test, relativieren sich diese Ergebnisse. Hauck und Donner weisen darauf hin den Likelihood-Quotienten-Test zu bevorzugen.

Für dieses Regressionsmodell lässt sich eine Erfolgsquote für die korrekte Einordnung in die Gruppen aus der Klassifikationstabelle von **36,9%** ermitteln.

Für die Gruppe mit den 40iger Minutenintervallen in der Kategorie „medizinisch induzierte Interventionen/-minuten“ ergeben sich folgende Ergebnisse:

Zur Erhebung der Modellanpassung kann aus der entsprechenden Tabelle ein Chi-Quadrat-Wert von 23,717 und ein Signifikanzniveau von 0,000 ersehen werden. Dies deutet auf eine nicht **befriedigende Anpassung** des Modells hin.

Die Güte der Anpassung zeigt äußerst schlechte Ergebnisse. Die Pearson-Statistik sowie die Devianz sind Null, d.h. eine Unabhängigkeit der Variablen kommt zum Ausdruck. Ein weiteres Indiz zur Festigung der Anerkennung der Nullhypothese deutet sich über das niedrigere Chi-Quadrat aus dem Likelihood-Quotienten-Test im Vergleich zur Chi-Quadrat Tabelle mit einem Wert von 55,475 ( $X^2$ -Statistik) an (siehe auch Backhaus et al., 2003: 440).

Die Pseudo R-Quadrat-Statistik ermittelt für den Cox- und Snell-Test ein inakzeptables Ergebnis von 0,053. Der Nagelkerke-Test errechnet einen Wert von 0,057 und damit eine Erklärungskraft von 5,7% für dieses Modell.

Der McFadden-Test zeigt ebenso schlechte Ergebnisse von 0,020 und repräsentiert darüber eine mangelhafte Trennkraft der CMS<sup>®</sup>-Variablen.

Aus der Tabelle der Parameterschätzung kann für alle LEP-Minutengruppen ein negativer Wert abgelesen werden. Diese Resultate deuten darauf hin, dass die Einordnung nicht korrekt erfolgte und die Beobachtungen eher der Referenzgruppe angehören. Ebenso zeigt der Wald-Test unbefriedigende Ergebnisse. Die entsprechende Irrtumswahrscheinlichkeit liegt zwischen 48,8% und 88,6% gleichsam in einem nicht akzeptablen Bereich.

Für dieses Regressionsmodell lässt sich eine Erfolgsquote für die korrekte Einordnung in die richtigen Gruppen aus der Klassifikationstabelle von **42,0%** ablesen.

Für die Gruppe mit den 100er Minutenintervallen in der Kategorie „Gesamtinterventionen/-minuten“ ergeben sich folgende Ergebnisse:

In diesem Regressionsmodell zeigen die Ergebnisse einen Chi-Quadrat-Wert von 214,672 auf, mit fünf Freiheitsgraden und einer Irrtumswahrscheinlichkeit von 0,0%. Dieses Chi-Quadrat aus dem Likelihood-Ratio-Test ist bedeutend höher als aus der  $X^2$ -Statistik mit 26,728. Dem entsprechend gibt es hierüber erste Hinweise auf eine **gute Modellanpassung**. Über die Auswertung der Güte der Anpassung wird deutlich, dass die Werte in einem ausreichendem Bereich liegen (siehe Backhaus et al., 2003:448). Die Cox- und Snell-Statistik zeigt mit 0,389 einen akzeptablen Wert.

Die Erklärungskraft des Modells liegt nach dem Nagelkerke-Test bei 41,9%. Der McFadden-Test mit einem Resultat von 0,187 ist nur knapp in einem nicht annehmbaren Bereich, da erst ab einem Wert von 0,2 von passablen Ergebnissen gesprochen werden kann (vgl.: ebd:447). Daraus lässt sich ableiten, dass die Trennkraft der unabhängigen Variablen nicht quantifiziert werden sollte.

Der Regressionskoeffizient (B), entnommen aus der Parameter-Schätzungstabelle, liegt in allen LEP-Minutengruppen im positiven Bereich. Dies weist auf eine Tendenz der korrekten Einordnung der Fälle in die richtige Gruppe hin. Die Wald-Statistik zeigt Werte zwischen 0,000 und 22,725 mit einem Signifikanzniveau von 0,000 bis 1,000. In der Spalte der Effektkoeffizienten (Exp (B)) liegen die Werte durchweg im positiven Bereich. Das Konfidenzintervall schließt Werte der Untergrenze unter Eins mit ein. Dies kann als Hinweis auf eine geringe Bedeutsamkeit der Variablen gedeutet werden.

Die Erfolgsquote für das korrekte Einordnen eines Falls in die richtige Gruppe wird mit **49,5%** angegeben.

Für die Gruppe mit den 100er Minutenintervallen in der Kategorie „*pflegerisch induzierte Interventionen/-minuten*“ ergeben sich folgende Ergebnisse:

Das Chi-Quadrat aus dem Likelihood-Quotienten-Test wird mit 252,448 und einem Signifikanzniveau von 0,000 bei entsprechenden fünf Freiheitsgraden angegeben. Dieser Wert ist erkennbar höher als das Chi-Quadrat der  $X^2$ -Statistik und deutet so auf eine **gute Modellanpassung** hin. Die Ergebnisse der Pseudo R-Quadrat-Statistik bestätigen die Tendenz einer guten Güte der Anpassung des Modells. Die Erklärungskraft des Modells liegt bei 52,6%. Die Trennkraft der unabhängigen Variablen nimmt in diesem Regressionsmodell den bisher höchsten Wert mit 0,320 an und verweist auf eine starke Trennkraft.

Die positiven B-Werte der Parameter Schätzung weisen darauf hin, dass mit einem hohen Beurteilungswert eine Beobachtung eher der eingeordneten Gruppe und nicht der Referenzgruppe angehört. Die Wald-Statistik zeigt Werte zwischen 0,000 und 0,001 mit einem Signifikanzniveau in der Nähe von 1,000. Aussagen zum Konfidenzintervall des Effektkoeffizienten können nicht getroffen werden, da keine aussagekräftigen Informationen identifiziert werden konnten.

Die Erfolgsquote für das korrekte Einordnen eines Falls in die richtige Gruppe wird mit **73,9%** angegeben.

Für die Gruppe mit den 100er Minutenintervallen in der Kategorie „medizinisch induzierte Interventionen/-minuten“ ergeben sich folgende Ergebnisse:

Die Ergebnisse des Likelihood-Ratio-Tests mit einem Chi-Quadrat von 20,491 und einer Signifikanz von 0,000 verweisen, verglichen mit dem Wert der Chi-Quadrat-Tabelle, auf ein **nicht signifikantes Modell** bezogen auf diesen Datensatz.

Die Ergebnisse der Pseudo R-Quadrat-Statistik festigen die Tendenz einer mangelnden Güte der Anpassung des Modells. Die Erklärungskraft des Modells liegt nach dem Nagelkerke-Test bei 7,1%. Der McFadden-Wert mit 0,045 verdeutlicht eine schlechte Trennkraft der CMS<sup>®</sup>-Variablen.

Dieser Eindruck wird mittels des Konfidenzintervalls der Effektkoeffizienten bestätigt. Die Werte liegen mit der Untergrenze bei Werten unter Eins und bei der Obergrenze deutlich über Eins. Nach Backhaus et al. (2003:466) sind dies Erkenntnisse die darauf hinweisen, dass eine geringe Bedeutung der Variablen vorliegt.

Die Trefferquote für das korrekte Einordnen eines Falls in die richtige Gruppe wird mit **79,6%** angegeben. Jedoch darf diese Erfolgsquote nicht isoliert betrachtet werden, da sie in Abhängigkeit zur Modellanpassung und Güte zu sehen ist.

## 6 Interpretation und Diskussion der Ergebnisse

Für die Regressionsmodelle aus der Gruppe mit den 40iger und 100er Minutenintervallen in der Kategorie „pflegerisch induzierte Interventionen/-minuten“ konnte in der vorliegenden Untersuchung ein Zusammenhang zwischen dem Patientenzustand und dem Pflegeaufwand nachgewiesen werden. In beiden Modellen zeigt sich eine gute bis sehr gute Modellanpassung und Güte. Des Weiteren kann eine klare Trennkraft zwischen den LEP-Minutengruppen ermittelt werden. In beiden Modellen zeigt sich ergänzend, anhand der Pearson-Statistik, ein sehr gutes Ergebnis bezogen auf den Variablenzusammenhang. Ebenso konnte eine gute Trennkraft der unabhängigen CMS<sup>®</sup>-Variablen mittels des McFadden-Tests in diesen besagten Modellen aufgezeigt werden.

Die Gruppe aus der Kategorie „pflegerisch induzierte Intervention/-minuten“ mit einem 40iger Minutenintervall wird trotz guter Modellanpassung und Trennkraft für weniger geeignet, im Vergleich zum Modell selbiger Einordnung mit 100er Minutenintervall, angesehen. Dies begründet sich in der Betrachtung der Erfolgsquote für eine korrekte Einordnung eines Falls in die richtige LEP-Minutengruppe. Mit einer 26,8%-igen Wahrscheinlichkeit der Erfüllung der Zielintention liegt dieses Modell vor dem Hintergrund der von Backhaus et al. (2003:448) geäußerten Kritik an den Klassifikationsergebnissen, unter einem akzeptablen Wert. Die Beanstandung richtet sich an eine „überhöhte Trefferquote“ (ebd:448) bei der Analyse der Klassifikationsergebnisse. Im Vergleich hierzu wurde für das Modell der 100er Intervalle der pflegerisch induzierten Interventionen/-minutenwerte eine Trefferquote von 73,9% identifiziert und kann vor diesem Hintergrund, unter Abwägung aller Einschränkungen dieses Untersuchungsdesigns, für ein geeignetes Modell zur Erklärung eines erhöhten Pflegeaufwands gedeutet werden. Jedoch ist dieses Modell zur Kalkulationsgrundlage ungeeignet, da es mit einem Intervall von 100 Minuten zu grob ist.

In den Resultaten der hier gerechneten Regressionsmodelle spiegeln sich Ergebnisse, die im Rahmen einer zuvor durchgeführten Literatursichtung<sup>18</sup> von Zusammenhangsstudien gewonnen werden konnten, wider.

Die verwendeten Instrumente zur Bildung der abhängigen und unabhängigen Variablen in den Zusammenhangsstudien weichen von den hier eingesetzten ab. Jedoch ist die Grundidee und letzten Endes das Ergebnis des Nachweises eines Zusammenhangs identisch. Die Studien, beispielhaft seien hier Fischer (2001), Baumberger (2001) und Molgard (2000) genannt, zeigen eine erhöhte Erklärungs-

---

<sup>18</sup> Siehe systematische Suchstrategie im Anhang Seite 168-169

kraft des verbrauchten Pflegeaufwands unter Verwendung von Daten mit Pflegebedürftigkeitsinformationen. Über diese gewonnenen Resultate, dass ein Zusammenhang zwischen dem Zustand eines Patienten und der verbrauchten Pflegezeit begründet existiert, können Ausführungen, wie zum Beispiel von Welton und Halloran et al. (2006:367), argumentativ vertreten werden. Welton und Halloran et al. veranschaulichen die Bestärkung der Nutzung elementarer Pflegedaten, wie zum Beispiel Aussagen zur pflegerischen Intensität, als Grundlage einer Pflegekostenkalkulation zu verwenden.

Auffällig bei der Betrachtung, der in der vorliegenden Arbeit erzielten Ergebnisse, ist der Erfolgsanteil von 79,6% im Modell der Kategorie „medizinisch induzierte Interventionen/-minuten“ mit einer 100er Minutenzeitspanne. Dieser Wert darf nicht für sich alleine stehen, sondern muss zwingend in Bezug zur Modellgüte betrachtet werden. Denn hier zeigt sich ganz deutlich eine nicht gegebene Signifikanz des Modells. In diesem Fall kann nicht von einem Zusammenhang zwischen der abhängigen und unabhängigen Variable gesprochen werden. Dieses Bild der ungenügenden bzw. schlechten Modellanpassung und Variablen trennkraft konnte in der vorliegenden Arbeit in allen Modellen aus der Kategorie „medizinisch induzierte Interventionen/-minuten“ aufgezeigt werden. Das generierte Ergebnis bestätigt eine im Vorfeld entwickelte mögliche Tendenz eines nicht gegebenen Zusammenhangs zwischen den hier untersuchten Variablen (siehe S. 37 „Scatterplot Gruppe 2“ und Anhang S. 94-96). Ein anderer Aspekt zur Begründung des hohen Klassifikationsergebnisses ist die Tatsache, dass in diesem Modell nur lediglich zwei LEP-Minutengruppen gebildet werden konnten. Somit stellt sich die Frage, ob die Wahrscheinlichkeit der korrekten Gruppenzuordnung tatsächlich höher ist als der Zufall. Im Rahmen der Interpretation der Ergebnisse, gerade in den Modellen der 100er Minutenintervalle, sollte zwingend berücksichtigt werden, dass mit Zunahme der Größe der Intervalle in den Minutenbereichen sich die Gruppenanzahl verringert und somit steigt die Wahrscheinlichkeit für eine korrekte Zuordnung der Fälle in die jeweiligen LEP-Gruppen. Das bedeutet, dass die Fehleinordnung sinkt bei geringer Gruppenanzahl. Daher muss die steigende prozentuale Wahrscheinlichkeit der Klassifikationsergebnisse in den hier gerechneten Modellen kritisch betrachtet werden.

Alle drei Modelle der Kategorie „Gesamtinterventionen/-minuten“ zeigen in der Ergebnisdarstellung eine passable Modellgüte. Einschränkungen dieser Modellentwürfe liegen zum Teil in der geringen Trennschärfe der Variablen und in den ausnahmslos sich unter 50% befindenden Klassifikationsergebnissen. Vor dem Hintergrund dieser ermittelten Resultate kann geschlussfolgert werden, dass die

Entscheidung für eine Trennung der LEP-Variablen in primär medizinisch ausgelöste Interventionen und in Interventionen, die durch eine Pflegebedürftigkeit bzw. Selbstpflegefähigkeitseinschränkung erzeugt werden, sinnvoll war. Dennoch sollte die vorgenommene Kategorisierung der LEP-Variablen überdacht werden, da Themen wie bspw. Beratungsgespräche des pflegerischen Personals sich weder aus dem ePA-AC<sup>®</sup> herleiten lassen noch nach ICF klassifiziert werden konnten. Somit konnten zum Teil sehr wichtige Aufgaben der Pflege nicht in der Kategorie „pflegerisch induzierte Intervention/-minuten“ berücksichtigt werden. Doch vor dem Hintergrund der adäquaten Abbildung der pflegerischen Kosten im DRG-System und unter der Überlegung, dass Beratungsparts in der Pflege möglicherweise in den Bereich der „Job-Anreicherung“ zu integrieren sind, muss zwingend geprüft werden, wie dieses Aufgabengebiet verarbeitet bzw. platziert werden kann.

Weiterhin muss in dieser Untersuchung die Einteilung des möglichen CMS<sup>®</sup>-Summenscores (10=schwer pflegebedürftiger Patient und 40=keine Einschränkungen in einer der ePA-AC<sup>®</sup> Kategorien) kritisch reflektiert werden, da die Gruppeneinteilung in CMS<sup>®</sup>-5er-Schritten keiner inhaltlichen Systematik folgte. Möglicherweise könnten die Ergebnisse bzw. die Regressionsmodelle optimiert werden, wenn die Gruppeneinteilung selbstpflegefähigkeitseingeschränkter Patienten aufgrund einer Merkmalshomogenität erfolgen würde. Somit könnte eine stärkere Einheitlichkeit in diesen Gruppen erzeugt werden. Diese Gruppeneinteilung bzw. die Identifikation homogener Merkmalsausprägungen bei diversen Patienten im klinischen Bereich lässt sich über eine hierarchische Clusteranalyse arrangieren. Bei bereits vorbestehenden Informationen zur denkbaren Anzahl der letztendlich verwendeten Gruppen empfiehlt sich eine Clusterzentrenanalyse. Diese Schritte gingen auch Isfort et al. (2004:220-224) bei dem Versuch der Konstruktion von pflegerelevanten Fallgruppen.

Das Nicht-Umsetzen der hier beschriebenen Vorgehensweise ist zum Teil des geringen Bearbeitungszeitraums dieser Arbeit geschuldet. Zeichnet sich dennoch dessen ungeachtet als unerlässliche Voraussetzung ab. Daher muss die hier gewählte Methode der CMS<sup>®</sup>-Gruppeneinteilung als ungenügend betrachtet werden. Generell sollte die Teilung der CMS<sup>®</sup>-Gruppen diskutiert werden. Im Verlauf der Ergebnisauswertung wurde der Versuch unternommen zu prüfen wie sich das Nicht-Verwenden einer Gruppeneinteilung auf die Regressionsergebnisse auswirkt. Hierfür wurde exemplarisch die Gruppe der 40-iger Minutenintervalle aller drei Gruppen heraus genommen. Hierbei zeigte sich, dass die Ergebnisse sich positiv veränderten. So konnten über diesen Weg bessere Ergebnisse erzielt werden. Bei dieser Berechnung sei jedoch angemerkt, dass in den Rechenmodellen eine Aus-

reißerbereinigung vorgenommen wurde. Die Ausreißer<sup>19</sup> wurden im Verlauf der Ergebnisanalyse in den zuvor konzipierten neun Modellen heraus gerechnet. Die Festsetzung der Grenzbereiche, d.h. ab wann ein Fall mit welchem Minutenwert als Ausreißer definiert wird, erfolgte gemäß einer von Fischer (2002:51) vorgeschlagenen Methode. Für jedes Modell wurde als Rechengrundlage der Mittelwert addiert bzw. subtrahiert mit 2x der Standardabweichung (Mittelwert + m x Standardabweichung bzw. Mittelwert – m x Standardabweichung). Zur Überprüfung der Einflussnahme der Ausreißerbereinigung auf die Ergebnisse im Rahmen der Regressionsanalyse wurde auch hier das Modell der 40-iger Minutenintervalle gewählt. Bei der Modellrechnung (siehe Anhang, S. 139ff.) mit Ausreißerbereinigung konnten sich alle drei Modelle in ihren Resultaten verbessern. Besonders auffällig bei der Betrachtung dieser Ergebnisse sind die deutlich verbesserten Resultate in der Wald-Statistik. Hierbei zeigte sich insbesondere für das Modell der pflegerisch induzierten Minutengruppe ein positiver Effekt. Es gelang über diesen Weg signifikante Ergebnisse in der Wald-Statistik zu generieren. Auch für das Konfidenzintervall des Effektkoeffizienten zeigten sich Verbesserungen. Die Spanne zwischen Ober- und Untergrenze ist immer noch sehr groß, jedoch konnten die zuvor erzielten Systemfehler in der Obergrenze bereinigt werden.

Eine unreflektierte und vollkommene Eliminierung der Extremfälle sollte dennoch kritisch betrachtet werden, da es selbstverständlich auch solche Fälle gibt bei denen ein hoher Minutenwert inhaltlich gerechtfertigt ist. In einem nächsten Schritt müsste vor diesem Hintergrund der Frage nach der Unterscheidung von Extremfällen zu den anderen Fällen nachgegangen werden. Ebenso ist die reine Verwendung des CMS<sup>®</sup>-Summenscores, statt einzelner Items, kritisch zu reflektieren.

Die Entscheidung der ausschließlichen Verwendung der CMS<sup>®</sup>-Items zur Bildung der unabhängigen Variable sollte ebenfalls durchdacht werden. Die Reduktion der ePA<sup>®</sup>-Items auf den Bereich der Selbstpflegefähigkeitseinschränkungen bedingt möglicherweise einen negativen Einfluss auf die Klassifikationsergebnisse der Modelle. Eine Reduktion der Items führt auch immer zu einem Verlust an Patientinformationen. In dieser Untersuchung kann somit nicht ausgeschlossen werden, dass dieser Informationsverlust sich wirkungslos auf die Wahrscheinlichkeit der korrekten Klassifizierung anhand der unabhängigen Variable verhält. Im Raum steht somit die Selbstkritik der Reduktion der eigenen Wirklichkeit von Pflegebedürftigkeit. Der verfolgte Gedanke ist die Gegebenheit, dass zum Beispiel Ableitungssysteme bzw. das Angewiesensein auf Hilfestellung beim Umgang mit Urinableitungssystemen eine Einschränkung in den Selbstpflegefähigkeiten eines Patienten

---

<sup>19</sup> Unter Ausreißer werden in dieser Arbeit die Fälle mit Extremwerte in den LEP-Minuten-Gruppen verstanden.

bewirken könnte und demnach auch Einfluss auf dessen Partizipation am gesellschaftlichen Leben hat. Die Verminderung der gesellschaftlichen Teilhabe muss nicht zwingend im Zusammenhang mit dem Verbrauch des Pflegeaufwands im akut-klinischen Setting stehen, dennoch muss bei der Deutung der vorliegenden Ergebnisse dieser Umstand berücksichtigt werden.

Ein anderer Aspekt der nicht ungeachtet bleiben soll, ist die Vorgehensweise bei der Bildung der Minutenintervalle der LEP-Variable. Allein die Tatsache der Minderung des Skalenniveaus in der Konzeption der abhängigen Variable bedingte den Verlust von Informationen. Ebenso verringert sich die Vielfalt der Datenauswertungsmöglichkeiten, wenn intervallskalierte Daten in nicht geordnete kategoriale Daten transformiert werden. Diese Vorgehensweise begründet sich aber in Anlehnung an die von Isfort et al. (2004:119) vorgeschlagenen Bedingungen für eine Fallgruppenkonstruktion. Isfort et al. sehen zur Abgrenzung der Gruppen mit identischem Pflegezeitverbrauch die Verwendung von Zeitkorridoren als vollkommen ausreichend. Zu überlegen ist jedoch hierbei die Setzung der Intervalle. In der vorliegenden Arbeit erfolgte der Minutenmodus in immer wiederkehrenden fixen Zeitspannen. Für zukünftige Modelle sollte überdacht werden ob eine Setzung der Minutenintervalle in Abhängigkeit zur Häufigkeitsverteilung der erfolgten Leistung für inhaltlich sinnvoller erachtet werden muss.

Ein Einschnitt, der in der Diskussion berücksichtigt werden muss, stellt die Gegebenheit dar, dass in dieser Studie lediglich der pflegerische Zeitaufwand (Minutenwerte nicht Leistungen) betrachtet werden. Schlussfolgerungen für jegliche Art der durchgeführten Interventionen im Sinne von detaillierten Leistungsbeschreibungen erfolgen nicht. Dieser Aspekt, der Bildung von homogenen Leistungsgruppen im Sinne von Leistungspaketen, sollte dringend in einem der nächsten Schritte erfolgen. Isfort et al. (2004:119) sehen die Notwendigkeit der Bildung von „zeitlich wie pflegerisch relativ homogenen Gruppen“, um das Wesen pflegerischer Fallgruppen zu beachten. Diesen Bedingungen kann in einer nachfolgenden Studie, auf der Grundlage des Datenmaterials der erfolgten Zusammenhangsstudie dieser Arbeit, Rechnung getragen werden.

Ein weiteres Moment der kritischen Auseinandersetzung mit den Ergebnissen stellt die Betrachtung der sich durch nahezu alle gerechneten Modelle (ohne Ausreißerbereinigung) ziehenden schlechten Resultate der Wald-Statistik aus der Tabelle der Parameterschätzung dar. Diese weniger guten Ergebnisse des statistischen Tests implizieren eine ungenügende Modellanpassung. Die Modellgüte wird ebenso durch den Likelihood-Ratio-Test überprüft. Hierbei wurde auffällig, dass die Modellkonstruktionen, die für eine logistische Regression gerechnet wurden,

ungenügende Resultate in der Wald-Statistik aufzeigten, im Vergleich dazu aber gute bis sehr gute Ergebnisse in den übrigen Auswertungen repräsentierten. Hauck und Donner (1977) bestätigen für den Wald-Test eine schnellere Berechnung, doch bescheinigen sie ihm schlechtere Testeigenschaften im Vergleich zum Likelihood-Quotienten-Test. Hauck und Donner vertreten den Ansatz den Likelihood-Ratio-Test der Wald-Statistik vorzuziehen.

Die Datenauswertung wird durch die Gegebenheit der „leeren Zellen bei der Gegenüberstellung der abhängigen Variablenstufen und den im Datensatz vorhandenen Kovariatenmustern“ (Backhaus et al., 2003:459) erschwert. Ein Kovariatenmuster bildet eine Zelle, welche durch eine Antwortkombination und eine bestimmte Einteilung der abhängigen Variable definiert ist. In SPSS werden leere Zellen über eine Warnmeldung kenntlich getan. Dieses Anzeichen verweist auf die Nicht-Erfüllung einer Grundvoraussetzung der logistischen Regression. Die mögliche Folge dieses Umstands kann eine Minderung der Ergebnisexaktheit sein (vgl.: ebd:459). In der im Rahmen der vorliegenden Arbeit durchgeführten Untersuchung konnte in keinem der gerechneten Regressionsmodelle diese Anforderung erfüllt werden. Vor diesem Hintergrund lässt sich das Problem fehlender Werte, insbesondere bei der Darstellung des Konfidenzintervalls für den Effektkoeffizienten (Exp (B)) erklären.

Die nach Backhaus et al. (2003:470) empfohlene Fallzahl von mindestens 25 pro Gruppe (bezogen auf die abhängige Variable) konnte nicht durchgehend Rechnung getragen werden. Die zum Teil zu geringe Stichprobenanzahl innerhalb der einzelnen Gruppen kann die generierten Ergebnisse in ihrer Aussagekraft gemindert haben. Dies sollte bei der Betrachtung der Resultate berücksichtigt werden.

Vor all diesen Ausführungen muss nachhaltig betont werden, dass die ermittelten Ergebnisse lediglich als Hinweise auf einen Zusammenhang gedeutet werden dürfen. Die sich in dieser Untersuchung gezeigten Anhaltspunkte decken sich mit bereits im Vorfeld durchgeführten Zusammenhangsstudien von Hunstein et al. (2007b). Hierbei konnte ebenfalls auf der Grundlage einer Regressionsanalyse ein deutlicher Zusammenhang zwischen dem Grad der Pflegebedürftigkeit und der erbrachten Pflegezeit für Leistungen, die in der Beziehung zu Fähigkeitsbeeinträchtigungen eines Patienten stehen, ermittelt werden. Gleichmaßen konnten in dieser Studie kongruente Ergebnisse zu der hier durchgeführten Untersuchung eruiert werden. Dies betrifft die Resultate für die Gruppe der medizinisch induzierten Interventionen und erbrachten Pflegeminuten. Hierbei wird ein nicht gegebener Zusammenhang zwischen eben diesen Pflegezeiten und des Zustands eines Patienten für diese Gruppe offensichtlich.

Im Rahmen der kritischen Auseinandersetzung mit der Untersuchungsanordnung sollte auch das Design reflektiert werden. Aufgrund des vorhandenen Datenmaterials musste sich für ein Querschnittsdesign entschieden werden, da die Daten nicht für jeden Tag in einer Aufenthaltsperiode eines Patienten auch eine tägliche ePA-AC<sup>®</sup> - Einschätzung vorwies. Folglich konnten periodische Betrachtungen ohne Interpolation der Daten nicht erfolgen. Eine Interpolierung war jedoch aufgrund der Datenbeschaffenheit nicht zulässig. Ebenso wenig konnte die Erhebungsqualität der Einschätzung sichergestellt werden, d.h. wurde tatsächlich immer dann ein Patientenzustand erhoben, wenn sich ein Wert im Verlauf verändert hat. Ein weiterer Aspekt der für die Wahl eines Querschnittsdesigns sprach, sind die sich zum Teil schnell wandelnden Krankheits- bzw. Zustandsveränderungen im akut-klinischen Bereich.

Das Querschnittsdesign nimmt eine der unteren Positionen in der Evidence-Hierarchiepyramide ein. Daher wird mit der Entscheidung, ein Querschnittsdesign zu wählen, implizit eine Reduktion der systematischen Kontrolle der Untersuchungsbedingungen in Kauf genommen.

Ein anderer Aspekt, der der Diskussion bedarf, ist die Tatsache der Verwendung schweizerischer Daten für diese Untersuchung. Die Frage der Übertragbarkeit der gewonnenen Ergebnisse auf deutsche Versorgungssituationen, vor dem Hintergrund, dass die Daten im Schweizer System erhoben wurden, tritt zu Tage. Die Gegebenheit differierender Zeitkontingente bei der Erfüllung pflegerischer Aufgaben bedingt durch einen unterschiedlichen Personalschlüssel zwischen dem deutschen und schweizerischen Setting wurde für diese Untersuchung vernachlässigt. Ungeachtet dessen steht prinzipiell die Einschränkung der Übertragbarkeit bzw. der Verallgemeinerbarkeit der Ergebnisse zur Diskussion aus. Mit einem anderen Datensatz und einem anderen Patientenkontext könnten diese Ergebnisse andere Klassifikationsresultate generieren. Die Untersuchungsergebnisse können somit lediglich als eine Tendenz auf dem Weg zur Bildung von pflegehomogenen Gruppen angesehen werden. Doch dieser Tatbestand negiert nicht die eruierten Ergebnisse, da dies erste notwendige Konstruktionsschritte darstellen und hier sich eine Kongruenz in der Vorgehensweise zur Entwicklung der Fallgruppen innerhalb der DRGs abzeichnet.

Wie in naher und ferner Zukunft Indikatoren zur Pflegekostenkalkulation ins DRG-System eingeschleust werden können, soll im nachfolgenden Kapitel skizziert werden. Des Weiteren wird auf bereits im internationalen Raum umgesetzte Modelle Bezug genommen.

## 7 Konklusion und die Relevanz für die Pflege

Neben dem Versuchen in Deutschland pflegerelevante Pflegefallgruppen zu bilden, existieren bereits in anderen Ländern eine Reihe von Ansätzen zur Bildung von Fallgruppen. Hierbei stehen nicht immer separate und isolierte Pflegefallgruppen im Vordergrund. Fischer (2002:102) merkt jedoch an, wenn es der Pflege gelänge zu belegen, „(...) das gewisse Patientenprobleme, welche die Pflege formuliert, relativ unabhängig von den ärztlichen Diagnosen sind (...)“ (Fischer, 2002:102), daher sollte man überdenken, ob das Herauslösen gesonderter Pflegefallgruppen im Gegensatz zu integrierten Lösungen nicht vorzuziehen ist.

Mehrere Möglichkeiten der Gestaltung von Klassifikationsmodellen werden in der Literatur erwähnt:

1. eine Pflegekostengruppe und eine DRG je Fall
2. mehrere Pflegekostengruppen und mehrere DRGs je Fall
3. mehrere Pflegekostengruppen tagesbezogen und eine DRG je Fall

In diesen durchdachten Modellen erfolgt die Gewichtung der Kosten einerseits über eine additive Methode im Sinne von Zu- und Abschlägen und andererseits über Faktoren durch multiplikatorische Verfahren (vgl.: Fischer, 2002:102-103).

Zur fachlichen und inhaltlich sinnvollen Konstruktion dieser Fallgruppen ist es notwendig, erklärende Variablen für die Streuung bzw. Vorhersage des Pflegeaufwands zu identifizieren. Nur über eindeutige Erklärungsmodelle können später präzise Gruppen mit einer hohen Trennschärfe entwickelt werden. Die Analyse von Variablen, die zur Klärung der Streuung des Pflegeaufwands beitragen können, wird im internationalen Raum verstärkt durchgeführt. O'Brien-Pallas et al. (1997:171) belegte in einer Schweizer Studie mittels eines Modells, welches neben anderen Faktoren auch die Information der Pflegekomplexität erfasst, eine Erklärungskraft von insgesamt 54%. Hunstein et al. (2007b) belegten für das ergebnisorientierte PflegeAssessment<sup>®</sup> eine Erklärungskraft bezüglich des Pflegeaufwands von 61%. Die von Hunstein et al. berechneten Werte basieren jedoch auf einer kleinen Fallzahl (N=433), so dass diese Ergebnisse lediglich eine Tendenz aufzeigen. Als ein Vergleich zu all diesen Angaben sei die Varianzbindung der G-DRGs von 55% zu nennen (vgl.: InEK, 2003:66). In einer amerikanischen Studie von Welton et al. (2006:367) wurden Daten eines Nursing Minimum Data Sets (NMDS) mit den Daten aus dem DRG-System verglichen. Die Ergebnisauswertung zeigte, dass unter Einbezug des NMDS eine optimierte Interpretation der DRGs ermöglicht werden konnte.

Als ein beispielhaftes Modell zur bereits erfolgreichen Umsetzung pflegerelevanter Fallgruppen in der Rehabilitation sei hier die „TAR-Pflegeklassifikation“ (Fischer, 1998:60) angeführt. Hierbei gelang die Ableitung der Pflegekosten von Kennzeichen der Patienten. Auf der Grundlage von ermittelten Patienteneinschränkungen aus den Bereichen „praktisch-motorische“ und „kognitive“ Fähigkeiten des FIM<sup>20</sup> konnte eine Erklärungskraft von 65% generiert werden (Fischer, 1998:60). Im Rahmen der Konstruktion der Kategorien zur Pflegekostenbestimmung wurden insgesamt sechs Gruppen aufgefunden. Die Studie zeigte das Ergebnis, dass der Zustand eines Patienten gemäß pflegerischer Definition den Pflegeaufwand im Reha-Setting in sehr nachdrücklicher Weise abbilden kann (vgl.: ebd).

Fischer (2003:108) verweist in der Diskussion über die Möglichkeiten der Konstruktion von pflegerelevanten Fallgruppen auf die Problematik der Berücksichtigung von Ausreißern. Fischer betont das Erfordernis der Definierung von Grenzwerten für den Pflegeaufwand bzgl. der Beachtung von übermäßigen Streuungen. Über diese Deklaration können Ausreißer gewichtet werden, mit dem Ziel Vergütungszuschläge zu gewähren.

Alternative Modelle die jedoch die Pflegeabbildung eher als integralen Bestandteil sehen, werden zum Beispiel in den USA umgesetzt. Die USA mit ihren weit längeren Erfahrungen in der Bearbeitung dieser Problematik zeigt einige Möglichkeiten zur Annäherung der Kostenhomogenität in den DRGs auf. Ein Versuch der Definition von Pflegekosten bzw. eine Möglichkeit Pflege zu berücksichtigen, lässt sich in beispielhafter Weise im State New York finden. Hier werden Pflegekosten über „nursing intensity weights“ (NIW) berücksichtigt. Die NIW repräsentieren eine relevante Kennzahl, um die Qualität und den Typ von pflegerischer Intervention innerhalb der verschiedenen DRGs zu berücksichtigen. Mit Hilfe einer modifizierten Delphi-Technik konnten die NIW über ein repräsentatives „Panel“ von Pflegefachkräften gebildet werden. Nach Aussagen von Knauf et al. (2006) ermöglicht diese Pflegeintensitätsgewichtung eine kostenorientierte Preisbildung, welche charakteristische Probleme von pauschalen Kostenmodellen minimiert. Des Weiteren kann den Ausführungen von Knauf et al. (2006) entnommen werden, dass ihre Untersuchungsergebnisse eine hohe Korrelation zwischen den NIW und den jeweiligen Pflegekosten pro Tag auswiesen (Knauf et al., 2006:281-289).

In einer Studie aus Chicago wurden ähnliche Überlegungen der Differenzierung im DRG-System auf der Grundlage von Patientenzustandsindikatoren getätigt. Lynk (2001:111) kommt zu dem Ergebnis, dass zwischen den Patienten differierende

---

<sup>20</sup> FIM = Functional Independence Measure („Funktionaler Selbstständigkeitsindex“) (IVAR, 1999:25)

Pflegekosten sich weniger über die Unterschiede zwischen den DRGs, als vielmehr durch „intra-DRG-Difference“ erklären. Die Differenzierung in einer DRG berücksichtigt beobachtbare Indikatoren der Patientenzustände. Mittels der „intra-DRG-Difference“ lässt sich zu 80% die Patientenkostenvarianz erklären. „Cross-DRG differences“ (Lynk, 2001:111) erklären lediglich die Hälfte der Preisvariationen. Diese Art der Klassifizierung findet aber keine Berücksichtigung im offiziellen DRG-System der USA, weist aber deutlich auf die Bedeutung der Berücksichtigung von Patientenzustandsindikatoren hin.

Peters-Alt (2005:39) schlägt ähnliche Wege der integrierten Pflegeabbildung für Deutschland vor. „Der Pflegefaktor soll neben der Hauptdiagnose, der Operation, dem Alter, dem Geschlecht, der Entlassungsart und den Zustandsmerkmalen als weiteres Fallmerkmal in den Eingruppierungsprozess integriert werden“ (Peters-Alt, 2005:43). Sie sieht in der Einführung eines Pflegefaktors in das DRG-System die Chance zur Steigerung der Kostenhomogenität. Der Pflegeaufwand soll nach Peters-Alt Informationen enthalten, über die der Pflegeaufwand beschrieben werden kann. Hierfür ist die Identifikation der beeinflussenden Faktoren des Pflegeaufwands notwendig. Als Einflussfaktoren sind die fünf folgenden Punkte genannt: „Auflistung der Tätigkeiten, Grund der Durchführung der Tätigkeit, Ziel der Pflege, Ergebnisse der Pflege (und) Aufwand der Tätigkeit“ (Peters-Alt, 2005:43).

All diese geforderten inhaltlichen Elemente des Pflegefaktors könnten durch den Einsatz von standardisierten Assessmentinstrumenten mit vergleichbaren Daten ausgestaltet werden.

Einer der am häufigsten identifizierten Indikatoren bezieht sich auf Informationen zum pflegerischen Zustand des Patienten. Das DRG-System in seiner Konstruktion folgt ähnlichen Ansätzen, auch hier werden Prozeduren und Diagnosen zur Gruppenbildung genutzt. Pflege kann in diesem Zusammenhang ihre Fähigkeiten im Bereich der Patientenzustandserfassung ausspielen. Dessen ungeachtet wird dies nur gelingen, wenn das Datenmaterial in einer standardisierten Form generiert wird. Da nur so die ermittelten pflegerischen Kennzahlen über Patientenfähigkeiten und Patienteneinschränkungen die nötige Transparenz zur Kalkulation, Qualitätssicherung, Evaluation etc. aufweisen können.

Die hier vorliegende Arbeit hat im Verlauf der Ausführungen heraus gearbeitet, dass Patientenzustandsinformationen, die von Seiten der Pflege ermittelt werden, eine hohe Erklärungskraft bzgl. der verbrauchten Pflegezeit haben. Weiterhin wurde in dieser Arbeit gezeigt, dass es möglich ist, Modelle zu entwickeln, mit denen Patienten auf der Grundlage von Zustandsdaten in Pflegeminutengruppen eingeordnet werden können. Mit diesen Ansätzen kann es zukünftig gelingen, der

Pflege ein System an die Hand zu geben, mit dem sie in der Lage sein wird, prospektive Aussagen über den erwarteten Pflegeaufwand zu bilden sowie den personellen und materiellen Mitteleinsatz zu steuern. Des Weiteren kann über diese Vorgehensweise eine stärkere Kostentransparenz hergestellt werden. Sowohl durch die Pflege, als auch für die Pflege und darüber hinaus für das Gesamtgebilde Krankenhaus.

Ein Punkt sei abschließend an dieser Stelle noch zu nennen, bei all der Forderung nach Kennzahlen und Standardisierung. Die unabdingbaren „kalkulativen Praktiken“ (Vormbusch, 2002:25) stellen mehr als nur Daten auf einem hohen Abstraktionsniveau dar. „Sie implizieren bestimmte Auffassungen über richtiges und falsches Verhalten, »gute« und »schlechte« Leistungen (...), und sie machen diese Auffassungen durch kontinuierliche Evaluationen verbindlich“ (ebd:25). Vormbusch merkt weiterhin an, dass Kennzahlen vor diesem Hintergrund nicht nur auf der abstrakten Ebene verweilen, sondern zu realen Handlungsaufforderungen transformiert werden, an dem sich das Handeln des Arbeitsalltags ausrichtet (vgl.: ebd:27). „Kennziffern implizieren deshalb die sukzessive Transformation des Handlungsfeldes im Lichte der durch sie induzierten Optimalitäts- und Rationalitätsvorstellungen“ (ebd:28).

Die Gefahr, die in diesen Mechanismen steckt, ist die tendenzielle Bewusstseinsausrichtung der Mitarbeiter zu kalkulierenden Individuen, mündend in der Frage an sich selbst: „Rentier` ich mich noch?“ (Schmidt, 2005; zitiert nach Vormbusch, 2006:147).

Bei aller Notwendigkeit der Kennzahlengenerierung sollte dies jedoch nicht in einer unreflektierten Kennzahlenverliebtheit enden. Das Individuum und die Konsequenzen für die Mitarbeiter dürfen nicht aus den Fokus der Betrachtungen rücken.

## 8 Empfehlungen

Die in dieser Arbeit ermittelten Ergebnisse konnten einen Zusammenhang zwischen Patientenzustand und Pflegeaufwand aufzeigen. Insbesondere konnte in dieser Arbeit belegt werden, dass es mit einer hohen Wahrscheinlichkeit möglich ist, Patientenfälle anhand ihres pflegerischen Zustands, ausgedrückt durch den CMS<sup>®</sup>, in die richtige LEP-Minutengruppe einzuordnen. Auch wenn diese Werte vor den Einschränkungen der Methodenunsicherheiten betrachtet werden müssen, wird gezeigt, dass dies den ersten Schritt in die Richtung einer Konstruktion von pflege-relevanten Fallgruppen darstellt. Die hier zugrunde gelegten Modelle sollten nun in einem weiterführenden Vorgehen optimiert und verfeinert werden. Auch das Einbeziehen einer höheren Fallzahl ist zwingend notwendig. Nur so kann es gelingen, die Übertragbarkeit auf größere Patientengruppen zu gewährleisten.

Weiterführend muss sich mit dem Thema der Auswahl des CMS<sup>®</sup> zur Zustandsbeschreibung von Patienten auseinander gesetzt werden. Die Überprüfung der prädiktiven Fähigkeiten des CMS<sup>®</sup> steht hierbei im Vordergrund. Es muss analysiert werden, ob alle im derzeitigen CMS<sup>®</sup> enthaltenen Items auch die relevanten Variablen sind, die die Vorhersage des definierten Ergebnisses herbeiführen. Schlarman (2007) untersuchte die prädiktive Fähigkeit des CMS<sup>®</sup> bezüglich einer nach dem Krankenhausaufenthalt fortbestehenden Pflegebedürftigkeit, unter der Fragestellung, ob die Aussagekraft auch bei einer reduzierten Itemliste in gleicher Weise gegeben ist. Dieser Ansatz kann mit Veränderung der abhängigen Variable auf die hier vorliegende Studie übertragen werden, um eine Verbesserung der konstruierten Regressionsmodelle zu erreichen.

Darüber hinaus können solche verdichteten Informationen zur Bildung eines Nursing Minimum Data Sets genutzt werden, um mit minimalem Dateneinsatz die Pflege maximal zu erklären (vgl. Welton et al. 2006:369).

Die Betrachtung der einzelnen Leistungen in den LEP-Minutengruppen sollte zur möglichen Identifikation von Leistungspaketen erfolgen. Somit könnte die zweite von Isfort et al. (2004:119) geforderte Bedingung einer pflegerischen Leistungshomogenität in den einzelnen Patientengruppen zur Entwicklung von pflegerelevanten Fallgruppen erfüllt werden.

Zur Gewährleistung einer Zustandshomogenität (Art der Pflegebedürftigkeit) in den Fallgruppen ist eine Clusteranalyse empfehlenswert.

Damit eine nachhaltige und akzeptierte Umsetzung der Integration von relevanten Fallgruppen in der Pflege gefördert wird, muss nach Isfort et al. (2004:129) ein Umdenken in der pflegerischen Handlungswirklichkeit erfolgen. „Die grundsätzliche Orientierung weg von »nur« tätigkeitsorientierten Beschreibungsformen und Berechnungen der Pflege hin zu einer Fallorientierung passt zudem in die derzeitige Entwicklung der Professionalisierung von Pflege“ (ebd, 2004:129).

Die Nutzung von klinisch relevanten Kennzahlen, unabhängig vom verfolgten Zweck und des Abstraktionsniveaus der Daten, bilden ein wichtiges Instrument der institutionellen Steuerung im Krankenhaus. „Sie nehmen (...) sowohl eine informierende als auch eine orientierende Funktion wahr (...)“ (Vormbusch, 2002:35).

Im Krankenhaus wird eine Menge von Kennzahlen generiert. Es muss jedoch in der Zukunft gelingen, diese Daten über intelligente Lösungen zu ordnen und sie für die unterschiedlichsten Zwecke nutzbar zu machen.

Insbesondere über in der Pflege genutzte standardisierte Assessmentinstrumente kann eine Selektion von Daten und deren Verdichtung hinsichtlich kalkulatorischer Dynamiken in klinischen Unternehmen erfolgen.

## Literaturverzeichnis

Backhaus, K., Erichson, B., Plinke, W., Weiber, R. (2003): Multivariate Analysemethoden. Eine anwendungsorientierte Einführung, Springer Verlag, Berlin, Heidelberg.

Baltes-Götz, B. (2006): Logistische Regressionsanalyse mit SPSS, Universitäts-Rechenzentrum Trier.

Bartholomeyczik, S. (2004): Pflegebedarf und Pflegebedürftigkeit – Konzeptentwicklung, Operationalisierung und Konsequenzen. In: PR-Internet, 7-8. 2004, S. 389-395.

Bartholomeyczik, S. (2003): Pflegediagnostik als Prinzip. In: Etzel B. S. (2003): Pflegediagnostik und Pflegeklassifikationssysteme, Entwicklung und Anwendung, Kohlhammer Verlag, Stuttgart.

Bartholomeyczik, S., Hunstein, D., Koch, V., Zegelin-Abt, A. (2001): Zeitrichtlinien zur Begutachtung des Pflegebedarfs – Evaluation der Orientierungswerte für die Pflegezeitbemessung, Marbuse-Verlag, Frankfurt am Main.

Bartholomeyczik, S., Hunstein, D (2000): Erforderliche Pflege – zu den Grundlagen einer Personalbemessung. In: Pflege & Gesellschaft, 5 Jg., 2000, Nr.4, S. 105-109.

Baukowitz A., Berker, Th., Boes A., Pfeiffer S., Schmiede R., Will M. (Hrsg.) (2006): Informatisierung der Arbeit – Gesellschaft im Umbruch, edition sigma Verlag, Berlin.

Baumberger, D. (2001): Pflegediagnosen als Indikator der Streuung des Pflegeaufwands in DRGs, Master Thesis im Studiengang 3 des Master of Nursing Science. Universität Maastricht NL, Aarau. [http://www.lep.ch/pdf/Masterarbeit\\_Baumberger.pdf](http://www.lep.ch/pdf/Masterarbeit_Baumberger.pdf), Stand: 27.02.2007.

Benninghaus, H. (2001): Deskriptive Statistik. Eine Einführung für Sozialwissenschaftler, Westdeutscher Verlag, Wiesbaden.

Blaudszun, A. (2000): Pflege im Spannungsfeld des Gesundheitssystems. Entwicklungschancen im Spannungsfeld zwischen Wirtschaftlichkeit und Humanität, Kohlhammer Verlag, Stuttgart, Berlin, Köln.

Brosius, F. (2006): SPSS 14, Mitp-Verlag, Heidelberg.

Brosziewski, A., Brügger, U (2001): Zur Wissenschaftlichkeit von Messinstrumenten im Gesundheitswesen: Am Beispiel der Methode LEP®. In: Pflege, 2001, Nr.14, S. 59-66, Hans Huber Verlag, Bern.

Dassen, T., Balzer, K., Bansemir, G., Kühne, P., Saborowski, R., Dijkstra, A. (2001): Die Pflegeabhängigkeitsskala, eine methodologische Studie. In: Pflege, 2001, Nr. 14, 123-127, Hans Huber Verlag, Bern.

Deutscher Pflegerat (2007): Projektinitiative des Deutschen Pflegerats – Adäquate Abbildung des Pflegaufwandes im G-DRG-System durch NRGs.

Deutsches Institut für Medizinische Dokumentation und Information, DIMDI (2005): Internationale Klassifikation der Funktionsfähigkeit, Behinderung und Gesundheit, ICF, World Health Organization, Genf.

Dörrie, U., Preißler, P. (2002): Grundlagen Kosten- und Leistungsrechnung, Oldenbourg Verlag.

Etzel, B. S. (2003): Pflegediagnostik und Pflegeklassifikationssysteme, Entwicklung und Anwendung, Kohlhammer Verlag, Stuttgart.

Fischer, W., Blanco, J., Mäder, M., Zangger, P., Conti, F., Bapst, L., Huwiler, B. (1998): Patientenklassifikation in der Rehabilitation – Das neue TAR-Projekt. In: KSK-aktuell, 1998:60-62.

Fischer, W. (1999): Die Bedeutung von Pflegediagnosen in Gesundheitsökonomie und Gesundheitsstatistik. In: PR-Internet Jg. 1, 1999, Nr.1, S. 5-17.

Fischer, W. (2001): Die Bedeutung von Pflegediagnosen in Gesundheitsökonomie und Gesundheitsstatistik, 3. Aufl., Wolfertswil: Zentrum für Informatik und wirtschaftliche Medizin.

Fischer, W. (2002): *Diagnosis Related Groups (DRGs) und Pflege. Grundlagen, Codierungssysteme, Integrationsmöglichkeiten*, Hans Huber Verlag, Bern.

Fischer, W. (2007): *Sprechen Sie G-DRG? – G-DRG 2007: Statistische Optimierung zu Lasten fachsprachlicher Verständlichkeit*, <http://www.fischer-zim.ch/streiflicht-pdf/GDRG-2007-Sprache-0701.pdf>, Stand 08.03.2007.

Fischer, W. (2005): *Paarweise Vergleich von Patientenklassifikationssystemen – Basis-DRGs, Fraktionierungskoeffizient und Belegungsdiagramme zur Begutachtung der relativen klinischen Homogenität von DRG-Systemen*, ZIM-Verlag, Wolfertswil.

Halloran, E. (1985): *Nursing workload, medical diagnosis related groups, and nursing diagnoses*. In: *Research Nursing Health*. 1995, 8(4):421-433.

Hauck, W., Donner, A. (1977): *Wald's Test as Applied to Hypotheses in Logit Analysis*. In: *Journal of the American Statistical Association*, 1977, 72(360):851-853.

Hosmer, D., Lemeshow, St. (2000): *Applied Logistic Regression*, John Wiley & Sons, New York.

Hunstein D., Fiebig M., Sippel B., Dintelmann Y. (2007a): *Clinical Testing of ePA-AC, a screeninginstrument designed to relevant assess signs and symptoms of nursing care dependency in acute care clinics*. In: Oud, N., Sheerin, F., Ehnfors M., Sermeus W. (Hrsg), (2007): *ACENDIO 2007 - 6<sup>th</sup> European Conference of ACENDIO, 2007*. Oud Consultancy, Amsterdam 190-196.

Hunstein, D., Fiebig, M., Sippel, B., Dintelmann, Y. (April 2007b): *Clinical Testing of ePA-AC© - Screening of relevant clinical indicators of nursing care dependency in acute care clinics*. Paper, presented at the 6th International Conference of ACENDIO, Amsterdam (NL), 19. - 21. April 2007.

Hunstein, D., Dintelmann, Y., Sippel, B. (2006): *Handbuch ergebnisorientiertes PflegeAssessment-AcuteCare (ePA-AC©) Version 1.0*, HSK Pflegeforschung/-entwicklung.

Hunstein, D., Dintelmann, Y., Sippel, B. (2005): Developing a screening instrument as a standardized assessment of signs and symptoms concerning basic nursing care needs in hospital nursing care. In: Oud N, Sermeus W, Ehnfors M (Hrsg.): ACENDIO 2005, Hans Huber, Bern: 396-402.

Hunstein, D., Sippel, B., Dintelmann, Y. (2005): Entwicklung eines Screeninginstruments zur Einschätzung von Pflegeanlässen im Akutkrankenhaus. Referat gehalten bei der 5. Internationalen Konferenz Pflege/Pflegewissenschaft (11.-13.9.2005) Nürnberg, Fürth.

Institut für das Entgeltsystem im Krankenhaus (InEK) (2006): G-DRG Version 2007 Definitionshandbuch, Kompaktversion,

Institut für das Entgeltsystem im Krankenhaus (InEK) (2003): Projektbericht.

IVAR (Hrsg.) (1999): Manual FIM – Funktionale Selbständigkeitsmessung.

Kantonsspital Uri (2003): Jahresbericht 2003. Altdorf (CH).

Knauf, R., Ballard, K., Mossman, P., Lichtig, L. (2006): Nursing cost by DRG: nursing intensity weights. In: Policy Polit Nurs Pract., 2006 Nov., 7 (4), S. 281-289.

König, P. (2003): Pflegediagnosen und ihre Bedeutung in Bezug zu den DRG. In: Etzel, B. S. (2003): Pflegediagnostik und Pflegeklassifikationssysteme, Entwicklung und Anwendung, Kohlhammer Verlag, Stuttgart.

Lynk, W. (2001): One DRG, one price? The effect of patient condition on price variation within DRGs and across hospitals. In: International Journal Health Care Finance Economic, 2001 Jun., 1 (2): 111-137.

Medizinischer Dienst der Spitzenverbände (1997): Richtlinien der Spitzenverbände der Pflegekassen zur Begutachtung von Pflegebedürftigkeit nach dem XI. Buch des Sozialgesetzbuches, DZS-Verlag, Essen.

Mølgard, E. (2000): Calculation of nursing costs in relation to the DRG-system (Abschlussbericht), Viborg: Council of Chief Nurse in Viborg County, 2000.

Mortensen R. (1996): The International Classification of Nursing Practice, Alpha Version, with Telenurse Introduction, Danish Institute for Health and Nursing Research, Kopenhagen.

O'Brien-Pallas, L., Irvine, D., Peerboom, E., Murray, M. (1997): Measuring Nursing Workload. Understanding the Variability. In: Nursing Economics, 1997 (15)4:171-182.

Oud, N., Sheerin, F., Ehnfors M., Sermeus W. (Hrsg), (2007): ACENDIO 2007 - 6<sup>th</sup> European Conference of ACENDIO, 2007. Oud Consultancy, Amsterdam

Oud, N., Sermeus W., Ehnfors M. (Hrsg), (2005): ACENDIO 2005, Hans Huber Verlag, Bern.

Peters-Alt, J. (2005): DRGs aus Sicht der Pflege – Notwendigkeit und Grenzen eines Pflegefaktors; Kohlhammer Verlag, Stuttgart.

Polit, D., Hungler, B. (1999): Nursing Research – Principles and Methods, Lippincott, Philadelphia.

Polit, D., Beck, Ch., Hungler, B. (1999) Lehrbuch Pflegeforschung – Methodik, Beurteilung und Anwendung, Huber Verlag, Bern.

Schlarmann, J. große (2007): Der CMS<sup>®</sup> im ePA – Verschiedene Qualitätsdimensionen eines Instruments. Eine empirische Analyse, Masterarbeit im Rahmen des Studiengangs Master of Science in Nursing, Universität Witten/Herdecke gGmbH.

Schmidt, A. (2005): Rentier`ich mich noch? Activity Based Costing und seine Wirksamkeit im Tun der Beschäftigten. In: Wagner, H. (Hrsg.) (2005): „Rentier`ich mich noch?“ Neue Steuerungskonzepte im Betrieb, Hamburg, S. 131-153.

Singleton-Green, B. (1993): If It Matters, Measure It! In: Accountancy, vol. 111, no. 5, pp. 52-53.

Tutz, G. (2000): Die Analyse kategorialer Daten. Anwendungsorientierte Einführung in Logit-Modellierung und kategoriale Regression, Oldenbourg Verlag, München, Wien.

Urner Kantonalbank (2005): Uri in Zahlen – Ausgabe 2005/2006.

Vormbusch, U. (2006): Accounting, Informatisierung und der Calculating Man. In: Baukowitz A., Berker, Th., Boes A., Pfeiffer S., Schmiede R., Will M. (Hrsg.) (2006): Informatisierung der Arbeit – Gesellschaft im Umbruch.

Vormbusch, U. (2002): Diskussion und Disziplin – Gruppenarbeit als kommunikative und kalkulative Praxis, Campus Verlag, Frankfurt, New York.

Wagner, H. (Hrsg.) (2005): „Rentier` ich mich noch?“ Neue Steuerungskonzepte im Betrieb, Hamburg.

Wieteck, Pia (2004): ENP – European Nursing care Pathways. Standardisierte Pflegefachsprache zur Abbildung von pflegerischen Behandlungspfaden, Recom Verlag, Bad Emstal.

Welton, J., Halloran, E., Zone-Smith, L. (2006): Nursing Intensity: In the Footsteps of John Thompson. In: Studies in Health Technology and Informatics. 2006, 122: 367-371.

Welton, J., Halloran E. (2005): Nursing Diagnoses, Diagnosis-Related Groups, and Hospital Outcomes. In: Journal of Nursing Administration. 2005, 35(12): 541-549.

## **Internetverzeichnis**

<http://destatis.de>; Stand: 21.02.2007

<http://www.health.gov.au/internet/wcms/publishing.nsf/Content/health-casemix-ardrg1.htm>, Stand: 08.03.2007.

<http://www.ksuri.ch/15.html>, Stand: 28.04.2007

[http://www.lrz-muenchen.de/~wlm/ilm\\_I7.htm](http://www.lrz-muenchen.de/~wlm/ilm_I7.htm), Stand: 18.04.2007

<http://db03.bmgs.de/gesetze/sgb11x014.htm>; Stand: 02.05.07

## Abkürzungsverzeichnis

ADL	activities of daily living
AR-DRG	Australian Refined Diagnosis Related Groups
CC	Comorbidity or Complication
CMS <sup>®</sup>	CaseManagementScore <sup>®</sup>
DIMDI	Deutsches Institut für Medizinische Dokumentation und Information
DRG	Diagnosis Related Groups
DPR	Deutscher Pflegerat
ENP <sup>®</sup>	European Nursing care Pathways <sup>®</sup>
ePA-AC <sup>®</sup>	ergebnisorientiertes PflegeAssessment-AcuteCare <sup>®</sup>
FIM	Functional Independence Measure
G-DRGs	German - Diagnosis Related Groups
ICF	Internationale Klassifikation der Funktionsfähigkeit, Behinderung und Gesundheit
InEK	Institut für das Entgeltsystem im Krankenhaus
LEP <sup>®</sup>	Leistungserfassung in der Pflege <sup>®</sup>
NANDA	North American Nursing Diagnosis Association
NIW	nursing intensity weights
NRG	Nursing Related Groups
RAI	Resident Assessment Instrument
SGB XI	Sozialgesetzbuch XI
PAS	Pflegeabhängigkeitsskala
PLAISIR <sup>®</sup>	PLAnification Informatisée des Soins Infirmiers Requis <sup>®</sup>
PRN	Projet de Recherche en Nursing
PPR	Pflegepersonalregelung
TAR	Leistungsbedarfsbezogenes Tarifsysteem für Rehabilitationskliniken

## Abbildungsverzeichnis

Abbildung 1: Variablendarstellung .....	23
Abbildung 2: Darstellung der Ausreißer bei der Verweildauer (N=436 Einschätzungen) .....	31
Abbildung 3: Verteilung CMS und LEP-Minuten in der Stichprobe (N=436) .....	32
Abbildung 4: Normalverteilungsdiagramm für die Variablen CMS und LEP- Minuten_Gesamt .....	32
Abbildung 5: Gegenüberstellung der LEP-Interventionsgruppen.....	37
Abbildung 6: Zusammenhänge der logistischen Regression (nach Backhaus et al., 2003:426) .....	39
Abbildung 7: Regressionsgleichung.....	40
Abbildung 8: Histogramm zur Tabelle 2 .....	43
Abbildung 9: Histogramm zur Tabelle 3 (N=436) .....	43
Abbildung 10: Pseudo R-Square.....	44
Abbildung 11: Likelihood-Quotienten-Test .....	44
Abbildung 12: Histogramm zur Tabelle 6 (N=436) .....	46
Abbildung 13: Histogramm zur Tabelle 11 (N=436) .....	48

## Tabellenverzeichnis

Tabelle 1: Unter- und Überlieger in der Stichprobe (N=436) .....	31
Tabelle 2: Verteilung in den CMS-Gruppen.....	42
Tabelle 3: Verteilung in den LEP-Minuten-Gruppen .....	43
Tabelle 4: Information zur Modellanpassung.....	43
Tabelle 5: Güte der Anpassung .....	44
Tabelle 6: Häufigkeitsverteilung .....	46
Tabelle 7: Information zur Modellanpassung.....	46
Tabelle 8: Güte der Anpassung .....	47
Tabelle 9: Pseudo R-Quadrat .....	47
Tabelle 10: Likelihood-Quotienten-Test .....	47
Tabelle 11: Häufigkeitsverteilung .....	48
Tabelle 12: Information zur Modellanpassung.....	48
Tabelle 13: Güte der Anpassung .....	48
Tabelle 14: Pseudo R-Quadrat .....	48
Tabelle 15: Likelihood-Quotienten-Test .....	48

## **Anhang**

Die ausführlichen Anhänge können bei Interesse bei der Autorin angefordert werden.

## **Persönliche Erklärung**

Ich versichere, dass ich die Arbeit ohne fremde Hilfe angefertigt und mich anderer als der von mir angegebenen Schriften und Hilfsmittel nicht bedient habe.

Darmstadt, 07.Mai 2007

---

Madlen Fiebig