

Aus dem Med. Zentrum für Operative Medizin  
Klinik für Unfall-, Hand- und Wiederherstellungschirurgie  
Direktor: Prof. Dr. med. Steffen Ruchholtz  
des Fachbereichs Medizin der Philipps-Universität Marburg



**Adjustierung und Optimierung des SAPS II  
anhand des Nationalregisters der DIVI**

Inaugural-Dissertation  
zur Erlangung des Doktorgrades der gesamten Humanmedizin

dem Fachbereich Medizin der Philipps-Universität Marburg

vorgelegt von

Erkan Duman  
aus Münster

Marburg, 2013

Angenommen vom Fachbereich Medizin  
der Philipps-Universität Marburg am  
14.05.2013

Gedruckt mit Genehmigung des Fachbereichs

**Dekan:**

Prof. Dr. med. Matthias Rothmund

**Referent:**

PD Dr. med. Raphael Stiletto

**Korreferentin:**

PD Dr. med. Caroline Rolfes

## Inhaltsverzeichnis

<b>1 Einleitung .....</b>	<b>3</b>
1.1 DIVI .....	3
1.2 Intensivmedizin .....	3
1.3 Qualitätssicherung .....	4
1.4 Vergleichbarkeit .....	5
1.5 Scoresysteme als Methode .....	7
1.6 Verschiedene Scoresysteme .....	8
1.7 Initiative der DIVI (Sektion „Qualität und Ökonomie in der Intensivmedizin“).....	10
1.8 Thema.....	11
<b>2 Material und Methode .....</b>	<b>13</b>
2.1 Allgemeine Beschreibung .....	13
2.2 Verwendete Scoresysteme.....	13
2.3 SAPS II.....	15
2.3.1 Hintergrund .....	15
2.3.2 Aufbau .....	15
2.3.3 Ein- und Ausschlusskriterien .....	16
2.3.4 Gründe für die Anwendung von SAPS II.....	16
2.3.5 Entwicklung der 17 Variablen.....	17
2.3.6 Kalibrierung .....	17
2.3.7 Diskriminierung .....	18
2.3.8 SAPS-II-Formel .....	18
2.3.9 Vergleich SAPS – SAPS II .....	18
2.3.10 Vorteile des SAPS II .....	19
2.3.11 Nachteile des SAPS II.....	19
2.3.12 Nutzung in unterschiedlichen Ländern .....	20
2.3.13 Vorhersagekraft für die Behandlungsdauer von Intensivpatienten .....	20
2.3.14 SAPS II bei Trauma-Patienten .....	21
2.3.15 SAPS II bei herzchirurgischen Patienten .....	21
2.3.16 SAPS II - angepasster SAPS II – erweiterter SAPS II.....	22
2.3.17 Zusammenfassung – SAPS II.....	23
2.3.18 SAPS 3 .....	23
2.4 Gütekriterien von Scores .....	25
2.4.1 Tests zur Überprüfung der Leistung von Scores .....	25
2.4.2 Bewertung von Scores.....	25
2.4.3 Reliabilität .....	25
2.4.4 Validität .....	26
2.4.5 Messbarkeit .....	26
2.4.6 Güte .....	26
2.4.7 Anwendbarkeit .....	26
2.4.8 Klinische Relevanz.....	27
2.4.9 Diskriminierung .....	27
2.4.10 Präzision .....	29
2.4.11 Kalibrierung .....	30
2.4.12 Konfidenzintervall.....	30
2.4.13 p-Wert.....	31
2.5 Statistik.....	32
2.5.1 Regressionsanalyse.....	32
2.5.2 Logistische Regression.....	33
2.6 Vorgehen .....	35

<b>3 Ergebnisse .....</b>	<b>37</b>
3.1 Generelle statistische Daten der untersuchten Patienten .....	37
3.2 SAPS-DIVI 2 .....	41
3.2.1 Aufbau .....	41
3.3 Analyse der einzelnen Parameter beim SAPS-DIVI 2 im Entwicklungsdatensatz.....	42
3.3.1 Regressionskoeffizienten .....	42
3.3.2 Konstante.....	44
3.3.3 ROC-Kurve .....	44
3.3.4 Fachrichtungen .....	46
3.3.5 Präzision .....	47
3.3.6 Zusammenfassung aller Variablen .....	48
3.4 SAPS-DIVI 2 (Phys) .....	51
3.4.1 Aufbau .....	51
3.5 Analyse der einzelnen Parameter beim SAPS-DIVI 2 (Phys) im Entwicklungsdatensatz .....	52
3.5.1 Regressionskoeffizienten .....	52
3.5.2 Konstante.....	53
3.5.3 ROC-Kurve .....	53
3.5.4 Fachrichtungen .....	55
3.5.5 Präzision .....	56
3.5.6 Zusammenfassung aller Variablen .....	56
3.6. SAPS-DIVI 2 im Validierungsdatensatz.....	59
3.6.1 Auswertung .....	59
3.6.2 Zusammenfassung aller Variablen .....	61
3.7 SAPS-DIVI 2 (Phys) im Validierungsdatensatz .....	63
3.7.1 Auswertung .....	63
3.7.2 Zusammenfassung aller Variablen .....	64
3.8 Vergleich SAPS II (Original) – SAPS-DIVI – SAPS-DIVI 2 .....	66
3.8.1 Einleitung .....	66
3.8.2 Vergleich der Gütekriterien.....	66
3.8.3 Diskriminierung .....	66
3.8.4 Präzision .....	67
3.8.5 Kalibrierung .....	67
3.8.6 Zusammenfassung .....	71
 <b>4 Diskussion .....</b>	 <b>72</b>
 <b>5 Zusammenfassung .....</b>	 <b>78</b>
 <b>6 Literaturverzeichnis.....</b>	 <b>79</b>
 <b>7 Abkürzungsverzeichnis.....</b>	 <b>87</b>
 <b>8 Anhang .....</b>	 <b>88</b>
8.1 Tabellen .....	88
8.2 Verzeichnis der akademischen Lehrer .....	107
8.3 Danksagung .....	108

## **1 Einleitung**

### **1.1 DIVI**

Die DIVI (Deutsche Interdisziplinäre Vereinigung für Intensiv- und Notfallmedizin) ist ein Zusammenschluss von persönlichen Mitgliedern, wissenschaftlichen Gesellschaften und Berufsfachverbänden mit dem Ziel der Förderung der Intensiv- und Notfallmedizin in Deutschland. Sie ist aus mehreren Sektionen zusammengesetzt, die jeweils ihre eigenen Aufgabenbereiche sowie Arbeitsschwerpunkte haben.

So gehört beispielsweise die IAG „Qualität und Ökonomie in der Intensivmedizin“ zu einer von diesen Sektionen. Dessen Zielsetzung ist es Qualitätsindikatoren zu entwickeln und einen externen Qualitätsvergleich zu schaffen. Zudem liegt der Fokus dieser Sektion im Qualitätsmanagement sowie in der Ökonomie der Intensivmedizin.

### **1.2 Intensivmedizin**

Bei der Intensivmedizin handelt es sich um einen medizinischen Fachbereich, der sich mit der Diagnostik und Therapie lebensbedrohlicher Zustände und Krankheiten befasst. Dies geschieht in der Regel auf besonders ausgestatteten Stationen (Intensivstationen), die durch verschiedene Fachdisziplinen wie die Anästhesie, Innere Medizin, Chirurgie, Pädiatrie oder Neurologie vertreten sind. Auf Intensivstationen ist aufgrund des hohen Betreuungsaufwandes eine Pflegekraft für 1-3 Patienten zuständig im Vergleich zu Normalstationen mit einem Verhältnis von 1:20.

Die Einführung der Intensivmedizin hat seinen Ursprung in der Anästhesiologie, da die Beatmung der Grund dafür war, dass derartig aufwendige Methoden und Einrichtungen geschaffen wurden.

Auf Intensivstationen werden Patienten eingewiesen, deren Zustand bedrohlich ist oder deren Zustand bedrohlich werden könnte. Folglich führen nicht nur schwere Krankheiten, sondern auch Zustände nach großen und stark eingreifenden Operationen zur intensivmedizinischen Überwachung und ggf. Behandlung.

Die Überwachung und Aufrechterhaltung von Organfunktionen mittels differenzierter medikamentöser und technischer Unterstützung bis hin zum vollständigen Organersatz bilden wesentliche Aspekte der Intensivmedizin. Durch die Einführung verschiedener Behandlungsoptionen im Verlaufe der Jahre wie der maschinellen Überdruckbeatmung, Nierenersatztherapie oder einer immer differenzierteren Versorgung konnten entscheidende Therapieerfolge erzielt und damit eine deutliche Verbesserung des Outcome der betroffenen

Patienten erreicht werden. Unter Outcome versteht man in der Intensivmedizin meist Überlebenszeiten oder die Dauer der Wirksamkeit einer spezifischen therapeutischen Intervention, was unter anderem an ihrer Letalität bemessen wird [35].

Neben geplanten Aufnahmen auf Intensivstationen, kommen auch Notaufnahmen von Nicht-Intensivstationen vor. Häufig fällt hierbei auf, dass eine frühere Aufnahme bzw. ein früheres Erkennen einer intensivmedizinischen Behandlungsnotwendigkeit mit hoher Wahrscheinlichkeit das Behandlungsergebnis nachhaltig hätte verbessern können [54].

Um rechtzeitig und mit einfachen Mittel zu erkennen, dass bei Patienten auf Nicht-Intensivstationen eine intensivmedizinische Behandlungsbedürftigkeit besteht, sind verschiedene Parameter und Scores eingeführt worden [16, 52-53].

Intention dieser Scores ist es, bei der Entscheidung einer potenziellen Übernahme des Patienten auf eine Intensivstation zu unterstützen. Des Weiteren gibt es Scores, mit denen die Letalität von Intensivpatienten vorhergesagt werden kann. Diese können als objektives Hilfsmittel bei der Beurteilung des Erkrankungsschweregrads von Intensivpatienten dienen [56].

Die Zunahme von Prozess-, Struktur- und Ergebnisqualität in der Intensivmedizin ist mit einem umfassenderen Einsatz der Methoden verbunden.

Ziel der Intensivmedizin ist die Wiederherstellung der völligen Gesundheit oder zumindest das Erreichen eines weitgehend autonomen Zustandes des Patienten.

### **1.3 Qualitätssicherung**

Unter Qualitätssicherung versteht man Maßnahmen und Instrumente, die geeignet erscheinen, die Beschaffenheit von Dienstleistungen, der intensivmedizinischen Versorgung, zur Erfüllung der in sie gesetzten Erwartungen zu stabilisieren oder zu verbessern.

Qualitätsmanagement sind die aufeinander abgestimmten Tätigkeiten zum Leiten und Lenken einer Organisation (Intensivstation) bezüglich Qualität. Dies beinhaltet üblicherweise das Festlegen der Qualitätsziele, die Qualitätsplanung, die Qualitätslenkung, die Qualitätssicherung und die Qualitätsverbesserung. Die Qualitätssicherung ist demnach ein Teilaspekt des übergeordneten Qualitätsmanagements.

Der erste Schritt im Qualitätsmanagement bildet die Definition eines Qualitätsziels, welches zugleich die Definition des Soll-Werts ist. Die Zieldefinierung bewegt sich im Spannungsfeld zwischen möglichst guter Effektivität (Zielerreichung unabhängig von den eingesetzten Mitteln) und möglichst hoher Effizienz (Zielerreichung mit möglichst geringem Einsatz von Mitteln). In der Qualitätsplanung werden dann Konzepte und Abläufe erarbeitet, um die

vorgegebenen Ziele zu erreichen, zu denen Struktur- oder Organisationsmaßnahmen sowie Weiterbildungs- und Trainingsaktivitäten gehören. In der anschließenden Phase werden die erarbeiteten Planungen umgesetzt (Qualitätslenkung). Das Auswerten der erreichten Leistungen (Ist-Zustand) und der Vergleich zwischen dem erreichten Ist-Zustand und dem definierten Soll-Wert stellen den Kern der Qualitätssicherung dar.

Für die Durchführung des Qualitätsmanagements hat sich die Form des Qualitätszirkels als äußerst nützliche Methode bestätigt. Dieser tagt regelmäßig, läuft nach definierten Regeln ab und steuert den gesamten Prozess (von der Zielvorgabe bis zur Analyse und den Schlussfolgerungen).

Die Bereiche, in denen Qualität erreicht werden kann, beziehen sich auf die Struktur und die Organisation, die Prozessabläufe und das Behandlungsergebnis. Während sich die Strukturqualität auf die Voraussetzungen, die für die Erbringung einer hohen Leistungsqualität erforderlich sind, bezieht, ist die Prozessqualität durch die Qualität des Behandlungsprozesses selbst definiert, die den Ablauf diagnostischer und therapeutischer Maßnahmen ebenso wie die Zusammenarbeit verschiedener Personengruppen oder Abteilungen innerhalb der Station oder des Krankenhauses beinhaltet. Die Ergebnisqualität hingegen bezieht sich auf die Qualität der Zielerreichung [91].

#### **1.4 Vergleichbarkeit**

Das Ergebnis einer Behandlung ist das wichtigste Kriterium für die Qualität der Versorgung. Alle anderen Maßnahmen, insbesondere auch in den Bereichen der Struktur- und Prozessqualität, zielen ausschließlich auf die Erreichung eines möglichst guten Ergebnisses ab. Im Gesamtergebnis ist allerdings ein weites Spektrum verschiedenster Ergebnisausprägungen enthalten.

Für die Bewertung jeglichen Ergebnisses ist grundsätzlich ein Vergleich notwendig. Ein Vorher-Nachher-Vergleich zeigt den Fortschritt, den die eigenen qualitätsverbessernden Maßnahmen bewirkt haben, am besten. Er ist das Herz des Qualitätszyklus. Es fehlt allerdings die Kalibrierung an einem übergeordneten allgemein gültigen Standard, um das Ergebnis im Spektrum dessen, was erwartet bzw. geleistet wird, einordnen zu können. Hierfür ist ein externer Qualitätsvergleich erforderlich.

Kernproblem aller Vergleiche ist die Frage, ob verschiedene Patientenkollektive überhaupt miteinander vergleichbar sind. So kann das Behandlungsergebnis vom Alter und Geschlecht des Patienten, seinen Vorerkrankungen, der Art und dem Stadium bzw. der Prognose der Grunderkrankung, der Dauer der akuten Erkrankung vor Aufnahme auf die Intensivstation,

dem Zustand des Patienten zum Zeitpunkt der Aufnahme auf die Intensivstation und vielen anderen Faktoren abhängen. Nur wenn bezüglich dieser und anderer Risikofaktoren eine Vergleichbarkeit zwischen Patientenkollektiven besteht, kann auf den Einfluss von Qualitätskriterien zurückgeschlossen werden.

Ein externer Qualitätsvergleich ist ohne eine Form der risikoadjustierten Schweregradbeschreibung (z.B. mit Scores) nicht möglich.

Für die Intensivmedizin existieren eine Reihe von gut validierten Prognosescores, wie z.B. der SAPS-II-Score.

Ein Ergebnis kann binominal (z.B. ja/nein; tot/lebend), in diskreten Kategorien (z.B. graduierte Schwere eines Organversagens) oder als kontinuierliche Variable (z.B. Liegedauern) dargestellt werden. Nicht selten werden komplexe Ergebnisse in Form eines Scores zusammengefasst. Aus Gründen der Vergleichbarkeit sollten publizierte und validierte Scores bevorzugt werden.

Für den Ergebnisvergleich ist der Vergleich mit den Ergebnissen anderer, vergleichbarer Intensivstationen von besonderem Interesse. Diese können, beispielsweise anhand publizierter Daten aus wissenschaftlichen Studien, als die möglicherweise besten erzielbaren Ergebnisse definiert werden. Dies hat den Nachteil, dass man sich mit Ergebnissen vergleicht, die zwar unter optimalen Bedingungen erreichbar sind, aber unter artifiziellen Studienbedingungen und Patientenselektionierungen entstanden sind und damit auf typische Patientenkollektive nicht unbedingt sinnhaft zu übertragen sind. Auch ist eine Vergleichbarkeit in der Datenerfassung und Datenqualität meist nicht sichergestellt.

Günstiger erscheint die Teilnahme an größeren Qualitätssicherungsprojekten mit definierter Datenerfassung, einheitlichen Definitionen und vorgegebenen Qualitätsindikatoren. Hierzu gehören unter anderem die Deutsche Interdisziplinäre Vereinigung für Intensivmedizin (DIVI), der Verein Österreichisches Zentrum für Dokumentation und Qualitätssicherung in der Intensivmedizin (ASDI), Krankenhaus-Infektions-Surveillance-System (KISS), Qualitätsmanagement IST der Landesärztekammer Thüringen und die Leapfrog Group (USA).

Zur Interpretation des Vergleichs ist die Definierung des angestrebten Ziels entscheidend. Allgemein gesprochen, kann es ein Ziel sein, eine bestimmte Anforderung immer zu erfüllen, schlechte Ergebnisse zu vermeiden bzw. zu minimieren oder im Durchschnitt eine bestimmte Leistung zu erreichen (benchmarking). Die angestrebte Leistung kann der durchschnittlichen Leistung vergleichbarer Intensivstationen entsprechen oder sich an den besten Werten orientieren. Der Vergleich mit dem Durchschnitt kann dazu führen, dass zwar



Problembereiche identifiziert werden können, dass jedoch nicht die Potenziale realisiert werden, die eine Spitzenleistung ermöglichen.

Andererseits zeigt eine Leistung wie der Durchschnitt eine adäquate Qualität an. Ein unterdurchschnittliches Ergebnis oder ein Ergebnis, das nicht den eigenen Erwartungen oder Anforderungen entspricht, gibt Anlass, neben möglichen Strukturdefiziten auch die eigenen Prozesse zu untersuchen, um das Verbesserungspotenzial auszuschöpfen [91].

### **1.5 Scoresysteme als Methode**

Anfang der 80er haben Intensivmediziner beschlossen den Erkrankungsschweregrad von Intensivpatienten in Form eines Scores wiederzugeben, um einerseits in der Lage zu sein verschiedene Patientengruppen zu vergleichen und andererseits die Leistung von Intensivstationen zu evaluieren. Das Outcome des Scores hängt dabei von mehreren Faktoren ab, die am ersten Tag auf der Intensivstation auftreten sowie von der Therapie, die im Verlauf des Aufenthalts den Patienten angeboten wird. Schweregrad-Scores umfassen in der Regel zwei Bereiche: den Score selbst und ein Wahrscheinlichkeitsmodell. Während der Score eine Punktzahl darstellt (höchste Punktzahl => größter Schweregrad), reflektiert das Wahrscheinlichkeitsmodell dagegen eine Gleichung, welche die Hospitalletalität von Patienten wiedergibt [37]. Scoresysteme werden somit auch zur Prognoseeinschätzung des Mortalitätsrisikos bei Intensivpatienten angewandt [76].

Der Begriff Score stammt aus dem Englischen und bedeutet Punktzahl. Man versucht mithilfe eines Scores einer komplexen klinischen Situation einen Punktwert zuzuordnen, also auf einer eindimensionalen Skala abzubilden. Ziel dieser Reduktion ist es übergreifende Aspekte (z.B. Schweregrad oder Prognose) zu vergleichen. Somit soll eine komplexe Situation durch die Konzentration auf das Wesentliche überschaubar gemacht und vergleichende Betrachtungen unter vielen Patienten ermöglicht werden. Auf Intensivstationen hingegen ist es interessant zu wissen, wie weit ein Patient davon entfernt ist die Intensivstation lebend oder sterbend zu verlassen. Diesen Ansatz versucht man anhand eines Scores zu quantifizieren [69-70].

Scores sind Systeme, die aus der Summation einzelner Punktwerte die Ausprägung eines Merkmals beschreiben [78]. Sie erfassen pathophysiologische Besonderheiten spezifischer Krankheitsbilder und sind somit gut geeignet die Ausprägung der Erkrankung zu quantifizieren. Die Anwendung prognostischer Scores in der Intensivmedizin ermöglicht zum einen die Quantifizierung des Krankheitsschweregrads und die Charakterisierung von Patientengruppen auf der Basis objektiver Kriterien, zum anderen tragen sie zum Qualitätsmanagement bei. Dies umfasst Qualitätsanalyse, Qualitätskontrolle,

Qualitätsverbesserung und Qualitätssicherung in den Bereichen der Struktur, des Prozesses und der Ergebnisqualität [77]. Darüber hinaus sind Scores ausreichend validiert und gut geeignet zur Evaluierung von klinischen Studien oder Kosten-Nutzen-Analysen durch Charakterisierung von Patientengruppen [6]. So ermöglichen Scoresysteme die Charakterisierung des behandelten Patientenguts einer Intensivstation, wodurch die entstehenden Kosten transparent gemacht werden, um die Effizienz der Therapie darzustellen und vergleichend zu evaluieren [86]. Sie bieten also eine größere Transparenz der intensivmedizinischen Leistungserbringung und damit Vergleichbarkeit zwischen unterschiedlichen Intensivstationen im Sinne des Qualitätsmanagements [1].

Die prognostische Qualität eines Scores kann man anhand der Korrelation zwischen Scorewert und Letalität erkennen. Je höher die lineare Übereinstimmung zwischen geschätzter und tatsächlicher Letalität, umso größer ist die diagnostische Qualität eines Scores [1]. Das resultierende Verhältnis von beobachtetem zu erwartetem Letalitätsrisiko verschafft einen Ansatz zur Einschätzung der Mortalität und ist zum „Goldstandard“ für die Bewertung der klinischen Leistung von Intensivstationen geworden [73]. Es beurteilt die Wirkung von Intensivstation relevanten Faktoren (z.B. Organisation und Management) auf das Outcome [65,72].

Ziel von krankheitsübergreifenden Scores ist die Ermöglichung Patientengruppen mit verschiedenen Erkrankungen zu vergleichen oder die Effektivität von Behandlungskonzepten zu evaluieren. So ist zur Qualitätssicherung oder zur Kosten-Nutzen-Effektivität in der Intensivmedizin ein objektives Kriterium notwendig, das dazu dient den Leistungsaufwand wiederzugeben [6]. Dieses ist auch der Grund dafür, warum es aus medizinischer, ethischer und ökonomischer Sicht gewünscht ist den Erfolg oder die Ineffektivität von Behandlungen vorzeitig und mit Bestimmtheit zu erkennen [86]. Andererseits jedoch werden häufig auch Entscheidungen über Therapiemaßnahmen subjektiv-emotional getroffen, weshalb hierbei Scoresysteme auf gewisse Art und Weise als objektive Entscheidungshilfe dienen können [1].

## **1.6 Verschiedene Scoresysteme**

Viele Schweregrad-Scores wurden publiziert, jedoch werden nur wenige benutzt. Die meisten Scores werden von Daten berechnet, die am ersten Tag der Intensivstation gesammelt werden. Zu diesen gehören Acute Physiology and Chronic Health Evaluation (APACHE), Simplified Acute Physiology Score (SAPS) und Mortality Prediction Model (MPM). Andere Systeme wie Organ System Failure (OSF) [32], Organ Dysfunction and Infection System (ODIN) [15], Sequential Organ Failure Assessment (SOFA) [88], Multiple Organs Dysfunction Score

(MODS) [51], Logistic Organ Dysfunction (LOD) [40] model und Three-Day Recalibrating ICU Outcomes (TRIOS) [84] hingegen sammeln die Daten während der Intensivverweildauer wiederholt an jedem Tag oder an den ersten drei Tagen [37].

Physiologische Scores betrachten nur die Intaktheit des Organismus anhand von physiologischen Parametern, unabhängig von der therapeutischen Einflussnahme. Hintergrund dieses Vorgehens ist, dass jedes Organ im Gesamtorganismus eine Aufgabe hat, wobei lediglich nur erfasst wird, inwieweit diese Aufgabe – gemessen an ausgewählten klinischen und Laborparametern – erfüllt wird [70]. Der SAPS-II-Score [41] zählt unter anderem zu den physiologischen Scores.

Die therapeutisch notwendige Unterstützung einer Organfunktion ist ein klinisch äußerst wichtiger Indikator für dessen Zustand. Ausschließlich auf therapeutischen, diagnostischen und pflegerischen Maßnahmen aufgebauter Score ist beispielsweise der TISS-28 [61].

Scores, die sich auf bestimmte Krankheitsbilder beziehen, betrachten nur eine für diese Situation spezifische Auswahl von Faktoren. Spezifische Scores haben den Vorteil, einzelne Aspekte einer Erkrankung deutlich stärker gewichten zu können oder spezielle Aspekte einzubeziehen, die nur bei diesem Krankheitsbild von Bedeutung sind. Bei homogenen Patientengruppen kann dies von Vorteil sein [70]. Beispiele sind LIS (Lung Injury Severity Score) [68] und ABSI (Abbreviated Burn Severity Index) [85].

Allgemeine bzw. krankheitsübergreifende Scores versuchen Aspekte zu kombinieren, die allgemeine Indikatoren von Gesundheit oder Krankheit sind. Bezogen auf die Intensivtherapie lassen sich solche Scores in der Regel auf alle Intensivpatienten anwenden. Liegt eine heterogene Patientenkohorte vor, macht es Sinn sich für ein krankheitsübergreifendes System zu entscheiden [70]. Dazu gehören die APACHE-Scores [28, 30, 34] sowie die SAPS-Scores [41, 59, 67].

Subjektive Scores wurden von einem Gutachterausschuss eingeführt, welche die Variablen aussuchen und deren Gewichtung festlegen, basierend auf ihre persönliche Meinung. Für jede Variable ist ein Normbereich definiert, bei dem der Score null innerhalb dieses Intervalls liegt. Umso abnormaler das Ergebnis, umso höher ist dessen Gewichtung, das zwischen null und vier liegt. Die Gesamtpunktzahl bildet den Score [37]. Das meistverwendete Score System ist APACHE II [30].

Objektive Scores dagegen benötigen zur Entwicklung eines Wahrscheinlichkeitsmodells eine umfangreiche Datenbank von möglichst vielen Intensivstationen. Die Variablen können im Allgemeinen in vier Gruppen klassifiziert werden: Alter, Komorbiditäten, physiologische Abweichungen und akute Diagnosen. Manche Systeme haben auch Variablen eingeführt, um

den lead-time bias zu mindern. Das Outcome jedes dieser Systeme ist der Vitalstatus bei Krankenhausentlassung, wobei aber auch andere Outcome Maße (beispielsweise Vitalstatus 28 Tage nach Krankenhausentlassung oder Lebensqualität bei Langzeitüberlebenden) gebildet werden können. Logistische Regressionsmodelltechniken [45, 47], smoothing Methoden [21] sowie die klinische Beurteilung werden verwendet, um Variablen auszuwählen und deren Gewichtung zu bestimmen sowie Intervalle für jeden einzelnen Parameter zu definieren [37]. Zu den objektiven Scores gehören unter anderem APACHE III [28], SAPS II [41] sowie MPM II [46, 48].

In dieser Studie wurde das Augenmerk auf den SAPS-II-Score gelegt. Bei SAPS II handelt es sich um einen Score, der zur Einschätzung der Erkrankungsschwere von Intensivpatienten auf der Basis physiologischer Daten sowie zur Berechnung des Letalitätsrisikos dient und sich gut zur Klassifizierung von Patientengruppen für klinische Studien, epidemiologische Untersuchungen oder für qualitätssichernde Maßnahmen eignet.

SAPS II setzt sich aus einer Formel zusammen, die es ohne großen Aufwand ermöglicht, jedem berechneten Scorewert ein individuelles Letalitätsrisiko zuzuordnen. Folglich kann man anhand einer mathematischen Gleichung das Letalitätsrisiko aus dem SAPS-II-Score berechnen. Dieses nimmt linear zum SAPS-II-Scorewert zu. Je höher der SAPS-II-Score umso größer ist der Schweregrad der Erkrankung und demzufolge ist auch gleichzeitig die Prognose schlechter.

### **1.7 Initiative der DIVI (Sektion „Qualität und Ökonomie in der Intensivmedizin“)**

Bei der Qualitätssicherung handelt es sich nicht nur um eine gesetzlich geforderte Maßnahme sondern auch für Ärzte um eine wichtige Rückmeldung über die Qualität der von ihnen geleisteten Behandlung. Die Arbeitsgruppe „Qualitätssicherung“ der DIVI beschäftigt sich mit der Qualität und Ökonomie in der Intensivmedizin. Ihre Aufgabe ist es in einem interdisziplinären Ansatz qualitätssichernde Maßnahmen für die Versorgung von intensivbehandlungs- oder überwachungsbedürftigen Patienten anhand von definierten Parametern zu entwickeln und in einem ersten Teilschritt ein nationales Register zum externen Qualitätsvergleich in der Intensivmedizin zu etablieren. Das Register basiert auf der anonymisierten Erhebung eines Datensatzes, der Informationen zur Struktur der beteiligten Intensivstationen, Patienten-Stammdaten, Verlaufsdaten zum Behandlungsaufwand und den Organfunktionen sowie Ergebniskriterien enthält. Anhand dieses Datensatzes können sich die beteiligten Intensivstationen mit einem nationalen Referenzwert vergleichen und im Rahmen des internen Qualitätsmanagements ggf. notwendige Verbesserungen implementieren. Das

Auftreten, die Inzidenz, die Schwere und der Verlauf von Organfunktionsstörungen wurden als die wesentlichen Verlaufsparameter und Tracer bzw. Indikatoren für den Qualitätsvergleich gewählt. Vorteile, die aus der Teilnahme an einem solchen Register für die beteiligte Intensivstation entstehen, können der Gewinn solider, unabhängiger Erkenntnisse über die eigene Qualität intensivmedizinischer Behandlung in Bezug auf Sterblichkeit und Multiorganversagen (Effektivität), der Gewinn von Informationen über die Effizienz der eigenen intensivmedizinischen Behandlung durch den Vergleich des Aufwandes und der damit verbundenen Kosten mit dem Behandlungsablauf und dessen Ergebnis, der Gewinn von Daten, die als Argumentationshilfe in den Verhandlungen mit Krankenhausverwaltungen, Kostenträgern und anderen verwendet werden können, die Erfüllung gesetzlicher Vorgaben, die Erfüllung von Voraussetzungen für Zertifizierungsverfahren und die Anwendung für Aspekte der Öffentlichkeitsarbeit sein [90].

### **1.8 Thema**

Ziel dieser Studie ist die spezifische Anpassung und Modifizierung des SAPS II an die Zwecke und Ziele der Sektion Qualitätssicherung der DIVI.

Beim SAPS II handelt es sich um einen Score, der sich aus physiologischen Parametern und Variablen wie Alter, Zuweisung, Vorerkrankungen sowie Glasgow Coma Scale (GCS) zusammensetzt, wobei die GCS aufgrund der hohen Beobachtersvariation bei sedierten und beatmeten Patienten in dieser Studie nicht ermittelt wurde und somit unberücksichtigt gelassen wurde. Zudem bezieht sich die mit dem originalen SAPS-II-Score ermittelte Prognose auf das Krankenhaus-Outcome, die mit unserem SAPS-II-Score berechnete Vorhersage hingegen auf das Intensiv-Outcome.

Die Variablen werden innerhalb der ersten 24 Stunden nach Aufnahme auf die Intensivstation bestimmt und die am weitesten von der Norm abweichenden Werte berücksichtigt. Mithilfe dieser Parameter kann man die Mortalität berechnen.

Da der SAPS II im Jahre 1993 publiziert wurde und mit der Zeit durch die Anwendung neuerer diagnostischer und therapeutischer Maßnahmen veraltet ist, spiegelt er nicht mehr die aktuellen Gegebenheiten wider. Aus diesem Grund wurde diese Studie durchgeführt, um den SAPS-II-Score an die heutige Situation anzupassen. Dazu wurde die DIVI-Datenbank verwendet und neue Variablen ergänzt sowie vorhandene Parameter angepasst.

Das Ziel ist somit durch das Hinzufügen von weiteren Parametern zum SAPS II die Spezifität und Sensitivität des SAPS II zu verbessern und damit die Vorhersagekraft der Letalität anhand der DIVI-Datenbank zu optimieren. Einerseits wurden weitere Variablen zum

originalen SAPS-II-Score hinzugefügt, andererseits zu einem SAPS-II-Score, der nur physiologische Parameter enthält. Bei beiden Varianten wurde die GCS unberücksichtigt gelassen.

## 2 Material und Methode

### 2.1 Allgemeine Beschreibung

Die Studie konzentriert sich auf die von der DIVI erhobenen SAPS-II-Daten. Seit 2000 wurden die Daten diesbezüglich bei 94838 Personen von 57 verschiedenen Kliniken bundesweit erhoben. Jedoch sind 37277 Personen aus der Studie ausgeschieden, weil bei diesen die Datenerhebung unvollständig war, sie jünger als 18 Jahre oder 100 Jahre und älter waren, nicht länger als 18 Stunden auf der Intensivstation verbracht haben oder in ein anderes Krankenhaus verlegt wurden. Somit sind letztendlich 57561 Patienten in die Studie aufgenommen worden. Diese wurden in zwei Kohorten unterteilt, eine Testgruppe (66 %) und eine Validierungsgruppe (34 %), wobei die Gruppenzuteilung nach dem Zufallsprinzip durchgeführt wurde.

Die Auswertung der Studie wurde mit SPSS 16.0 durchgeführt.

Auf der einen Seite wurden zum SAPS II erneut die Variablen Alter, Vorerkrankungen und Zuweisung sowie neue Parameter hinzugefügt, andererseits wurden nur die physiologischen Parameter im SAPS II berücksichtigt und daneben die Variablen Alter, Vorerkrankungen und Zuweisung sowie neue Parameter betrachtet. Zu den neuen Variablen gehören Aufnahme von Extern, Geschlecht, Trauma, Fachbereiche sowie Neuro.

Die normalerweise im originalen SAPS-II-Score enthaltene GCS wurde auf Grund der hohen Beobachtersvariation bei sedierten und beatmeten Patienten nicht ermittelt und somit unberücksichtigt gelassen. Stattdessen wurde die Variable neurologische bzw. neurochirurgische Patienten verwendet, um den neurologischen Status einzubeziehen.

### 2.2 Verwendete Scoresysteme

Zu den in der DIVI verwendeten Scoresystemen zählen unter anderem SAPS II, SOFA sowie der TISS-28 Score.

SOFA steht für Sepsis-related Organ Failure Assessment. Er ist ein von Experten der European Society for Intensive Care Medicine (ESICM) festgelegter Konsens zur objektiven Beschreibung der Organ(-dys)funktion – nicht nur bei Sepsis. Grundlage der täglichen Erhebung (24-Stunden-Zeiträume) sind die jeweils schlechtesten Werte für jedes Organsystem. Zu diesen gehören Lunge, Niere, Leber, Herz/Kreislauf, Blut und das zentrale Nervensystem.

Beim TISS-28 hingegen handelt es sich um eine Revision des bekannten TISS (Therapeutic Intervention Scoring System) von Cullen und Keene. Er erfasst über therapeutische,

diagnostische und pflegerische Maßnahmen den Zustand des Patienten. Die im TISS enthaltenen 28 Maßnahmen werden auch als „Items“ bezeichnet. Eigentlich sind es nur 23 verschiedene Items, von denen fünf in ihrer Intensität zusätzlich abgestuft sind (z.B. Verbandswechsel: keiner / Routine / häufig). Jedes Item wird mit einem Punktwert versehen, und die Summe dieser Punkte ergibt den eigentlichen TISS-28-Wert. Die TISS-Erhebung erfolgt täglich, wobei jeweils der Zeitraum von 0.00 bis 24.00 Uhr zugrunde gelegt wird. Jede Maßnahme, die auch nur an Teilen des ganzen Tages geleistet wurde, ist anzukreuzen. Schwerpunkt dieser Studie ist der SAPS-II-Score. Er wird auf den darauf folgenden Seiten genauer erläutert.



## 2.3 SAPS II

### 2.3.1 Hintergrund

SAPS II wurde von Le Gall et al. im Jahre 1993 eingeführt, wobei der SAPS-II-Score in einer multizentrischen Studie an über 13000 Patienten unterschiedlicher Intensivstationen entwickelt und validiert wurde. Er wurde von einer europäischen/nordamerikanischen Studie errichtet, an der Patienten von medizinischen, chirurgischen und gemischten Intensivstationen aus zehn europäischen und zwei nordamerikanischen Ländern teilnahmen. Er dient zur Einschätzung der Erkrankungsschwere von Intensivpatienten auf der Basis physiologischer Daten sowie zur Berechnung des Letalitätsrisikos und eignet sich gut zur Klassifizierung von Patientengruppen für klinische Studien, epidemiologische Untersuchungen oder für qualitätssichernde Maßnahmen [1].

Das besondere Feature des SAPS II, bei dem es sich um ein krankheitsübergreifendes Klassifikationssystem handelt, liegt darin, dass dieser Score überwiegend an europäischen Patientenkollektiven entwickelt wurde und somit europäische Gegebenheiten besser reflektieren kann [1].

### 2.3.2 Aufbau

Die primäre Intention des SAPS, der 1984 entwickelt wurde, war es das APACHE-System durch Reduktion der Variablen von 34 auf 13 zu vereinfachen [89]. 1993 wurde von Le Gall eine Formel vorgestellt, die es ohne großen Aufwand ermöglicht, jedem berechneten Scorewert ein individuelles Letalitätsrisiko zuzuordnen [1]. Folglich kann man anhand einer mathematischen Gleichung das Letalitätsrisiko aus dem SAPS-II-Score berechnen. Dieses nimmt linear zum SAPS-II-Scorewert zu.

SAPS II ist aus 17 Variablen zusammengesetzt: zwölf physiologische Variablen, das Alter, den Aufnahmemodus (geplant chirurgisch, ungeplant chirurgisch oder medizinisch) sowie drei zugrunde liegende Krankheitsvariablen (metastasiertes Tumorleiden, maligne hämatologische Erkrankung, AIDS) (s. Anhang S. 88, Tabelle 1). Zu den physiologischen Parametern gehören die Herzfrequenz, der systolische Blutdruck, die Körpertemperatur, das Verhältnis von arteriellem Sauerstoffpartialdruck zur fraktionellen inspiratorischen Sauerstoffkonzentration bei Beatmung ( $\text{PaO}_2/\text{FIO}_2$ -Ratio), das Urinvolumen pro Tag, der Serumharnstoffwert, der Leukozytenwert im Blut, der Serumkaliumwert, der Serumnatriumwert, der Serumbicarbonatwert, der Serumbilirubinwert sowie der neurologische Status, wofür die GCS angewandt wird.

In den SAPS II gehen die Scorewerte ein, die innerhalb der ersten 24 Stunden nach Intensivaufnahme am weitesten von der Norm abweichen und somit von der Punktzahl her am höchsten sind [41]. Je höher der SAPS-II-Score umso größer ist der Schweregrad der Erkrankung und demzufolge ist auch gleichzeitig die Prognose schlechter.

### **2.3.3 Ein- und Ausschlusskriterien**

Der SAPS II und die Hospitalletalität wurden von 137 medizinischen und/oder chirurgischen Intensivstationen in zwölf Ländern erhoben. Patienten, die jünger als 18 Jahre waren, Patienten mit Verbrennungen sowie kardiologische Patienten (Koronar- und herzchirurgische Patienten) wurden ausgeschlossen. Patienten mit Verbrennungen wurden ausgeschlossen, weil sie einerseits oft auf anderen Stationen als allgemeine Intensivstationen behandelt wurden und andererseits bereits spezifische prognostische Systeme für diese entwickelt worden sind. Außerdem wurden kardiologische Patienten bei der Entwicklung und Validierung des SAPS II nicht mitberücksichtigt, obwohl es einige Hinweise gibt, dass dieses System auch für diese Patienten eine gute Performance aufweist [62]. Von 14745 eingeschriebenen Patienten wurden 1593 Patienten aufgrund der Ausschlusskriterien nicht berücksichtigt. Somit blieben 13152 Patienten für die Entwicklung und Validierung des Modells übrig, wobei nach dem Zufallsprinzip 65 % der Patienten in die Entwicklungsgruppe und 35 % der Patienten in die Validierungsgruppe eingeteilt wurden. Bei Patienten, bei denen Informationen bezüglich des Aufnahmegrunds (geplant chirurgisch, ungeplant chirurgisch oder medizinisch) fehlten, konnten nicht in der Analyse mitberücksichtigt werden, weil sie nicht richtig bei der Punktevergabe kategorisiert werden konnten. Auf die gleiche Weise wurden auch Patienten, bei denen Informationen bezüglich der Beatmung zur Berechnung des Verhältnisses von arteriellem Sauerstoffpartialdruck zur fraktionellen inspiratorischen Sauerstoffkonzentration ( $\text{PaO}_2/\text{FIO}_2$ -Ratio) nicht vorhanden waren, ausgeschlossen. Dieser Ausschluss nahm 155 Fälle von der Datenbank heraus, so dass der endgültige SAPS II an 12997 Patienten entwickelt und validiert worden ist, davon 8369 im Entwicklungs- und 4628 im Validierungskollektiv [41].

### **2.3.4 Gründe für die Anwendung von SAPS II**

Die europäische/nordamerikanische Studie wurde durchgeführt, um auf der einen Seite einen neuen vereinfachten akuten physiologischen Score, den SAPS II, vorzustellen, und auf der anderen Seite eine Methode zu entwickeln, die den Score in eine Wahrscheinlichkeit der Krankenhausmortalität überführt. Logistische Regressionsanalyse [8, 45-47] wurde dafür verwendet, um das Auswählen der Variablen, die den SAPS II bilden werden, das Festlegen

von angemessenen Gruppierungen und die Punkteverteilung für jede Variable sowie das Umformen des SAPS-II-Score in eine Krankenhausletalität zu erleichtern [41].

### **2.3.5 Entwicklung der 17 Variablen**

Von den 37 Variablen, die gesammelt wurden, um den SAPS II zu bilden, wurden nur 17 Variablen zum Schluss berücksichtigt, die im SAPS-II-Score enthalten sind. Variablen wurden ausgeschlossen, wenn sie sich nicht auf die Krankenhausletalität in den bivariaten Analysen bezogen oder nicht notwendig waren, weil andere Variablen in dem multivariaten Modell enthalten waren. Die Datenerhebung enthielt demografische Informationen des Patienten, alle Variablen, die erforderlich zur Berechnung des originalen SAPS sind, eine Reihe an neuen Variablen, die möglicherweise Bestandteil des SAPS II werden und den Vitalstatus bei Krankenhausentlassung. Die neuen Variablen wurden vor Studienbeginn aufgrund von klinischen Gründen gewählt. Bei den physiologischen Variablen wurde der schlechteste Wert innerhalb der ersten 24 Stunden nach Intensivaufnahme berücksichtigt. Der schlechteste Wert war als der Wert definiert, der die höchste SAPS Punktzahl im originalen SAPS erhalten hätte. Punkte, die jeder Variable zugewiesen wurden, variierten von null bis drei (für die Temperatur) bis auf null bis 26 (für die Glasgow Coma Scale). Eine arterielle Blutprobe ist nicht notwendig, wenn der Patient nicht ventilert ist oder eine CPAP-Beatmung (Continuous Positive Airway Pressure) erhält. Für sedierte Patienten wurde die GCS vor der Sedierung berücksichtigt.

Jede der möglichen Variablen wurde unabhängig in Verbindung mit der Krankenhausmortalität beurteilt. Diese bivariaten Analysen wurden benutzt, um die unabhängigen Variablen genauer zu untersuchen, um eine kleinere Teilmenge zu ermitteln, die im Zusammenhang mit der Krankenhausletalität steht. Jede statistisch signifikante Variable wurde dann gegen den Vitalstatus bei Krankenhausentlassung aufgetragen und die LOWESS (locally weighted least squares) smoothing Funktion [13] wurde verwendet, um Intervalle für jede Variable zu bestimmen [41].

### **2.3.6 Kalibrierung**

Um die Leistung des Systems zu bewerten, wurden goodness-of-fit Tests (Hosmer-Lemeshow) [25-27, 43] sowohl bei der Entwicklungs- als auch Validierungsreihe durchgeführt, um die Kalibrierung zu beurteilen sowie die Fläche unter der ROC-Kurve bei beiden Datensätzen ermittelt, um die Diskriminierung zu bewerten. Der p-Wert des goodness-of-fit Tests für den Entwicklungsdatensatz war 0,883. Dieser p-Wert ist groß ( $>0,5$ ), daher

reflektiert das Modell sehr genau die wahre Mortalitätserfahrung im Entwicklungsdatensatz. Im Validierungsdatensatz war der p-Wert für den goodness-of-fit Test 0,104, was nahe legt, dass das Modell die Mortalität in einer Patientengruppe unabhängig von derjenigen reflektiert, für die das Modell entwickelt worden ist [41].

### 2.3.7 Diskriminierung

Die Flächen unter der ROC-Kurve [24] für den SAPS II waren 0,88 (95 % Konfidenzintervall, 0,87 bis 0,90) im Entwicklungsdatensatz und 0,86 (95 % Konfidenzintervall, 0,84 bis 0,88) im Validierungsdatensatz. Flächen dieser Größe sind im Allgemeinen anerkannt zu einer exzellenten Diskriminierung beizutragen [41].

### 2.3.8 SAPS-II-Formel

Aus dem Entwicklungsdatensatz von 8369 Patienten wurde eine Gleichung, die auf das multiple logistische Regressionsmodell basiert, entwickelt, um den SAPS-II-Score in eine Wahrscheinlichkeit der Hospitalmortalität umzuformen. Sie enthält zwei erklärende Variablen, SAPS-II-Score und  $\ln(\text{SAPS-II-Score} + 1)$  [21].

Der erste Schritt, um die Krankenhausletalität zu berechnen, ist den logit (Patientendaten) wie folgt zu berechnen, wobei  $\ln$  den natürlichen Logarithmus darstellt:

$$\text{logit} = \beta_0 + \beta_1 * (\text{SAPS-II-Score}) + \beta_2 * [\ln(\text{SAPS-II-Score} + 1)] = -7,7631 + 0,0737 * (\text{SAPS-II-Score}) + 0,9971 * [\ln(\text{SAPS-II-Score} + 1)].$$

Dieser logit ist zu einer Wahrscheinlichkeit der Hospitalmortalität mit der folgenden Gleichung konvertiert:

$\text{Pr}(y = 1/\text{logit}) = e^{\text{logit}} / (1 + e^{\text{logit}})$ , wobei Pr (Predicted Death Rate) die vorhergesagte Mortalitätswahrscheinlichkeit darstellt und  $e$  eine mathematische Konstante von 2,7182818 ist, die die Basis des natürlichen Logarithmus repräsentiert [41].

### 2.3.9 Vergleich SAPS – SAPS II

Der SAPS II (17 Variablen) wurde mit dem alten SAPS (14 Variablen) in der Validierungsgruppe für jeden Patienten verglichen. Diese Gruppe wurde gewählt, weil diese Patienten bei der Entwicklung des SAPS II nicht verwendet wurden. Die Ergebnisse weisen darauf hin, dass der SAPS II eine signifikante Verbesserung in jeder Hinsicht anbietet und

somit bedeutende Modifikationen vorgenommen worden sind, obwohl die meisten Variablen dieselben sind. Außerdem wurden die Fläche unter der ROC-Kurve für den alten SAPS und SAPS II sowie die Korrelation zwischen beiden Scores berechnet. Während die Fläche unter der ROC-Kurve 0,80 für den alten SAPS war, war sie beim SAPS II 0,86. Weiterhin war der Korrelationskoeffizient zwischen dem alten und neuen SAPS 0,79, was nahe legt, dass nur 62 % der Variabilität im SAPS II durch den alten SAPS erklärt werden kann. Dies deutet darauf hin, dass der SAPS II eine signifikante Verbesserung über das originale SAPS System hinaus darstellt [41].

### **2.3.10 Vorteile des SAPS II**

Das Sammeln der Daten, die zur Berechnung des SAPS-II-Score notwendig sind, ist sehr einfach und schnell. Es dauert weniger als fünf Minuten pro Patient. Alle Variablen des SAPS II sind schnell verfügbar und erfordern weder spezielle venöse noch arterielle Blutentnahmen, wobei alle Parameter genau definiert sind (s. Anhang S. 88, Tabelle 2).

Eine der Ziele dieses Prognosemodells war es ein pures Physiologie basiertes System zu behalten. Jedoch wurden die Kriterien für die Kalibrierung und Diskriminierung durch das Einschließen der chronischen Vorerkrankungen beträchtlich verbessert. Wie mit einem beliebigen System, das auf klinische Messungen basiert, stellen fehlende Werte ein Problem dar [50]. Dieses Problem ist im Allgemeinen durch die Annahme gelöst worden, dass Werte, die bei der medizinischen Aufnahme nicht protokolliert bzw. erhoben wurden, innerhalb des Normbereichs liegen. Dieselben Regeln wurden auch für den SAPS II befolgt, weil beispielsweise in manchen Ländern das Serumbilirubin nicht systematisch gemessen wird. Ähnlich ist es auch bei den Blutgasen, die nicht typischerweise bei allen nicht ventilierten Patienten gemessen werden [41].

### **2.3.11 Nachteile des SAPS II**

Manche Systeme erfordern eine einzige Diagnose, um die Sterbewahrscheinlichkeit zu schätzen. Beim SAPS II wird die Sterbewahrscheinlichkeit aus dem Score durch das Anwenden einer logistischen Regressionsgleichung direkt berechnet, ohne Punkte hinzuzufügen oder irgendeine Art der Korrektur für die Diagnose durchzuführen. Diese Entscheidung wurde vor Studienbeginn getroffen, mit der Annahme, dass die Auswahl einer einzigen Diagnose zu schwierig für die meisten Intensivpatienten ist [3]. Während manche Patienten entsprechend einer spezifischen, einfachen und einzigen Diagnose kategorisiert werden können, ist dies in der Regel nicht der Fall. In der Tat war es in der Studie nur bei

37% der Patienten möglich sie nur in eine Diagnosekategorie zuzuordnen. Wenn Intensivpatienten mehrere Diagnosen haben, ist es oft problematisch die Wichtigste zu wählen. Andere Systeme verlangen, dass die wichtigste Diagnose festgelegt wird, wodurch sich das Letalitätsrisiko als Ergebnis demzufolge, entsprechend der gewählten Kategorie, verändert.

Wahrscheinlichkeiten machen nur dann Sinn, wenn sie als gesamtes Risikomaß verwendet werden. Sie können nicht genau sagen, welche Individuen im Krankenhaus versterben werden, nur dass ein Patientenanteil mit derselben Wahrscheinlichkeit voraussichtlich sterben wird [49]. Der Wert des SAPS-II-Score und die Wahrscheinlichkeit sind nicht primär dazu da das Letalitätsrisiko für einen einzelnen bzw. bestimmten Patienten zu schätzen. Dies wäre nur in der Theorie interessant, weil das beobachtete Outcome für einen individuellen Patienten nur null (überlebt) oder eins (verstorben) sein kann. Während es vielleicht nützlich für Familienangehörige sein könnte zu wissen, wie hoch die geschätzte Sterbewahrscheinlichkeit ist, wäre es selten angemessen solche Informationen zu benutzen, um Behandlungsentscheidungen zu beeinflussen [41].

### **2.3.12 Nutzung in unterschiedlichen Ländern**

Es gibt große Unterschiede in der Nutzung eines gemeinsamen Score Systems, wenn man es für Patienten in verschiedenen Ländern anwendet [58, 66, 87]. Es wird vermutet, dass Faktoren, die sich nicht direkt auf den Erkrankungsschweregrad beziehen und auch nicht im SAPS II berücksichtigt werden, die prognostische Leistung beeinflussen [63]. Dieses würde dafür sprechen, dass die Leistung auch von der Zusammensetzung der Bevölkerung abhängt sowie, dass sich die Unterschiede der beobachteten zu den erwarteten Verstorbenen eher auf die ungemessenen Unterschiede als auf die klinische Leistung der Intensivstationen bezieht. Denn in Europa gibt es große Unterschiede unter den Intensivstationen, sei es in strukturellen oder organisatorischen Aspekten, was den SAPS-II-Score beeinflusst. Deshalb sollte man mit großer Vorsicht und Sorgfalt die klinische Leistung von Intensivstationen verschiedener Länder bewerten. Viele sind der Ansicht, dass die Leistung von diesen Vorhersagemodellen, vor allem wenn man sie in verschiedenen Bevölkerungen anwendet, stark variieren kann [63, 66].

### **2.3.13 Vorhersagekraft für die Behandlungsdauer von Intensivpatienten**

Bei kritisch kranken Patienten tragen die Dauer der Intensivtherapie und Verweildauer im Krankenhaus zu den Krankenhauskosten bei [81]. Neben der Erkrankungsschwere und dem

dadurch bedingten diagnostischen und therapeutischen Aufwand [14] wird der Kostenaufwand auch von der Krankenhausverweildauer beeinflusst.

Für Verstorbene ist die Behandlungsdauer auf der Intensivstation signifikant länger, wohingegen die Gesamtverweildauer im Krankenhaus signifikant kürzer ist. Jedoch hat man in Studien festgestellt, dass die Gesamtkrankenhauskosten der verstorbenen Intensivpatienten nur geringfügig und nicht signifikant über denen der Überlebenden sind. SAPS II ist zur Outcome-Evaluation geeignet [29, 31, 71, 92] und scheint ebenfalls zur Beurteilung der Effizienz von Intensivstationen in der Lage zu sein [10, 71].

#### **2.3.14 SAPS II bei Trauma-Patienten**

Mehrere Score Systeme wurden zur Evaluation von Trauma-Patienten entwickelt. Dazu gehören Trauma and Injury Severity Score (TRISS) [12], Injury Severity Score (ISS) [4] sowie Revised Trauma Score (RTS) [11].

SAPS II ist jedoch nicht in der Lage die Erkrankungsschwere bei Trauma-Patienten zu beschreiben und ist deswegen für diese ungeeignet [49]. Er ist ein exzellentes System um den Erkrankungsschweregrad bei allgemeinen Intensivstation-Patienten abzuschätzen, jedoch nicht bei Patienten mit größerem Trauma. Er weist zwar eine zufriedenstellende diskriminierende Leistung bei Trauma-Patienten auf, jedoch ist die Kalibrierung bei den meisten Trauma-Patienten unzureichend [60, 73], insbesondere bei denjenigen, die operiert wurden und bei Patienten, die unter 60 Jahre alt sind, denn er überschätzt bei diesen die Mortalität. Infolgedessen starben weniger Patienten als vorhergesagt wurde. Dagegen wird die Mortalität bei Trauma-Patienten, die nicht operiert und somit konservativ behandelt wurden und bei Patienten, die älter als 60 Jahre sind, nicht überschätzt, wobei die physiologischen Merkmale im Vordergrund stehen und nicht das Trauma selbst. Dieses liefert die Tatsache, dass bei älteren Patienten die physiologischen Merkmale und nicht das Trauma selbst das Ergebnis entscheidend beeinflussen [73].

Die Ergänzung des SAPS II zum TRISS hingegen verbessert die Risikoadjustierung und somit Beurteilung kritisch kranker Trauma-Patienten [73].

#### **2.3.15 SAPS II bei herzchirurgischen Patienten**

Es gibt auch Studien, die Patienten mit Myokardinfarkt sowie herzchirurgische Patienten berücksichtigt haben, obwohl diese eigentlich nicht für das originale SAPS-II-Modell geeignet sind, bei denen die Leistung des SAPS II aber hervorragend war [75, 79]. Bei der Analyse durch den Ausschluss dieser Patientengruppe hat man keine Verbesserung der

Kalibrierung festgestellt. Im Gegenteil, in der Subgruppen-Analyse lieferten die Herzpatienten wesentlich bessere Ergebnisse in der Kalibrierung und Diskriminierung als viele andere Patientengruppen. Somit besitzt der SAPS II auch die Fähigkeit Patienten mit kardiovaskulären Erkrankungen genauso gut zu bewerten und einzuschätzen [57].

### **2.3.16 SAPS II - angepasster SAPS II – erweiterter SAPS II**

Man hat eine retrospektive Studie durchgeführt, welche die Kalibrierung und Diskriminierung des originalen SAPS II, eines angepassten SAPS II sowie eines erweiterten SAPS II ermittelt hat. Bei dem erweiterten SAPS II wurden sechs Variablen hinzugefügt (s. Anhang S. 90, Tabelle 3), die routinemäßig auf den Intensivstationen gesammelt werden ohne den Arbeitsaufwand zu erhöhen.

Es handelt sich dabei um Parameter, die leicht zu bestimmen sind und genauso wie die anderen Variablen innerhalb der ersten 24h nach Intensivaufnahme gesammelt werden. Zu diesen gehören unter anderem das Alter, das Geschlecht, die Krankenhausverweildauer vor der Intensivaufnahme, der Patienten-Standort vor Intensivaufnahme (Notfallaufnahme, auf Station im selben Krankenhaus, anderes Krankenhaus), die klinische Kategorie (medizinischer Patient oder anderer) sowie ob eine Intoxikation vorliegt.

Weiterhin wurden keine Diagnosen berücksichtigt, weil auf der einen Seite die meisten Intensivpatienten mehrere Diagnosen hatten und auf der anderen Seite man ein angemessenes Modell entwickeln wollte, das zur Evaluierung der Performance von Intensivstationen für Patienten mit spezifischen Diagnosen dient.

Die beiden Prognosemodelle sind vom originalen SAPS II hergeleitet. Der angepasste SAPS II ist durch das folgende logit charakterisiert:

$$\text{logit} = -8,1834 + 0,0467 * \text{SAPS II} + 1,3287 * \log(\text{SAPS II} + 1)$$

Das logit des erweiterten SAPS-II-Modells ist wie folgt:

$$\text{logit} = -14,4761 + 0,0844 * (\text{erweiterten SAPS II}) + 6,6158 * \log(\text{erweiterten SAPS II} + 1)$$

Alle drei SAPS-II-Modelle haben ordentliche zufriedenstellende Flächen unter der ROC-Kurve (0,858-0,879) geliefert. Trotz allem war die Diskriminierung mit dem erweiterten SAPS II signifikant besser als mit den anderen beiden Modellen. Es gab jedoch unter den drei Systemen größere Unterschiede bei der Kalibrierung (s. Anhang S. 91, Tabelle 4).



Das originale SAPS-II-Modell überschätzte deutlich die Mortalität und wies eine unzureichende einheitliche Anpassung auf. Die Anpassung des SAPS II verbesserte zwar die Kalibrierung ( $p=0,78$ ), aber nicht die einheitliche Anpassung. Eine gute einheitliche Anpassung wurde durch den erweiterten SAPS II erreicht. Mit diesem Modell waren auch die SMR-Werte für die Patientensubgruppen nicht signifikant unterschiedlicher als eins [39].

### **2.3.17 Zusammenfassung – SAPS II**

Der SAPS II, basierend auf eine große Stichprobe an Patienten, liefert eine Vorhersage der Sterbewahrscheinlichkeit ohne eine Aufnahmediagnose zu bestimmen, was die Evaluation der Leistung von Intensivstationen ermöglicht [41].

Die Anwendung des SAPS II als Prognosesystem zur Ermittlung der Hospitalletalität für internistische [80] und chirurgische [1] Intensivpatienten wurde in Studien bestätigt. Es eignet sich somit zur Klassifikation der Erkrankungsschwere bei Intensivpatienten sowie zur Klassifizierung von Patientengruppen für klinische Studien, epidemiologische Untersuchungen oder qualitätssichernde Maßnahmen [1].

Im Laufe der Zeit veralten Vorhersagemodelle wie der SAPS II [18], so dass sie an die aktuelle Intensivbevölkerung angepasst werden müssen [60, 63]. Dieses liegt unter anderem daran, dass der SAPS II dann nicht mehr gut kalibriert [57]. Er muss erneut kalibriert werden, um die Verwendbarkeit für die Qualitätsversorgung auf Intensivstationen zu optimieren [9].

Um die Leistung von Score Systemen zu verbessern gibt es zwei Methoden, die Anpassung eines Modells [42, 60, 63] sowie das Hinzufügen von weiteren Variablen zum Score System [33].

### **2.3.18 SAPS 3**

Der SAPS 3 stellt eine Weiterentwicklung des SAPS-II-Score dar.

In der SAPS 3 Studie wurden Daten prospektiv über Risikofaktoren und Outcomes in einer heterogenen Kohorte von Intensivpatienten bzw. internationalen multizentrischen Kohorte von kritisch kranken Patienten gesammelt. Das Ziel des SAPS 3 ist es das SAPS-II-Modell zu aktualisieren und somit ein neues, verbessertes Prognosemodell zur Risikoeinschätzung bei kritisch kranken Patienten zu entwickeln. Er soll die Erkrankungsschwere abschätzen und den Vitalstatus bei Krankenhausentlassung vorhersagen, was auf die Daten der Intensivaufnahme basiert [59].

Der Grund warum man diese Studie durchgeführt hat, lag unter anderem daran, dass die Ausgangslinie, die von SAPS II verwendet wurde, von einer Datenbank hergeleitet wurde, die

in den frühen 90ern erhoben wurde. Seit dieser Zeit gab es Veränderungen in der Prävalenz von größeren Krankheiten und in der Verfügbarkeit und Anwendung größerer diagnostischer und therapeutischer Methoden, die mit einer Verlagerung der unzureichenden Kalibrierung älterer Modelle wie SAPS II verbunden waren. Außerdem wurde SAPS II von einer Datenbank entwickelt, die ausschließlich von Patienten aus Europa und Nordamerika aufgebaut worden war. Dieses Kollektiv kann nicht repräsentativ für die Patientenvielfalt und ärztliche Praxis sein, die die Realität der Intensivmedizin in der restlichen Welt darstellt, wo strukturelle und organisatorische Unterschiede wahrscheinlich Einfluss auf das Outcome haben [5]. Weiterhin sollte der Anwender in der Lage sein zwischen den Datenbanken zu wählen, d.h. eine globale Datenbank, welche einen größeren Vergleich der örtlichen Gegebenheiten liefert und eine regionale Datenbank, welche einen besseren Vergleich mit den Intensivstationen in geografischer Nähe, aber zu Lasten abnehmender Vergleichbarkeit mit Intensivstationen in anderen Teilen der Welt liefert. Darüber hinaus ist man durch den Einsatz von Computern in der Lage ein neues Vorhersagemodell zu entwickeln, was auf eine digitale Datenanschaffung und -analyse mit nur geringen Unterschieden bei den Definitions- und Anwendungskriterien basiert [67].

Auf statistischer Ebene gibt es einen Bedarf nach einer Änderung, um die hierarchische Beschaffenheit der Daten zu berücksichtigen [19-20]. Derzeitig gebräuchliche allgemeine Vorhersagemodelle beachten nicht die Existenz bzw. das Vorhandensein von klinischen und nichtklinischen Faktoren auf dem Level der Intensivstation, welche eine wichtige Auswirkung auf die Prognose haben können. Stattdessen nehmen sie an, dass diese Faktoren entweder nicht wichtig oder zufällig bei großen Patientenkollektiven aufgeteilt sind und die Variation zwischen den Intensivstationen klein ist. Wenn die Variation zwischen den Intensivstationen jedoch nicht vernachlässigbar ist, wird es die Beständigkeit der Gleichungen gefährden, die zur Berechnung der Vorhersagemortalität dienen. Weiterhin betrachten die veröffentlichten Modelle die Beziehung zwischen Leistung und Erkrankungsschwere als konstant und das mag nicht zutreffen, da die Performance innerhalb von Intensivstationen entsprechend des Erkrankungsschweregrades der Patienten variieren kann [72, 83]. Um dieses Problem zu bewältigen hat man sich entschieden eine neue Strategie für die Entwicklung des SAPS 3 einzuschlagen und statistische Modelltechniken anzuwenden, die die Clusterbildung von Patienten innerhalb von Intensivstationen kontrollieren, anstatt die Unabhängigkeit von Beobachtungen anzunehmen [67].

## 2.4 Gütekriterien von Scores

### 2.4.1 Tests zur Überprüfung der Leistung von Scores

Die Leistung eines Modells muss in einem Patientenkollektiv demonstriert werden, welches unabhängig von dem ist, das zur Entwicklung des Systems verwendet wurde. Validierungsgruppen wurden entweder durch das Erheben der Daten bei einer neuen Patientenkohorte oder durch das zufällige Aufteilen einer vorhandenen Datenbank in zwei Teile zusammengestellt – eins zur Entwicklung des Modells und das andere zur Validierung [37, 44].

Für die Anwendung von Scores, die sich auf den generellen Erkrankungsschweregrad beziehen, müssen deren Leistung zunächst durch Tests wie Diskriminierung und Kalibrierung nachgewiesen werden [42]. Während sich die Diskriminierung auf die Fähigkeit bezieht zwischen Überlebenden und Versterbenden zu unterscheiden, konzentriert sich die Kalibrierung auf die Genauigkeit der Vorhersage der Mortalität, wenn man die Zahl an vorhergesagten und beobachteten Verstorbenen vergleicht [55, 57]. Dieses ist sehr wichtig, denn eine falsche Kalibrierung kann zu falschen Schlussfolgerungen führen. Außerdem ist es sinnvoll, eine Anpassung von Scores durchzuführen, die den Erkrankungsschweregrad angeben, wenn die Kalibrierung des Score Systems unzureichend ist, insbesondere für die Anwendung bei spezifischen Patientengruppen [42, 63-64, 75].

### 2.4.2 Bewertung von Scores

Wie „gut“ ein Score ist oder ob er in einer klinischen Situation geeignet ist, wird anhand der Eigenschaften des Scores als Messinstrument (Reliabilität, Validität, Güte) und an der Anwendbarkeit in der betrachteten Situation (Messbarkeit, Anwendbarkeit) geprüft.

### 2.4.3 Reliabilität

Reliabilität beschreibt die Genauigkeit mit der eine Scoremessung durchgeführt werden kann, d.h. wie genau ein Score wiedergegeben werden kann. Somit geht es bei der Reliabilität darum, ob der Score das, was er misst, genau und verlässlich misst. Reliabilität ist darauf fokussiert, wie etwas gemessen wird und nicht was die Messung bedeutet. Ein Score kann sehr zuverlässig sein ohne irgendeine Bedeutung zu haben. Kriterien für eine gute Reliabilität sind folgende: gut definierte Punkte, klare und eindeutige Auswahl an Messwerten sowie festgelegte Anweisungen im Falle von fehlenden Parametern. Um die Reliabilität von einem Messinstrument zu beweisen, kann man Test-Retest-Untersuchungen durchführen (Intra-

Rater-Reliability) oder die Anwendung des Scores durch verschiedene Personen vergleichen (Inter-Rater-Reliability) [69-70].

#### **2.4.4 Validität**

Ein Score ist valide, wenn er tatsächlich das misst, was er beabsichtigt zu messen. Ein Beweis für Validität ist die sogenannte „Face Validity“, was bedeutet, dass die verwendeten Parameter offensichtlich mit dem Ziel des Scores (Prognose, Schweregrad, Therapieaufwand) korrespondieren. Eine mehr formale Methode, um die Gültigkeit von Schweregrad-Scores zu überprüfen, ist es Untergruppen mit steigenden Scoreparametern zu bilden. Diese sollte im Durchschnitt einen linearen Anstieg der Morbidität und Mortalität zeigen. Andererseits sollten Patienten mit unterschiedlichem klinischen Outcome (z.B. Überlebende und Verstorbene) entsprechende Unterschiede bei den initialen Werten des Schweregrad-Scores aufweisen [69-70].

#### **2.4.5 Messbarkeit**

Die gute Messbarkeit von einem Score ist von der Verfügbarkeit der notwendigen Parameter abhängig und von der Zeit, die erforderlich ist den Score zu ermitteln. Umso komplexer Variablen in einem Score sind, und aufwendige Untersuchungen wie spezielle Labortests oder Berechnungen notwendig sind, umso mehr verschlechtert sich die Messbarkeit [69-70].

#### **2.4.6 Güte**

Die Güte von prognostizierenden Scoresystemen kann man vor allem daran erkennen, ob sie in der Lage sind zwischen Überlebenden und Versterbenden gut zu differenzieren (Diskriminierung), und zwar in allen Risikobereichen (Kalibrierung). Andererseits sollte auch die Gesamtprognose gut mit dem tatsächlich beobachteten Outcome übereinstimmen (Präzision) [69-70].

#### **2.4.7 Anwendbarkeit**

Bei der Nutzung eines Scores ist es wichtig zu wissen, ob die betreffende Erkrankung oder die betrachtete Patientengruppe mit dem übereinstimmt, welches die Entwickler des Scores betrachtet hatten.

Wenn die Patienten oder das Krankheitsbild nicht ähnlich sind, müssen Validierungsuntersuchen der Anwendung des Scores vorausgehen. Wenn die Anwendung eines Scores in einer bestimmten Patientengruppe oder bei einem gewissen Krankheitsbild in

der Literatur nicht berichtet wurde, muss eine eigene Validitätsstudie durchgeführt werden, wodurch möglicherweise die Aussagekraft des Scores eingeschränkt wird [69-70].

#### **2.4.8 Klinische Relevanz**

Bei der Auswahl eines Scores zur Patientenbeschreibung oder als Zielgröße in klinischen Studien ist es bedeutsam, dass die Ergebnisse klinisch gut interpretierbar und eventuell beobachtete Unterschiede klinisch relevant sind. Ein Scorewert ist nur dann verwendbar und verständlich für jemanden, der sich intensiv mit diesem Score auseinandergesetzt hat. Deshalb sollte ein im Allgemeinen akzeptierter und häufig angewendeter Score einer Neuentwicklung vorgezogen werden, weil er somit Vergleiche zu anderen Untersuchungen ermöglicht [69-70].

#### **2.4.9 Diskriminierung**

Die Qualität eines Score-Systems lässt sich mit mehreren Kriterien prüfen. Dazu gehören die Diskriminierung, Präzision sowie Kalibrierung.

Unter Diskriminierung versteht man die Fähigkeit eines Scores zwischen bestimmten Patientengruppen unterscheiden zu können, so z.B. zwischen Überlebenden und Versterbenden.

Dazu zählt unter anderem die mittlere Differenz im Scorewert zwischen Überlebenden und Verstorbenen. Je größer der Unterschied zwischen Überlebenden und Verstorbenen desto besser ist die Diskriminierungsfähigkeit eines prognostischen Scores.

Ein weiteres wichtiges Score-Kriterium ist die Richtigkeit der Score-basierten Prognose. Dazu muss der Scorewert in eine ja/nein-Aussage transformiert werden, was mit Hilfe eines Cut-off-Punktes erreicht wird. Bei Score-Systemen, deren Wert direkt als Überlebens- oder Versterbewahrscheinlichkeit aufgefasst werden kann, liegt der Cut-off-Punkt in der Regel bei 0,5 oder 50 %, wobei jedoch auch jeder andere Grenzwert ebenso möglich ist. Anhand des tatsächlich eingetretenen Outcome kann man die Richtigkeit der Vorhersage ermitteln.

Im Hinblick auf die Vorhersage der Letalität eines Scores sind folgende Definitionen von Bedeutung:

Die Sensitivität gibt den Anteil der vorhergesagten Verstorbenen (Richtigkeit der Prognose) bezogen auf alle tatsächlich verstorbenen Patienten an, d.h. ob der Score in der Lage ist die verstorbenen Patienten zu erfassen.

Die Spezifität dagegen gibt den Anteil der vorhergesagten Überlebenden (Richtigkeit der Prognose) bezogen auf alle tatsächlich überlebenden Patienten an, d.h. ob der Score in der

Lage ist die überlebenden Patienten zu registrieren und möglichst wenige Überlebende fälschlicherweise als „versterbend“ zu prognostizieren.

Der positive Vorhersagewert gibt den Anteil der als richtig vorhergesagten Verstorbenen (Richtigkeit der Prognose) bezogen auf alle vom Score als „versterbend“ Vorausgesagten an, d.h. wie oft ist die vom Score vorhergesagte Mortalitätsprognose zutreffend.

Der negative Vorhersagewert gibt hingegen den Anteil der als richtig vorhergesagten Überlebenden (Richtigkeit der Prognose) bezogen auf alle vom Score als „überlebend“ prognostizierten Patienten an, d.h. wie oft ist eine Score-basierte Prognose des Überlebens richtig.

Der Cut-off-Wert beeinflusst die Sensitivität und Spezifität. So verschlechtert sich bei hohen Cut-off-Werten die Sensitivität zugunsten der Spezifität, d.h. es werden immer weniger tatsächlich Verstorbene erfasst, wohingegen immer mehr Überlebende (unterhalb des Cut-off-Punktes) ermittelt werden. Bei niedrigen Cut-off-Werten ist die Tendenz gegenläufig.

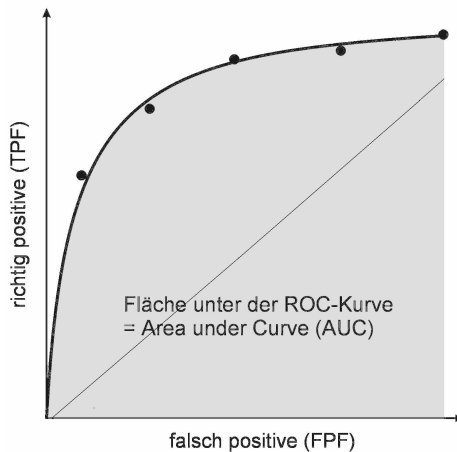
Wenn der Cut-off-Punkt beispielsweise bei 0,35 liegt, dann würde vorhergesagt werden, dass alle Patienten deren Mortalitätswahrscheinlichkeit 0,35 oder größer ist, versterben, wohingegen bei Patienten deren Wahrscheinlichkeit kleiner als 0,35 ist die Prognose „Überleben“ bedeuten würde.

Die sogenannte ROC-Kurve (Receiver Operating Characteristic Curve) [23-24, 93] erhält man, indem man für jeden Cut-off-Punkt Sensitivität und Spezifität in ein Diagramm einträgt und diese miteinander verbindet. Während die Abszisse den Anteil der Falsch-Positiven darstellt, repräsentiert die Ordinate den Anteil der Richtig-Positiven (Sensitivität). Je höher die richtig-positive Rate im Verhältnis zur falsch-positiven Rate ist, umso größer ist die Fläche unter der ROC-Kurve (s. Bild 1).

Der Punkt links unten im Diagramm entspricht dem kleinstmöglichen Cut-off-Wert, der alle Patienten als „versterbend“ prognostiziert (100 % sensitiv, 0 % spezifisch), der Punkt rechts oben im Diagramm hingegen entspricht dem größten Cut-off-Wert, d.h. alle Patienten werden als „überlebend“ bezeichnet (0 % sensitiv, 100 % spezifisch). Der Verlauf der Kurve zwischen diesen Extremen stellt zusammenfassend die prognostische Qualität eines Scores dar. Ein Score diskriminiert umso stärker, je weiter sich seine ROC-Kurve in die obere linke Ecke bewegt (hohe Sensitivität und hohe Spezifität). Es resultiert im günstigen Fall eine gekrümmte, parabelartige Kurve. Ein Score ohne jegliche prognostische Information würde dagegen einen diagonalen Verlauf aufweisen. Dieses würde auf einen Zufallsprozess hindeuten.

Interessant ist auch die Fläche unter der ROC-Kurve (AUC = Area Under the Curve). Sie kann Werte zwischen 0,5 und 1 annehmen, wobei größere Werte mit einer besseren Güte einhergehen. Wenn der Wert nahe 0,50 liegt, ist die Leistung des Modells nicht besser als ein Münzwurf. Modellentwickler sind dann normalerweise nicht zufrieden außer wenn die Fläche unter der ROC-Kurve des Modells 0,70 überschreitet.

**Bild 1** ROC-Kurve



#### 2.4.10 Präzision

Bei der Präzision handelt es sich um ein Kriterium, das die tatsächlich beobachtete Mortalitätsrate mit der vorhergesagten Mortalitätsrate (Mittelwert der Einzelprognosen) vergleicht und somit die Übereinstimmung der Prognose eines Scores mit dem tatsächlichen Outcome beschreibt. Es ist ein Maß für die Übereinstimmung zwischen unabhängigen Messergebnissen unter festen Bedingungen. Liegen also mehrere Messwerte dicht beieinander, so hat die Messmethode eine hohe Präzision.

Ein Score, der jedem einzelnen Patienten die gleiche, durchschnittlich zu erwartende Gesamtmortalität prognostiziert, wäre sehr präzise, da sich im Mittel genau die Gesamtmortalität als Prognose ergibt, würde aber nicht zwischen Überlebenden und Verstorbenen differenzieren.

Häufig beschreibt man das Verhältnis von beobachteter zu erwarteter Mortalität auch mit Hilfe eines Quotienten, der Standardisierten Mortalitätsrate (SMR) [38, 74]. Die Division der tatsächlichen Letalitätsrate durch die erwartete Rate ergibt die SMR. Diese ermöglicht es verschiedene Fachabteilungen mit unterschiedlich schwer erkrankten Patienten untereinander zu vergleichen. Wenn  $SMR = 1$  ist, so entspricht die Beobachtung genau der Prognose, ist sie dagegen größer als 1, sind mehr Patienten verstorben als erwartet, und bei Werten unter 1 liegt die Prognose höher als die tatsächliche Mortalitätsrate.

### 2.4.11 Kalibrierung

Kalibrierung beurteilt den Grad der Übereinstimmung der geschätzten Mortalitätswahrscheinlichkeit, welche mit einem Modell ermittelt worden ist und der eigentlichen Mortalität. Sie kann statistisch durch die Anwendung von goodness-of-fit Tests nach Hosmer-Lemeshow [25-27, 43] evaluiert werden [25]. Wenn die beobachtete Anzahl an Versterbenden der durch das Modell vorhergesagten Anzahl weitgehend entspricht, dann kann die Kalibrierung als zufriedenstellend angesehen werden. Um die Kalibrierung zu prüfen werden Patienten entsprechend ihrer Letalität in Gruppen geordnet. Typischerweise werden zehn solcher Gruppen gebildet, wobei jede ungefähr dieselbe Anzahl an Patienten (sogenannte „Risikodezilen“) enthält. Um die vorhergesagte Anzahl an Versterbenden in einer Gruppe zu erhalten, werden die Mortalitätswahrscheinlichkeiten für alle Patienten in dieser Gruppe zusammengefasst. Goodness-of-fit Tests vergleichen die beobachtete mit der vorhergesagten Anzahl an Versterbenden sowie die beobachtete mit der vorhergesagten Anzahl an Überlebenden in jeder Patientengruppe. Der resultierende Wert kann zur Entscheidung verwendet werden, ob die kombinierte Diskrepanz zwischen beobachtetem und vorhergesagtem Outcome über alle Gruppen innerhalb der Variabilität der Kohorte liegt. Wenn große Unterschiede vorliegen, reflektiert das Modell das Outcome in dieser Patientenkohorte nicht korrekt [37].

### 2.4.12 Konfidenzintervall

Das Konfidenzintervall [2, 17, 22] ist der Unsicherheitsbereich für die Schätzung eines bestimmten, nicht bekannten Parameters. Ein 95 %-Konfidenzintervall beispielsweise enthält den gesuchten Parameter mit einer Wahrscheinlichkeit von 95 %. Aus dem Konfidenzintervall lassen sich Schlüsse bezüglich der statistischen Signifikanz ziehen [7]. Dazu muss man sich den Standardfehler, p-Wert sowie das 95 %-Konfidenzintervall anschauen:

$$\text{KI (95 \%)} = \text{Wert/Koeffizient} \pm 1,96 * \text{S.E.}$$

Wenn die Zahl „0“ nicht im 95 %-Konfidenzintervall vorkommt, dann ist der p-Wert  $< 0,05$  (5 %) und es liegt somit eine statistische Signifikanz vor.

Das 95 %-Konfidenzintervall ist derjenige Bereich in dem der wahre Messwert mit einer Wahrscheinlichkeit von 95 % liegt. Der wahre Wert ist eigentlich unbekannt und würde sich nur durch „unendlich“ viele Messungen bestimmen lassen. Der wahre Messwert wird aber



approximiert durch den sog. „Point estimate“ (=aktueller Messwert), der im Zentrum des Konfidenzintervalls liegt. Das Konfidenzintervall beschreibt die Unsicherheit über die Zuverlässigkeit der Aussage zur Effektgröße. Die Breite des Konfidenzintervalls hängt u.a. von der Zahl der in die Studie eingeschlossenen Patienten und der Standardabweichung [36] ab und wird mit zunehmender Patientenzahl und abnehmender Standardabweichung enger, d.h. die Effektgröße kann präziser geschätzt werden. Der Vorteil, die Genauigkeit von Messresultaten mit Hilfe von Konfidenzintervallen anzugeben, liegt darin, dass die Verlässlichkeit der Resultate quantifiziert werden kann. Der p-Wert, besagt lediglich, ob ein Resultat statistisch signifikant ist oder nicht, lässt aber keine Aussage über die quantitativen Unterschiede zu.

#### **2.4.13 p-Wert**

Die statistische Signifikanz (p-Wert,  $p = \text{probability}$ ) sagt aus, wie groß die Wahrscheinlichkeit ist, dass der gemessene Unterschied dem Zufall entspringt.  $p < 0.05$  bedeutet, dass die Wahrscheinlichkeit kleiner als 5 % ist, dass der gemessene Unterschied zwischen den zwei Vergleichsgruppen nur auf Zufall beruht. Die statistische Signifikanz ist von der Größe der untersuchten Population abhängig. Wenn die Population genügend groß ist, erlangen schon kleine Unterschiede zwischen zwei Gruppen statistische Signifikanz. Der p-Wert ist ein rein statistischer Ausdruck und sagt nichts aus über die klinische Relevanz des Ergebnisses. Wenn die Signifikanz von Effekten interpretiert wird, sollten p-Werte immer im Zusammenhang mit Konfidenzintervallen verwendet werden.

## 2.5 Statistik

### 2.5.1 Regressionsanalyse

Die Regressionsanalyse ist ein komplexes Verfahren, das zu den am häufigsten verwendeten statistischen Analyseverfahren zählt. Sie wertet nicht nur einzelne Variablen aus oder untersucht den einfachen Zusammenhang zwischen zwei Variablen, sondern betrachtet mehrere Variablen gleichzeitig. Ziel ist es, die Zusammenhänge zwischen einer (abhängigen) Variablen auf der einen Seite und mehreren (erklärenden) Variablen auf der anderen Seite aufzuzeigen, d.h. inwieweit kann man anhand der Werte aus den erklärenden Variablen auf den Wert der abhängigen Variablen schließen. Das Ergebnis der Regressionsanalyse ist eine Gleichung, mit der man aus den Werten der erklärenden Variablen einen (bestmöglichen) Schätzwert für die abhängige Variable berechnen kann.

Mit der Regressionsanalyse wird ein Modell untersucht, in dem eine abhängige Variable durch eine oder mehrere andere Variablen erklärt wird. Unbekannt ist hierbei jedoch zumeist, wie die einzelnen Angaben miteinander verknüpft werden müssen, um einen vernünftigen Schätzwert für die Zielgröße zu erhalten. Genau dieses soll mit Hilfe der Regressionsanalyse beantwortet werden.

Der Zusammenhang zwischen den unabhängigen Parametern auf der einen Seite und der abhängigen Variable (Zielgröße) auf der anderen Seite lässt sich auch als mathematische Formel in einer so genannten Regressionsgleichung darstellen:

$$a + b_1 * \text{unabhängige Var}_1 + \dots + b_n * \text{unabhängige Var}_n = \text{abhängige Var (Zielgröße)}$$

Die Zielgröße ergäbe sich danach aus einem konstanten Sockelbetrag ( $a$ ), der sich je nach unabhängiger Variable um einen bestimmten Faktor erhöht oder verringert. Wie stark dabei genau der Einfluss der einzelnen erklärenden Variablen ist, hängt in dieser Formel von den zunächst noch unbekanntem Koeffizienten  $b_n$  ab. Die Aufgabe der Regressionsanalyse besteht nun darin, genau diese Koeffizienten und auch den konstanten Sockelbetrag zu schätzen.

Natürlich wird der in dem Modell unterstellte Zusammenhang in Wirklichkeit niemals genau so zutreffen, wie er in der Regressionsanalyse formuliert ist. Selbst wenn ein sehr enger Zusammenhang zwischen den einzelnen erklärenden Parametern auf der einen und der Zielgröße auf der anderen Seite besteht, wird es selbstverständlich niemals möglich sein, aus den unabhängigen Variablen exakt die Zielgröße zu berechnen. Vielmehr kann man im besten Fall eine Schätzung für die Zielgröße erhalten, die hoffentlich einigermaßen zuverlässig ist

und die tatsächliche Zielgröße möglichst genau trifft. Genau darin besteht auch die zentrale Aufgabe der Regressionsanalyse, nämlich die Parameter  $a$  und  $b_n$  so zu berechnen, dass die mit der resultierenden Regressionsgleichung geschätzten Zielgrößen möglichst nahe an den tatsächlichen Zielwerten liegen.

### 2.5.2 Logistische Regression

Bei der logistischen Regression [8, 45-47] handelt es sich um ein Verfahren zur multivariaten Analyse diskreter abhängiger Variablen, was bei der Entwicklung und Anpassung von Score Systemen eine Schlüsselrolle spielt. Sie unterstellt einen nicht-linearen Zusammenhang zwischen einer nominal skalierten, dichotomen abhängigen Variablen und mindestens einer unabhängigen Variablen als Modellprämisse und dient somit der statistischen Beurteilung dieses Zusammenhangs. Die Nicht-Linearität des Zusammenhangs kann als eine Art "Sättigungs"-Effekt verstanden werden, der darin besteht, dass Änderungen der unabhängigen Variablen in den Extrembereichen noch nicht bzw. nicht mehr zu wesentlichen Änderungen der Eintrittswahrscheinlichkeit führen. Folglich wird der Effekt der unabhängigen Variablen auf die abhängige Variable untersucht. Typische Beispiele für abhängige Variablen sind Parameter, die das Eintreten eines Ereignisses erfassen. Solche Variablen haben nur zwei mögliche, sich ausschließende Ausprägungen, wie z.B. "Ereignis findet statt" ( $Y = 1$ ) und "das Ereignis findet nicht statt" ( $Y = 0$ ). Nun interessiert der Einfluss der jeweiligen unabhängigen Variablen auf diese Eintrittswahrscheinlichkeit. Weiter könnte die konkrete Wahrscheinlichkeit für das Eintreten des Ereignisses  $p$  ( $Y = 1$ ) interessieren, wenn man die Ausprägungen der unabhängigen Variablen kennt.

Die Zielgröße wird als Linearkombination der unabhängigen Variablen ( $V_1$  bis  $V_n$ ) dargestellt, wobei die Koeffizienten ( $k_1$  bis  $k_n$ ) mit dem jeweiligen Wert der Variablen multipliziert werden. Da die abhängige Variable dichotom ist, kann sie nur die Werte 0 oder 1 annehmen, je nachdem ob das Zielereignis eintritt (1) oder nicht eintritt (0). Deshalb verwendet man eine stetige Hilfsvariable ( $x$ ) (Log-Odds oder Logit), die beliebige Werte annehmen kann. Der Grund für die Benutzung einer Hilfsvariablen liegt darin, dass man nicht direkt die Zielgröße (0 oder 1) betrachtet, sondern die Wahrscheinlichkeit ( $p$ ) für das Auftreten des Zielereignisses, die Werte zwischen 0 und 1 annimmt.

$$x = k_1 * V_1 + k_2 * V_2 + \dots + k_n * V_n$$

Jedoch wird genauer gesagt nicht die Wahrscheinlichkeit sondern die Odds Ratio betrachtet.

$$e^x = p/(1-p)$$

Odds Ratios können Werte zwischen 0 und  $+\infty$  annehmen, mit dem Wert 1 bei  $p = 0,5$ . Verwendet man dagegen den Log-Odds, indem man den natürlichen Logarithmus darauf anwendet, erhält man einen Wert zwischen  $-\infty$  und  $+\infty$ , mit dem Wert 0 bei  $p = 0,5$ . Und genau dies spiegelt die Hilfsvariable wider.

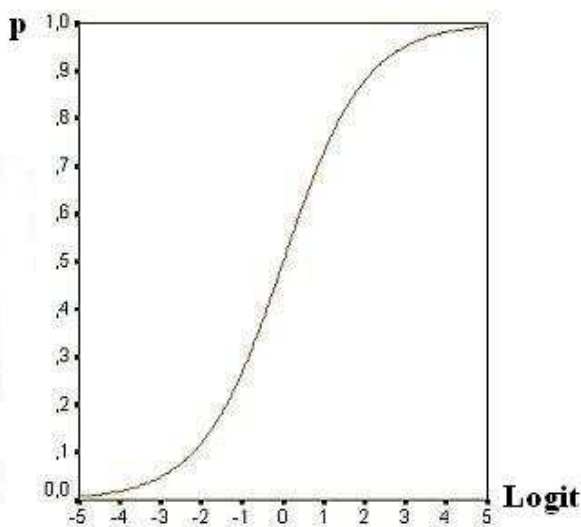
$$x = \ln(p/(1-p))$$

In der logistischen Regression wird dann die Regressionsgleichung geschätzt. Es werden Regressionsgewichte bestimmt, nach denen die geschätzten Logits für eine gegebene Matrix von unabhängigen Variablen  $X$  berechnet werden können.

$$\text{Logit} = \beta_0 + \beta_1 X_1 + \dots + \beta_n X_n$$

Die logistische Funktionskurve hat eine geschwungene, S-förmige Gestalt und verläuft punktsymmetrisch, wobei die Symmetrie auf den jeweiligen Wendepunkt der Kurve bezogen ist (s. Bild 2). Dieser Wendepunkt liegt immer bei  $p = 0,5$ . Abgetragen ist hier jeweils die Wahrscheinlichkeit des Eintretens der Ausprägung ( $Y = 1$ ) der abhängigen Variable  $Y$  (also  $p$ ) gegen die unabhängige Variable  $X$ . Die Werte für  $p$  nähern sich diesen Grenzen asymptotisch bei sehr extremen  $X$ -Werten an, überschreiten sie aber nicht.

**Bild 2** Logistische Funktionskurve



Den Wert  $p$ , also die Wahrscheinlichkeit für das Auftreten des Zielereignisses, erhält man über Odds Ratio und Log-Odds.

$$p = e^x / (1 + e^x) = 1 / (1 + e^{-x})$$

Anhand der oben angegebenen Formeln kann man aus den Variablenwerten für jeden einzelnen Patienten die Wahrscheinlichkeit für das Auftreten des Zielereignisses ermitteln.

Der „Fehler“ in dieser Prognose ergibt sich aus der Differenz zwischen dem erwarteten Outcome, welches mittels der oben angegebenen Formel ermittelt wird, und dem tatsächlich beobachteten Outcome. Bei einem Wert von 1 ist das Ereignis tatsächlich eingetreten und bei 0 dagegen nicht. Ziel eines Prognosemodells ist es diese Differenz so weit wie möglich zu minimieren, so dass die berechnete Prognose weitestgehend dem beobachteten Outcome entspricht.

Odds Ratios sind multivariat adjustiert für die übrigen Variablen im Modell. Insbesondere bei Indikatorvariablen ist also  $e^{\text{Koeffizient}}$  als adjustierte Odds Ratio zu interpretieren, die angibt, um welchen Faktor sich die Wahrscheinlichkeit bei Vorliegen dieses Ereignisses ändert. Bei stetigen Variablen gibt  $e^{\text{Koeffizient}}$  an, um wie viel sich die Prognose bei Änderung um jeweils eine Einheit der stetigen Variable ändert.

Da die Regressionskoeffizienten der logistischen Regression nicht einfach zu interpretieren sind, bildet man häufig sog. Effektkoeffizienten durch Bildung des Antilogarithmus, wodurch sich die Regressionsgleichung auf die Odds Ratios (OR) bezieht.

$$OR = e(\beta_0 + \beta_1 X_1 + \dots + \beta_n X_n)$$

Effektkoeffizienten kleiner 1 weisen einen negativen Einfluss auf die Odds Ratios, ein positiver Einfluss ist gegeben, wenn  $\exp(\beta_n) > 1$  ist.

## 2.6 Vorgehen

Ziel dieser Studie ist es einen anhand von der DIVI erhobenen SAPS-II-Daten modifizierten bzw. adjustierten SAPS-II-Score zu entwickeln, der eine bessere Aussagekraft im Gegensatz zum originalen SAPS-II-Score besitzt. Dazu wurden einerseits nur die im SAPS-II-Score enthaltenen physiologischen Parameter berücksichtigt. Die Variablen Alter, Vorerkrankungen sowie Zuweisung wurden aus dem SAPS-II-Score genommen und genauso wie die neuen Parameter Trauma, neurologische bzw. neurochirurgische Patienten, Aufnahme von Extern,

Geschlecht und Fachbereiche, die normalerweise nicht im SAPS-II-Score vorkommen, neben dem SAPS-II-Score betrachtet. Andererseits wurden die Parameter Alter, Vorerkrankungen sowie Zuweisung im SAPS-II-Score gelassen und zusätzlich neben den oben genannten neuen Variablen neben dem SAPS-II-Score berücksichtigt.

Die normalerweise im originalen SAPS-II-Score enthaltene GCS wurde auf Grund der hohen Beobachtersvariation bei sedierten und beatmeten Patienten nicht ermittelt und somit unberücksichtigt gelassen.

Die von Le Gall et al. im Jahre 1993 eingeführte SAPS-II-Formel lautet folgendermaßen:

$$p = \frac{1}{1 + e^{-x}} \quad \text{mit } x = -7,763 + 0,074 * \text{SAPS-II-Score} + 0,997 * \ln(\text{SAPS II} + 1)$$

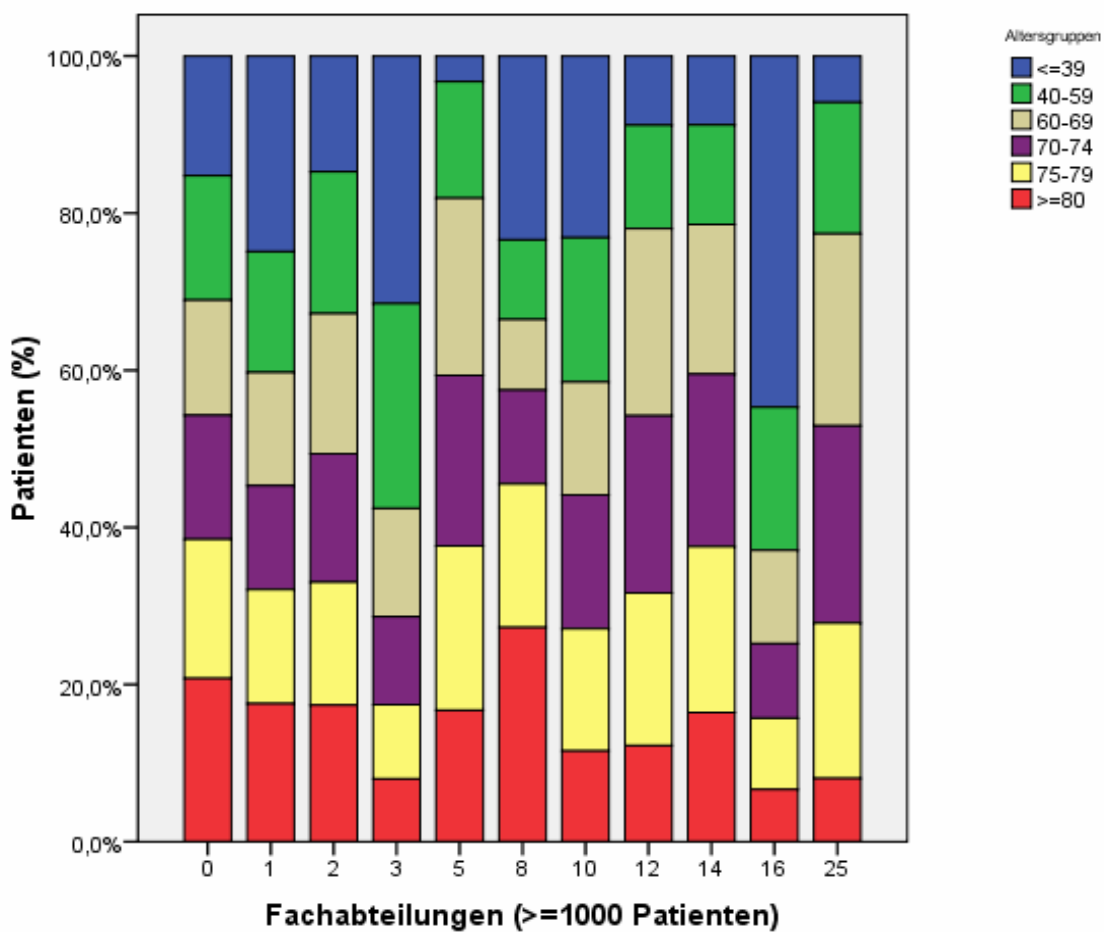
Als Methode wurde zur Entwicklung einer „neuen“ SAPS-II-Formel die logistische Regression angewendet. Dazu wurde als abhängige Variable „VERST“ (verstorben = 1, nicht verstorben = 0) genommen und als Kovariaten zunächst nur der SAPS-II-Score und  $\ln(\text{SAPS II} + 1)$  und danach wurden diese beiden Parameter um einer einzelnen der oben aufgeführten Variablen ergänzt. Außerdem wurde zu den einzelnen Parametern eine ROC-Kurve erstellt, wobei als Zustandsvariable „VERST=1“ genommen wurde und als Testvariable den aus der logistischen Regression ermittelten p-Wert. Zum Abschluss wurde die logistische Regression und ROC-Kurve für alle Variablen zusammen durchgeführt, als abhängige Variable VERST, als Kovariaten alle oben aufgeführten Parameter und als Methode Vorwärts: LR angewendet.

## 3 Ergebnisse

### 3.1 Generelle statistische Daten der untersuchten Patienten

Die Studie umfasst insgesamt 57561 Patienten, die nach dem Zufallsprinzip in zwei Kohorten unterteilt wurden, nämlich in die Entwicklungs- und Validierungsgruppe. Während in der Entwicklungsgruppe 66 % aller Patienten (37990 Patienten) enthalten sind, besteht die Validierungsgruppe aus 19571 Patienten, was 34 % aller Patienten entspricht. Die Daten wurden seit 2000 in 57 verschiedenen Kliniken gesammelt. Das durchschnittliche Alter liegt bei  $65 \pm 15,38$  (18-99) Jahren, wobei 26,4 % der Patienten zwischen 60 und 69 Jahren sind (s. Bild 3, s. Anhang S. 91, Tabelle 5). Die Anzahl der Verstorbenen in den verschiedenen Altersgruppen nimmt in beiden Kohorten mit zunehmendem Alter linear zu (s. Anhang S. 92, Tabelle 6-7).

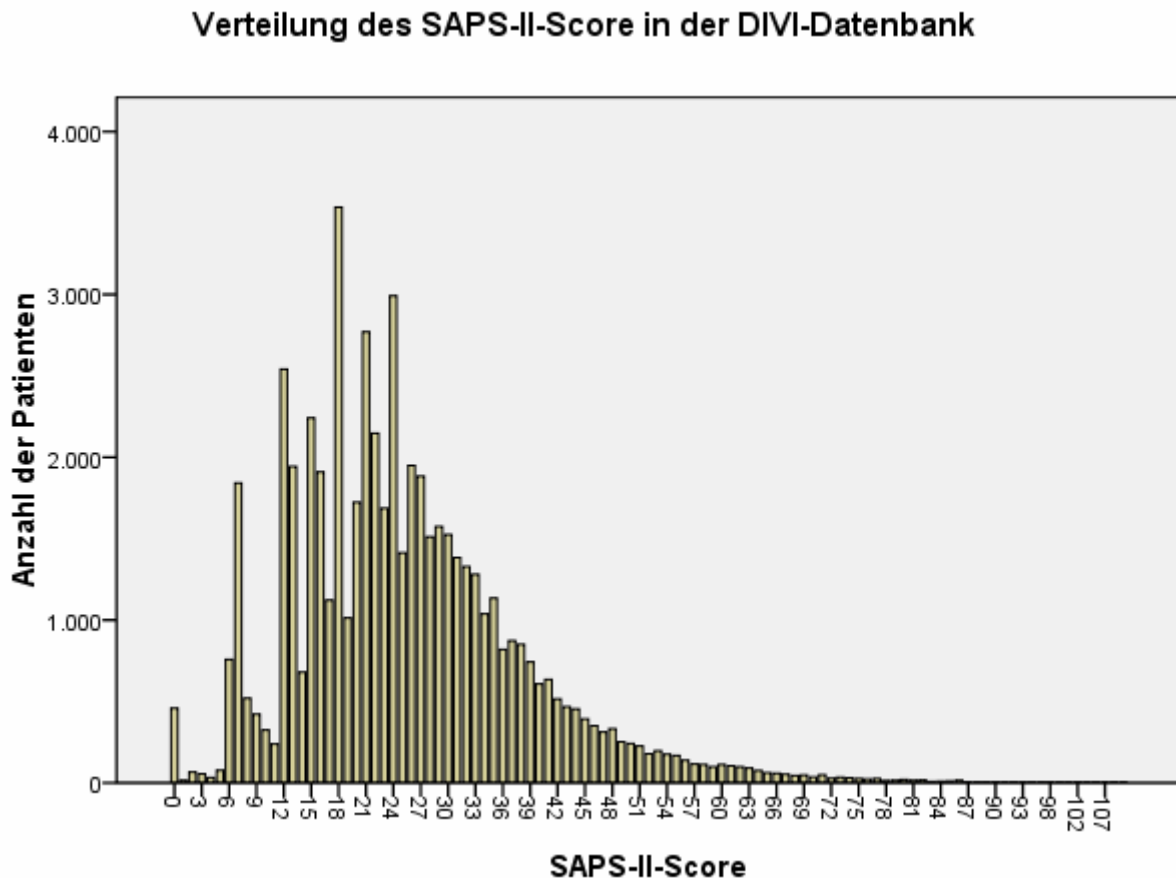
**Bild 3** Altersverteilung in den verschiedenen Fachabteilungen



\* 0 = Innere Medizin, 1 = Anästhesie/Intensiv-/Notfallmedizin, 2 = Allgemeinchirurgie, 3 = Neurochirurgie, 5 = Gefäßchirurgie, 8 = Unfall-Chirurgie, 10 = Neurologie, 12 = Urologie, 14 = Orthopädie, 16 = Gynäkologie, 25 = HTG-Chirurgie

45,2 % (N=26010) der Patienten sind weiblichen Geschlechts. Der SAPS-II-Score kann theoretisch von einem Minimum von null Punkten bis zu einem Maximum von 163 Punkten variieren. Der kleinste Wert, der beobachtet wurde, ist null und der größte Wert 113 mit einem Mittelwert von  $25,5 \pm 12,8$  und einem Median von 24 (17-32), wobei die GCS nicht berücksichtigt wurde (s. Bild 4).

**Bild 4** Verteilung des SAPS-II-Score in der DIVI-Datenbank



Bei den neuen Variablen, die zum SAPS II ergänzt wurden, handelt es sich um Parameter, die mit der Erwartung hinzugefügt wurden, die Vorhersagekraft des SAPS-II-Modell zu verbessern. Dazu gehören Trauma, neurologische bzw. neurochirurgische Patienten, Aufnahme von Extern, Geschlecht sowie Fachbereiche. Bei den Variablen Trauma (ja/nein), neurologische bzw. neurochirurgische Patienten (ja/nein), Aufnahme von Extern (ja/nein) sowie Geschlecht (m/w) handelt es sich um dichotome Variablen. Der Parameter Fachbereiche hingegen ist in verschiedene Fachabteilungen unterteilt, um zu überprüfen, ob es gewisse Fachrichtungen gibt, in denen der SAPS II eine bedeutende Rolle einnimmt.

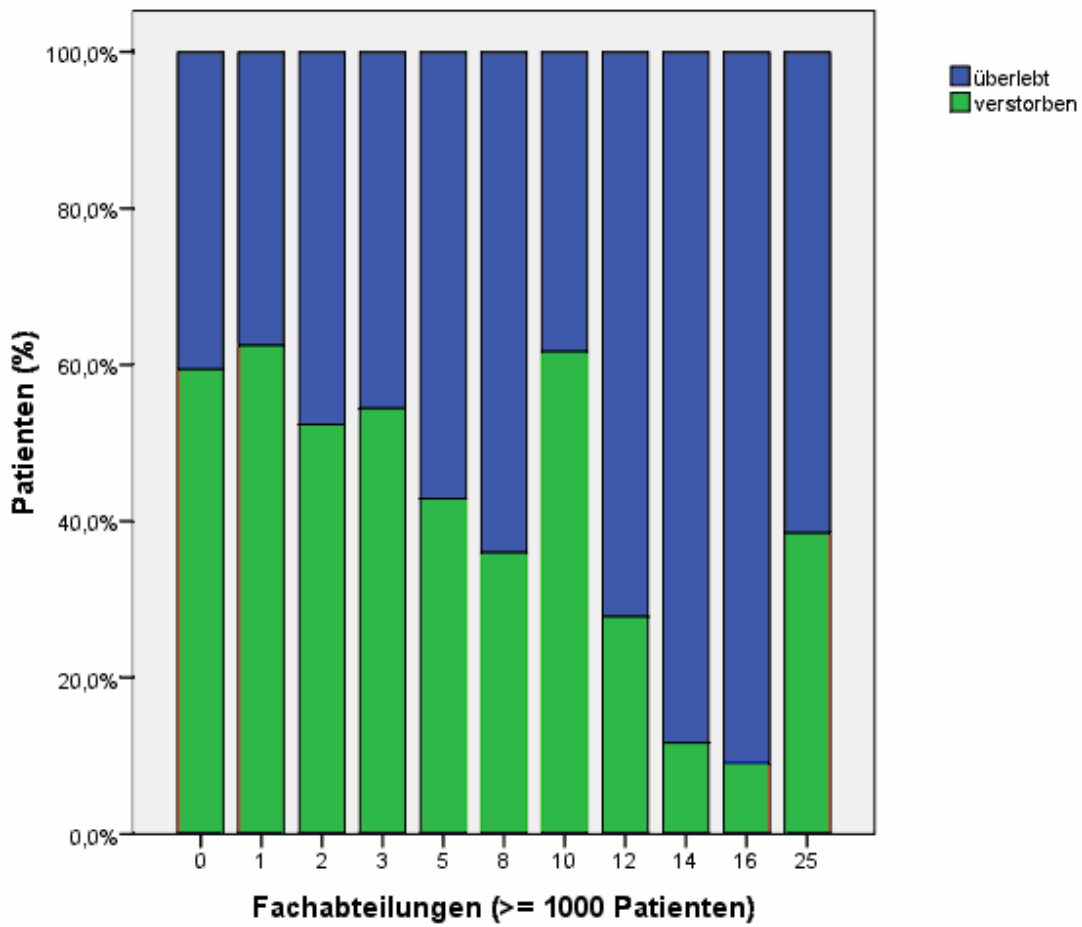
In der Entwicklungsgruppe handelt es sich bei 3312 Patienten (8,7 %) und in der Validierungsgruppe bei 1695 Patienten (8,7 %) um neurologische bzw. neurochirurgische



Patienten. 2402 Patienten (6,3 %) dagegen weisen in der Entwicklungsgruppe und 1238 Patienten (6,3 %) in der Validierungsgruppe ein Trauma auf. Zugewiesen aus einem anderen Krankenhaus, also von Extern aufgenommen, wurden in der Entwicklungsgruppe 2666 Patienten (7,0 %) und in der Validierungsgruppe 1314 Patienten (6,7 %). Während in der Entwicklungsgruppe bei 18396 Patienten (48,4 %) elektiv chirurgisch, bei 13219 Patienten (34,8 %) medizinisch (ohne OP) und bei 6375 Patienten (16,8 %) ungeplant chirurgisch der Aufnahmegrund war, wurden in der Validierungsgruppe 9436 Patienten (48,2 %) elektiv chirurgisch, 6817 Patienten (34,8 %) medizinisch (ohne OP) sowie 3318 Patienten (17,0 %) ungeplant chirurgisch aufgenommen (s. Anhang S. 92, Tabelle 8).

Bei Zuordnung der Patienten in Fachbereiche fällt auf, dass die meisten Patienten aus dem allgemeinchirurgischen Gebiet (15469 Patienten) stammen, danach folgen der Reihenfolge nach unter anderem Innere Medizin mit 13279 Patienten, Herz-Thorax-Gefäß-Chirurgie mit 7184 Patienten, Unfallchirurgie mit 4797 Patienten, Neurochirurgie mit 3917 Patienten, Orthopädie mit 1995 Patienten, Gefäßchirurgie mit 1909 Patienten, Urologie mit 1624 Patienten sowie Gynäkologie mit 1601 Patienten. Auffallend bei den Fachgebieten ist bei Betrachtung der Anzahl der Verstorbenen, dass ca. 30 % aus dem anästhesiologischen (10,3 %), neurologischen (10 %) sowie internistischen (9,2 %) Bereich stammen (s. Bild 5). Dieses mag unter anderem daran liegen, dass in der Anästhesie auch die Intensiv- und Notfallmedizin berücksichtigt ist sowie am Alter, da die meisten Intensivpatienten in diesen drei Bereichen zwischen 40 und 69 Jahren sind. Die niedrigste Letalität liegt hingegen in den Fachrichtungen Gynäkologie (0,7 %) und Orthopädie (0,9 %), was unter anderem darauf zurückzuführen ist, dass es sich bei den gynäkologischen Patienten im Vergleich zu den anderen Fachgebieten überwiegend um junge Patienten handelt (<=39 J.: 23,9 %, 40-59 J.: 28 %, 60-69 J.: 22 %) (s. Anhang S. 94, Tabelle 9).

**Bild 5** Anzahl der Verstorbenen und Überlebenden in den verschiedenen Fachabteilungen



\* 0 = Innere Medizin, 1 = Anästhesie/Intensiv-/Notfallmedizin, 2 = Allgemeinchirurgie, 3 = Neurochirurgie, 5 = Gefäßchirurgie, 8 = Unfall-Chirurgie, 10 = Neurologie, 12 = Urologie, 14 = Orthopädie, 16 = Gynäkologie, 25 = HTG-Chirurgie

## 3.2 SAPS-DIVI 2

### 3.2.1 Aufbau

Beim SAPS-DIVI 2 handelt es sich um einen SAPS-II-Score, bei dem neben den im SAPS-II-Score enthaltenen üblichen Parametern die Variablen Alter, Vorerkrankungen und Zuweisung erneut sowie weitere neue Variablen wie Trauma, Neuro, Aufnahme von Extern, Geschlecht und Fachbereiche neben den Koeffizienten SAPS-II-Score und  $\ln(\text{SAPS-II-Score} + 1)$  berücksichtigt wurden. Diese wurden auf der einen Seite jeweils einzeln zum SAPS-II-Score und  $\ln(\text{SAPS-II-Score} + 1)$  ergänzt und auf der anderen Seite wurden alle Variablen zusammen bewertet.

Die normalerweise im originalen SAPS-II-Score enthaltene GCS wurde auf Grund der hohen Beobachtersvariation bei sedierten und beatmeten Patienten nicht ermittelt und somit unberücksichtigt gelassen. Stattdessen wurde die Variable neurologische bzw. neurochirurgische Patienten verwendet, um den neurologischen Status einzubeziehen. Zudem bezieht sich die mit dem originalen SAPS-II-Score ermittelte Prognose auf das Krankenhaus-Outcome, die mittels SAPS-DIVI 2 berechnete Vorhersage hingegen auf das Intensiv-Outcome.

### 3.3 Analyse der einzelnen Parameter beim SAPS-DIVI 2 im Entwicklungsdatensatz

#### 3.3.1 Regressionskoeffizienten (s. Anhang S. 94, Tabelle 10)

Anhand des Regressionskoeffizienten lässt sich der Beitrag einer unabhängigen Variablen für die Erklärung oder Prognose der abhängigen Variable ablesen. Der Regressionskoeffizient misst den Einfluss der unabhängigen Variablen auf die Zielvariable in einer Regressionsanalyse.

Auffällig ist hierbei, dass die Regressionskoeffizienten der Variablen Alter und Vorerkrankungen ein negatives Vorzeichen besitzen, wobei es bei den Vorerkrankungen jedoch auch wichtig ist die Patientenzahl der jeweiligen Vorerkrankung zu berücksichtigen, denn in der Entwicklungsgruppe haben von insgesamt 37990 Patienten 4816 ein metastasierendes Karzinom, 315 maligne hämatologische Erkrankungen und 20 Patienten AIDS. Demzufolge sollte man das negative Vorzeichen des Regressionskoeffizienten mit Vorsicht interpretieren, da man bei dieser Patientenzahl den Einfluss auf den SAPS-II-Score nicht wirklich bewerten kann. Stattdessen wäre es der Einfachheit halber sinnvoller alle drei Vorerkrankungen als dichotome Variable (Vorerkrankungen ja/nein) zusammenzufassen und nicht nach den drei verschiedenen Vorerkrankungen zu unterscheiden, da man dadurch eine höhere Patientenzahl hat (5151, ca. 16 % der Patienten in der Entwicklungsgruppe) und somit einen zuverlässigeren Effekt auf den SAPS-II-Score verzeichnen kann. Dieser Ansatz wurde in dieser Studie verfolgt und zeigte ein ähnliches Resultat. Das Vorzeichen des Regressionskoeffizienten blieb unverändert negativ bei -0,496. Der Regressionskoeffizient liegt dagegen bei Betrachtung der einzelnen Vorerkrankungen bei metastasierendem Karzinom bei -0,521, bei maligne hämatologische Erkrankungen beträgt er -0,277 sowie bei AIDS -0,520. Somit liegt er genau zwischen diesen drei Regressionskoeffizienten. Jedoch kann man nicht anhand des Vorzeichens des Regressionskoeffizienten eine Aussage über die Signifikanz machen. Dazu muss man sich den Standardfehler, den p-Wert sowie das 95 %-Konfidenzintervall anschauen. Wenn die Zahl „0“ nicht im 95 %-Konfidenzintervall vorkommt, dann ist der p-Wert  $< 0,05$  (5 %) und es liegt somit eine statistische Signifikanz vor.

Des Weiteren sind die Koeffizienten von Alter und Vorerkrankungen nicht als alleiniger Effekt fehlzuinterpretieren, da im SAPS-II-Score bereits Punkte für die Parameter Alter und Vorerkrankungen enthalten sind. Somit sind die negativen Koeffizienten nur zusammen mit der Gewichtung im SAPS-II-Score zu interpretieren. In Abschnitt 3.5.1 (s. S. 52) hingegen

werden im SAPS-II-Score nur physiologische Parameter berücksichtigt und daneben unter anderem Variablen wie Alter und Vorerkrankungen betrachtet. Somit enthalten diese Parameter dort keine Punkte, wie es hier der Fall ist.

Bei der Variablen Vorerkrankungen fällt auf, dass die Erkrankung maligne hämatologische Erkrankungen einen Regressionskoeffizienten aufweist, der fast so groß wie der Standardfehler ist (bei AIDS ist der Standardfehler sogar größer) und somit die Signifikanz bei beiden fraglich ist, wobei der p-Wert jeweils bei 0,130 und 0,452 liegt. Jedoch muss man auch die geringe Patientenzahl berücksichtigen, denn die statistische Signifikanz ist von der Größe der untersuchten Population abhängig.

Bei der Variablen Alter liegt in jeder Altersstufe eine Signifikanz des Regressionskoeffizienten vor, wobei der Standardfehler in jeder Altersklasse im Großen und Ganzen gleich groß ist und ungefähr bei 0,128 liegt, der Regressionskoeffizient dagegen nimmt von der Altersgruppe 40-59 bis  $\geq 80$  linear zu. Folglich nimmt die Signifikanz mit zunehmendem Alter zu.

Bei allen anderen Parametern hingegen, d.h. bei Neuro, Aufnahme von Extern, Trauma-Patienten, Geschlecht sowie Zuweisung weist der Regressionskoeffizient ein positives Vorzeichen auf.

Unter der Variablen Neuro versteht man, ob es sich um einen neurologischen bzw. neurochirurgischen Patienten handelt. Der Regressionskoeffizient liegt bei 1,038 und der Standardfehler bei 0,076. Wenn man hierfür das 95 %-Konfidenzintervall berechnet, kann man erkennen, dass die Zahl „0“ nicht im Konfidenzintervall vorkommt und man von ihr sehr weit entfernt ist.

Ähnliches gilt auch für die Variablen Aufnahme von Extern und Zuweisung. Mit Aufnahme von Extern ist gemeint, ob der Patient von Extern aufgenommen, also von einem anderen Krankenhaus zugewiesen worden ist. Mit Zuweisung hingegen ist die Aufnahmeart gemeint, die in elektiv, nicht operativer Notfall und operativer Notfall unterteilt ist. Der Regressionskoeffizient sowie der Standardfehler bei Aufnahme von Extern (RK = 0,695; SF = 0,075) und medizinisch (ohne OP) (RK = 0,659; SF = 0,062) sind ungefähr gleich groß. Auffallend bei der Variablen Zuweisung ist außerdem, dass der Regressionskoeffizient positiv ist und eine statistische Signifikanz aufweist, obwohl diese Variable bereits im SAPS-II-Score berücksichtigt ist. Bei der Subvariable „ungeplant chirurgisch“ des Parameters Zuweisung liegt der Regressionskoeffizient bei 0,491 und der Standardfehler bei 0,071 und ist somit ebenfalls statistisch signifikant.

Bei den Variablen Trauma (ja/nein) und Geschlecht (weiblich/männlich) ist der Regressionskoeffizient zwar statistisch signifikant, jedoch ist er bei Betrachtung des 95 %-Konfidenzintervalls relativ nah an der „0“, so dass diese beiden Variablen einen nicht allzu sehr großen Einfluss auf die mit dem SAPS-II-Score ermittelte Letalität haben. Während der Regressionskoeffizient und Standardfehler bei der Variablen Trauma jeweils bei 0,299 und 0,093 liegen, betragen sie bei dem Parameter Geschlecht 0,131 und 0,047.

Bei einem großen Datensatz, wie es hier der Fall ist, erlangen schon kleine Unterschiede statistische Signifikanz, so dass die Signifikanz mit Vorsicht interpretiert werden muss.

Die Regressionskoeffizienten hingegen liegen bei nur Betrachtung der Variablen SAPS-II-Score und  $\ln(\text{SAPS-II-Score} + 1)$  beim originalen SAPS-II-Score (0,0737 und 0,9971) und bei unserem SAPS-II-Score (0,065 und 1,042) dicht beieinander.

### 3.3.2 Konstante (s. Anhang S. 94, Tabelle 10)

An dieser Stelle sei noch einmal die originale SAPS-II-Formel vorgestellt:

$$\text{Logit} = -7,7631 + 0,0737 * (\text{SAPS II}) + 0,9971 * \ln((\text{SAPS II}) + 1)$$

$$\text{Predicted Death Rate} = e^{(\text{Logit})} / (1 + e^{(\text{Logit})})$$

Eine Konstante ist ein fester, unveränderlicher Wert.

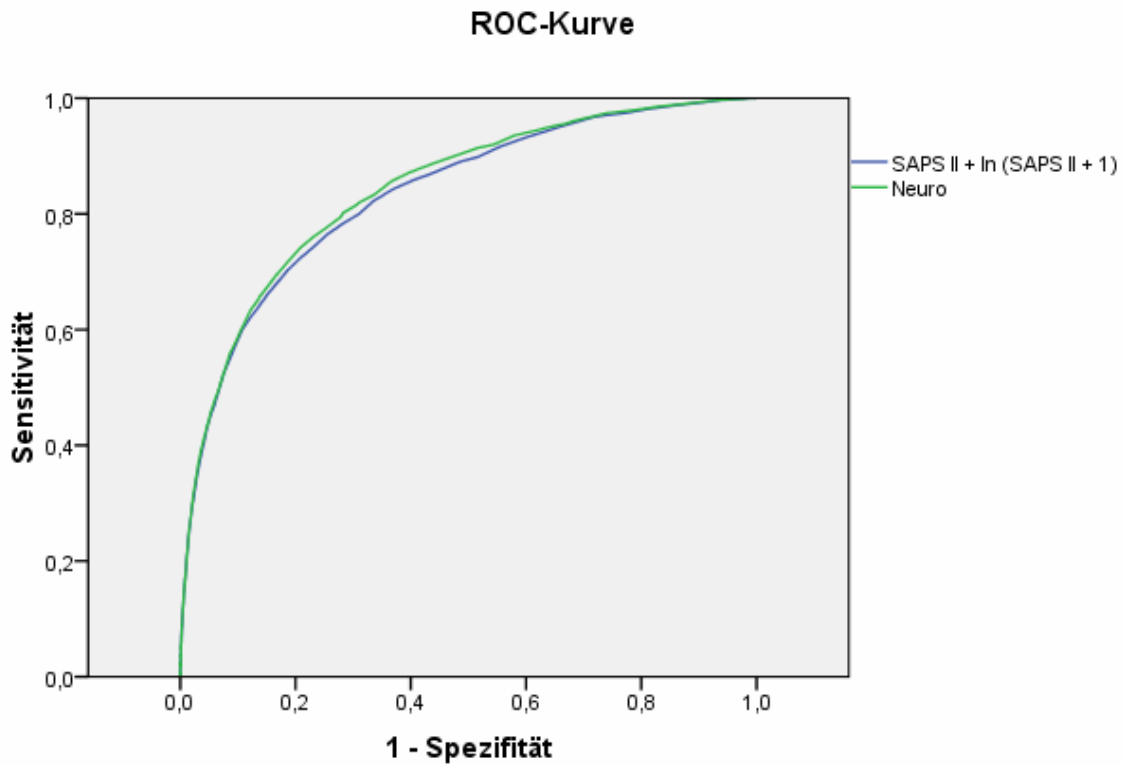
In unserer Studie liegt die Konstante bei -8,421, wenn man nur den SAPS-II-Score und den natürlichen Logarithmus von SAPS II + 1 wie in der originalen Formel berücksichtigt, wobei jedoch in unserem SAPS-II-Score die GCS nicht enthalten ist. Wenn man noch zusätzlich eine weitere der oben genannten Variablen zum SAPS-II-Score und  $\ln(\text{SAPS-II-Score} + 1)$  hinzufügt, verändert sich die Konstante bei den meisten Variablen nicht im Wesentlichen. Bei den Parametern Neuro (-8,900) und Alter (-9,148) hingegen ist die Konstante erheblich negativer bzw. kleiner, dagegen wird die Konstante bei der Variablen Zuweisung größer (-8,098).

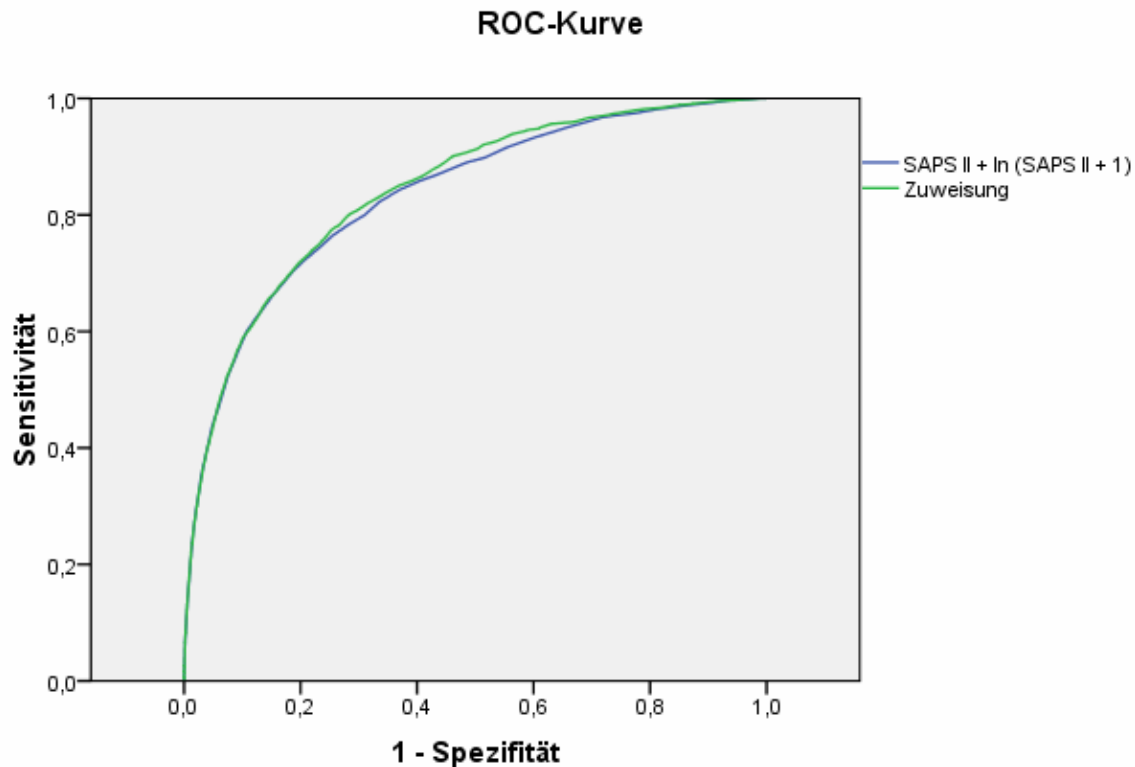
### 3.3.3 ROC-Kurve (s. Anhang S. 95, Tabelle 11)

Die Receiver Operating Characteristic (ROC) – Kurve ist eine Methode zur Bewertung und Optimierung von Analyse-Strategien. Die ROC-Kurve stellt visuell die Abhängigkeit der Effizienz mit der Fehlerrate für verschiedene Parameterwerte dar.

Wenn man sich die ROC-Kurven der jeweiligen Parameter anschaut, so fällt auf, dass die Fläche unter der Kurve bei den „Ausgangsvariablen“ SAPS-II-Score und  $\ln(\text{SAPS-II-Score} + 1)$  des SAPS-DIVI 2 bei 0,834 liegt und durch das Hinzufügen von nur jeweils einer weiteren Variable größer wird, am größten bei den Variablen Neuro (0,842) (s. Bild 6) und Zuweisung (0,841) (s. Bild 7).

**Bild 6** ROC-Kurve - Neuro



**Bild 7** ROC-Kurve - Zuweisung

Eine kleine Zunahme der Fläche unter der ROC-Kurve ist hingegen bei den Variablen Trauma (0,835) und Vorerkrankungen (0,836) zu verzeichnen und eine etwas größere Zunahme bei den Parametern Alter (0,838) und Aufnahme von Extern (0,839). Bei der Variablen Geschlecht (m) (0,834) dagegen ist kein Zuwachs zu verzeichnen.

Je größer die Fläche unter der ROC-Kurve ist, umso besser kann ein Score diskriminieren, d.h. zwischen Überlebenden und Versterbenden differenzieren. Folglich resultiert eine höhere Sensitivität und Spezifität vor allem bei den Parametern Neuro und Zuweisung.

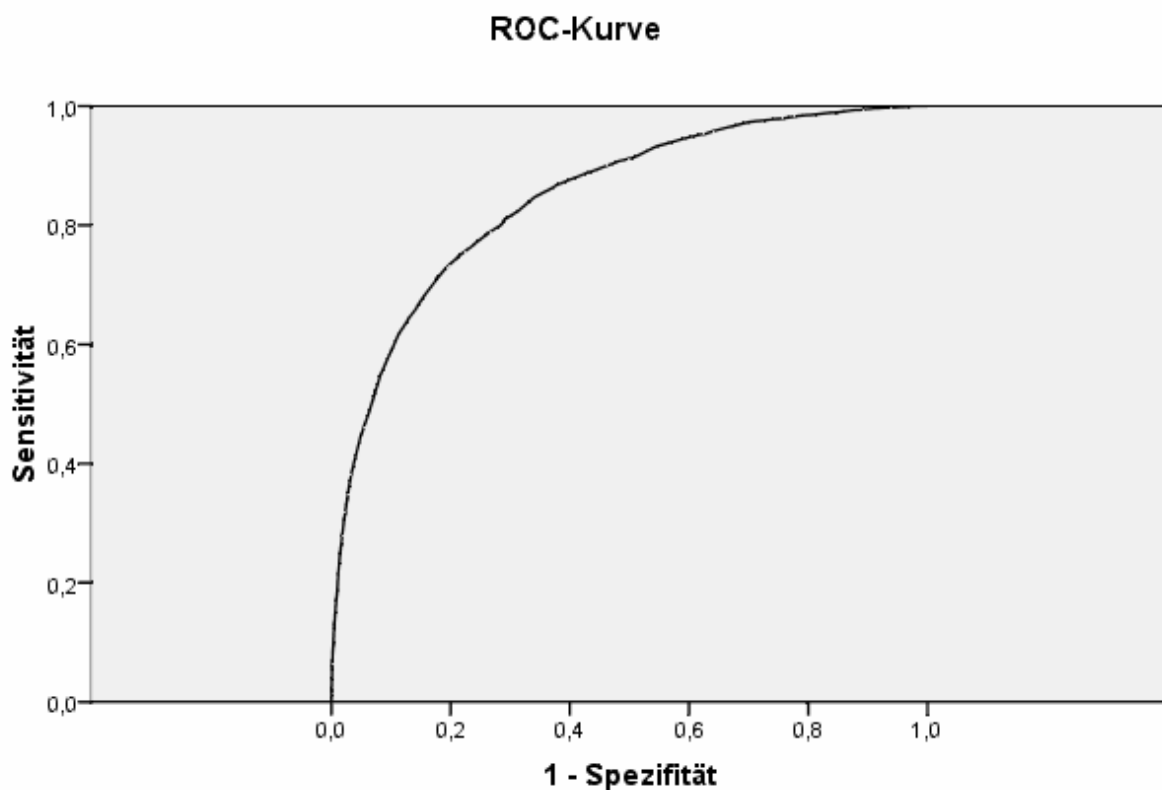
### 3.3.4 Fachrichtungen (s. Anhang S. 96, Tabelle 12-13)

Wenn man die Variable Fachrichtungen zum SAPS-II-Score +  $\ln(\text{SAPS-II-Score} + 1)$  hinzufügt, so fallen bei den Fachgebieten mit positivem Regressionskoeffizienten die Bereiche Neurologie und Neurochirurgie sowie bei den Fachgebieten mit negativem Regressionskoeffizienten MKG-Chirurgie, HNO, Urologie, Orthopädie und Gynäkologie ins Augenmerk. All diese Subparameter besitzen eine statistische Signifikanz, da ihr Koeffizient mehr als doppelt so groß als ihr Standardfehler ist. Jedoch muss man auch die Patientenzahl in den jeweiligen Fachbereichen berücksichtigen, denn der Entwicklungsdatensatz umfasst beispielsweise in der Neurologie 732 Patienten, in der HNO 409 Patienten und in der MKG-Chirurgie 342 Patienten. Bei dieser Patientenzahl muss man die Signifikanz mit Vorsicht



interpretieren, da man deren Effizienz zum SAPS-II-Score bei dieser Größenordnung noch nicht wirklich bewerten kann, erst bei größeren Patientengruppen kann man eine aussagekräftigere Beurteilung abgeben. Ungeachtet dessen ist es sinnvoll die Variablen mit positivem Regressionskoeffizienten und die Parameter mit negativem Koeffizienten jeweils zusammenzufassen, um einerseits eine größere Anzahl an Patienten zu erhalten und andererseits die Effektivität als eine einzige Variable der Übersicht halber zusammengefasst mit positivem bzw. negativem Regressionskoeffizienten zu evaluieren. So weisen die jeweiligen Koeffizienten eine eindeutige Signifikanz auf, da sie um ein Vielfaches größer als ihr Standardfehler sind und von der „0“ im 95 %-Konfidenzintervall weit entfernt liegen. Des Weiteren beträgt die Fläche unter der ROC-Kurve bei Zusammenfassung der Variablen 0,846 (s. Bild 8).

**Bild 8** ROC-Kurve - Fachrichtungen



### 3.3.5 Präzision (s. Anhang S. 92, Tabelle 8)

Die Präzision ist ein Maß für die Übereinstimmung zwischen unabhängigen Messergebnissen unter festen Bedingungen. Liegen also mehrere Messwerte dicht beieinander, so hat die Messmethode eine hohe Präzision.

Während die vorhergesagte Letalität mit unserem adjustierten SAPS-II-Score 6,3 % beträgt, ist die tatsächlich beobachtete Mortalitätsrate bei 6,3 %, was die Übereinstimmung der Prognose des Scores mit dem tatsächlichen Outcome wiedergibt.

### 3.3.6 Zusammenfassung aller Variablen (s. Anhang S. 97, Tabelle 14)

Durch das Hinzufügen aller Variablen zum SAPS-DIVI 2 werden bei der logistischen Regression bis auf die Parameter Trauma und Geschlecht alle Variablen berücksichtigt. Dabei nehmen insbesondere das Alter und die Fachbereiche eine nicht zu unterschätzende Rolle ein. Während das Alter in verschiedene Altersgruppen unterteilt ist, unterscheidet man bei den Fachbereichen nach positivem und negativem Regressionskoeffizient, wobei Neurologie und Neurochirurgie zu der Gruppe mit positivem Koeffizienten gehören und MKG-Chirurgie, HNO, Urologie, Orthopädie sowie Gynäkologie in die Gruppe mit negativem Regressionskoeffizienten. Der Regressionskoeffizient nimmt in der Altersgruppe von 40-59 bis  $\geq 80$  linear von -0,529 auf -1,066 zu und der Standardfehler bleibt im Mittel bei 0,131. Somit nimmt die statistische Signifikanz mit zunehmendem Alter zu, da der Regressionskoeffizient mehr als doppelt so groß wie der Standardfehler ist. Bei der Variablen Fachbereiche ist der Koeffizient ebenfalls um ein Vielfaches größer als sein Standardfehler und somit statistisch signifikant. Die Variablen Zuweisung, Vorerkrankungen (alle) und Aufnahme von Extern hingegen weisen einen kleineren Regressionskoeffizienten auf, sind jedoch ebenfalls statistisch signifikant.

#### SAPS-DIVI 2 (Entwicklungsdatensatz / Validierungsdatensatz):

$$p = \frac{1}{1 + e^{-x}} \quad \text{mit} \quad x = -9,380 + 0,055 * \text{SAPS II} + 1,593 * \ln(\text{SAPS II} + 1) - 0,529 * \text{Alter (40-59 J.)} - 0,821 * \text{Alter (60-69 J.)} - 0,949 * \text{Alter (70-74 J.)} - 0,957 * \text{Alter (75-79 J.)} - 1,066 * \text{Alter } (\geq 80 \text{ J.)} + 0,341 * \text{Zuweisung (medizinisch (ohne OP))} + 0,233 * \text{Zuweisung (ungeplant chirurgisch)}$$

- 0,324 \* Vorerkrankungen (alle)

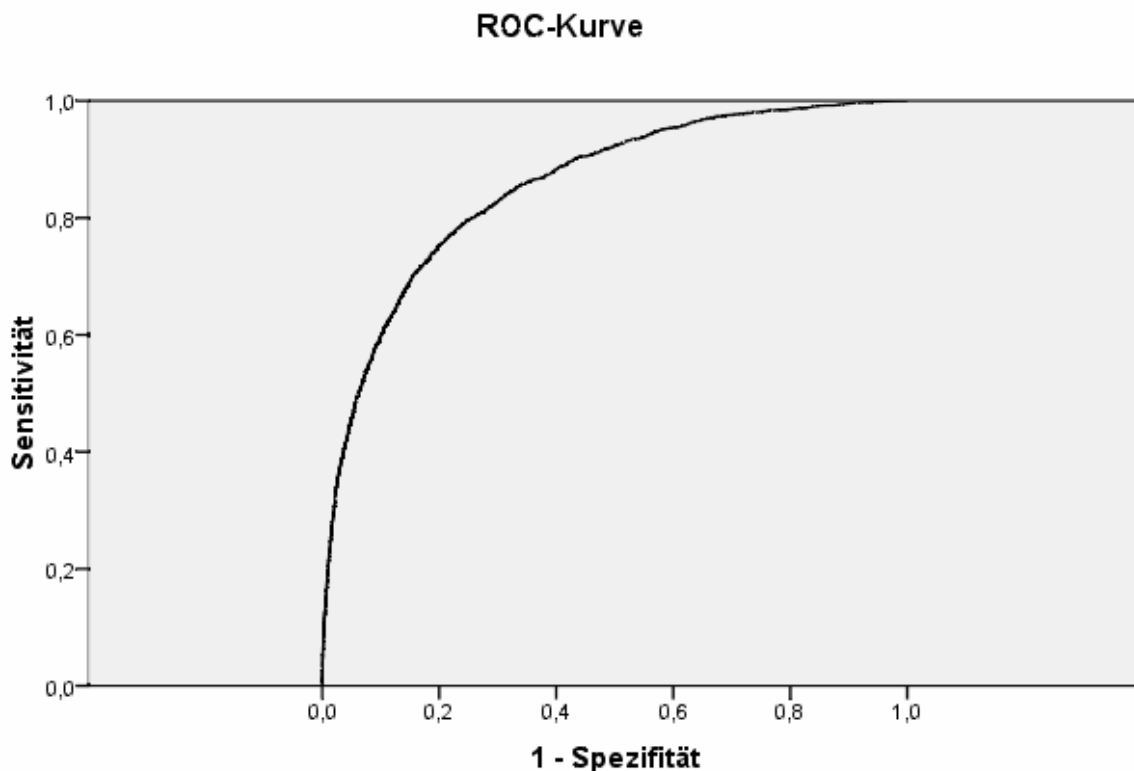
+ 0,340 \* Aufnahme von Extern

+ 0,809 \* Plus-Fächer (Neurologie, Neurochirurgie)

- 0,874 \* Minus-Fächer (MKG-Chirurgie, HNO,  
Urologie, Orthopädie, Gynäkologie)

Die Fläche unter der ROC-Kurve beträgt bei Berücksichtigung aller Variablen beim SAPS-DIVI 2 0,854 (s. Bild 9) und ohne einer dieser Variablen, also bei SAPS-II-Score + ln(SAPS-II-Score + 1) 0,834.

**Bild 9** ROC-Kurve – Zusammenfassung aller Variablen

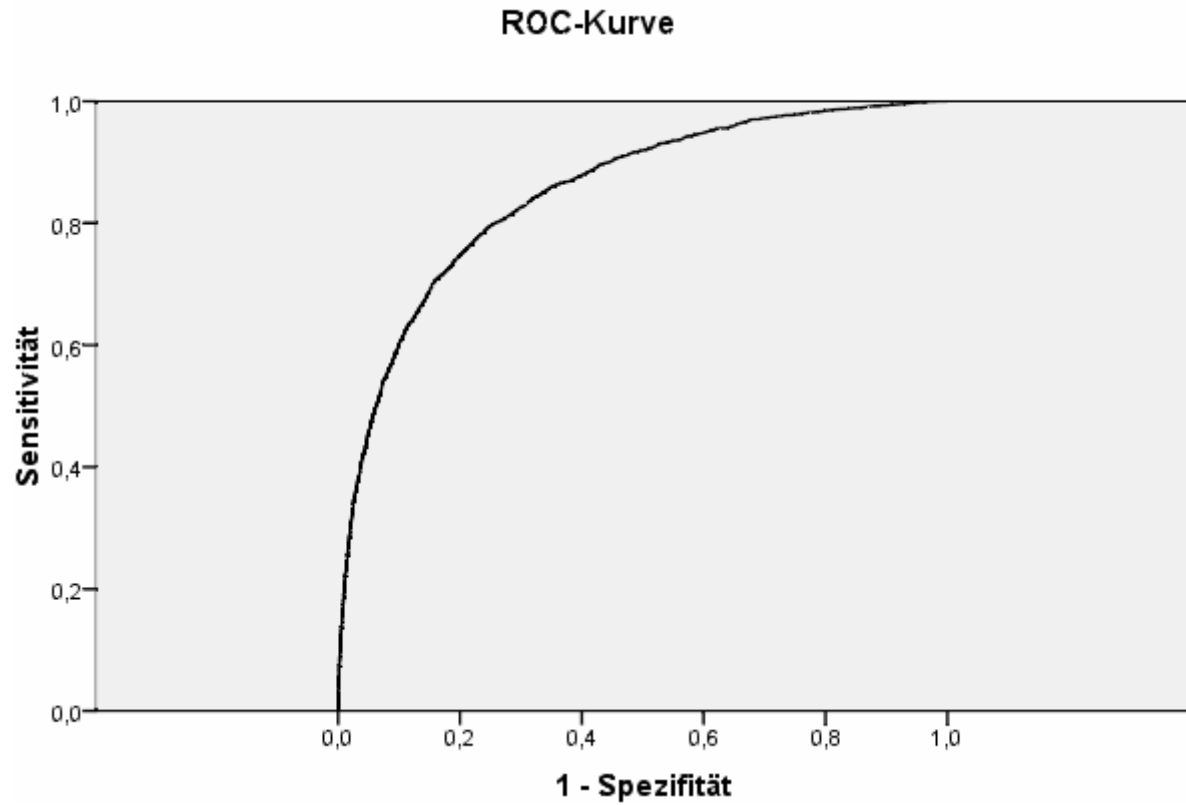


All diese Variablen leisten bei der Ermittlung der Intensivletalität mit dem SAPS-II-Score einen entscheidenden Beitrag, was an den ermittelten Ergebnissen kenntlich wird. An neuen Parametern im Vergleich zum originalen SAPS-II-Score sind Aufnahme von Extern, Fachbereiche sowie Vorerkrankungen (alle) hinzugekommen.

Wenn man anstelle der Variablen Fachbereiche den Parameter Neuro anwendet, da die Variable Neuro die Subvariable „Plus-Fächer“ (Neurologie, Neurochirurgie) widerspiegelt, kommt es nicht zu einer bemerkenswerten Veränderung, wobei die Fläche unter der ROC-

Kurve minimal abnimmt und bei 0,851 liegt (s. Bild 10). Der Regressionskoeffizient liegt bei 0,855 und der Standardfehler bei 0,079 (s. Anhang S. 97, Tabelle 15).

**Bild 10** ROC-Kurve – Zusammenfassung aller Variablen (Neuro statt Fachbereiche)



## 3.4 SAPS-DIVI 2 (Phys)

### 3.4.1 Aufbau

Der SAPS-DIVI 2 (Phys) enthält im Gegensatz zum originalen SAPS-II-Score und SAPS-DIVI 2 nur physiologische Parameter. Die Variablen Alter, Vorerkrankungen sowie Zuweisung wurden jeweils einzeln neben SAPS-II-Score und  $\ln(\text{SAPS-II-Score} + 1)$  betrachtet. Außerdem wurden hier genauso wie beim SAPS-DIVI 2 die neuen Variablen Trauma, neurologische bzw. neurochirurgische Patienten, Aufnahme von Extern, Geschlecht sowie Fachbereiche jeweils einzeln neben SAPS-II-Score und  $\ln(\text{SAPS-II-Score} + 1)$  berücksichtigt. Zum Schluss wurden dann alle Variablen zum SAPS-II-Score und  $\ln(\text{SAPS-II-Score} + 1)$  hinzugefügt. Auch beim SAPS-DIVI 2 (Phys) wurde die GCS nicht beachtet. Stattdessen wurde desgleichen die Variable neurologische bzw. neurochirurgische Patienten angewendet, um den neurologischen Status einzubeziehen. Ebenfalls bezieht sich die mit SAPS-DIVI 2 (Phys) ermittelte Prognose nicht auf das Krankenhaus-Outcome, sondern auf das Intensiv-Outcome wie bei SAPS-DIVI 2.

### **3.5 Analyse der einzelnen Parameter beim SAPS-DIVI 2 (Phys) im Entwicklungsdatensatz**

#### **3.5.1 Regressionskoeffizienten (s. Anhang S. 98, Tabelle 16)**

Bei den Regressionskoeffizienten fallen vor allem die Parameter Neuro (Regressionskoeffizient 0,912; Standardfehler 0,075), Aufnahme von Extern (Regressionskoeffizient 0,653; Standardfehler 0,076), Alter und Zuweisung ins Augenmerk. Während bei der Variablen Alter der Regressionskoeffizient von 40-59 bis  $\geq 80$  Jahren von 0,221 auf 0,806 zunimmt, bleibt der Standardfehler im Großen und Ganzen im Mittel bei 0,121. Folglich steigt das Letalitätsrisiko mit zunehmendem Lebensalter an.

Bei der Variablen Zuweisung hingegen liegt der Regressionskoeffizient bei der Subvariablen medizinisch (ohne OP) bei 1,040 und der Standardfehler bei 0,061 sowie bei ungeplant chirurgisch jeweils bei 1,101 und 0,069, was deren überragende Rolle bei der Ermittlung des Letalitätsrisikos mittels SAPS II unterstreicht, denn im 95 %-Konfidenzintervall ist man von der Zahl „0“ relativ weit entfernt und die Regressionskoeffizienten sind um ein Vielfaches größer als ihr Standardfehler.

Die Variable Trauma besitzt einen Regressionskoeffizienten von 0,294 und einen Standardfehler von 0,093, was zwar eine Signifikanz darstellt, die jedoch im Vergleich zu den zuvor erwähnten Parametern nicht so ausgeprägt ist.

Ähnliches gilt auch für die Variable Vorerkrankungen, wobei man hierbei jedoch berücksichtigen sollte, dass von insgesamt 37990 Patienten „nur“ 4816 ein metastasierendes Karzinom (Regressionskoeffizient: 0,223; Standardfehler 0,068), 315 maligne hämatologische Erkrankungen (Regressionskoeffizient: 0,579; Standardfehler: 0,187) und 20 Patienten AIDS (Regressionskoeffizient: 0,386; Standardfehler: 0,686) haben und somit die Aussagekraft nur eingeschränkt ist. Vernünftiger wäre es, so wie es in dieser Studie gemacht wurde, nämlich alle drei Vorerkrankungen zusammenzufassen, wodurch man eine höhere Patientenzahl und somit zugleich auch eine bessere Aussagekraft über die Effizienz dieser Variablen machen kann (Regressionskoeffizient: 0,259; Standardfehler: 0,064).

Die Variable Geschlecht hingegen weist einen Regressionskoeffizienten von 0,001 und einen Standardfehler von 0,047 auf, was bedeutet, dass Frauen und Männer dieselbe Ausgangsprognose haben und somit keines von beiden Geschlechtern bei der Ermittlung der Letalität mittels SAPS II benachteiligt ist. Folglich spielt die Variable Geschlecht im SAPS-II-Score zur Ermittlung der Letalität keine Rolle.

Die Regressionskoeffizienten hingegen liegen bei nur Betrachtung der Variablen SAPS-II-Score und  $\ln(\text{SAPS-II-Score} + 1)$  beim originalen SAPS-II-Score jeweils bei 0,0737 und 0,9971 und hier bei 0,093 (SAPS-II-Score) und 0,145 ( $\ln(\text{SAPS-II-Score} + 1)$ ). Sie weichen stark voneinander ab, was auf die Wichtigkeit der Parameter Alter, Vorerkrankungen und Zuweisung hindeutet, die im SAPS-DIVI 2 (Phys) nicht enthalten sind.

In Bezug auf das 95 %-Konfidenzintervall fällt auf, dass der Regressionskoeffizient bei den Parametern neurologische bzw. neurochirurgische Patienten, Aufnahme von Extern sowie Aufnahmeart von der „0“ relativ weit entfernt ist, was zugleich bedeutet, dass diese Variablen einen entscheidenden Beitrag zur Ermittlung der Letalität mittels SAPS II leisten. Bei der Variablen Geschlecht hingegen liegt der Regressionskoeffizient im 95 %-Konfidenzintervall im Bereich der „0“ und ist somit statistisch nicht signifikant. Das Geschlecht spielt also keine wesentliche Rolle bei der Berechnung der Mortalitätswahrscheinlichkeit mittels SAPS II.

### **3.5.2 Konstante (s. Anhang S. 98, Tabelle 16)**

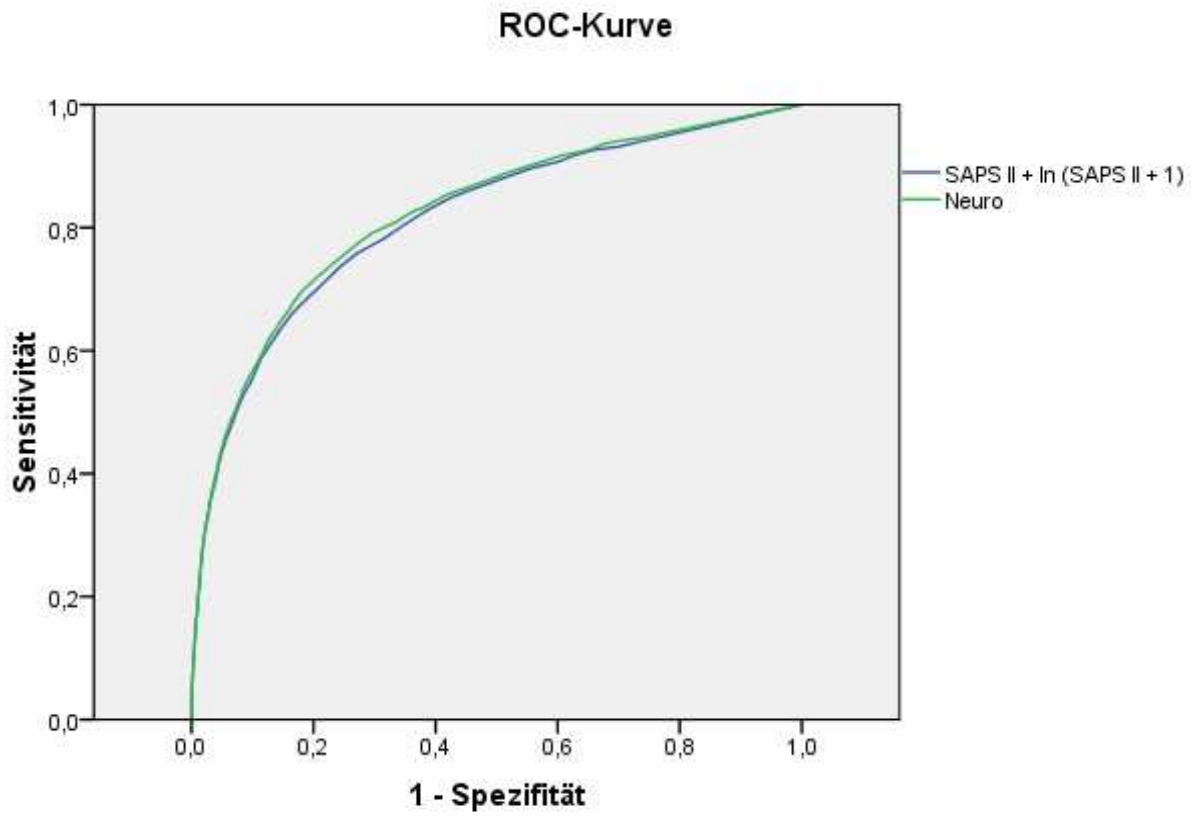
Die Konstante liegt hier bei -4,366, wenn man wie im originalen SAPS-II-Score nur den SAPS-II-Score und den natürlichen Logarithmus von  $\text{SAPS-II-Score} + 1$  berücksichtigt, wobei die GCS nicht enthalten ist.

Während sich die Konstante bei Ergänzung der Variablen Neuro (-4,540), Alter (-4,790) und Aufnahmemodus (-4,970) zunehmend verkleinert, wird sie beim Parameter Aufnahme von Extern (-4,438), Vorerkrankungen (alle) (-4,404) sowie Vorerkrankungen (-4,402) geringfügig kleiner. Bei der Variablen Geschlecht (-4,366) dagegen ist die Konstante genauso groß und bei dem Parameter Trauma (-4,388) minimal kleiner.

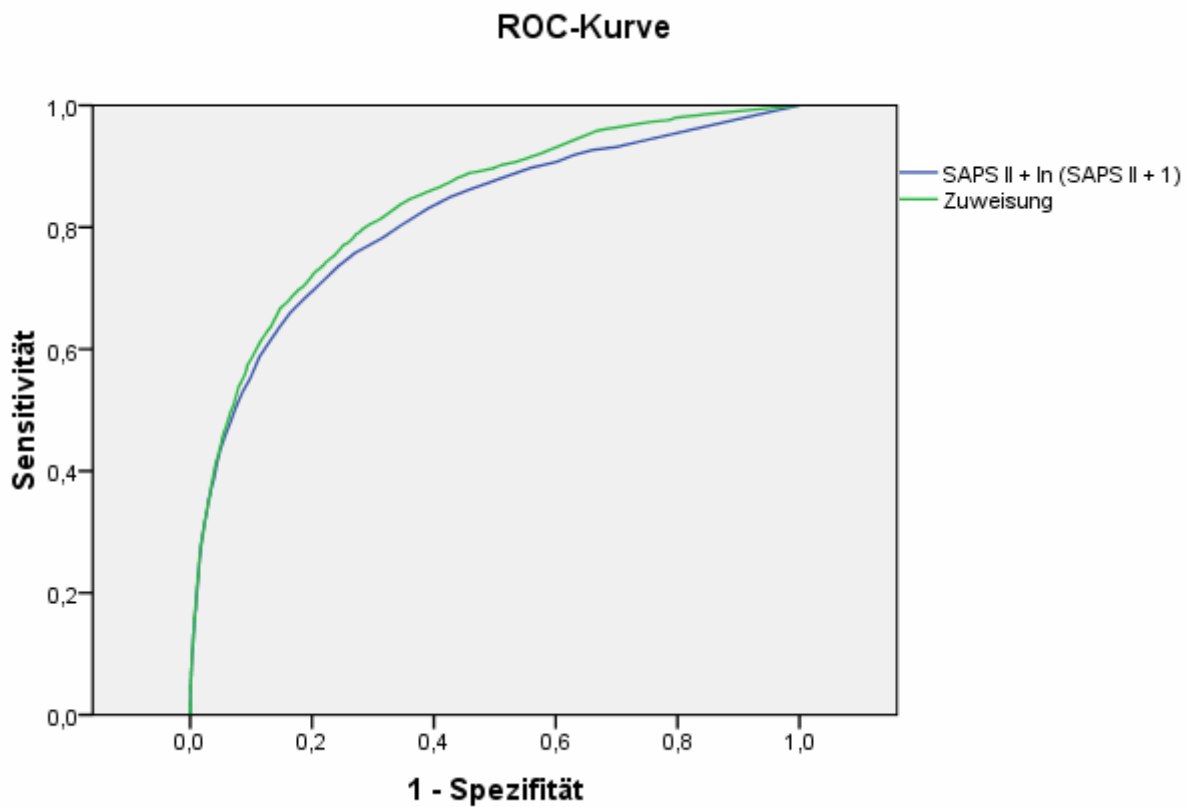
### **3.5.3 ROC-Kurve (s. Anhang S. 99, Tabelle 17)**

Die Fläche unter der ROC-Kurve ist am größten, wenn man zu den „Ausgangsvariablen“ SAPS-II-Score und  $\ln(\text{SAPS-II-Score} + 1)$  des SAPS-DIVI 2 (Phys) die Variable Neuro (0,823) (s. Bild 11) oder Zuweisung (0,837) (s. Bild 12) hinzufügt.

**Bild 11** ROC-Kurve - Neuro



**Bild 12** ROC-Kurve - Zuweisung



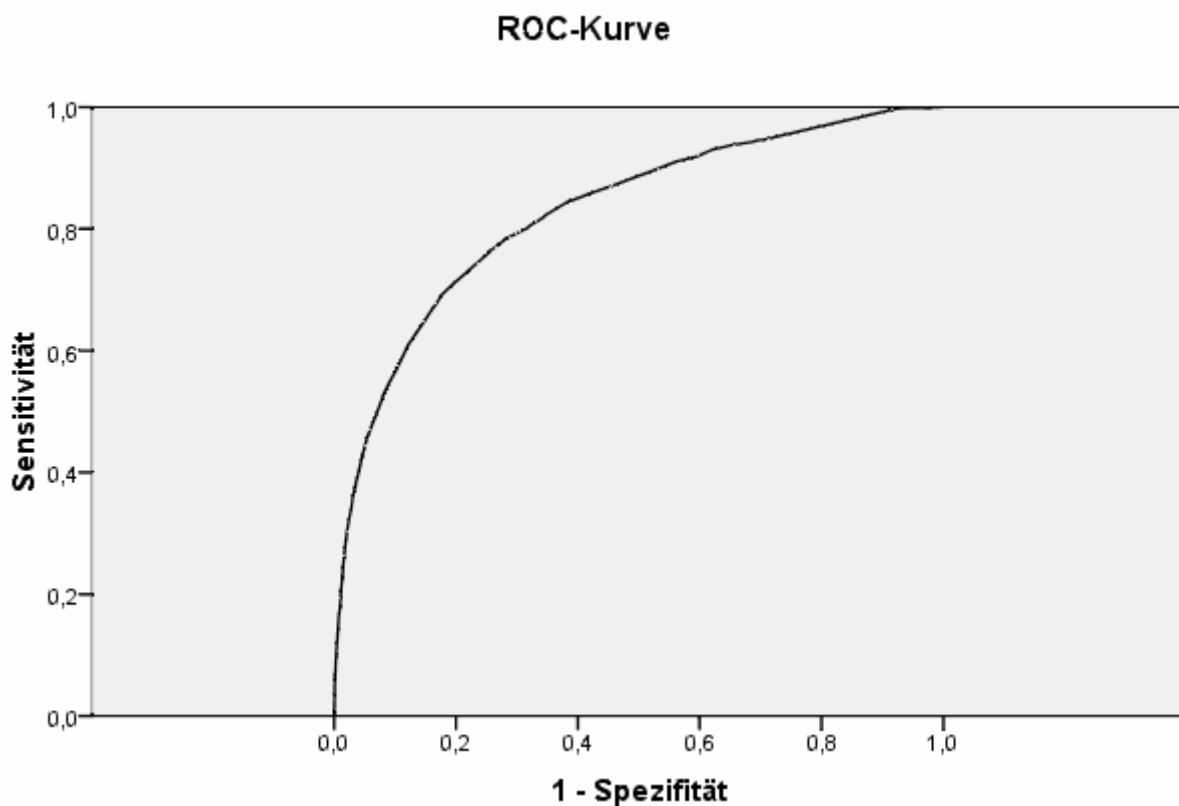


Die Fläche unter der ROC-Kurve beträgt bei alleiniger Anwendung von SAPS-II-Score und  $\ln(\text{SAPS-II-Score} + 1)$  auf die Entwicklungsgruppe 0,815. Die neuen Variablen Trauma sowie Geschlecht (m) dagegen und der bereits im SAPS-II-Score bekannte Parameter Vorerkrankungen weisen nur eine minimale Zunahme der Fläche von 0,816 auf. Während die Fläche bei der Variablen Vorerkrankungen (alle) genauso groß ist, ergibt sich bei den Parametern Aufnahme von Extern (0,820) und Alter (0,821) eine etwas größere Zunahme der Fläche unter der ROC-Kurve.

### 3.5.4 Fachrichtungen (s. Anhang S. 99-100, Tabelle 18-19)

Beim SAPS-DIVI 2 (Phys) wurden dieselben Fachrichtungen wie beim SAPS-DIVI 2 berücksichtigt. Sie weisen ebenfalls eine statistische Signifikanz auf, da ihr Regressionskoeffizient größer als ihr Standardfehler ist. Durch das Zusammenfassen der Variablen hat man ein ähnliches Resultat wie beim SAPS-DIVI 2 erzielt. Jedoch beträgt die Fläche unter der ROC-Kurve hier 0,828 (s. Bild 13).

**Bild 13** ROC-Kurve - Fachrichtungen



### 3.5.5 Präzision

Die vorhergesagte Letalität liegt hier bei 6,3 % und entspricht ebenfalls der beobachteten Mortalitätsrate von 6,3 %.

### 3.5.6 Zusammenfassung aller Variablen (s. Anhang S. 100, Tabelle 20)

Der SAPS-DIVI 2 (Phys) weist bei Berücksichtigung aller Variablen im Großen und Ganzen ein ähnliches Resultat wie der SAPS-DIVI 2 auf. Ebenso werden bei Zusammenfassung aller Variablen die Parameter Trauma und Geschlecht unberücksichtigt gelassen. Auch hier fallen insbesondere das Alter und die Fachbereiche sowie der Aufnahmegrund ins Augenmerk. Während bei der Variablen Alter der Regressionskoeffizient von 0,309 auf 0,899 kontinuierlich zunimmt, liegt der Standardfehler im Durchschnitt bei 0,123. Dieses spiegelt die überragende Bedeutung dieser Variablen im SAPS-II-Score wider.

Bei dem Parameter Fachbereiche hingegen, der in zwei Subvariablen unterteilt ist, liegt bei den „Plus-Fächern“ (Neurologie, Neurochirurgie) der Koeffizient bei 0,785 und der Standardfehler bei 0,079 und bei den „Minus-Fächern“ (MKG-Chirurgie, HNO, Urologie, Orthopädie, Gynäkologie) der Koeffizient bei -0,903 und der Standardfehler bei 0,136. Beide Subparameter weisen somit eine statistische Signifikanz auf.

Die größte Bedeutung scheint die Variable Zuweisung zu besitzen, denn die Subvariablen medizinisch (ohne OP) (Koeffizient: 0,974; Standardfehler: 0,063) und ungeplant chirurgisch (Koeffizient: 1,064; Standardfehler: 0,070) besitzen einen Koeffizienten, der um ein Vielfaches größer als ihr Standardfehler ist.

Die Variablen Vorerkrankungen (alle) und Aufnahme von Extern dagegen besitzen einen kleineren Regressionskoeffizienten, sind jedoch ebenfalls statistisch signifikant.

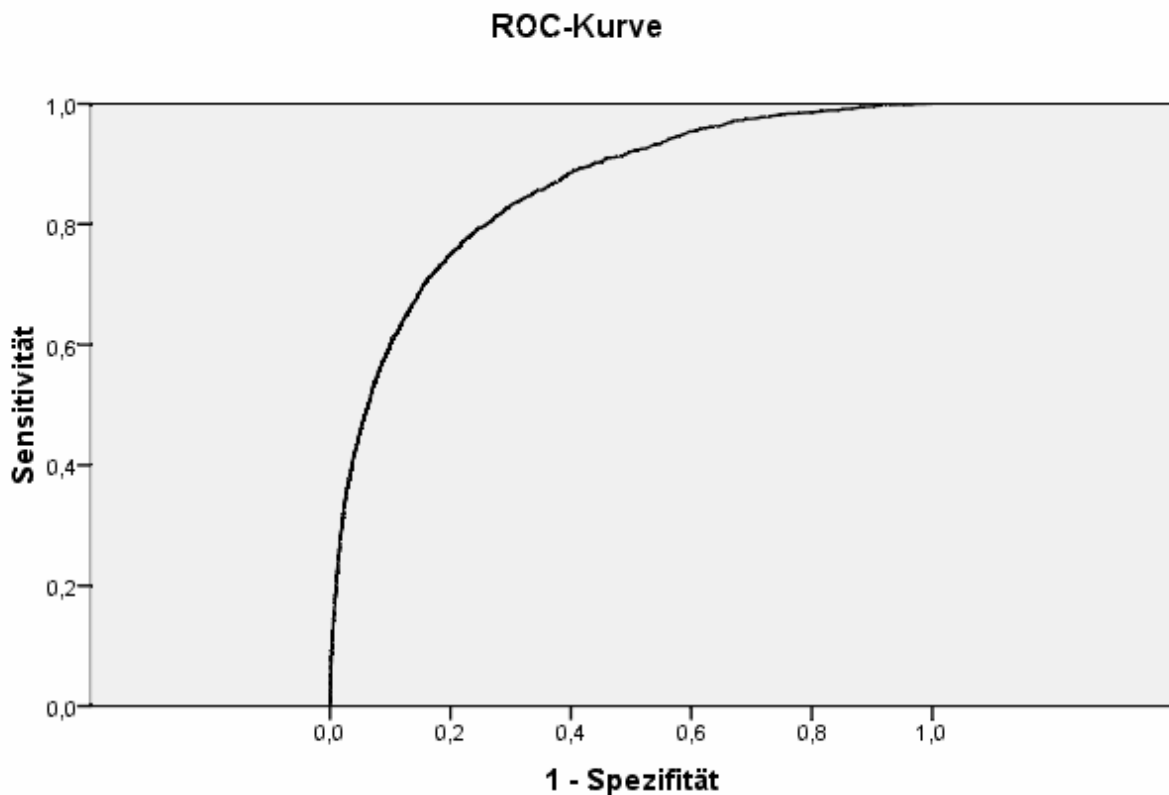
### SAPS-DIVI 2 (Phys) (Entwicklungsdatensatz / Validierungsdatensatz):

$$p = \frac{1}{1 + e^{-x}} \quad \text{mit} \quad x = -5,695 + 0,086 * \text{SAPS II (Phys)} + 0,152 * \ln(\text{SAPS II (Phys)} + 1) + 0,309 * \text{Alter (40-59 J.)} + 0,553 * \text{Alter (60-69 J.)} + 0,726 * \text{Alter (70-74 J.)} + 0,819 * \text{Alter (75-79 J.)} + 0,899 * \text{Alter (}\geq 80 \text{ J.)}$$

- + 0,974 \* Zuweisung (medizinisch (ohne OP))
- + 1,064 \* Zuweisung (ungeplant chirurgisch)
  
- + 0,582 \* Vorerkrankungen (alle)
  
- + 0,359 \* Aufnahme von Extern
  
- + 0,785 \* Plus-Fächer (Neurologie, Neurochirurgie)
- 0,903 \* Minus-Fächer (MKG-Chirurgie, HNO,  
Urologie, Orthopädie, Gynäkologie)

Beim Hinzufügen aller Variablen zum SAPS-II-Score und  $\ln(\text{SAPS-II-Score} + 1)$  beträgt die Fläche bei SAPS-DIVI 2 (Phys) unter der ROC-Kurve 0,853 (s. Bild 14), wobei sie ohne das Ergänzen einer Variablen bei 0,815 liegt.

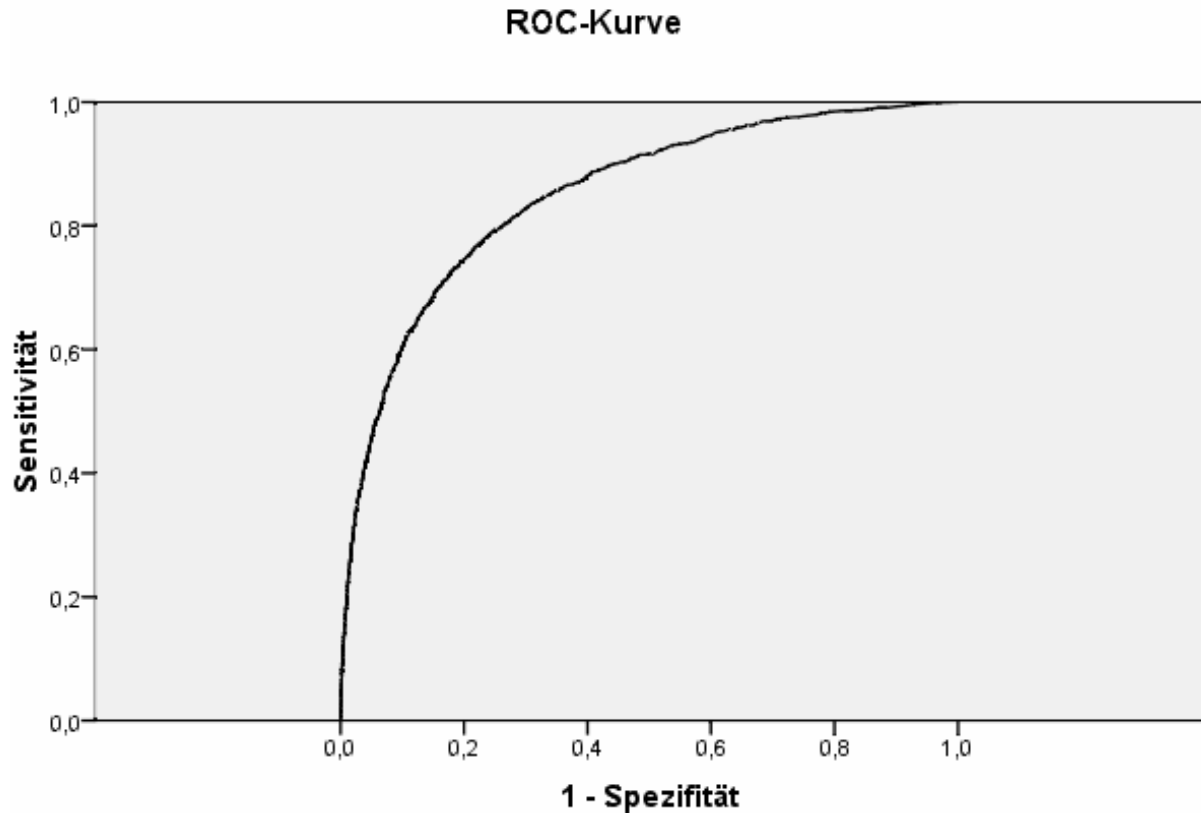
**Bild 14** ROC-Kurve – Zusammenfassung aller Variablen



Wenn man die Variable Fachbereiche durch den Parameter Neuro ersetzt, kommt es nicht zu einer wesentlichen Veränderung, der Regressionskoeffizient liegt bei 0,833 und der

Standardfehler bei 0,079 (s. Anhang S. 101, Tabelle 21). Die Fläche unter der ROC-Kurve nimmt geringfügig ab und liegt bei 0,851 (s. Bild 15).

**Bild 15** ROC-Kurve – Zusammenfassung aller Variablen (Neuro statt Fachbereiche)

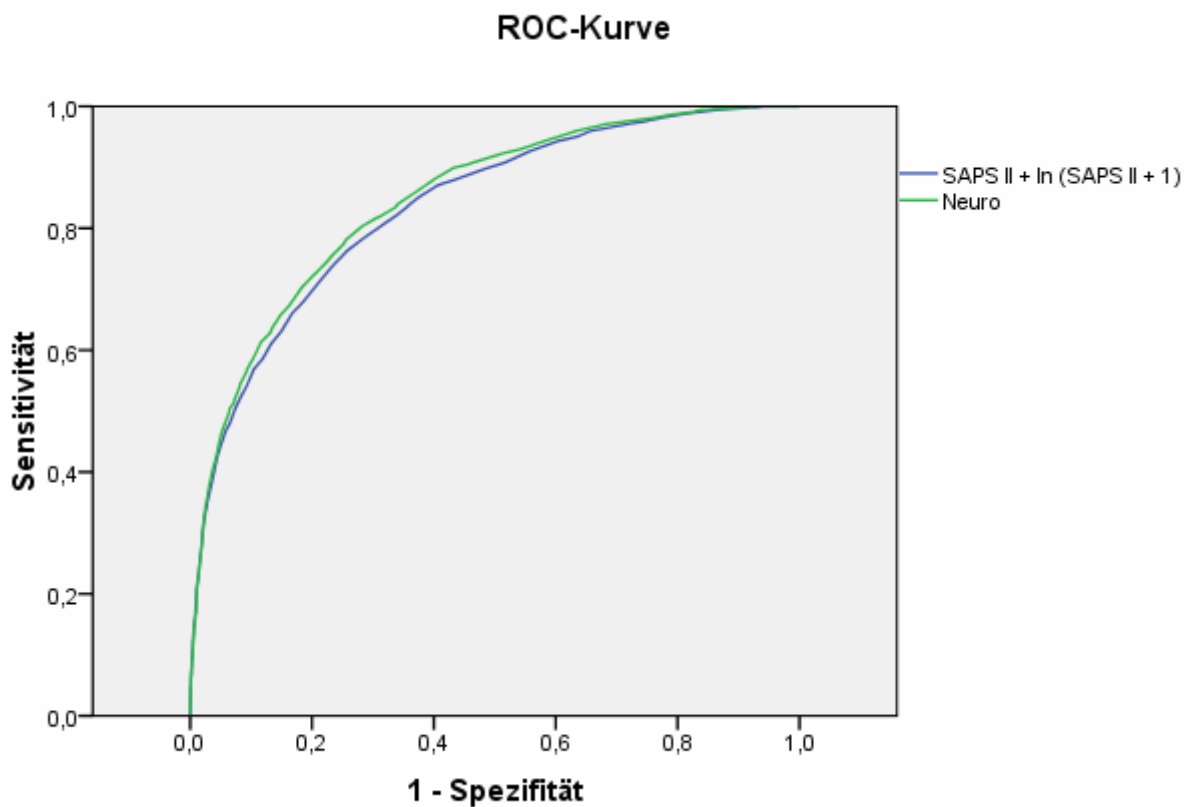


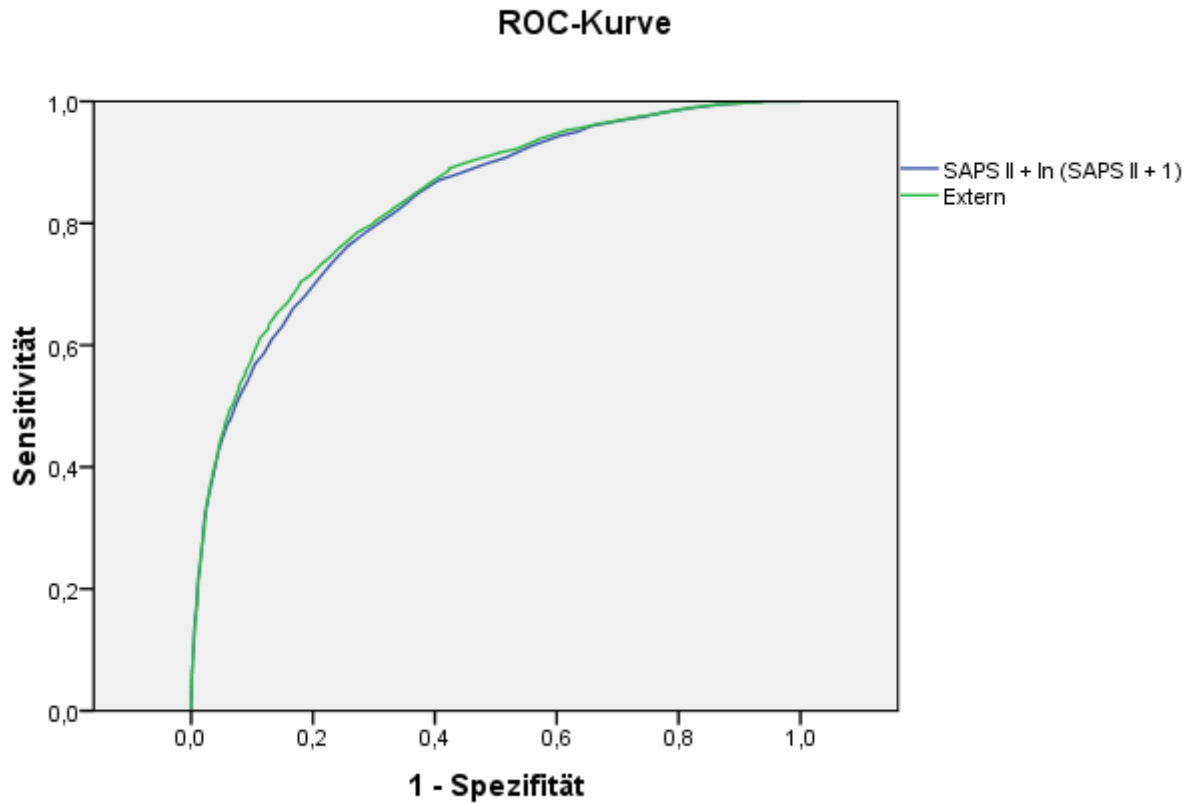
### 3.6 SAPS-DIVI 2 im Validierungsdatensatz

#### 3.6.1 Auswertung (s. Anhang S. 101, Tabelle 22)

Bei Betrachtung der ROC-Kurven beträgt die Fläche unter der ROC-Kurve bei den „Ausgangsvariablen“  $\text{SAPS-II-Score} + \ln(\text{SAPS-II-Score} + 1)$  des SAPS-DIVI 2 0,835. Bis auf der Variablen Geschlecht (0,834) ist es bei allen anderen Parametern zu einer Zunahme der Fläche unter der ROC-Kurve gekommen. So verzeichnen im Gegensatz zur Entwicklungsgruppe die Variablen neurologische bzw. neurochirurgische Patienten (0,846) (s. Bild 16) und Aufnahme von Extern (0,843) (s. Bild 17) den größten Zuwachs.

**Bild 16** ROC-Kurve - Neuro



**Bild 17** ROC-Kurve – Aufnahme von Extern

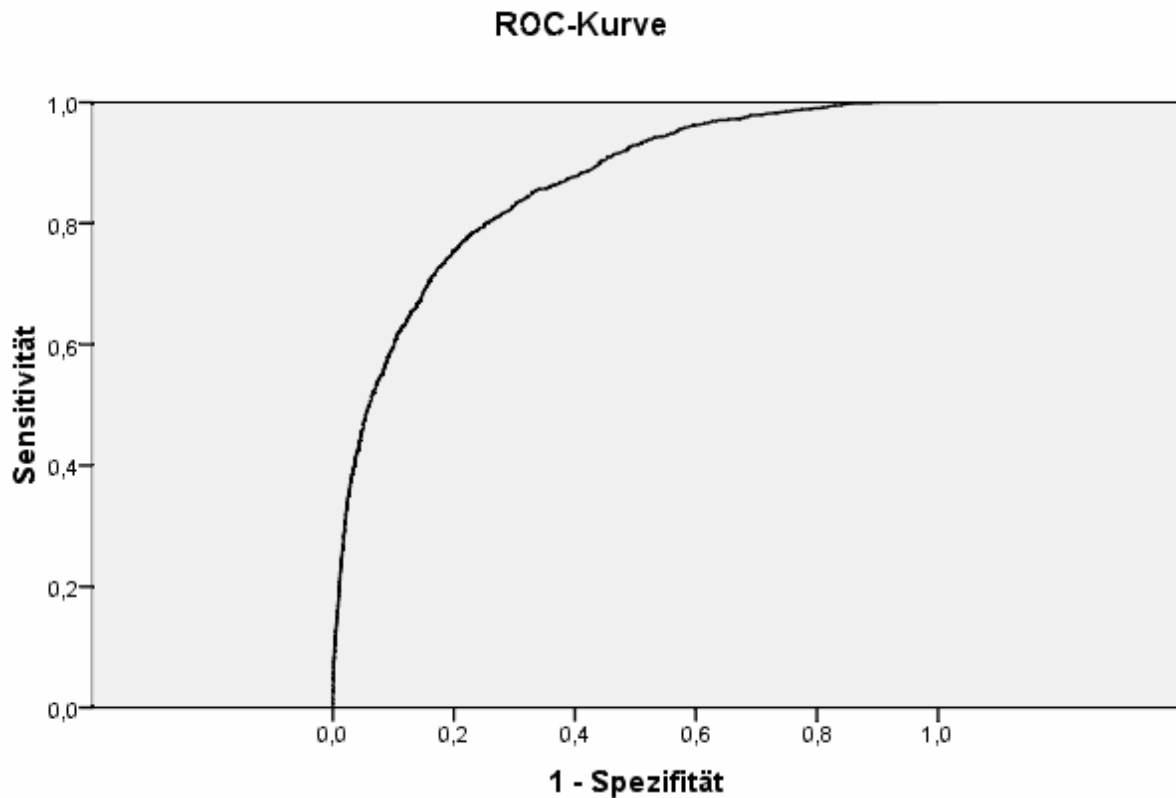
Zu nur einer kleinen Vergrößerung der Fläche kommt es hingegen bei den Parametern Trauma (0,836) sowie Vorerkrankungen (0,837) und einer etwas größeren Zunahme bei Alter (0,838) und Zuweisung (0,838).

Die erwartete Mortalitätsrate liegt hier bei 6,3 % und ist somit genauso hoch wie im Entwicklungsdatensatz. Sie entspricht somit ebenfalls der tatsächlich beobachteten Mortalitätsrate (6,3 %).

### 3.6.2 Zusammenfassung aller Variablen

Bei Betrachtung aller Variablen zusammen, beträgt die Fläche im Validierungsdatensatz beim SAPS-DIVI 2 unter der ROC-Kurve 0,856 (s. Bild 18).

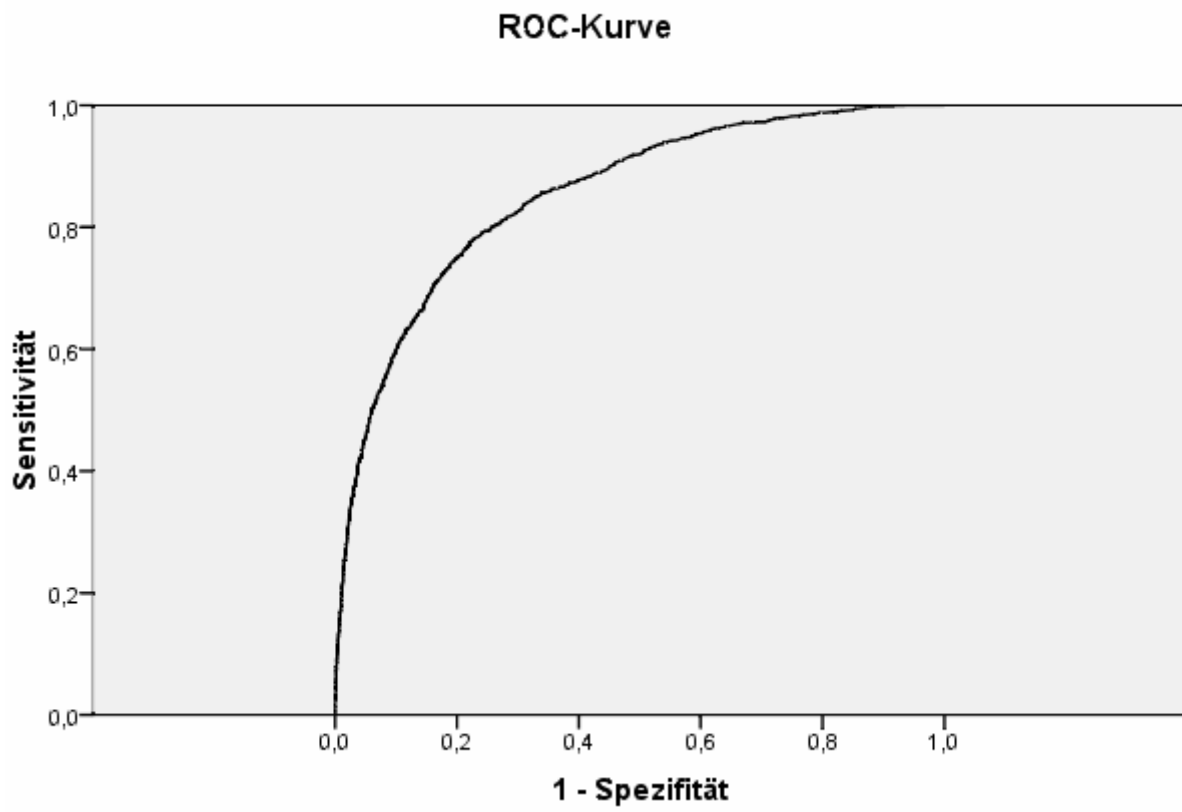
**Bild 18** ROC-Kurve - Zusammenfassung aller Variablen



Sie ist etwas größer als im Entwicklungsdatensatz, wo sie 0,854 beträgt. Somit bestätigt SAPS-DIVI 2 im Validierungsdatensatz das im Entwicklungsdatensatz ermittelte Ergebnis.

Bei Anwendung der Variablen Neuro statt Fachbereiche nimmt die Fläche unter der ROC-Kurve leicht ab und liegt bei 0,853 (s. Bild 19).

**Bild 19** ROC-Kurve - Zusammenfassung aller Variablen (Neuro statt Fachbereiche)



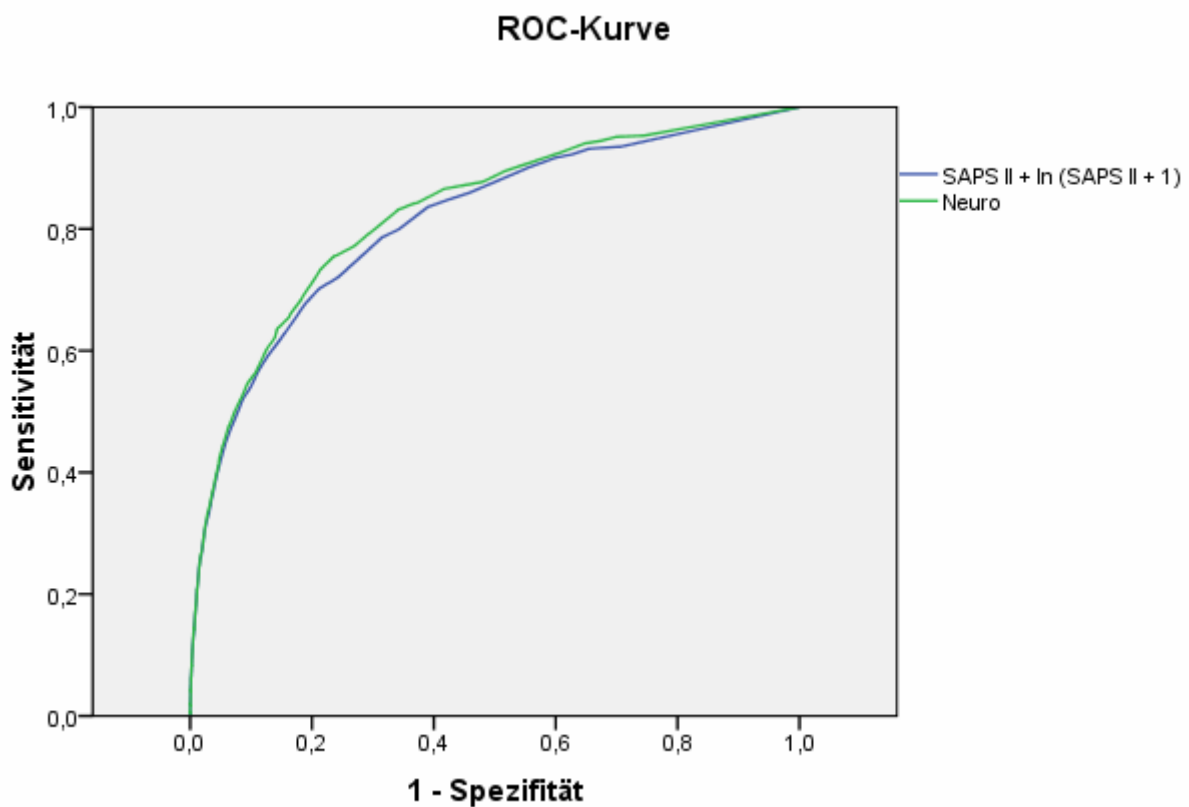


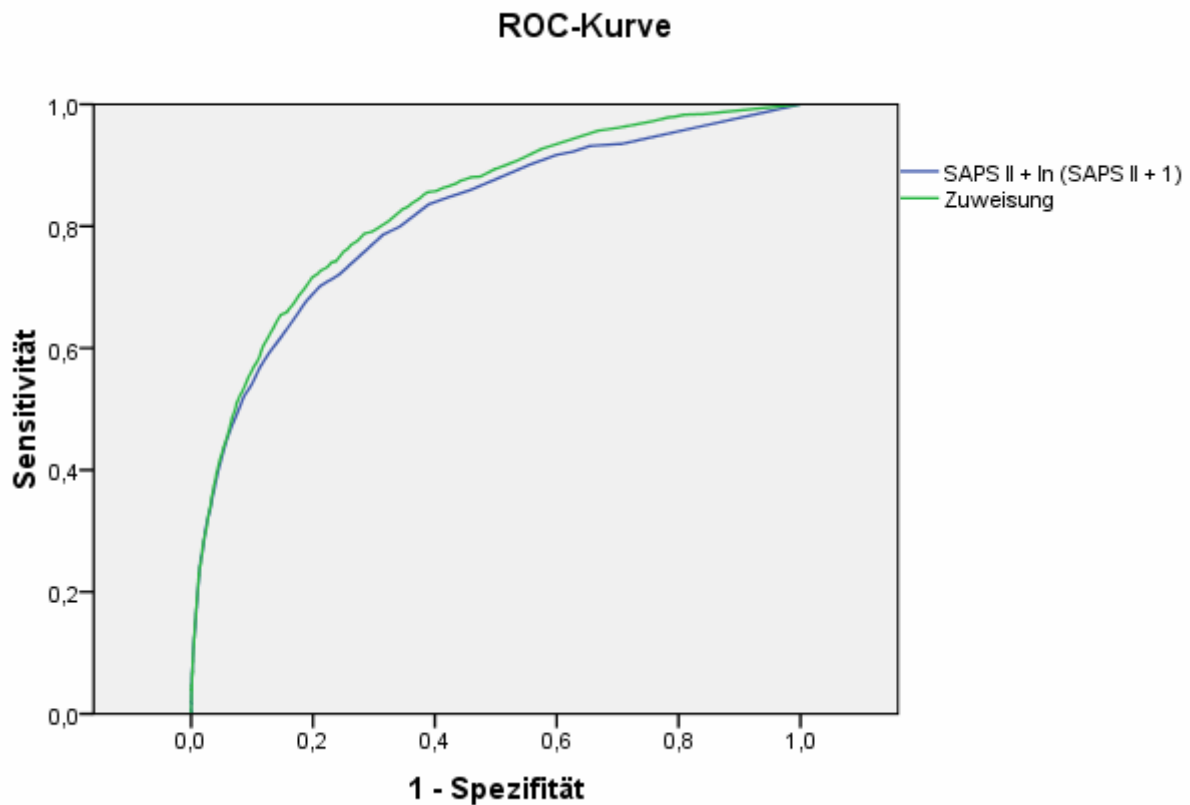
### 3.7 SAPS-DIVI 2 (Phys) im Validierungsdatensatz

#### 3.7.1 Auswertung (s. Anhang S. 102, Tabelle 23)

Die Fläche unter der ROC-Kurve beträgt bei Betrachtung der „Ausgangsvariablen“ SAPS-II-Score +  $\ln(\text{SAPS-II-Score} + 1)$  des SAPS-DIVI 2 (Phys) im Validierungsdatensatz 0,813. Bis auf der Variablen Geschlecht (m) (0,813) ist es bei allen anderen Parametern zu einer Zunahme der Fläche unter der ROC-Kurve gekommen. Folglich verzeichnen die Variablen Neuro (0,826) (s. Bild 20) und Zuweisung (0,832) (s. Bild 21) den größten Zuwachs, wobei es hingegen zu nur einer kleinen Vergrößerung der Fläche bei den Parametern Trauma (0,814), Vorerkrankungen (alle) (0,814) sowie Vorerkrankungen (0,815) und einer etwas größeren Zunahme bei Alter (0,822) und Aufnahme von Extern (0,822) gekommen ist.

**Bild 20** ROC-Kurve - Neuro



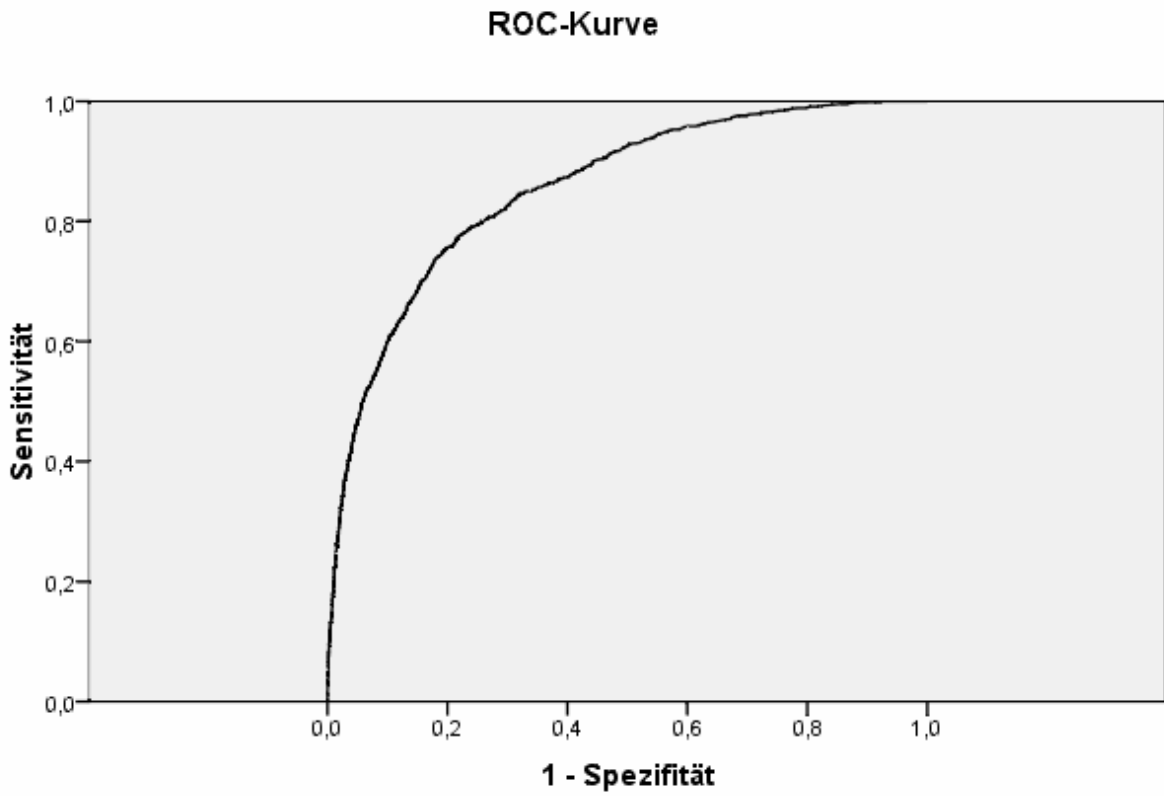
**Bild 21** ROC-Kurve - Zuweisung

Bei Anwendung des Kriteriums Präzision fällt ebenfalls eine Übereinstimmung zwischen der erwarteten (6,3 %) und beobachteten Letalität (6,3 %) auf, da sich im Mittel genau die Gesamtmortalität als Prognose ergibt, wobei aber nicht zwischen Überlebenden und Verstorbenen unterschieden wird.

### 3.7.2 Zusammenfassung aller Variablen

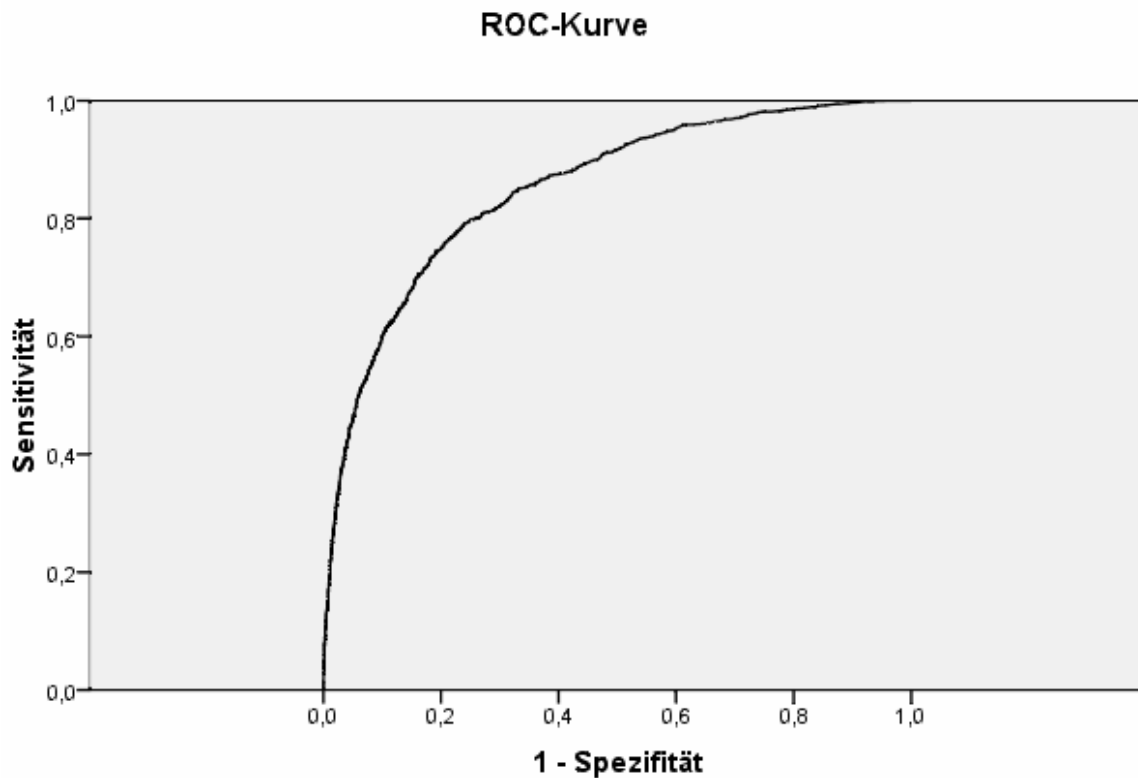
Die Fläche unter der ROC-Kurve ist ebenfalls wie im Validierungsdatensatz beim SAPS-DIVI 2 größer. Sie beträgt hier 0,855 (s. Bild 22) im Vergleich zum Entwicklungsdatensatz, wo sie bei 0,853 liegt. Somit werden die mit dem SAPS-DIVI 2 (Phys) im Entwicklungsdatensatz ermittelten Ergebnisse bestätigt.

**Bild 22** ROC-Kurve – Zusammenfassung aller Variablen



Bei Anwendung der Variablen Neuro statt Fachbereiche kommt es zu einer leichten Abnahme der Fläche unter der ROC-Kurve auf 0,851 (s. Bild 23).

**Bild 23** ROC-Kurve - Zusammenfassung aller Variablen (Neuro statt Fachbereiche)



## **3.8 Vergleich SAPS II (Original) – SAPS-DIVI – SAPS-DIVI 2**

### **3.8.1 Einleitung**

Um die Aussagekraft eines Score-Systems zu beurteilen, ist es wichtig diesen mit anderen Scores zu vergleichen. Zum Vergleich ist es dabei sinnvoll Gütekriterien anzuwenden wie z.B. Diskriminierung, Präzision sowie Kalibrierung.

In dieser Studie wurde der von Le Gall et al. im Jahre 1993 publizierte SAPS-II-Score mit dem SAPS-DIVI-Score sowie nun anhand der DIVI-Datenbank weiterentwickeltem SAPS-DIVI-2-Score verglichen.

Insgesamt wurden vier SAPS-DIVI-2-Formeln erarbeitet, dabei wurden im Allgemeinen weitere Variablen zum bisherigen SAPS-II-Score ergänzt. Einerseits wurden im Score nur die physiologischen Parameter berücksichtigt, alle weiteren Parameter wie unter anderem Alter, Vorerkrankungen und Zuweisung wurden daneben betrachtet. Andererseits wurde anstelle der Variablen „Neurologie/Neurochirurgie“ der Parameter „Fachbereiche“ angewendet, um zu überprüfen, ob es möglicherweise neben der Neurologie/Neurochirurgie Fachgebiete mit einer aussagekräftigen Signifikanz gibt, die bei der Ermittlung der Prognose berücksichtigt werden sollten. Es ist jedoch zu erwähnen, dass die Variable „Neurologie/Neurochirurgie“ eigentlich nicht im originalen SAPS-II-Score vorkommt, diese wurde jedoch im DIVI-Score anstelle des Parameters „GCS“ aufgrund der hohen Beobachtersvariation bei der Bestimmung der GCS bei sedierten und beatmeten Patienten erfasst. Zudem bezieht sich die mit der SAPS-II-Formel nach Le Gall et al. berechnete Prognose auf das Krankenhaus-Outcome, die mittels des SAPS-DIVI-Score bzw. SAPS-DIVI-2-Score ermittelte Vorhersage hingegen auf das Intensiv-Outcome.

### **3.8.2 Vergleich der Gütekriterien**

Die Gütekriterien wurden jeweils in einem Entwicklungs- sowie Validierungsdatensatz verglichen. Dabei wurde ein Augenmerk auf die Diskriminierung, Präzision sowie Kalibrierung geworfen.

### **3.8.3 Diskriminierung (s. Anhang S. 103, Tabelle 24)**

Beim Vergleich des originalen SAPS II mit SAPS-DIVI sowie SAPS-DIVI 2 fällt auf, dass die Fläche unter der ROC-Kurve beim SAPS II in der Entwicklungsgruppe 0,834 beträgt. Beim SAPS-DIVI hingegen beträgt die Fläche unter der ROC-Kurve 0,842 und beim SAPS-DIVI 2 0,854.

Wenn man anstelle der Variablen Fachbereiche den Parameter Neuro berücksichtigt, beträgt die Fläche im Entwicklungsdatensatz unter der ROC-Kurve 0,851. Bei Berücksichtigung nur der physiologischen Parameter beträgt die Fläche unter der ROC-Kurve dagegen 0,853, bei Anwendung der Variablen Neuro statt Fachbereiche bei der physiologischen Variante 0,851. Bei Betrachtung nur der Parameter SAPS-II-Score sowie  $\ln(\text{SAPS-II-Score} + 1)$  beim SAPS-DIVI 2, also ohne das Hinzufügen weiterer Variablen, beträgt die Fläche im Entwicklungsdatensatz unter der ROC-Kurve 0,834, beim SAPS-DIVI 2 (Phys) hingegen im Entwicklungsdatensatz 0,815.

Anhand der ROC-Kurve kann man die Bedeutung der Ergänzung weiterer neuer Variablen zum bisherigen SAPS-II-Score erkennen. Durch das Hinzufügen weiterer Parameter nimmt die Fläche unter der ROC-Kurve zu, was zugleich bedeutet, dass die Fähigkeit zu diskriminieren, also zwischen Überlebenden und Verstorbenen zu differenzieren, besser wird. Folglich besitzt der SAPS-DIVI 2 im Vergleich zum originalen SAPS II und SAPS-DIVI die beste Fähigkeit zu diskriminieren.

Die im Entwicklungsdatensatz ermittelten Ergebnisse werden im Validierungsdatensatz bestätigt. Die Fläche unter der ROC-Kurve in der jeweiligen Validierungsgruppe ist im Wesentlichen gleich groß.

Je größer die Differenz zwischen Überlebenden und Verstorbenen ist, umso besser kann ein Score diskriminieren. Diese nimmt jedoch vom originalen SAPS-II-Score zum SAPS-DIVI 2 hin ab, was bedeutet, dass der originale SAPS-II-Score in diesem Aspekt besser abschneidet als SAPS-DIVI und SAPS-DIVI 2.

Schließlich kommt es durch das Ergänzen weiterer Variablen zum SAPS II zu einer Verbesserung der Diskriminierung, was sich durch eine Zunahme der Fläche unter der ROC-Kurve äußert.

#### **3.8.4 Präzision (s. Anhang S. 103, Tabelle 24)**

Bezüglich der Präzision fällt ins Augenmerk, also dem Verhältnis tatsächlich beobachteter Mortalität zur vorhergesagten Mortalität, dass die Differenz zwischen beobachteter Mortalität und Score-Prognose vom originalen SAPS II zum SAPS-DIVI 2 hin kleiner wird. Folglich nähert sich die Übereinstimmung der Prognose des SAPS-DIVI-2-Scores dem tatsächlichen Outcome.

### 3.8.5 Kalibrierung

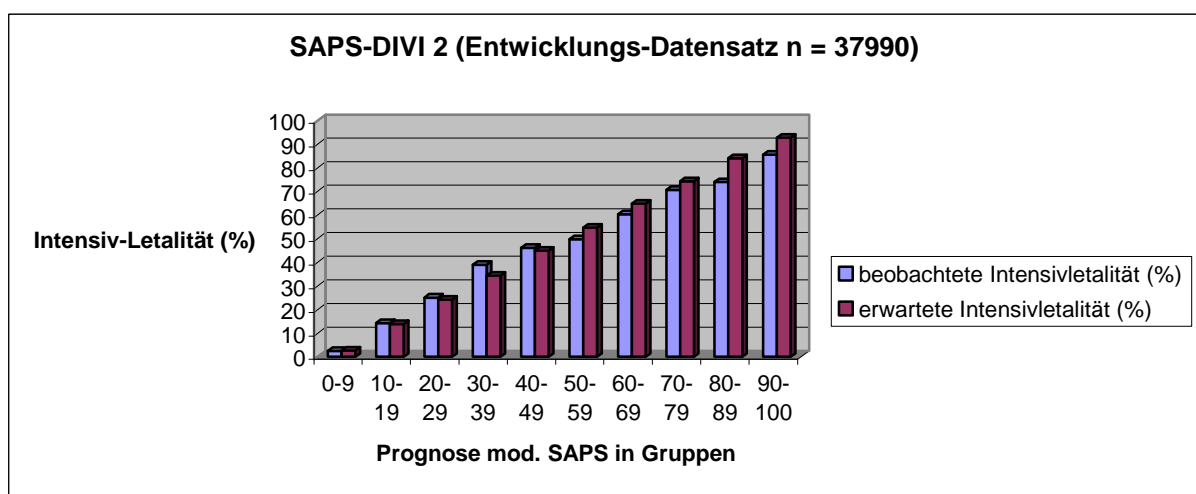
Die Präzision legt lediglich die Kongruenz von Prognose und Outcome im gesamten Kollektiv dar. Wichtig ist jedoch auch zu untersuchen, ob die Prognosen über den gesamten Bereich der Vorhersagewahrscheinlichkeiten gleichmäßig genau sind. Diese kann man mittels der „Goodness-of-Fit“ Statistik nach Hosmer und Lemeshow ermitteln. Hierbei wird das Patientenkollektiv nach der vorhergesagten Wahrscheinlichkeit sortiert und in zehn Gruppen aufgeteilt, wobei entweder nach der Prognosewahrscheinlichkeit oder nach der Fallzahl zehn etwa gleich große Kohorten geformt werden.

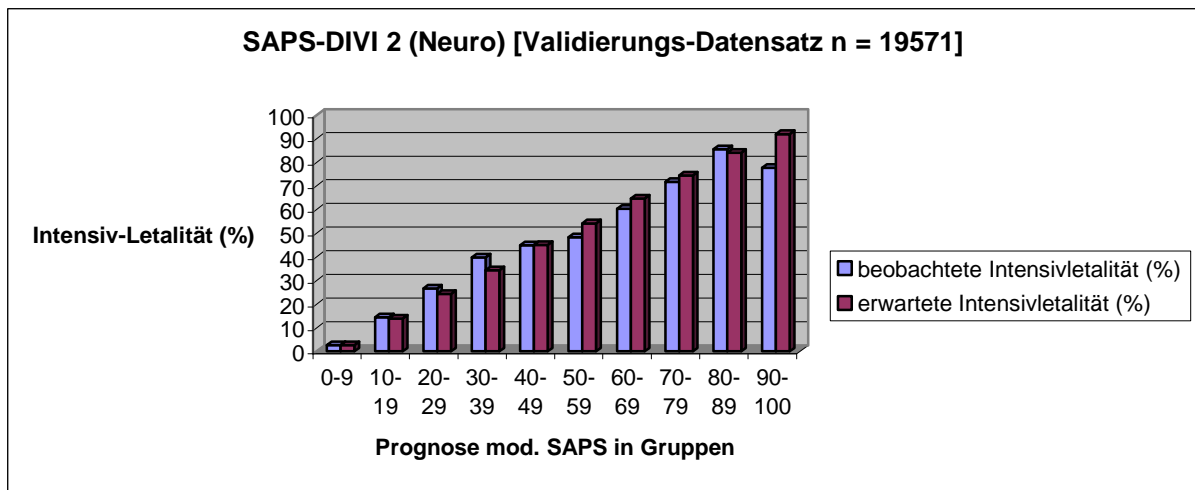
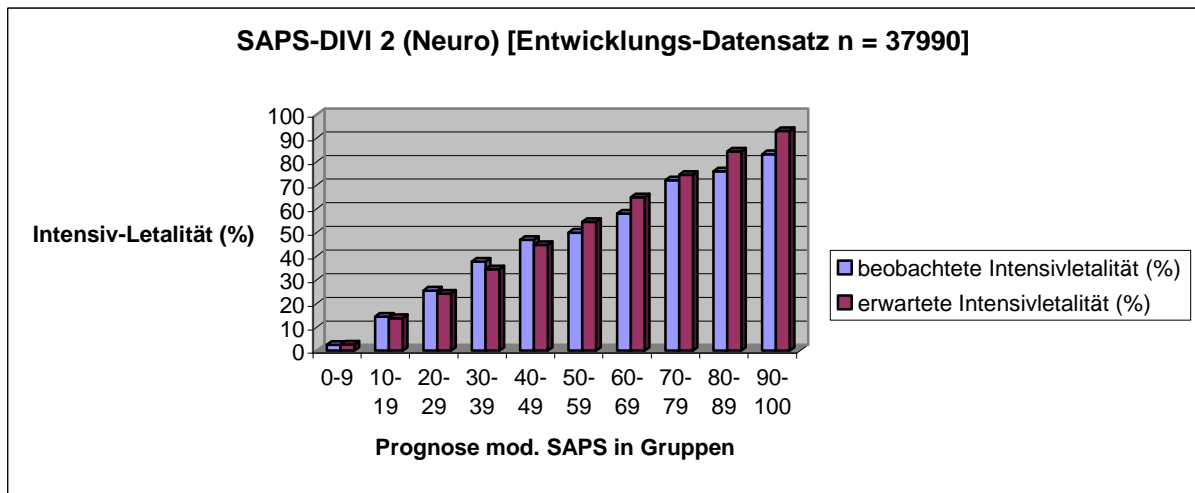
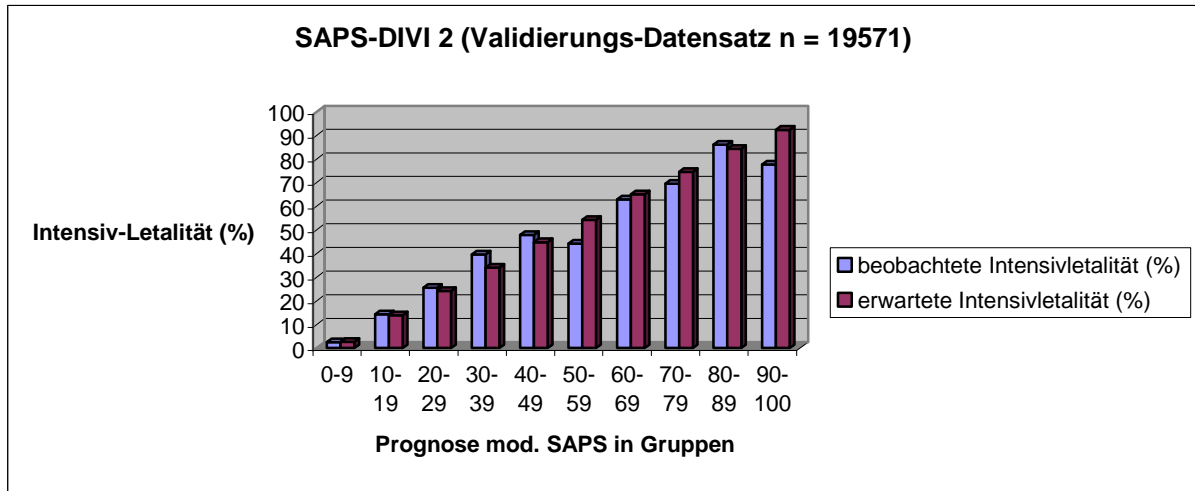
In dieser Studie wurden zehn Gruppen entsprechend der Prognosewahrscheinlichkeit gebildet. Anschließend wurde der Mittelwert der vorhergesagten Letalität in jeder Gruppe ermittelt und gemeinsam mit der tatsächlich beobachteten Letalität in jeder Gruppe grafisch dargestellt.

Wegen der großen Fallzahl in dieser Studie wurde bewusst auf eine Kalibrierung nach Hosmer und Lemeshow verzichtet, denn geringe Unterschiede stellen sich bereits als signifikant dar. Daher wurde die Kalibrierung lediglich grafisch dargestellt.

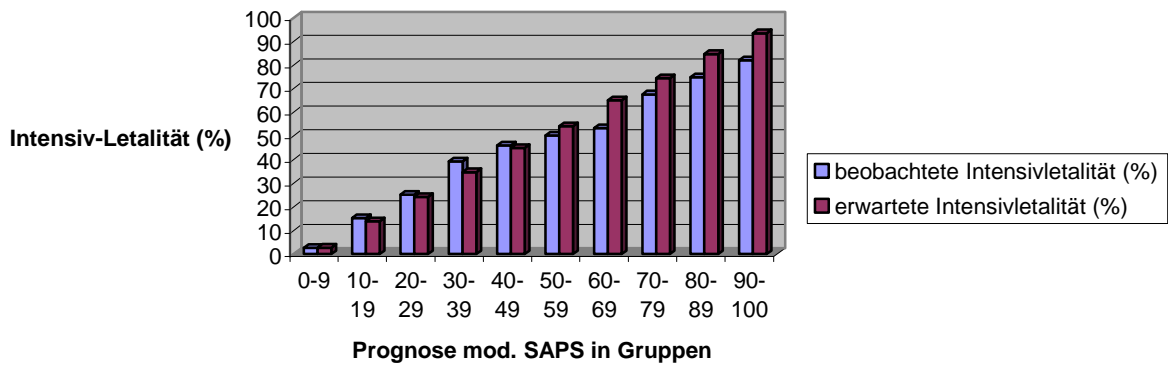
Hierbei fällt auf, dass die vorhergesagte sowie tatsächlich beobachtete Mortalität im Großen und Ganzen linear zunimmt. Des Weiteren stellen sich Abweichungen zwischen vorhergesagter und tatsächlicher Letalität besonders in den Gruppen mit hoher Prognosewahrscheinlichkeit dar. Die Kalibrierung ist somit besonders in den höheren Risikodezilen unzureichend, es kommt zu einer Überschätzung der Mortalität, da die erwartete Intensivletalität hier über der beobachteten Letalität liegt.

**Bild 24** Grafische Darstellung der Kalibrierung beim SAPS-DIVI 2 und SAPS-DIVI 2 (Phys)

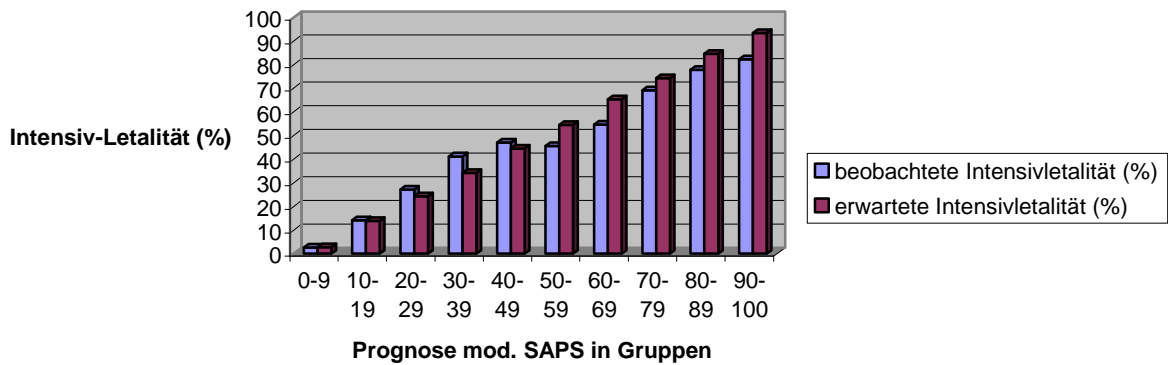




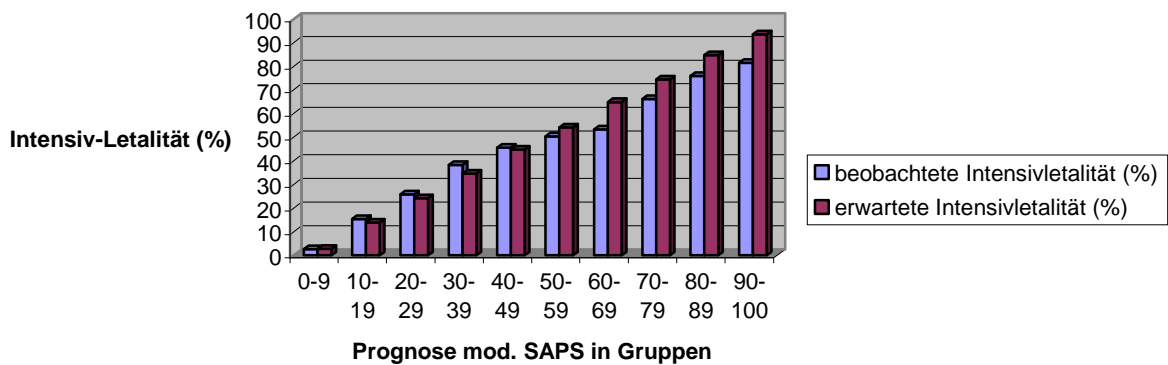
**SAPS-DIVI 2 (Phys) [Entwicklungs-Datensatz n = 37990]**



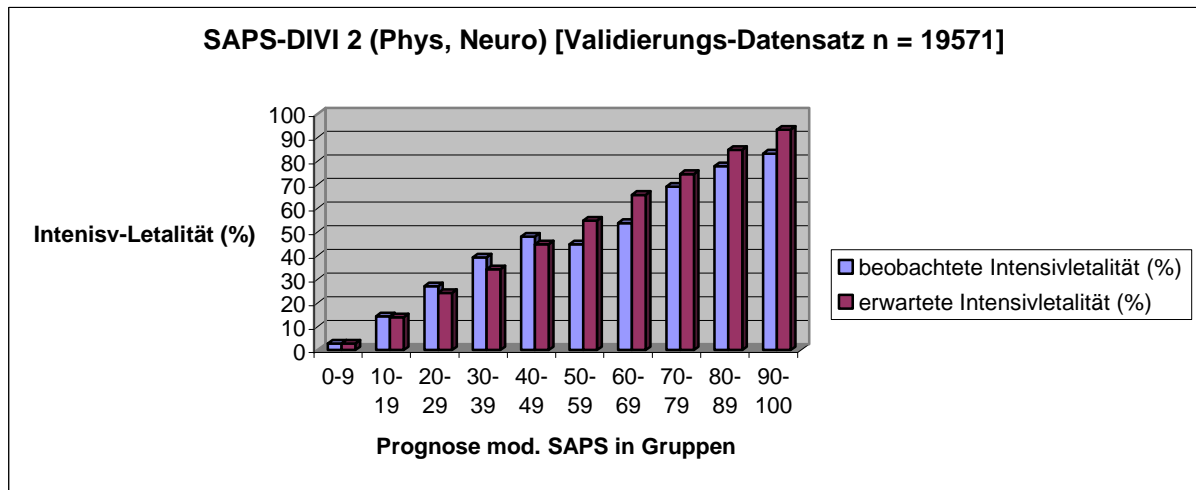
**SAPS-DIVI 2 (Phys) [Validierungs-Datensatz n = 19571]**



**SAPS-DIVI 2 (Phys, Neuro) [Entwicklungs-Datensatz n = 37990]**







### 3.8.6 Zusammenfassung

Der SAPS II ist im Verlaufe der Zeit seit Publizierung veraltet. Durch Modifizierung des SAPS II, also durch das Hinzufügen sowie Anpassen vorhandener Variablen kann man eine wesentlich bessere Diskriminierung erreichen. Diese muss jedoch ständig durchgeführt werden, um das Scoresystem an die aktuellen Gegebenheiten anzupassen und zu optimieren. Die Bedeutung neuer Variablen sowie die Angleichung existierender Parameter an die aktuellen Gegebenheiten wird vor allem deutlich, wenn man  $\text{SAPS-II-Score} + \ln(\text{SAPS-II-Score} + 1)$  ohne das Ergänzen weiterer Faktoren betrachtet. Durch das Hinzufügen von weiteren Variablen erkennt man schon einen enormen Anstieg der Fläche unter der ROC-Kurve, also eine effektivere Vorhersagekraft der Letalität, die mit einer höheren Spezifität sowie Sensitivität einhergeht.

Die Intention des SAPS II ist es die Mortalitätswahrscheinlichkeit wiederzugeben. Diese Absicht sollte auch weiterhin verfolgt werden, jedoch mit dem Hintergrund die Vorhersagekraft durch das Abstimmen an die aktuellen Umstände ständig zu verbessern. Dieses kann einerseits, wie in dieser Studie, durch das Beifügen weiterer neuer Variablen oder durch die Anpassung vorhandener Parameter geschehen, andererseits durch eine radikalere Überarbeitung des Scoresystems mit Einführung einer neuen Version, nämlich wie bereits geschehen, den SAPS 3.

## 4 Diskussion

In dieser Studie wurde versucht die Vorhersagekraft des SAPS II durch eine spezifische Anpassung an die DIVI-Datenbank zu optimieren. Dies erfolgte durch die Adjustierung des SAPS II sowie durch das Ergänzen weiterer Variablen zum SAPS II. Einerseits wurden nur die physiologischen Parameter im SAPS II berücksichtigt und alle anderen sowie neue Variablen daneben berücksichtigt, andererseits wurden die bereits im SAPS II vorkommenden Parameter Vorerkrankungen, Zuweisung sowie Alter erneut und neue Variablen neben SAPS-II-Score +  $\ln(\text{SAPS-II-Score} + 1)$  betrachtet. Die normalerweise im SAPS II vorkommende GCS wurde auf Grund der hohen Beobachtervariation bei sedierten und beatmeten Patienten nicht ermittelt und somit unberücksichtigt gelassen. Stattdessen wurde in dieser Studie der neurologische Status durch die Variable Neuro erfasst.

Zu den neuen Variablen gehören Trauma, Geschlecht, Fachbereiche, Neuro sowie Aufnahme von Extern. Außerdem bezieht sich die mit dem originalen SAPS-II-Score ermittelte Prognose auf das Krankenhaus-Outcome, die mittels SAPS-DIVI 2 berechnete Vorhersage hingegen auf das Intensiv-Outcome.

Beim SAPS-DIVI 2 handelt es sich um einen SAPS-II-Score, der neben den im originalen SAPS-II-Score vorkommenden Parametern erneut die Variablen Alter, Vorerkrankungen und Zuweisung sowie weitere neue Variablen wie Trauma, Neuro, Aufnahme von Extern, Geschlecht und Fachbereiche neben den Koeffizienten SAPS-II-Score und  $\ln(\text{SAPS-II-Score} + 1)$  berücksichtigt.

Während beim SAPS-DIVI 2 im Entwicklungsdatensatz besonders die Variablen Neuro und Zuweisung ins Augenmerk fallen (s. Anhang S. 95, Tabelle 11), sind es im Validierungsdatensatz vor allem die Parameter Neuro sowie Aufnahme von Extern (s. Anhang S. 101, Tabelle 22). Somit bekräftigt der Validierungsdatensatz des SAPS-DIVI 2 die im Entwicklungsdatensatz ermittelte Signifikanz dieser Variablen und unterstreicht deren Bedeutung beim SAPS-II-Score. Sie haben einen entscheidenden Einfluss bei der Ermittlung der Intensivletalität und sind somit nicht zu vernachlässigende Faktoren, die eine erhebliche Wirkung auf die mit dem SAPS-II-Score ermittelte Mortalitätswahrscheinlichkeit haben.

Bei der Variablen Aufnahme von Extern handelt es sich um Patienten, die von außerhalb wahrscheinlich aufgrund ihrer komplizierten Krankheitslage sowie der begrenzten Diagnose- und Therapiemöglichkeiten aufgenommen wurden. Infolge der zeitlichen Verzögerung kommt es zu einer negativen Wirkung auf das Outcome der Patienten. Diese Patienten besitzen somit eine wesentlich schlechtere Ausgangssituation als solche Patienten, die ohne

Verlegung direkt im selben Krankenhaus bzw. Klinikum behandelt werden, was sich auch beim SAPS-DIVI 2 anhand der statistischen Signifikanz bemerkbar macht.

Von großem Interesse ist auch der Parameter Zuweisung, denn diese Variable ist bereits im SAPS-II-Score enthalten und wurde wie die anderen Variablen neben SAPS-II-Score und  $\ln(\text{SAPS-II-Score} + 1)$  betrachtet. Der Regressionskoeffizient ist positiv, obwohl dieser Parameter bereits im SAPS-II-Score enthalten ist. Die Variable weist beim SAPS-DIVI 2 im Entwicklungsdatensatz eine statistische Signifikanz auf. Dieses bedeutet, dass man den Subparametern medizinisch (ohne OP) und ungeplant chirurgisch der Variablen Zuweisung mehr Punkte vergeben könnte, da die Unterscheidung in medizinisch (ohne OP) und operativer Notfall anscheinend eine wesentlich größere Rolle auf die Letalität spielt als es im Hinblick auf die Punktevergabe im SAPS-II-Score widerspiegelt wird. Die Aufnahmeart scheint in der Intensivmedizin eine wichtige Funktion zu haben, denn der Schweregrad der invasiven Eingriffe (OP ja/nein) in der Initialphase scheint das Outcome des Patienten entscheidend zu beeinflussen. Hierbei könnten der OP-Eingriff und die OP-Dauer interessant sein, was in der Initialphase des Patienten sowohl aus diagnostischer als auch aus therapeutischer Sicht sicherlich einen entscheidenden Einfluss auf das Outcome des Patienten hat, wobei jedoch auch nicht vergessen werden darf, dass im SAPS-II-Score bereits Punkte für die Variable Zuweisung enthalten sind und der Parameter daher mit Vorsicht interpretiert werden muss.

Weiterhin ist auffällig, dass der Parameter Alter eine Signifikanz aufweist, die mit zunehmendem Alter zunimmt, wobei der Regressionskoeffizient ein negatives Vorzeichen besitzt, obwohl das Alter bereits im SAPS-II-Score berücksichtigt ist. Dieses bedeutet, dass das Alter überbewertet wird und heutzutage nicht mehr so eine große Rolle spielt wie es z.B. vor mehr als 19 Jahren der Fall war, was daran liegen mag, dass sich die Intensivmedizin im Verlaufe der Jahre weiterentwickelt und Fortschritte bei der Diagnostik und Therapie erzielt hat und folglich mit älteren Patienten besser zurecht kommt. Somit ist die Punktevergabe nach dem originalen SAPS-II-Score mittlerweile zu hoch eingestuft, so dass man bei der Punkteverteilung bei der Variablen Alter weniger Punkte vergeben könnte oder zumindest erst bei einer höheren Altersstufe beginnt Punkte zu vergeben. Auch hier muss allerdings wie bei dem Parameter Zuweisung berücksichtigt werden, dass im SAPS-II-Score bereits Punkte für die Variable Alter enthalten sind.

Die Variable Vorerkrankungen weist auch einen Regressionskoeffizienten mit negativem Vorzeichen auf, wobei dieser Parameter jedoch aufgrund der niedrigen Fallzahl mit Vorsicht interpretiert werden sollte.

Mit den noch vor ungefähr 19 Jahren von Le Gall et al. im SAPS-II-Score eingeführten Vorerkrankungen metastasierendes Karzinom, maligne hämatologische Erkrankungen und AIDS kommt man heutzutage wahrscheinlich aufgrund der zur Verfügung stehenden besseren Diagnose- und Therapiemöglichkeiten besser zurecht, so dass man dieser Variable weniger Punkte vergeben könnte oder der Einfachheit halber sogar als dichotome Variable zusammenfasst (Vorerkrankungen ja/nein= 1/0), was auch in dieser Studie gemacht wurde. Dadurch erhält man eine höhere Patientenzahl, so dass eine bessere Aussagekraft über die Signifikanz dieser Variablen gemacht werden kann. Alternativ zur Zusammenfassung dieses Parameters als dichotome Variable, könnte die Punktevergabe bei den einzelnen Subvariablen verändert oder neue Krankheitsbilder eingebaut bzw. die im SAPS-II-Score enthaltenen Vorerkrankungen durch andere Vorerkrankungen ersetzt werden.

Bei den Variablen Trauma (ja/nein) und Geschlecht (weiblich/männlich) ist der Regressionskoeffizient beim SAPS-DIVI 2 im Entwicklungsdatensatz zwar statistisch signifikant, jedoch ist er bei Betrachtung des 95 %-Konfidenzintervalls relativ nah an der „0“, so dass diese beiden Variablen einen nicht allzu sehr großen Einfluss auf die mit dem SAPS-II-Score ermittelte Letalität haben. Hier sollte auch die große Fallzahl nicht vergessen werden, denn kleine Unterschiede führen schon bei einem großen Datensatz zu einer statistischen Signifikanz. Folglich haben Männer eine etwas schlechtere Ausgangssituation als Frauen, ebenso Patienten, die mit einem Trauma auf die Intensivstation kommen im Vergleich zu Patienten ohne Trauma. Beim SAPS-DIVI 2 im Validierungsdatensatz hingegen hat sich bei der Variablen Geschlecht die statistische Signifikanz nicht bestätigt, so dass Frauen und Männer dieselbe Ausgangsprognose haben und somit keines von beiden Geschlechtern bei der Ermittlung der Letalität mittels SAPS II benachteiligt ist. Aus diesem Grund spielt die Variable Geschlecht im SAPS-II-Score zur Ermittlung der Letalität von Intensivpatienten keine Rolle. Bei allen anderen Parametern dagegen ist es im Validierungsdatensatz beim SAPS-DIVI 2 zu einer Zunahme der Fläche unter der ROC-Kurve gekommen, was die im Entwicklungsdatensatz mittels SAPS-DIVI 2 ermittelte Signifikanz bekräftigt.

Die von uns ermittelte Konstante beim SAPS-DIVI 2 im Entwicklungsdatensatz (-8,421) ist negativer und somit kleiner als die Konstante im originalen SAPS-II-Score (-7,7631). Dieses würde bedeuten, dass die Patienten von vornherein eine bessere Ausgangssituation haben als nach der originalen SAPS-II-Formel, da z.B. bei der Annahme von einem SAPS-II-Score von 0 Punkten die Mortalitätswahrscheinlichkeit niedriger liegt. Dieses würde ebenfalls darauf hindeuten, dass die Intensivmedizin Fortschritte erzielt hat, wodurch sich eine bessere Ausgangssituation für die Patienten ergibt.

Mit Ausgangssituation ist gemeint, dass gewisse Patientengruppen wie beispielsweise neurologische oder alte Patienten von vornherein ein schlechteres Outcome haben als Patienten, die keine neurologischen Auffälligkeiten aufweisen oder als Patienten, die jung sind.

Neben dem Hinzufügen nur einer einzelnen Variablen zum SAPS-II-Score wurden auch alle Variablen gemeinsam beim SAPS-DIVI 2 im Entwicklungsdatensatz betrachtet. Dabei wurden die beiden Parameter Trauma und Geschlecht aufgrund ihrer Bedeutungslosigkeit im Gesamtkollektiv im SAPS-II-Score nicht berücksichtigt. Hervortreten dagegen tun die Variablen Alter und Fachrichtungen (s. Anhang S. 97, Tabelle 14). Während die Signifikanz beim Parameter Alter mit zunehmendem Alter steigt, ragen die Fachbereiche Neurologie und Neurochirurgie mit positivem Koeffizienten sowie die Fachgebiete MKG-Chirurgie, HNO, Urologie, Orthopädie und Gynäkologie mit negativem Regressionskoeffizienten heraus. Damit spielen gewisse Fachrichtungen eine wichtigere Rolle bei der Ermittlung der Intensivletalität mittels SAPS II als andere Fachbereiche.

Die Variablen Zuweisung, Vorerkrankungen (alle) und Aufnahme von Extern hingegen weisen bei Betrachtung aller Variablen zusammen einen kleineren Regressionskoeffizienten auf, sind jedoch ebenfalls statistisch signifikant.

Wenn man anstelle der Variablen Fachbereiche den Parameter Neuro anwendet, da die Variable Neuro die Subparameter „Plus-Fächer“ (Neurologie, Neurochirurgie) widerspiegelt, kommt es nicht zu einer bemerkenswerten Veränderung (s. Anhang S. 97, Tabelle 15).

All diese Variablen leisten bei der Ermittlung der Intensivletalität mit dem SAPS-II-Score einen entscheidenden Beitrag, was an den ermittelten Ergebnissen kenntlich wird.

An neuen Parametern im Vergleich zum originalen SAPS-II-Score sind Aufnahme von Extern, Fachbereiche sowie Vorerkrankungen (alle) hinzugekommen. Es handelt sich dabei um mit geringem Aufwand zu erhebende Parameter, die keinen zusätzlichen Zeitaufwand benötigen und deshalb zur Berechnung der SAPS-II-Letalität mit einbezogen werden sollten.

Der SAPS-DIVI 2 (Phys) enthält nur physiologische Parameter. Die Variablen Vorerkrankungen, Zuweisung sowie Alter wurden genauso wie die neuen Parameter Trauma, neurologische bzw. neurochirurgische Patienten, Aufnahme von Extern, Geschlecht sowie Fachbereiche jeweils einzeln und zusammen neben SAPS-II-Score und  $\ln(\text{SAPS-II-Score} + 1)$  betrachtet.

Beim SAPS-DIVI 2 (Phys) fallen sowohl im Entwicklungs- (s. Anhang S. 99, Tabelle 17) als auch im Validierungsdatensatz (s. Anhang S. 102, Tabelle 23) vor allem die Parameter Zuweisung und Neuro auf.

Diese Variablen nehmen eine überragende Rolle bei der Bestimmung der Intensivletalität ein, der Regressionskoeffizient ist um ein Vielfaches größer als ihr Standardfehler und im 95 %-Konfidenzintervall ist man von der Zahl „0“ relativ weit entfernt.

Die Regressionskoeffizienten weichen bei nur Betrachtung der Variablen SAPS-II-Score und  $\ln(\text{SAPS-II-Score} + 1)$  beim originalen SAPS-II-Score und beim SAPS-DIVI 2 (Phys) im Entwicklungsdatensatz stark voneinander ab, was auf die Wichtigkeit der Parameter Alter, Vorerkrankungen und Zuweisung hindeutet, die im SAPS-DIVI 2 (Phys) nicht enthalten sind. Die alleinige Anwendung von physiologischen Parametern genügt nicht, um die Letalität von Intensivpatienten präzise anzugeben.

Die Konstante beim SAPS-DIVI 2 (Phys) im Entwicklungsdatensatz beträgt -4,366 und ist somit positiver als die Konstante beim originalen SAPS-II-Score mit -7,7631, was zur Folge hat, dass die Prognose bei Annahme von einem SAPS-II-Score von 0 Punkten schlechter ist als beim originalen SAPS-II-Score.

Dementsprechend haben Patienten schon vor Aufnahme auf die Intensivstation eine schlechtere Ausgangssituation als nach der originalen SAPS-II-Formel. Folglich hat die alleinige Berücksichtigung der physiologischen Parameter im SAPS-II-Score nicht so eine große Aussagekraft zur Ermittlung der Mortalitätswahrscheinlichkeit wie mit dem originalen SAPS-II-Score. Die drei Variablen Alter, Vorerkrankungen und Zuweisung sind somit nicht zu vernachlässigende Faktoren, da sie einen nicht zu unterschätzenden Einfluss auf den SAPS-II-Score haben, was auch für die Variable Neuro gilt, die anstelle der GCS verwendet wurde und den neurologischen Status berücksichtigt.

Von all diesen Variablen hebt sich vor allem der Parameter Aufnahmegrund hervor. Er scheint den größten Einfluss auf die SAPS-II-Letalität zu haben, was unter anderem daran liegen mag, dass er „indirekt“ den Schweregrad der Erkrankung wiedergibt. Denn ein Patient, der ungeplant operiert wird, weist eine größere Erschütterung des Organismus auf als ein Patient, bei dem konservative Maßnahmen ausreichend sind und eine geringere Invasivität notwendig ist. Durch die Erschütterung des Organismus werden neben dem physiologischen Elektrolythaushalt auch zahlreiche weitere Faktoren in ihrer Funktion beeinträchtigt.

Bei Zusammenfassung aller Variablen werden wie beim SAPS-DIVI 2 die Variablen Trauma und Geschlecht nicht berücksichtigt. Ebenso fallen hier vor allem die Parameter Alter, Fachbereiche und der Aufnahmegrund ins Augenmerk (s. Anhang S. 100, Tabelle 20).

Wenn man anstelle der Variablen Fachbereiche den Parameter Neuro anwendet, kommt es nicht zu einer wesentlichen Veränderung (s. Anhang S. 101, Tabelle 21).

Der SAPS-II-Score wurde 1993 von Le Gall et al. eingeführt und ist im Laufe der Zeit veraltet. Er spiegelt nicht mehr die aktuellen Gegebenheiten wider und muss kontinuierlich angepasst und aktualisiert werden, um die Vorhersagekraft der Letalität aufrechtzuerhalten und zu optimieren. Dies kann auf der einen Seite durch eine Modifizierung der im SAPS II enthaltenen Parameter (Ersetzen von Variablen bzw. Anpassung der Punktevergabe bei den einzelnen Variablen) geschehen und/oder durch das Hinzufügen weiterer Variablen zum SAPS-II-Score.

In dieser Studie wurde eine spezifische Anpassung des SAPS II anhand der DIVI-Datenbank durch das Hinzufügen von weiteren Variablen durchgeführt. Dabei stellte sich fest, dass das Ergänzen nicht jeder Variablen wie z.B. Trauma oder Geschlecht zum SAPS II sinnvoll erscheint, denn diese Parameter hatten bei Zusammenfassung aller Variablen keinen Einfluss auf die Vorhersagekraft der Mortalitätswahrscheinlichkeit von Intensivpatienten. Darüber hinaus fiel auf, dass Parameter wie beispielsweise das Alter eine Adjustierung durch eine Veränderung der Punktzahl erfahren sollten, indem entweder erst ab einem höheren Alter Punkte vergeben werden oder die Punktzahl in den verschiedenen Altersklassen modifiziert wird. Außerdem spielt die Variable Fachrichtungen neben der Variablen Alter bei Zusammenfassung aller Variablen sowohl beim SAPS-DIVI 2 als auch beim SAPS-DIVI 2 (Phys) eine wichtige Rolle, die einen nicht zu unterschätzenden Einfluss auf die mit dem SAPS II ermittelte Intensivletalität aufweist.

Schließlich bildet der SAPS II einen Score, der zur Beurteilung der Erkrankungsschwere dient und somit ein objektives Hilfsmittel darstellt, der bei medizinischen Entscheidungen unterstützen kann. Aus diesem Grund sollte er regelmäßig auf seine Gütekriterien hin überprüft und im Laufe der Jahre angepasst werden, um die Vorhersagekraft des SAPS II aufrechtzuerhalten und zu verbessern. Dieses kann einerseits, wie in dieser Studie, durch das Beifügen weiterer neuer Variablen oder durch die Anpassung vorhandener Parameter geschehen, andererseits durch eine radikalere Überarbeitung des Scoresystems mit Einführung einer neuen Version, nämlich wie bereits geschehen, den SAPS 3.

Der SAPS 3 ist in drei Teile gegliedert. Er berücksichtigt was der Patient „mitbringt“ (s. Anhang S. 106, Tabelle 25 [Box 1]), was zur Intensivaufnahme führte (s. Anhang S. 106, Tabelle 25 [Box 2]) sowie den aktuellen Zustand (Physiologie) des Patienten (s. Anhang S. 106, Tabelle 25 [Box 3]). Abhängig davon werden dem Patienten jeweils Punkte vergeben, um somit die Letalität zu ermitteln.

## 5 Zusammenfassung

In dieser Arbeit geht es um die spezifische Anpassung des SAPS II (Simplified Acute Physiology Score II) an die Zwecke und Ziele der Sektion Qualitätssicherung der DIVI (Deutsche Interdisziplinäre Vereinigung für Intensiv- und Notfallmedizin).

Der SAPS-II-Score stellt die Nachfolge von SAPS dar und dient zur Einschätzung der Erkrankungsschwere auf der Basis physiologischer Daten sowie zur Berechnung des Letalitätsrisikos.

Seit der Einführung des SAPS II sind 19 Jahre vergangen, er ist im Laufe der Jahre veraltet und muss an die aktuellen Gegebenheiten angepasst werden. Die Anpassung des SAPS II erfolgte in dieser Studie durch das Hinzufügen neuer Parameter wie Aufnahme von externen Krankenhäusern, Trauma, Geschlecht, Neuro sowie Fachbereiche zum SAPS II. Die normalerweise im originalen SAPS-II-Score vorkommende GCS (Glasgow Coma Scale) wurde auf Grund der hohen Beobachtersvariation bei sedierten und beatmeten Patienten nicht ermittelt und somit unberücksichtigt gelassen. Stattdessen wurde die Variable neurologische bzw. neurochirurgische Patienten verwendet, um den neurologischen Status einzubeziehen. Des Weiteren bezieht sich die mit dem originalen SAPS-II-Score ermittelte Prognose auf das Krankenhaus-Outcome, die mittels SAPS-DIVI 2 berechnete Vorhersage hingegen auf das Intensiv-Outcome.

Bei den neuen Variablen fiel insbesondere der Parameter Aufnahme von Extern ins Augenmerk, da er einen nicht zu unterschätzenden Einfluss auf die Letalität hat und daher im SAPS II berücksichtigt werden sollte. Außerdem spielen auch die beiden Parameter Alter und Zuweisung eine große Rolle bei der Bestimmung der Letalität mittels SAPS II.

Die Variablen Trauma und Geschlecht hingegen spielen keine wesentliche Rolle bei der Ermittlung der Mortalitätswahrscheinlichkeit beim SAPS II.

Der SAPS II ist ein Score, der die Erkrankungsschwere wiedergibt und sich somit als nützliches objektives Hilfsmittel bei subjektiv-emotionalen Entscheidungen eignet. Er kann also bei medizinischen Entscheidungen unterstützend eingesetzt werden. Damit seine Aussagekraft auch im Laufe der Jahre bestehen bleibt, muss er regelmäßig an die aktuellen Gegebenheiten angepasst werden, indem die Gütekriterien des SAPS II kontinuierlich verbessert werden oder eine gravierendere Modifizierung des Score-Systems erfolgt mit Einführung eines neuen Score-Modells, nämlich SAPS 3.



## **Abstract**

This dissertation is dealing with the specific adaption of SAPS II (Simplified Acute Physiology Score II) to the aims and purposes of the section quality assurance of DIVI (Deutsche Interdisziplinäre Vereinigung für Intensiv- und Notfallmedizin).

SAPS II represents the succession of SAPS and serves for disease severity based on physiologic information and calculates the probability of death.

Since, for the first introduction of SAPS II 19 years before, it has to be adapted to the current circumstances. In this study SAPS II was modified. Among others new parameter like admission from external clinics, trauma, sex, neuro and faculty added to SAPS II. The GCS (Glasgow Coma Scale) which was originally included in SAPS II was ignored because of high observer variation at sedated and ventilated patients. Instead of GCS neurologic respectively neurosurgical patients were considered to take neurologic status into account. Moreover original SAPS II computed prognosis refers to hospital-outcome, in contrast to SAPS-DIVI 2 calculating refers to intensive care unit outcome.

From the new variables especially admission from external clinics was associated with high influence on mortality, and therefore it should be integrated in SAPS II. Furthermore, age and type of admission are playing an important role in mortality of SAPS II.

The variables trauma and sex however do not play an important role in estimating death in SAPS II.

SAPS II is a score which reflects the degree of disease severity and for this reason it is a useful tool for subjective-emotional decisions, therefore it can be used as a support for medical decisions.

To uphold the validity of SAPS II over the years it has to be regularly adapted to the current circumstances by continuous improving quality criteria or by modification of the score-system with introducing a new score-model, namely SAPS 3.

## 6 Literaturverzeichnis

- [1] Agha A, Bein T, Fröhlich D, Höfler S, Krenz D, Jauch KW (2002). "[Simplified Acute Physiology Score" (SAPS II) in the assessment of severity of illness in surgical intensive care patients]." Chirurg 73(5): 439-442
- [2] Altman DG (1996). "Use of confidence intervals to indicate uncertainty in research findings." Evidence-Based Med 1: 102-104
- [3] Bahloul F, Le Gall JR, Loirat P, Alperovitch A, Patois E (1988). "[Prognostic factors in resuscitation]." Presse Med 17(34): 1741-1744
- [4] Baker SP, O'Neill B, Haddon W Jr, Long WB (1974). "The injury severity score: a method for describing patients with multiple injuries and evaluating emergency care." J Trauma 14(3): 187-196
- [5] Bastos PG, Knaus WA, Zimmerman JE, Magalhães A JR, Sun X, Wagner DP (1996). "The importance of technology for achieving superior outcomes from intensive care. Brazil APACHE III Study Group." Intensive Care Med 22(7): 664-669
- [6] Bein T, Unertl K (1993). "[Potentialities and limitations of the score system in intensive medicine]." Anesthesiol Intensivmed Notfallmed Schmerzther 28(8): 476-483
- [7] Bender R, Lange S (2007). "[What is a confidence interval?]." Dtsch Med Wochenschr 132 Suppl 1: e17-18
- [8] Bender R, Ziegler A, Lange S (2007). "[Logistic regression]." Dtsch Med Wochenschr 132 Suppl 1: e33-35
- [9] Bosman RJ, de Keizer NF (1999). "Prognostic performance and customization of the SAPS II." Intensive Care Med 25(8): 878-879
- [10] Chalfin DB, Cohen IL, Lambrinos J (1995). "The economics and cost-effectiveness of critical care medicine." Intensive Care Med 21(11): 952-961
- [11] Champion HR, Sacco WJ, Copes WS, Gann DS, Gennarelli TA, Flanagan ME (1989). "A revision of the Trauma Score." J Trauma 29(5): 623-629
- [12] Champion HR, Sacco WJ, Hunt TK (1983). "Trauma severity scoring to predict mortality." World J Surg 7(1): 4-11
- [13] Cleveland WS (1979). "Robust locally weighted regression and smoothing scatterplots." J Am Stat Assoc 74: 829-836
- [14] Cullen DJ (1977). "Results and costs of intensive care." Anesthesiology 47(2): 203-216

- [15] Fagon JY, Chastre J, Novara A, Medioni P, Gibert C (1993). "Characterization of intensive care unit patients using a model based on the presence or absence of organ dysfunctions and/or infection: the ODIN model." Intensive Care Med 19(3): 137-144
- [16] Frost P (2003). "In response to 'Effect of introducing the Modified Early Warning score on clinical outcomes, cardio-pulmonary arrests and intensive care utilisation in acute medical admissions', Subbe CP et al., *Anaesthesia* 2003; 58: 797-802." Anaesthesia 58(11): 1154.
- [17] Gardner MJ, Altman DG (1986). "Confidence intervals rather than P values: estimation rather than hypothesis testing." Br Med J (Clin Res Ed) 292(6522): 746-750
- [18] Glance LG, Osler TM, Dick A (2002). "Rating the quality of intensive care units: is it a function of the intensive care unit scoring system?" Crit Care Med 30(9): 1976-1982
- [19] Goldstein H (1995). "Multilevel statistical models." Arnold, London
- [20] Goldstein H, Spiegelhalter DJ (1996). "League tables and their limitations: statistical issues in comparisons of institutional performance." J R Stat Soc A 159: 385-443
- [21] Guerro VM, Johnson RA (1982). "Use of the Box-Cox transformation with binary response models." Biometrika 69(2): 309-314
- [22] Guyatt G, Jaeschke R, Heddle N, Cook D, Shannon H, Walter S (1995). "Basic statistics for clinicians: 2. Interpreting study results: confidence intervals." CMAJ 152(2): 169-173
- [23] Hanley JA, McNeil BJ (1983). "A method of comparing the areas under receiver operating characteristic curves derived from the same cases." Radiology 148(3): 839-843
- [24] Hanley JA, McNeil BJ (1982). "The meaning and use of the area under a receiver operating characteristic (ROC) curve." Radiology 143(1): 29-36
- [25] Hosmer DW, Lemeshow S (1989). "Applied Logistic Regression." John Wiley & Sons, New York
- [26] Hosmer DW, Lemeshow S (1989). "Applied Logistic Regression 2<sup>nd</sup> edition." John Wiley & Sons, New York
- [27] Hosmer DW, Lemeshow S (1995). "Confidence interval estimates of an index of quality performance based on logistic regression models." Stat Med 14(19): 2161-2172
- [28] Knaus WA, Wagner DP, Draper EA, Zimmerman JE, Bergner M, Bastos PG, Sirio CA, Murphy DJ, Lotring T, Damiano A, et al. (1991). "The APACHE III prognostic

- system. Risk prediction of hospital mortality for critically ill hospitalized adults." Chest 100(6): 1619-1636
- [29] Knaus WA, Draper EA, Wagner DP, Zimmerman JE (1986). "An evaluation of outcome from intensive care in major medical centers." Ann Intern Med 104(3): 410-418
- [30] Knaus WA, Draper EA, Wagner DP, Zimmerman JE (1985). "APACHE II: a severity of disease classification system." Crit Care Med 13(10): 818-829
- [31] Knaus WA, Draper EA, Wagner DP, Zimmerman JE, Birnbaum ML, Cullen DJ, Kohles MK, Shin B, Snyder JV (1982). "Evaluating outcome from intensive care: a preliminary multihospital comparison." Crit Care Med 10(8): 491-496
- [32] Knaus WA, Draper EA, Wagner DP, Zimmerman JE (1985). "Prognosis in acute organ-system failure." Ann Surg 202(6): 685-693
- [33] Knaus WA, Harrell FE, Fisher CJ Jr, Wagner DP, Opal SM, Sadoff JC, Draper EA, Walawander CA, Conboy K, Grasela TH (1993). "The clinical evaluation of new drugs for sepsis. A prospective study design based on survival analysis." JAMA 270(10): 1233-1241
- [34] Knaus WA., Zimmerman JE, Wagner DP, Draper EA, Lawrence DE (1981). "APACHE-acute physiology and chronic health evaluation: a physiologically based classification system." Crit Care Med 9(8): 591-597
- [35] Kuhlen R, Quintel M (2007). "Möglichkeiten und Grenzen der Intensivmedizin." In: Burchardi H, Larsen R, Kuhlen R, Jauch K.-W., Schölmerich J. (eds) Die Intensivmedizin. Springer, Berlin Heidelberg New York, pp 4-8
- [36] Lange S, Bender R (2007). "[Measures of variability]." Dtsch Med Wochenschr 132 Suppl 1: e5-6
- [37] Le Gall JR (2005). "The use of severity scores in the intensive care unit." Intensive Care Med 31(12): 1618-1623
- [38] Le Gall JR, Loirat P (1995). "Can we evaluate the performance of an intensive care unit?." Curr Opin Crit Care 1: 219-220
- [39] Le Gall JR, Neumann A, Hemery F, Bleriot JP, Fulgencio JP, Garriques B, Gouzes C, Lepage E, Moine P, Villers D (2005). "Mortality prediction using SAPS II: an update for French intensive care units." Crit Care 9(6): R645-652
- [40] Le Gall JR, Klar J, Lemeshow S, Saulnier F, Alberti C, Artigas A, Teres D (1996). "The Logistic Organ Dysfunction system. A new way to assess organ dysfunction in the intensive care unit. ICU Scoring Group." JAMA 276(10): 802-810

- [41] Le Gall JR, Lemeshow S, Saulnier F (1993). "A new Simplified Acute Physiology Score (SAPS II) based on a European/North American multicenter study." JAMA 270(24): 2957-2963
- [42] Le Gall JR, Lemeshow S, Leleu G, Klar J, Huillard J, Rué M, Teres D, Artigas A (1995). "Customized probability models for early severe sepsis in adult intensive care patients. Intensive Care Unit Scoring Group." JAMA 273(8): 644-650
- [43] Lemeshow S, Hosmer DW Jr (1982). "A review of goodness of fit statistics for use in the development of logistic regression models." Am J Epidemiol 115(1): 92-106
- [44] Lemeshow S, Le Gall JR (1994). "Modeling the severity of illness of ICU patients. A systems update." JAMA 272(13): 1049-1055
- [45] Lemeshow S, Teres D, Pastides H, Avrunin JS, Steingrub JS (1985). "A method for predicting survival and mortality of ICU patients using objectively derived weights." Crit Care Med 13(7): 519-525
- [46] Lemeshow S, Teres D, Klar J, Avrunin JS, Gehlbach SH, Rapoport J (1993). "Mortality Probability Models (MPM II) based on an international cohort of intensive care unit patients." JAMA 270(20): 2478-2486
- [47] Lemeshow S, Teres D, Avrunin JS, Gage RW (1988). "Refining intensive care unit outcome prediction by using changing probabilities of mortality." Crit Care Med 16(5): 470-477
- [48] Lemeshow S, Klar J, Teres D, Avrunin JS, Gehlbach SH, Rapoport J, Rué M (1994). "Mortality probability models for patients in the intensive care unit for 48 or 72 hours: a prospective, multicenter study." Crit Care Med 22(9): 1351-1358
- [49] Lemeshow S, Klar J, Teres D (1995). "Outcome prediction for individual intensive care patients: useful, misused, or abused?" Intensive Care Med 21(9): 770-776
- [50] Lemonnier E, Loirat P, Kleinknecht D, Brivet F, Landais P (1992). "Translation ambiguity and inter-observer variability of severity scoring systems." Int Care Med 20: 581
- [51] Marshall JC, Cook DJ, Christou NV, Bernard GR, Sprung CL, Sibbald WJ (1995). "Multiple organ dysfunction score: a reliable descriptor of a complex clinical outcome." Crit Care Med 23(10): 1638-1652
- [52] Matukaitis J, Stillman P, Wykpiasz E, Ewen E (2005). "Appropriate admissions to the appropriate unit: a decision tree approach." Am J Med Qual 20(2): 90-97.
- [53] McArthur-Rouse F (2001). "Critical care outreach services and early warning scoring systems: a review of the literature." J Adv Nurs 36(5): 696-704.

- [54] McQuillan P, Pilkington S, Allan A, Taylor B, Short A, Morgan G, Nielsen M, Barrett D, Smith G, Collins CH (1998). "Confidential inquiry into quality of care before admission to intensive care." BMJ 316(7148): 1853-1858.
- [55] Meade MO, Cook DJ (1995). "A critical appraisal and systematic review of illness severity scoring systems in the intensive care unit." Curr Opin Crit Care 1: 221
- [56] Meier-Hellmann A (2007). "Die Intensivmedizin in der Versorgungskette." In: Burchardi H, Larsen R, Kuhlen R, Jauch K.-W., Schölmerich J. (eds) Die Intensivmedizin. Springer, Berlin Heidelberg New York, pp 110-113
- [57] Metnitz PG, Valentin A, Vesely H, Alberti C, Lang T, Lenz K, Steltzer H, Hiesmayr M (1999). "Prognostic performance and customization of the SAPS II: results of a multicenter Austrian study. Simplified Acute Physiology Score." Intensive Care Med 25(2): 192-197
- [58] Metnitz PG, Pokorny H, Valentin A, Popow C, Steltzer H (1998). "A national database for intensive care medicine in a European country (ASDI): first results." Crit Care Med 26(01): 68A
- [59] Metnitz PG, Moreno RP, Almeida E, Jordan B, Bauer P, Campos RA, Lapichino G, Edbrooke D, Capuzzo M, Le Gall JR (2005). "SAPS 3--From evaluation of the patient to evaluation of the intensive care unit. Part 1: Objectives, methods and cohort description." Intensive Care Med 31(10): 1336-1344
- [60] Metnitz PG, Lang T, Vesely H, Valentin A, Le Gall JR (2000). "Ratios of observed to expected mortality are affected by differences in case mix and quality of care." Intensive Care Med 26(10): 1466-1472
- [61] Miranda DR, de Rijk A, Schaufeli W (1996). "Simplified Therapeutic Intervention Scoring System: the TISS-28 items--results from a multicenter study." Crit Care Med 24(1): 64-73
- [62] Moreau R, Soupison T, Vauquelin P, Derrida S, Beaucour H, Sicot C (1989). "Comparison of two simplified severity scores (SAPS and APACHE II) for patients with acute myocardial infarction." Crit Care Med 17(5): 409-413
- [63] Moreno R, Apolone G (1997). "Impact of different customization strategies in the performance of a general severity score." Crit Care Med 25(12): 2001-2008
- [64] Moreno R, Miranda DR, Fidler V, Van Schilfgaarde R (1998). "Evaluation of two outcome prediction models on an independent database." Crit Care Med 26(1): 50-61

- [65] Moreno R, Miranda DR, Matos R, Fevereiro T (2001). "Mortality after discharge from intensive care: the impact of organ system failure and nursing workload use at discharge." Intensive Care Med 27(6): 999-1004
- [66] Moreno R, Apolone G, Miranda DR (1998). "Evaluation of the uniformity of fit of general outcome prediction models." Intensive Care Med 24(1): 40-47
- [67] Moreno RP, Metnitz PG, Almeida E, Jordan B, Bauer P, Campos RA, Lapichino G, Edbrooke D, Capuzzo M, Le Gall JR (2005). "SAPS 3--From evaluation of the patient to evaluation of the intensive care unit. Part 2: Development of a prognostic model for hospital mortality at ICU admission." Intensive Care Med 31(10): 1345-1355
- [68] Murray JF, Matthay MA, Luce JM, Flick MR (1988). "An expanded definition of the adult respiratory distress syndrome." Am Rev Respir Dis 138(3): 720-723
- [69] Neugebauer EA, Lefering R (2002). "Severity scores in surgery: what for and who needs them? An introduction: definition, aims, classification and evaluation." Langenbecks Arch Surg 387(1): 55-58
- [70] Neugebauer E, Lefering R (2007). "Scores." In: Burchardi H, Larsen R, Kuhlen R, Jauch K.-W., Schölmerich J. (eds) Die Intensivmedizin. Springer, Berlin Heidelberg New York, pp 54-63
- [71] Rapoport J, Teres D, Lemeshow S, Gehlbach S (1994). "A method for assessing the clinical performance and cost-effectiveness of intensive care units: a multicenter inception cohort study." Crit Care Med 22(9): 1385-1391
- [72] Reis Miranda D, Ryan DW (1997). "Organization and management of Intensive Care: a prospective study in 12 European countries." Springer, Berlin Heidelberg New York
- [73] Reiter A, Mauritz W, Jordan B, Lang T, Pölzl A, Pelinka L, Metnitz PG (2004). "Improving risk adjustment in critically ill trauma patients: the TRISS-SAPS Score." J Trauma 57(2): 375-380
- [74] Ridley S (1998). "Severity of illness scoring systems and performance appraisal." Anaesthesia 53(12): 1185-1194
- [75] Sarmiento X, Rué M, Guardiola JJ, Toboso JM, Soler M, Artigas A (1997). "Assessment of the prognosis of coronary patients: performance and customization of generic severity indexes." Chest 111(6): 1666-1671
- [76] Schönhofer B, Lefering R, Suchi S, Köhler D (2002). "Survey of evaluation of scoring systems by intensive care physicians." Intensivmed 39(3):240-245
- [77] Schuster FP, Schuster HP (1995). "Kontrolle der Ergebnisqualität einer Intensivstation mittels des Simplified Acute Physiology Score (SAPS) II." Intensivmed 32: 682

- [78] Schuster HP (1991). "Score-Systeme: Methoden und Zielsetzungen des Scoring in der Intensivmedizin." Intensivmed 28: 206
- [79] Schuster HP, Schuster FP, Ritschel P, Wilts S, Bodmann KF (1997). "The ability of the Simplified Acute Physiology Score (SAPS II) to predict outcome in coronary care patients." Intensive Care Med 23(10): 1056-1061
- [80] Schuster HP, Wilts S, Ritschel P (1996). "Analyse einer Ergebnisqualitätskontrolle in der Intensivmedizin mittels des Simplified Acute Physiology Score II." Med Klin 91: 343
- [81] Schuster HP, Wilts S, Ritschel P, Schuster FP (1996). "[Predictive value of score parameters of the Simplified Acute Physiology Score (SAPS)-II for the duration of treatment of intensive care patients]." Wien Klin Wochenschr 108(15): 462-466
- [82] Sicignano A, Giudici D (1997). "Probability model of hospital death for severe trauma patients based on the Simplified Acute Physiology Score I: development and validation. Archivio Diagnostico." J Trauma 43(4): 585-589
- [83] Teres D, Lemeshow S (1993). "Using severity measures to describe high performance intensive care units." Crit Care Clin 9(3): 543-554
- [84] Timsit JF, Fosse JP, Troché G, De Lassence A, Alberti C, Garrouste-Orgeas M, Azoulay E, Chevret S, Moine P, Cohen Y (2001). "Accuracy of a composite score using daily SAPS II and LOD scores for predicting hospital mortality in ICU patients hospitalized for more than 72 h." Intensive Care Med 27(6): 1012-1021
- [85] Tobiasen J, Hiebert JH, Edlich RF (1982). "Prediction of burn mortality." Surg Gynecol Obstet 154(5): 711-714
- [86] Unertl K, Kottler BM (1997). "[Prognostic scores in intensive care]." Anaesthesist 46(6): 471-480
- [87] Vesely H, Gregor T (1998). "[Application of scoring systems for quality assurance and cost estimation in intensive care medicine]." Anesthesiol Intensivmed Notfallmed Schmerzther 33(1): 55-58
- [88] Vincent JL, de Mendonça A, Cantraine F, Moreno R, Takala J, Suter PM, Sprung CL, Colardyn F, Blecher S (1998). "Use of the SOFA score to assess the incidence of organ dysfunction/failure in intensive care units: results of a multicenter, prospective study. Working group on "sepsis-related problems" of the European Society of Intensive Care Medicine." Crit Care Med 26(11): 1793-1800



- [89] Wagner DP, Draper EA, Abizanda Campos R, Nikki P, Le Gall JR, Loirat P, Knaus WA (1984). "Initial international use of APACHE. An acute severity of disease measure." Med Decis Making 4(3): 297-313
- [90] Waydhas C (2000). "Vorschlag für ein nationales Register zum externen Qualitätsvergleich in der Intensivmedizin." Intensivmed 37(5): 454-460
- [91] Waydhas C, Mörer O (2007). "Leistungserfassung und Qualitätssicherung." In: Burchardi H, Larsen R, Kuhlen R, Jauch K.-W., Schölmerich J. (eds) Die Intensivmedizin. Springer, Berlin Heidelberg New York, pp 80-86
- [92] Zimmerman JE, Shortell SM, Rousseau DM, Duffy J, Gillies RR, Knaus WA, Devers K, Wagner DP, Draper EA (1993). "Improving intensive care: observations based on organizational case studies in nine intensive care units: a prospective, multicenter study." Crit Care Med 21(10): 1443-1451
- [93] Zweig MH, Campbell G (1993). "Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine." Clin Chem 39(4): 561-577

## 7 Abkürzungsverzeichnis

ABSI	Abbreviated Burn Severity Index
APACHE	Acute Physiology and Chronic Health Evaluation
CPAP	Continuous Positive Airway Pressure
DIVI	Deutsche Interdisziplinäre Vereinigung für Intensiv- und Notfallmedizin
ESICM	European Society for Intensive Care Medicine
GCS	Glasgow Coma Scale
ISS	Injury Severity Score
KISS	Krankenhaus-Infektions-Surveillance-System
LIS	Lung Injury Severity Score
LOD	Logistic Organ Dysfunction
MODS	Multiple Organs Dysfunction Score
MPM	Mortality Prediction Model
ODIN	Organ Dysfunction and Infection System
OSF	Organ System Failure
ROC	Receiver Operating Characteristic
RTS	Revised Trauma Score
SAPS	Simplified Acute Physiology Score
SMR	Standardisierte Mortalitätsrate
SOFA	Sequential Organ Failure Assessment/ Sepsis-related Organ Failure Assessment
TISS	Therapeutic Intervention Scoring System
TRIOS	Three-Day Recalibrating ICU Outcomes
TRISS	Trauma and Injury Severity Score

## 8 Anhang

### 8.1 Tabellen

Tabelle 1 SAPS II

#### SAPS II

maßgeblich sind die schlechtesten Werte (höchste Punktzahl) in einem 24-Stunden-Zeitraum seit dem zur Intensiv-Aufnahme führenden Ereignis (Unfall, OP, Komplikation)

	Punkte bei abnormal niedrigen Werten				„normal“	Punkte bei abnormal hohen Werten				
Alter (Jahre)					<40	40-59	60-69	70-74	75-79	≥80
						<b>7</b>	<b>12</b>	<b>15</b>	<b>16</b>	<b>18</b>
Herzfrequenz (pro Min.)			<40	40-69	70-119	120-159	≥160			
			<b>11</b>	<b>2</b>		<b>4</b>	<b>7</b>			
Blutdruck (syst., mmHg)			<70	70-99	100-199	≥200				
			<b>13</b>	<b>5</b>		<b>2</b>				
Temperatur (°C)					<39,0	≥39,0				
						<b>3</b>				
nur bei Beatmung o. Pulmonalis-Katheder PaO <sub>2</sub> (mmHg) / FiO <sub>2</sub>		<100	100-199	≥200	-					
		<b>11</b>	<b>9</b>	<b>6</b>						
Urinausscheidung (L/Tag)			<0.5	0.5-0.99	≥1,0					
			<b>11</b>	<b>4</b>						
Harnstoff (mg/dl) oder Harnstoff-Stickstoff (mg/dl)					<60	60-179	≥180			
					<28	28-83	≥84			
						<b>6</b>	<b>10</b>			
Leukozyten (10 <sup>3</sup> /mm <sup>3</sup> )				<1.0	1.0-19,9	≥20				
				<b>12</b>		<b>3</b>				
Kalium (mmol/L)				<3	3.0-4.9	≥5.0				
				<b>3</b>		<b>3</b>				
Natrium (mmol/L)				<125	125-144	≥145				
				<b>5</b>		<b>1</b>				
Serum Bikarbonat (mEq/L)			<15	15-19	≥20					
			<b>6</b>	<b>3</b>						
Bilirubin (mg/dl)					<4.0	4.0-5.9	≥6.0			
						<b>4</b>	<b>9</b>			
Glasgow Coma Scale (vor Sedierung)	<6	6-8	9-10	11-13	14-15					
	<b>26</b>	<b>13</b>	<b>7</b>	<b>5</b>						
Vorerkrankungen					-	metastas. Karzinom	maligne hämatolog. Erkr.		AIDS	
						<b>9</b>	<b>10</b>		<b>17</b>	
Zuweisung auf ICU					elektiv chirurgisch	medizinisch (ohne OP)	ungeplant chirurgisch			
						<b>6</b>	<b>8</b>			

Le Gall et al. „A new Simplified Acute Physiology Score (SAPS II) based on a European/North American multicenter study.“ JAMA (1993) 270: 2967-83

Tabelle 2 Definitionen der SAPS-II-Variablen

Variable	Definition
Alter	Verwende das Patientenalter (in Jahren) seit seinem letzten Geburtstag
Herzfrequenz	Verwende den schlechtesten Wert der ersten 24 Stunden, entweder die niedrigste oder höchste Herzfrequenz; wenn es von Asystolie (11 Punkte) bis extremer Tachykardie (7 Punkte) variiert, werden 11 Punkte vergeben
Systolischer Blutdruck	Verwende dieselbe Methode wie bei der Herzfrequenz; wenn der systolische Blutdruck z.B. zwischen 60 mmHg und 205 mmHg schwankt, werden 13 Punkte vergeben
Körpertemperatur	Verwende die höchste Temperatur in Grad Celsius oder Fahrenheit

Variable	Definition
PaO <sub>2</sub> /FIO <sub>2</sub> -Ratio	Verwende den niedrigsten Wert bei CPAP oder wenn der Patient ventiliert wird
Urinvolumen	Wenn der Patient auf der Intensivstation für weniger als 24 Stunden verbleibt, berechne es für 24 Stunden auf: z.B. 1l in 8 Stunden = 3l in 24 Stunden
Harnstoff- oder Harnstoff-Stickstoffserum	Verwende den höchsten Wert für Serumharnstoff in mmol/l oder g/l, in mg/dl für Serum-Harnstoff-Stickstoff
Leukozytengehalt	Verwende den schlechtesten (höchsten oder niedrigsten) Wert, entsprechend der Score-Tabelle
Serumkalium	Verwende den schlechtesten (höchsten oder niedrigsten) Wert in mmol/l, entsprechend der Score-Tabelle
Serumnatrium	Verwende den schlechtesten (höchsten oder niedrigsten) Wert in mmol/l, entsprechend der Score-Tabelle
Serumbikarbonat	Verwende den niedrigsten Wert in mEq/l
Bilirubin	Verwende den höchsten Wert in µmol/l oder mg/dl
Glasgow Coma Scale (GCS)	Verwende den niedrigsten Wert; wenn der Patient sediert ist, benutze die Glasgow Coma Scale vor Sedierung
Einweisungsgrund	notfallchirurgische (†), elektiv chirurgische (‡) oder medizinische (§) Aufnahme
AIDS	Ja, wenn HIV-positiv mit klinischen Komplikationen wie Pneumocystis carinii pneumonia, Kaposi-Sarkom, Lymphom, Tuberkulose oder Toxoplasmose
Maligne hämatologische Erkrankung	Ja, wenn Lymphom, akute Leukämie oder multiples Myelom
Metastasiertes Tumorleiden	Ja, wenn die Metastase chirurgisch, mit CT oder einer anderen Methode nachgewiesen wurde

† Patienten, die innerhalb der ersten 24 Stunden seit Intensivaufnahme operiert wurden

‡ Patienten, deren OP mindestens 24 Stunden im Voraus geplant war

§ Patienten, die innerhalb einer Woche seit Intensivaufnahme nicht operiert wurden

Le Gall, J. R., S. Lemeshow, et al. (1993). "A new Simplified Acute Physiology Score (SAPS II) based on a European/North American multicenter study." *JAMA* 270(24): 2957-2963

**Tabelle 3** Erweiterter SAPS II

<b>Variable</b>	<b>Punkte</b>
<b>Originaler SAPS II Score</b>	0,0742 x SAPS II
<b>Alter</b>	
<40 Jahre	0
40-59 Jahre	0,1639
60-69 Jahre	0,2739
70-79 Jahre	0,3690
>79 Jahre	0,6645
<b>Geschlecht</b>	
m	0,2083
w	0
<b>Krankenhausverweildauer vor Intensivaufnahme</b>	
<24h	0
1 Tag	0,0986
2 Tage	0,1944
3-9 Tage	0,5284
>9 Tage	0,9323
<b>Patientenaufenthalt vor Intensivaufnahme</b>	
Notfallaufnahme oder Krankenwagen	0
Station im selben Krankenhaus	0,2606
Anderes Krankenhaus	0,3381
<b>Klinische Kategorie</b>	
Internistischer Patient	0,6555
Anderer Patient	0
<b>Intoxikation</b>	
Nein	1,6693
Ja	0

---

\*  $\text{Logit} = -14,4761 + 0,0844 \times \text{Score} + 6,6158 \times \log(\text{Score}+1)$

Le Gall, J. R., A. Neumann, et al. (2005). "Mortality prediction using SAPS II: an update for French intensive care units." *Crit Care* 9(6): R645-652

**Tabelle 4** Kalibrierung und Diskriminierung

Modell	Interne Validierung (fünffache Vergleichsprüfung des Entwicklungsdatensatzes)			
	p-Wert des Hosmer-Lemeshow Test		Fläche unter ROC-Kurve	
	MW	SA	MW	SA
Original SAPS II	0.001	0.001	0.8591	0.0058
Angepasster SAPS II	0.6280	0.1422	0,8562	0.0058
Erweiterter SAPS II	0.2754	0.3832	0.8797	0.0054

Modell	Externe Validierung (der Validierungsgruppe)			
	p-Wert des Hosmer-Lemeshow Test		Fläche unter ROC-Kurve	
	Test Statistik	C	p-Wert	
Original SAPS II	1162.9		<0.0001	0.8575
Angepasster SAPS II	6.41		0.7794	0.8575
Erweiterter SAPS II	6.04		0.8116	0.8787

MW=Mittelwert, SA=Standardabweichung

Le Gall, J. R., A. Neumann, et al. (2005). "Mortality prediction using SAPS II: an update for French intensive care units." *Crit Care* 9(6): R645-652

**Tabelle 5** Altersverteilung in den verschiedenen Fachbereichen (%)

Fachrichtungen	<=39	40-59	60-69	70-74	75-79	>=80
Innere Medizin	7,0	20,9	23,4	14,2	15,2	19,3
Anästhesie (Intensiv-/Notfallmedizin)	12,0	21,3	24,1	12,5	13,1	17,1
Allgemein-Chirurgie	6,5	23,1	27,5	14,2	13,0	15,6
Neurochirurgie	15,0	35,7	22,7	10,4	8,4	7,7
MKG-Chirurgie	17,4	35,9	25,1	9,4	6,8	5,4
Gefäßchirurgie	1,4	17,8	32,6	17,8	16,3	14,1
Unfall-Chirurgie	11,9	14,8	15,9	11,9	17,4	28,0
Neurologie	10,9	25,0	23,6	15,8	13,8	11,0
HNO	8,6	40,8	29,7	10,2	5,5	5,2
Urologie	3,8	16,1	35,2	18,8	15,5	10,5
Orthopädie	3,9	16,0	29,1	18,9	17,4	14,6
Gynäkologie	23,9	28,0	22,0	9,9	9,1	7,2
Thorax-Chirurgie	7,5	31,2	30,4	15,1	11,3	4,5
HTG-Chirurgie	2,4	19,9	35,1	20,4	15,3	6,8
Transplantation	14,5	46,6	27,4	6,5	3,5	1,4
Sonstige	23,8	25,2	20,5	10,2	9,4	10,8

**Tabelle 6** Anzahl der Verstorbenen in den verschiedenen Altersgruppen der Entwicklungsgruppe

Altersgruppen	Überlebende		Verstorbene		Gesamt
	n	%	n	%	
<=39	2901	96,7	100	3,3	3001
40-59	8244	95,4	402	4,6	8646
60-69	9394	94,3	568	5,7	9962
70-74	5188	92,9	397	7,1	5585
75-79	4821	92,2	410	7,8	5231
>=80	5045	90,7	520	9,3	5565
<b>Gesamt</b>	<b>35593</b>	<b>93,7</b>	<b>2397</b>	<b>6,3</b>	<b>37990</b>

**Tabelle 7** Anzahl der Verstorbenen in den verschiedenen Altersgruppen der Validierungsgruppe

Altersgruppen	Überlebende		Verstorbene		Gesamt
	n	%	n	%	
<=39	1487	97,5	38	2,5	1525
40-59	4178	95,2	211	4,8	4389
60-69	4935	94,7	275	5,3	5210
70-74	2624	92,7	206	7,3	2830
75-79	2493	92,0	218	8,0	2711
>=80	2624	90,3	282	9,7	2906
<b>Gesamt</b>	<b>18341</b>	<b>93,7</b>	<b>1230</b>	<b>6,3</b>	<b>19571</b>

**Tabelle 8** Allgemeine statistische Daten

	Entwicklungsgruppe		Validierungsgruppe	
	n	%	n	%
<b>Patientenanzahl</b>	37990	100	19571	100
<b>Geschlecht</b>				
weiblich	17142	54,9	8868	45,3
männlich	20848	45,1	10703	54,7
<b>Alter (in Jahren) (Median, Quartile)</b>	67	56-76	67	56-76
<b>Altersgruppen</b>				
<=39	3001	7,9	1525	7,8
40-59	8646	22,8	4389	22,4
60-69	9962	26,2	5210	26,6
70-74	5585	14,7	2830	14,5
75-79	5231	13,8	2711	13,9
>=80	5565	14,6	2906	14,8
<b>Trauma</b>	2402	6,3	1238	6,3

	<b>Entwicklungsgruppe</b>		<b>Validierungsgruppe</b>	
	n	%	n	%
<b>Neuro</b>	3312	8,7	1695	8,7
<b>Aufnahme von Extern</b>	2666	7,0	1314	6,7
<b>Vorerkrankungen</b>				
<b>keine</b>	32839	86,4	16928	86,5
<b>metastas. CA</b>	4816	12,7	2479	12,7
<b>maligne hämatolog. Erkr.</b>	315	0,8	157	0,8
<b>AIDS</b>	20	0,1	7	0,0
<b>Vorerkrankungen (zusammengefasst)</b>	5151	13,6	2643	13,5
<b>Aufnahmegrund</b>				
<b>elektiv chirurgisch</b>	18396	48,4	9436	48,2
<b>medizinisch (ohne OP)</b>	13219	34,8	6817	34,8
<b>ungeplant chirurgisch</b>	6375	16,8	3318	17,0
<b>SAPS-II-Score (Median, Quartile)</b>	24	17-32	24	17-32
<b>Intensivletalität (erwartet, beobachtet, in %)</b>	6,31	6,30	6,26	6,30
<b>Fachrichtungen</b>				
<b>Innere Medizin</b>	8707	22,9	4572	23,4
<b>Anästhesie (Intensiv-/Notfallmedizin)</b>	921	2,4	507	2,6
<b>Allgemein-Chirurgie</b>	10145	26,7	5324	27,2
<b>Neurochirurgie</b>	2580	6,8	1337	6,8
<b>MKG-Chirurgie</b>	342	0,9	159	0,8
<b>Gefäßchirurgie</b>	1262	3,3	647	3,3
<b>Unfall-Chirurgie</b>	3183	8,4	1614	8,2
<b>Neurologie</b>	732	1,9	358	1,8
<b>HNO</b>	409	1,1	241	1,2
<b>Urologie</b>	1071	2,8	553	2,8
<b>Orthopädie</b>	1321	3,5	674	3,4
<b>Gynäkologie</b>	1091	2,9	510	2,6
<b>Thorax-Chirurgie</b>	672	1,8	318	1,6
<b>HTG-Chirurgie</b>	4802	12,6	2382	12,2
<b>Transplantation</b>	509	1,3	257	1,3
<b>Sonstige</b>	243	0,6	118	0,6

\* MKG-Chirurgie = Mund-Kiefer-Gesichts-Chirurgie, HTG-Chirurgie = Herz-Thorax-Gefäß-Chirurgie



**Tabelle 9** Anzahl der Verstorbenen und Überlebenden in den verschiedenen Fachrichtungen

Fachrichtungen	Überlebende		Verstorbene	
	n	%	n	%
<b>Innere Medizin</b>	12061	90,8	1218	9,2
<b>Anästhesie</b>	1281	89,7	147	10,3
<b>(Intensiv-/Notfallmedizin)</b>				
<b>Allgemein-Chirurgie</b>	14379	93,0	1090	7,0
<b>Neurochirurgie</b>	3619	92,4	298	7,6
<b>MKG-Chirurgie</b>	496	99,0	5	1,0
<b>Gefäßchirurgie</b>	1815	95,1	94	4,9
<b>Unfall-Chirurgie</b>	4618	96,3	179	3,7
<b>Neurologie</b>	981	90,0	109	10,0
<b>HNO</b>	639	98,3	11	1,7
<b>Urologie</b>	1582	97,4	42	2,6
<b>Orthopädie</b>	1977	99,1	18	0,9
<b>Gynäkologie</b>	1590	99,3	11	0,7
<b>Thorax-Chirurgie</b>	961	97,1	29	2,9
<b>HTG-Chirurgie</b>	6887	95,9	297	4,1
<b>Transplantation</b>	714	93,2	52	6,8
<b>Sonstige</b>	334	92,5	27	7,5

\* MKG-Chirurgie = Mund-Kiefer-Gesichts-Chirurgie, HTG-Chirurgie = Herz-Thorax-Gefäß-Chirurgie

**Tabelle 10** SAPS-DIVI 2 [Entwicklungsdatensatz]

Variablen	RK	SF	Sig	Exp	95% - KI
ln(SAPS II + 1)	1,042	0,237	<0,001	2,834	1,781 –4,509
SAPS II	0,065	0,006	<0,001	1,068	1,054 –1,081
Konstante	-8,421	0,622	<0,001	0,000	
ln(SAPS II + 1)	1,148	0,238	<0,001	3,150	1,978 –5,019
SAPS II	0,066	0,006	<0,001	1,068	1,055 –1,081
<b>Neuro</b>	1,038	0,076	<0,001	2,823	2,431 –3,279
Konstante	-8,900	0,625	<0,001	0,000	
ln(SAPS II + 1)	1,052	0,238	<0,001	2,864	1,798 –4,564
SAPS II	0,065	0,006	<0,001	1,067	1,054 –1,081
<b>Extern</b>	0,695	0,075	<0,001	2,003	1,730 –2,318
Konstante	-8,513	0,624	<0,001	0,000	
ln(SAPS II + 1)	1,032	0,237	<0,001	2,808	1,765 –4,468
SAPS II	0,066	0,006	<0,001	1,068	1,055 –1,082
<b>Trauma</b>	0,299	0,093	0,001	1,348	1,123 –1,619
Konstante	-8,428	0,622	<0,001	0,000	
ln(SAPS II + 1)	1,052	0,237	<0,001	2,864	1,799 –4,560
SAPS II	0,065	0,006	<0,001	1,067	1,054 –1,081
<b>Geschlecht</b>	0,131	0,047	0,006	1,140	1,039 –1,250
Konstante	-8,523	0,623	<0,001	0,000	

<b>Variablen</b>	<b>RK</b>	<b>SF</b>	<b>Sig</b>	<b>Exp</b>	<b>95% - KI</b>
ln(SAPS II + 1)	1,628	0,241	<0,001	5,093	3,177 –8,165
SAPS II	0,054	0,006	<0,001	1,056	1,043 –1,069
<b>Alter</b>					
<=39			<0,001		
40-59	-0,596	0,127	<0,001	0,551	0,429 –0,707
60-69	-0,965	0,125	<0,001	0,381	0,298 –0,487
70-74	-1,080	0,130	<0,001	0,340	0,263 –0,439
75-79	-1,082	0,130	<0,001	0,339	0,262 –0,437
>=80	-1,162	0,129	<0,001	0,313	0,243 –0,403
Konstante	-9,148	0,622	<0,001	0,000	
ln(SAPS II + 1)	1,141	0,238	<0,001	3,131	1,963 –4,995
SAPS II	0,064	0,006	<0,001	1,067	1,053 –1,080
<b>Vorerkrankungen</b>					
-			<0,001		
metastas. Karzinom	-0,521	0,069	<0,001	0,594	0,519 –0,680
maligne häm. Erkr.	-0,277	0,183	0,130	0,758	0,530 –1,085
AIDS	-0,520	0,691	0,452	0,594	0,153 –2,304
Konstante	-8,653	0,625	<0,001	0,000	
ln(SAPS II + 1)	1,127	0,238	<0,001	3,087	1,937 –4,918
SAPS II	0,065	0,006	<0,001	1,067	1,054 –1,081
<b>Vorerkrankungen (zusammengefasst)</b>	-0,496	0,065	<0,001	0,609	0,536 –0,692
Konstante	-8,621	0,624	<0,001	0,000	
ln(SAPS II + 1)	0,835	0,233	<0,001	2,304	1,458 –3,640
SAPS II	0,065	0,006	<0,001	1,067	1,054 –1,080
<b>Zuweisung</b>					
elektiv chirurgisch			<0,001		
medizinisch (ohne OP)	0,659	0,062	<0,001	1,933	1,711 –2,184
ungeplant chirurgisch	0,491	0,071	<0,001	1,633	1,421 –1,878
Konstante	-8,098	0,610	<0,001	0,000	

\* RK = Regressionskoeffizient, SF = Standardfehler, Sig = Signifikanz, Exp = Exponent, KI = Konfidenzintervall

**Tabelle 11** ROC-Kurve (SAPS-DIVI 2 [Entwicklungsdatensatz])

<b>Variable</b>	<b>Fläche</b>	<b>Standardfehler</b>	<b>95% - KI</b>
<b>SAPS II (original)</b>	0,834	0,004	0,825 – 0,843
<b>SAPS II + ln(SAPS II + 1)</b>	0,834	0,004	0,825 – 0,843
<b>Neuro</b>	0,842	0,004	0,834 – 0,851
<b>Extern</b>	0,838	0,004	0,830 – 0,847
<b>Trauma</b>	0,835	0,004	0,826 – 0,844
<b>Geschlecht (m)</b>	0,834	0,004	0,826 – 0,843
<b>Alter</b>	0,839	0,004	0,830 – 0,847
<b>Vorerkrankungen</b>	0,836	0,004	0,828 – 0,845
<b>Vorerkrankungen (zusammengefasst)</b>	0,836	0,004	0,828 – 0,845

Variable	Fläche	Standardfehler	95% - KI
Zuweisung	0,841	0,004	0,833 – 0,850

Tabelle 12 SAPS-DIVI 2 [Entwicklungsdatensatz]

Variablen	RK	SF	Sig	Exp	95% - KI
ln(SAPS II + 1)	1,087	0,236	<0,001	2,966	1,867 - 4,712
SAPS II	0,065	0,006	<0,001	1,067	1,054 –1,080
Fachrichtungen			<0,001		
Anästhesie	-0,056	0,130	0,666	0,945	0,733 –1,220
Allgemein-Chirurgie	0,011	0,060	0,858	1,011	0,899 –1,137
Neurochirurgie	0,908	0,095	<0,001	2,480	2,061 –2,985
MKG-Chirurgie	-1,276	0,520	0,014	0,279	0,101 –0,774
Gefäßchirurgie	-0,148	0,151	0,327	0,862	0,642 –1,159
Unfall-Chirurgie	-0,452	0,111	<0,001	0,636	0,511 –0,792
Neurologie	0,820	0,143	<0,001	2,270	1,716 –3,003
HNO	-0,998	0,407	0,014	0,368	0,166 –0,818
Urologie	-0,985	0,201	<0,001	0,374	0,252 –0,554
Orthopädie	-0,920	0,257	<0,001	0,398	0,241 –0,660
Gynäkologie	-1,850	0,387	<0,001	0,157	0,074 –0,336
Thoraxchirurgie	-0,452	0,265	0,088	0,636	0,378 –1,070
HTG-Chirurgie	-0,284	0,088	0,001	0,753	0,634 –0,894
Transplantation	0,153	0,207	0,461	1,165	0,776 –1,750
Sonstige	0,263	0,299	0,379	1,301	0,724 –2,338
Konstante	-8,518	0,623	<0,001	0,000	

\* RK = Regressionskoeffizient, SF = Standardfehler, Sig = Signifikanz, Exp = Exponent, KI = Konfidenzintervall

Tabelle 13 SAPS-DIVI 2 [Entwicklungsdatensatz]

Variablen	RK	SF	Sig	Exp	95% - KI
ln(SAPS II + 1)	1,081	0,237	<0,001	2,947	1,854 - 4,686
SAPS II	0,066	0,006	<0,001	1,069	1,055 –1,082
Fach_d_v			<0,001		
+	0,967	0,076	<0,001	2,629	2,263 –3,054
-	-1,060	0,133	<0,001	0,347	0,267 –0,450
Konstante	-8,622	0,622	<0,001	0,000	

\* RK = Regressionskoeffizient, SF = Standardfehler, Sig = Signifikanz, Exp = Exponent, KI = Konfidenzintervall, + = Neurologie, Neurochirurgie, - = MKG-Chirurgie, HNO, Urologie, Orthopädie, Gynäkologie

**Tabelle 14** Zusammenfassung - SAPS-DIVI 2 [Entwicklungsdatensatz]

<b>Variablen</b>	<b>RK</b>	<b>SF</b>	<b>Sig</b>	<b>Exp</b>	<b>95% - KI</b>
ln(SAPS II + 1)	1,593	0,244	<0,001	4,921	3,049 - 7,941
SAPS II	0,055	0,006	<0,001	1,056	1,043 - 1,070
<b>Altersgruppen</b>					
<=39			<0,001		
40-59	-0,529	0,129	<0,001	0,589	0,458 - 0,758
60-69	-0,821	0,128	<0,001	0,440	0,343 - 0,565
70-74	-0,949	0,133	<0,001	0,387	0,298 - 0,502
75-79	-0,957	0,133	<0,001	0,384	0,296 - 0,499
>=80	-1,066	0,131	<0,001	0,344	0,266 - 0,445
<b>Zuweisung</b>					
elektiv chirurgisch			<0,001		
medizinisch (ohne OP)	0,341	0,066	<0,001	1,407	1,236 - 1,601
ungeplant chirurgisch	0,233	0,074	0,002	1,263	1,093-1,458
Vorerkrankungen (alle)	-0,324	0,069	<0,001	0,723	0,631 - 0,828
Aufnahme von Extern	0,340	0,077	<0,001	1,405	1,208 - 1,635
Innere Medizin + alle nicht "relevanten" Fächer			<0,001		
+ (Plus)	0,809	0,079	<0,001	2,245	1,922 - 2,623
- (Minus)	-0,874	0,135	<0,001	0,417	0,320 - 0,544
Konstante	-9,380	0,627	<0,001	0,000	

\* RK = Regressionskoeffizient, SF = Standardfehler, Sig = Signifikanz, Exp = Exponent, KI = Konfidenzintervall, + = Neurologie, Neurochirurgie, - = MKG-Chirurgie, HNO, Urologie, Orthopädie, Gynäkologie

**Tabelle 15** Zusammenfassung (Neuro) - SAPS-DIVI 2 [Entwicklungsdatensatz]

<b>Variablen</b>	<b>RK</b>	<b>SF</b>	<b>Sig</b>	<b>Exp</b>	<b>95% - KI</b>
ln(SAPS II + 1)	1,644	0,245	<0,001	5,177	3,205 - 8,364
SAPS II	0,054	0,006	<0,001	1,056	1,043 - 1,069
<b>Altersgruppen</b>					
<=39			<0,001		
40-59	-0,507	0,128	<0,001	0,602	0,469 - 0,775
60-69	-0,802	0,127	<0,001	0,448	0,349 - 0,576
70-74	-0,935	0,133	<0,001	0,393	0,303 - 0,509
75-79	-0,940	0,133	<0,001	0,391	0,301 - 0,507
>=80	-1,044	0,131	<0,001	0,352	0,272 - 0,455
<b>Zuweisung</b>					
elektiv chirurgisch			<0,001		
medizinisch (ohne OP)	0,392	0,066	<0,001	1,480	1,300 - 1,684
ungeplant chirurgisch	0,240	0,074	0,001	1,271	1,100-1,468
Vorerkrankungen (alle)	-0,387	0,069	<0,001	0,679	0,594 - 0,777
Aufnahme von Extern	0,347	0,077	<0,001	1,415	1,216 - 1,647

Variablen	RK	SF	Sig	Exp	95% - KI
Neuro	0,855	0,079	<0,001	2,352	2,014 –2,748
Konstante	-9,624	0,629	<0,001	0,000	

\* RK = Regressionskoeffizient, SF = Standardfehler, Sig = Signifikanz, Exp = Exponent, KI = Konfidenzintervall, + = Neurologie, Neurochirurgie, - = MKG-Chirurgie, HNO, Urologie, Orthopädie, Gynäkologie

**Tabelle 16** SAPS-DIVI 2 (Phys) [Entwicklungsdatensatz]

Variablen	RK	SF	Sig	Exp	95% - KI
ln(SAPS II + 1)	0,145	0,047	0,002	1,156	1,053 –1,269
SAPS II	0,093	0,004	<0,001	1,098	1,090 –1,106
Konstante	-4,366	0,080	<0,001	0,013	
ln(SAPS II + 1)	0,172	0,048	<0,001	1,188	1,082 –1,304
SAPS II	0,095	0,004	<0,001	1,099	1,091 –1,107
<b>Neuro</b>	0,912	0,075	<0,001	2,490	2,149 –2,885
Konstante	-4,540	0,083	<0,001	0,011	
ln(SAPS II + 1)	0,159	0,048	0,001	1,172	1,068 –1,287
SAPS II	0,092	0,004	<0,001	1,096	1,089 –1,104
<b>Extern</b>	0,653	0,076	<0,001	1,922	1,657 –2,229
Konstante	-4,438	0,081	<0,001	0,012	
ln(SAPS II + 1)	0,143	0,048	0,003	1,154	1,052 –1,267
SAPS II	0,094	0,004	<0,001	1,098	1,091 –1,106
<b>Trauma</b>	0,294	0,093	0,002	1,342	1,119 –1,611
Konstante	-4,388	0,081	<0,001	0,012	
ln(SAPS II + 1)	0,145	0,048	0,002	1,156	1,053 –1,269
SAPS II	0,093	0,004	<0,001	1,098	1,090 –1,106
<b>Geschlecht</b>	0,001	0,047	0,981	1,001	0,913 –1,098
Konstante	-4,366	0,084	<0,001	0,013	
ln(SAPS II + 1)	0,121	0,047	0,011	1,128	1,028 –1,238
SAPS II	0,094	0,004	<0,001	1,099	1,091 –1,106
<b>Alter</b>					
<=39			<0,001		
40-59	0,221	0,122	0,070	1,248	0,982 –1,585
60-69	0,367	0,118	0,002	1,444	1,145 –1,822
70-74	0,546	0,123	<0,001	1,726	1,356 –2,196
75-79	0,647	0,123	<0,001	1,911	1,502 –2,430
>=80	0,806	0,120	<0,001	2,240	1,770 –2,834
Konstante	-4,790	0,131	<0,001	0,008	
ln(SAPS II + 1)	0,145	0,048	0,002	1,156	1,053 –1,268
SAPS II	0,093	0,004	<0,001	1,098	1,090 –1,106
<b>Vorerkrankungen</b>					
-			<0,001		
metastas. Karzinom	0,223	0,068	0,001	1,250	1,094 –1,428

<b>Variablen</b>	<b>RK</b>	<b>SF</b>	<b>Sig</b>	<b>Exp</b>	<b>95% - KI</b>
maligne häm. Erkr.	0,579	0,187	0,002	1,784	1,236 –2,576
AIDS	0,386	0,686	0,574	1,471	0,383 –5,649
Konstante	-4,402	0,081	<0,001	0,012	
ln(SAPS II + 1)	0,143	0,048	0,003	1,154	1,051 –1,267
SAPS II	0,094	0,004	<0,001	1,098	1,090 –1,106
<b>Vorerkrankungen (zusammengefasst)</b>	0,259	0,064	<0,001	1,295	1,142 –1,470
Konstante	-4,404	0,081	<0,001	0,012	
ln(SAPS II + 1)	0,155	0,047	0,001	1,167	1,064 –1,281
SAPS II	0,085	0,004	<0,001	1,088	1,081 –1,096
<b>Zuweisung</b>					
elektiv chirurgisch			<0,001		
medizinisch (ohne OP)	1,040	0,061	<0,001	2,829	2,512 –3,187
ungeplant chirurgisch	1,101	0,069	<0,001	3,007	2,627 –3,441
Konstante	-4,970	0,089	<0,001	0,007	

\* RK = Regressionskoeffizient, SF = Standardfehler, Sig = Signifikanz, Exp = Exponent, KI = Konfidenzintervall

**Tabelle 17** ROC-Kurve (SAPS-DIVI 2 (Phys) [Entwicklungsdatensatz])

<b>Variable</b>	<b>Fläche</b>	<b>Standardfehler</b>	<b>95% - KI</b>
<b>SAPS II (original)</b>	0,815	0,005	0,805 – 0,825
<b>SAPS II + ln(SAPS II + 1)</b>	0,815	0,005	0,805 – 0,825
<b>Neuro</b>	0,823	0,005	0,813 – 0,832
<b>Extern</b>	0,820	0,005	0,810 – 0,830
<b>Trauma</b>	0,816	0,005	0,806 – 0,826
<b>Geschlecht (m)</b>	0,816	0,005	0,806 – 0,825
<b>Alter</b>	0,821	0,005	0,811 – 0,830
<b>Vorerkrankungen</b>	0,816	0,005	0,806 – 0,826
<b>Vorerkrankungen (zusammengefasst)</b>	0,816	0,005	0,806 – 0,826
<b>Zuweisung</b>	0,837	0,004	0,829 – 0,846

**Tabelle 18** SAPS-DIVI 2 (Phys) [Entwicklungsdatensatz]

<b>Variablen</b>	<b>RK</b>	<b>SF</b>	<b>Sig</b>	<b>Exp</b>	<b>95% - KI</b>
ln(SAPS II + 1)	0,213	0,048	<0,001	1,238	1,126 - 1,360
SAPS II	0,089	0,004	<0,001	1,093	1,085 –1,101
Fachrichtungen			<0,001		
Anästhesie	-0,189	0,132	0,152	0,828	0,639 –1,072
Allgemein-Chirurgie	0,056	0,061	0,357	1,058	0,939 –1,191
Neurochirurgie	0,622	0,093	<0,001	1,863	1,552 –2,236
MKG-Chirurgie	-1,484	0,511	0,004	0,227	0,083 –0,617
Gefäßchirurgie	-0,305	0,150	0,041	0,737	0,550 –0,988
Unfall-Chirurgie	-0,514	0,111	<0,001	0,598	0,481 –0,744
Neurologie	0,807	0,143	<0,001	2,242	1,695 –2,965

Variablen	RK	SF	Sig	Exp	95% - KI
HNO	-1,206	0,401	0,003	0,299	0,136 –0,657
Urologie	-0,847	0,199	<0,001	0,428	0,290 –0,633
Orthopädie	-1,182	0,255	<0,001	0,307	0,186 –0,505
Gynäkologie	-1,948	0,385	<0,001	0,142	0,067 –0,303
Thoraxchirurgie	-0,674	0,265	0,011	0,510	0,303 –0,857
HTG-Chirurgie	-0,774	0,087	<0,001	0,461	0,389 –0,547
Transplantation	-0,426	0,205	0,038	0,653	0,437 –0,976
Sonstige	0,118	0,296	0,689	1,126	0,630 –2,010
Konstante	-4,314	0,091	<0,001	0,013	

\* RK = Regressionskoeffizient, SF = Standardfehler, Sig = Signifikanz, Exp = Exponent, KI = Konfidenzintervall

**Tabelle 19** SAPS-DIVI 2 (Phys) [Entwicklungsdatensatz]

Variablen	RK	SF	Sig	Exp	95% - KI
ln(SAPS II + 1)	0,167	0,048	<0,001	1,182	1,077 –1,297
SAPS II	0,093	0,004	<0,001	1,098	1,090 –1,105
Fach_d_v_phys			<0,001		
+	0,837	0,075	<0,001	2,309	1,992 –2,676
-	-1,040	0,132	<0,001	0,353	0,273 –0,458
Konstante	-4,436	0,083	<0,001	0,012	

\* RK = Regressionskoeffizient, SF = Standardfehler, Sig = Signifikanz, Exp = Exponent, KI = Konfidenzintervall, + = Neurologie, Neurochirurgie, - = MKG-Chirurgie, HNO, Urologie, Orthopädie, Gynäkologie

**Tabelle 20** Zusammenfassung - SAPS-DIVI 2 (Phys) [Entwicklungsdatensatz]

Variablen	RK	SF	Sig	Exp	95% - KI
ln(SAPS II + 1)	0,152	0,048	0,001	1,164	1,060 –1,278
SAPS II	0,086	0,004	<0,001	1,090	1,082 –1,097
Altersgruppen					
<=39			<0,001		
40-59	0,309	0,123	0,012	1,362	1,070 –1,735
60-69	0,553	0,120	<0,001	1,739	1,374 –2,200
70-74	0,726	0,125	<0,001	2,066	1,619 –2,638
75-79	0,819	0,124	<0,001	2,269	1,779 –2,893
>=80	0,899	0,121	<0,001	2,457	1,936 –3,117
Zuweisung					
elektiv chirurgisch			<0,001		
medizinisch (ohne OP)	0,974	0,063	<0,001	2,649	2,340 –2,998
ungeplant chirurgisch	1,064	0,070	<0,001	2,897	2,524 –3,325
Vorerkrankungen (alle)	0,582	0,068	<0,001	1,789	1,565 –2,045
Aufnahme von Extern	0,359	0,077	<0,001	1,432	1,230 –1,666

Variablen	RK	SF	Sig	Exp	95% - KI
Innere Medizin + alle nicht "relevanten" Fächer			<0,001		
+ (Plus)	0,785	0,079	<0,001	2,192	1,878 –2,559
- (Minus)	-0,903	0,136	<0,001	0,406	0,311 –0,529
Konstante	-5,695	0,142	<0,001	0,003	

\* RK = Regressionskoeffizient, SF = Standardfehler, Sig = Signifikanz, Exp = Exponent, KI = Konfidenzintervall, + = Neurologie, Neurochirurgie, - = MKG-Chirurgie, HNO, Urologie, Orthopädie, Gynäkologie

**Tabelle 21** Zusammenfassung (Neuro) - SAPS-DIVI 2 (Phys) [Entwicklungsdatensatz]

Variablen	RK	SF	Sig	Exp	95% - KI
ln(SAPS II + 1)	0,153	0,048	0,001	1,166	1,062 –1,280
SAPS II	0,087	0,004	<0,001	1,090	1,083 –1,098
Altersgruppen					
<=39			<0,001		
40-59	0,340	0,123	0,006	1,405	1,104 –1,788
60-69	0,586	0,120	<0,001	1,797	1,421 –2,272
70-74	0,757	0,124	<0,001	2,131	1,671 –2,719
75-79	0,853	0,124	<0,001	2,347	1,842 –2,991
>=80	0,940	0,121	<0,001	2,559	2,019 –3,244
Zuweisung					
elektiv chirurgisch			<0,001		
medizinisch (ohne OP)	1,034	0,063	<0,001	2,811	2,485 –3,180
ungeplant chirurgisch	1,079	0,070	<0,001	2,942	2,564 –3,376
Vorerkrankungen (alle)	0,525	0,068	<0,001	1,690	1,480 –1,930
Aufnahme von Extern	0,366	0,077	<0,001	1,442	1,239 –1,678
Neuro	0,833	0,079	<0,001	2,301	1,971 –2,686
Konstante	-5,818	0,141	<0,001	0,003	

\* RK = Regressionskoeffizient, SF = Standardfehler, Sig = Signifikanz, Exp = Exponent, KI = Konfidenzintervall, + = Neurologie, Neurochirurgie, - = MKG-Chirurgie, HNO, Urologie, Orthopädie, Gynäkologie

**Tabelle 22** ROC-Kurve (SAPS-DIVI 2 [Validierungsdatensatz])

Variable	Fläche	Standardfehler	95% - KI
SAPS II (original)	0,835	0,006	0,823 – 0,847
SAPS II + ln(SAPS II + 1)	0,835	0,006	0,823 – 0,847
Neuro	0,846	0,006	0,835 – 0,857
Extern	0,843	0,006	0,831 – 0,854
Trauma	0,836	0,006	0,825 – 0,848
Geschlecht (m)	0,834	0,006	0,822 – 0,846
Alter	0,838	0,006	0,827 – 0,850
Vorerkrankungen	0,837	0,006	0,826 – 0,849
Vorerkrankungen (zusammengefasst)	0,837	0,006	0,825 – 0,849



<b>Variable</b>	<b>Fläche</b>	<b>Standardfehler</b>	<b>95% - KI</b>
<b>Zuweisung</b>	0,838	0,006	0,827 – 0,850

**Tabelle 23** ROC-Kurve (SAPS-DIVI 2 (Phys) [Validierungsdatensatz])

<b>Variable</b>	<b>Fläche</b>	<b>Standardfehler</b>	<b>95% - KI</b>
<b>SAPS II (original)</b>	0,813	0,007	0,799 – 0,827
<b>SAPS II + ln(SAPS II + 1)</b>	0,813	0,007	0,799 – 0,827
<b>Neuro</b>	0,826	0,007	0,813 – 0,839
<b>Extern</b>	0,822	0,007	0,809 – 0,835
<b>Trauma</b>	0,814	0,007	0,800 – 0,828
<b>Geschlecht (m)</b>	0,813	0,007	0,799 – 0,827
<b>Alter</b>	0,822	0,007	0,809 – 0,834
<b>Vorerkrankungen</b>	0,815	0,007	0,801 – 0,828
<b>Vorerkrankungen (zusammengefasst)</b>	0,814	0,007	0,801 – 0,828
<b>Zuweisung</b>	0,832	0,006	0,820 – 0,844

Tabelle 24 Zusammenfassung der Ergebnisse

**SAPS II (Original)**

Entwicklungs-Datensatz n = 37990		Validierungs-Datensatz n = 19571	
	SAPS II (Original)		SAPS II (Original)
<b><i>Diskriminierung</i></b>			
Mittlere Differenz zw. Überlebenden u. Verst.	9,6 – 36,3 <b>26,7</b>		9,6 – 36,0 <b>26,4</b>
Sensitivität / Spezifität (%)	30,7 / 97,8		32,4 / 97,7
Gesamtrichtigkeit (%)	<b>93,5</b>		<b>93,6</b>
AUC der ROC Kurve mit 95% CI	<b>,834</b> ,825 - ,843		<b>,835</b> ,823 - ,847
<b><i>Präzision</i></b>			
beobachtete Mortalität (%)	6,3		6,3
Score-Prognose (%)	<b>11,3</b>		<b>11,2</b>

**SAPS-DIVI**

Entwicklungs-Datensatz n = 37990		Validierungs-Datensatz n = 19571	
	SAPS- DIVI		SAPS- DIVI
<b><i>Diskriminierung</i></b>			
Mittlere Differenz zw. Überlebenden u. Verst.	6,0 – 25,3 <b>19,3</b>		6,0 – 25,2 <b>19,2</b>
Sensitivität / Spezifität (%)	14,8 / 99,4		14,9 / 99,3
Gesamtrichtigkeit (%)	<b>94,0</b>		<b>94,0</b>
AUC der ROC Kurve mit 95% CI	<b>,842</b> ,834 - ,851		<b>,846</b> ,834 - ,857
<b><i>Präzision</i></b>			
beobachtete Mortalität (%)	6,3		6,3
Score-Prognose (%)	<b>7,2</b>		<b>7,2</b>

**SAPS-DIVI 2**

<b>Entwicklungs-Datensatz</b> n = 37990		<b>Validierungs-Datensatz</b> n = 19571	
	<b>SAPS-DIVI II</b>		<b>SAPS-DIVI II</b>
<b><i>Diskriminierung</i></b>			
Mittlere Differenz zw. Überlebenden u. Verst.	5,1 – 24,9 <b>19,8</b>	5,0 – 24,7 <b>19,7</b>	
Sensitivität / Spezifität (%) Gesamtrichtigkeit (%)	15,4 / 99,3 <b>94,0</b>	15,1 / 99,3 <b>94,0</b>	
AUC der ROC Kurve mit 95% CI	<b>,854</b> ,846 - ,862	<b>,856</b> ,846 - ,867	
<b><i>Präzision</i></b>			
beobachtete Mortalität (%)	6,3	6,3	
Score-Prognose (%)	<b>6,3</b>	<b>6,3</b>	

**SAPS-DIVI 2 (Neuro)**

<b>Entwicklungs-Datensatz</b> n = 37990		<b>Validierungs-Datensatz</b> n = 19571	
	<b>SAPS-DIVI II (Neuro)</b>		<b>SAPS-DIVI II (Neuro)</b>
<b><i>Diskriminierung</i></b>			
Mittlere Differenz zw. Überlebenden u. Verst.	5,1 – 24,8 <b>19,7</b>	5,0 – 24,1 <b>19,1</b>	
Sensitivität / Spezifität (%) Gesamtrichtigkeit (%)	15,4 / 99,3 <b>94,0</b>	14,7 / 99,4 <b>94,0</b>	
AUC der ROC Kurve mit 95% CI	<b>,851</b> ,843 - ,860	<b>,853</b> ,842 - ,864	
<b><i>Präzision</i></b>			
beobachtete Mortalität (%)	6,3	6,3	
Score-Prognose (%)	<b>6,2</b>	<b>6,2</b>	

**SAPS-DIVI 2 (Phys)**

Entwicklungs-Datensatz n = 37990		Validierungs-Datensatz n = 19571	
	SAPS-DIVI II (Phys)		SAPS-DIVI II (Phys)
<b>Diskriminierung</b>			
Mittlere Differenz zw. Überlebenden u. Verst.	5,0 – 25,2 <b>20,2</b>		5,0 – 25,0 <b>20,0</b>
Sensitivität / Spezifität (%)	16,6 / 99,3		16,4 / 99,2
Gesamtrichtigkeit (%)	<b>94,0</b>		<b>94,0</b>
AUC der ROC Kurve mit 95% CI	<b>,853</b> ,845 - ,861		<b>,855</b> ,844 - ,865
<b>Präzision</b>			
beobachtete Mortalität (%)	6,3		6,3
Score-Prognose (%)	<b>6,3</b>		<b>6,3</b>

**SAPS-DIVI 2 (Phys, Neuro)**

Entwicklungs-Datensatz n = 37990		Validierungs-Datensatz n = 19571	
	SAPS-DIVI II (Phys, Neuro)		SAPS-DIVI II (Phys, Neuro)
<b>Diskriminierung</b>			
Mittlere Differenz zw. Überlebenden u. Verst.	5,0 – 25,1 <b>20,1</b>		5,0 – 24,9 <b>19,9</b>
Sensitivität / Spezifität (%)	16,4 / 99,3		16,1 / 99,2
Gesamtrichtigkeit (%)	<b>94,0</b>		<b>94,0</b>
AUC der ROC Kurve mit 95% CI	<b>,851</b> ,842 - ,859		<b>,851</b> ,840 - ,862
<b>Präzision</b>			
beobachtete Mortalität (%)	6,3		6,3
Score-Prognose (%)	<b>6,3</b>		<b>6,3</b>

**Tabelle 25** SAPS 3

### **Box 1** Was bringt der Patient mit?

Informationen über die Patientenmerkmale vor Intensivaufnahme:

- Alter
- bisheriger Gesundheitszustand
- Komorbiditäten
- Ort vor Intensivaufnahme
- Krankenhausverweildauer vor Intensivaufnahme
- Anwendung größerer Therapiemöglichkeiten vor Intensivaufnahme

### **Box 2** Was führte zur Intensiv-Aufnahme?

Informationen über die Umstände der Intensivaufnahme:

- Gründe für die Intensivaufnahme
- anatomische Seite der Chirurgie (falls angebracht)
- geplante oder ungeplante Intensivaufnahme
- chirurgischer Status
- Infektion bei Intensivaufnahme

### **Box 3** Wie ist der aktuelle Zustand (Physiologie)?

Informationen über die Anwesenheit und den Grad der physiologischen Störung bei Intensivaufnahme (innerhalb einer Stunde vor oder nach Aufnahme):

- Glasgow Coma Scale (GCS)
- Herzfrequenz
- systolische Blutdruck
- Bilirubin
- Körpertemperatur
- Kreatinin (*neu*)
- Leukozyten
- Thrombozyten (*neu*)
- Wasserstoffionenkonzentration (pH) (*neu*)
- Atemhilfe und Oxygenierung

Moreno, R. P., P. G. Metnitz, et al. (2005). "SAPS 3--From evaluation of the patient to evaluation of the intensive care unit. Part 2: Development of a prognostic model for hospital mortality at ICU admission." *Intensive Care Med* 31(10): 1345-1355

## 8.2 Verzeichnis der akademischen Lehrer

Meine akademischen Lehrer waren die Damen und Herren:

### In Marburg:

Alter, Arnold, Aumüller, Bals, Barth, Basler, Baum, Behr, Berger, Bien, Bertalanffy, Bette, Christiansen, Cetin, Czubayko, Daut, Duda, El Farra, Engenhard-Cabillic, Feuser, Funk, Görg, Görg, Grundmann, Grimm, Griss, Groß, Hasilik, Hassan, Happle, Herzum, Hesse, Hoffmann, Hofmann, Höffken, Jungclas, Kalder, Kann, Kill, Kimberger, Klaus, Klenk, Klose, Köhler, König, Koolman, Krause, Kretschmer, Krieg, Kroll, Kuhlmann, Lange, Leonhard, Lohoff, Löffler, Maisch, Mandric, Moll, Moosdorf, Mueller, Neubauer, Oertel, Ramaswamy, Renz, Rausch, Remschmidt, Richter, Röhm, Rosenow, Rothmund, Ruchholtz, Rüper, Schäfer, Schmidt, Schmitz, Schnabel, Schneider, Seitz, Sesterhenn, Sommer, Sprinzl, Steiniger, Stiletto, Sure, Teymoortash, Vogelmeier, Voigt, Wagner, Weihe, Werner, Westermann, Wilhelm, Wulf, Wenimut, Wündisch, Zettl, Zwiork

### **8.3 Danksagung**

Mein besonderer Dank gilt Herrn PD Dr. med. Raphael Stiletto für die Überlassung des interessanten Themas, für seine Unterstützung sowie konstruktive Hilfestellung bei dieser Arbeit.

Ganz herzlich bedanke ich mich auch bei Herrn Prof. Dr. Rolf Lefering für seine sehr gute Betreuung, für die stetige Hilfsbereitschaft sowie großartige Unterstützung bei der Auswertung dieser Arbeit. Durch seine Korrekturen und Anregungen, im Rahmen der Fertigstellung dieser Dissertation bin ich ihm zu großem Dank verpflichtet.

Bedanken möchte ich mich weiterhin bei Herrn Prof. Dr. med. Ali Canbay für die Hilfe bei der Korrektur dieser Arbeit.

Bei meinen Geschwistern möchte ich mich aufrichtig bedanken, da sie in vielfältiger Weise zum Gelingen dieser Arbeit beigetragen haben.

Der größte Dank gebührt jedoch meinen Eltern, die mir immer und überall bei allen Problemen mit Rat und Tat zur Seite standen, die für meine Geschwister und mich alles taten, um aus uns gute Ärzte und erfolgreiche Wissenschaftler zu machen.