

## 5. Diskussion

In diesem Kapitel werden mehrere für die Arbeit relevante Punkte überprüft. Es wird diskutiert, ob die Auswahl der Patienten oder die Art der Durchführung der Studie eine Beeinflussung der Ergebnisse herbeiführt. Es folgt eine Diskussion der Ergebnisse der Analysen: Ereignis gegen Ereignis\* und RDI gegen RDI\*. Der Stellenwert der Testmethode wird durch den Vergleich mit den Resultaten anderer Studien zu Messverfahren der SBAS Erkennung eingeschätzt. Anschließend werden die Qualitäten der transthorakalen Impedanz erörtert. In den Schlußfolgerungen sind die Ergebnisse des Kapitels zusammengefaßt.

### 5.1 Beurteilung der Patientenauswahl und der Durchführung der Studie

Die Studie unterliegt äußeren Einflüssen, welche zu Veränderungen der Ergebnisse führen können. Mögliche Faktoren sind die Auswahl der Studienteilnehmer und die Durchführung der Studie.

#### 5.1.1 Beurteilung der Patientenauswahl

Die Gestaltung der Studiengruppe sollte eine möglichst hohe Wahrscheinlichkeit für das Auftreten von respiratorischen Ereignissen während des Schlafes gewährleisten. Die aufgestellten Kriterien für die Teilnahme von Patienten an der Studie haben Auswirkungen auf die Art der untersuchten Gruppe.

- Die teilnehmenden Patienten der Studie unterliegen einer Vorauswahl durch die schlafmedizinischen Ambulanz der Klinik. Bei allen Patienten ist der Aufenthalt im Schlaflabor zumindest durch eine subjektive Beschwerdesymptomatik oder durch eine objektive NLM System Untersuchung begründet. Es ist davon auszugehen, dass der Patientenanteil ohne objektive Symptome von SBAS (RDI <5 E/h) sehr gering ist.
- Die durch die Klinik erfolgte Selektion hat Auswirkungen auf verschiedene Merkmale der Patientengruppe:
  - Die Altersstruktur der Patientengruppe zeigt eine typische Häufigkeitsverteilung für SBAS erkrankte Patientengruppen ( $\bar{x} 57,9 \pm 11,5$ ; 33-80 J.). Eine ähnliche Altersverteilung ist auch in anderen Studiengruppen mit SBAS zu finden. Für die Normalbevölkerung wird ein Häufigkeitsgipfel der Prävalenz für Patienten im Alter zwischen 45-65 Jahren angegeben. 77% der an der Studie teilnehmenden Patienten liegen in diesem Altersbereich. [82, 18, 3, 62]
  - Das Verhältnis von Männern zu Frauen mit SBAS liegt in der Studie bei 5:1. Das in der Studie vorliegende Verhältnis beruht auf Zufall. Es spiegelt den geschlechtsbedingten Unterschied in der Prävalenz von SBAS wider.
  - Viele der untersuchten Patienten der Studiengruppe haben einen erhöhten

BMI ( $\bar{x}$  29,3  $\pm$  4,4; 20,8-38,7). Im Vergleich zur Normalbevölkerung ist in der Studiengruppe der Anteil an übergewichtigen Personen mit einem BMI > 25 kg/m<sup>2</sup> sowohl bei den männlichen Patienten (Norm: 68%, Studie: 84%) als auch bei den weiblichen Patienten (Norm: 49%, Studie: 83%) stark erhöht. Der Anteil an Personen mit sehr starkem Übergewicht, also einem BMI > 30 kg/m<sup>2</sup>, ist bei Männern mehr als doppelt so hoch (Norm: 17,4%, Studie 41%) und bei Frauen sogar viermal so hoch (Norm: 15,2%, Studie 67%) wie die Norm. Diese starke Erhöhung der BMIs findet sich auch in anderen Studiengruppen von SBAS kranken Patienten. Es handelt sich letztendlich um ein zu erwartendes Ergebnis bei dieser untersuchten Studiengruppe. Die Ergebnisqualität bei Studienpatienten mit normalgewichtigem BMI weist keinen relevanten Unterschied zu der Ergebnisqualität bei Patienten mit hohem BMI auf. [98, 82, 62]

- Die Patientenauswahl kann durch die Entscheidungsfindung des Patienten beeinflusst werden, da dieser freiwillig an der Studie teilnimmt. Kein Patient lehnte die Teilnahme an der Studie ab. Die Gruppengröße und -zusammensetzung wird hierdurch also nicht beeinflusst.
- Die Teilnahme der Patienten an der Studie wird nicht beschränkt durch Teilnahme Kriterien, welche bei Vorliegen von bestimmten Erkrankungen zu einem Ausschluss des Patienten von der Studie führt. Die Studie ist dadurch sehr realistisch und entspricht den zukünftigen Einsatzbedingungen des Untersuchungsverfahrens. Nachteilig ist, dass andere Erkrankungen Messparameter beeinflussen können. Erkrankungen mit starken Veränderungen des thorakalen Gas- und Flüssigkeitsverhältnisses nehmen Einfluss auf die Impedanz. Ebenso ist die Veränderung des Hautwiderstandes eine mögliche Quelle für fehlerhafte Messungen. Auftreten können diese Widerstandsveränderungen z.B. durch Temperaturveränderungen oder Veränderungen des Hautmilieus.

Zusammenfassend zeigt sich, dass die hier vorliegende Patientengruppe in ihren Eigenschaften einer repräsentativen SBAS Studiengruppe entspricht. Die Werte und die Verteilung charakteristischer Merkmale (Alter, Geschlecht und BMI) ähneln denen vergleichbarer Studien. Die Studiengruppe ist gut geeignet, um das Studienziel, eine Einschätzung der Qualität der transthorakalen Atemimpedanzmessung zur Erkennung von SBAS typischen Atemmustern, zu erreichen. Die Patientengruppe ist nicht für die Gesamtbevölkerung, sondern nur für SBAS kranke Gruppen repräsentativ.

### 5.1.2 Beurteilung der Studiendurchführung

Die Art und Weise der Studiendurchführung kann zu einer Beeinflussung der Ergebnisse führen.

#### (1) Einfluss durch die stationäre Untersuchungen der Patienten

Die Patienten werden außerhalb ihrer gewohnten häuslichen Umgebung untersucht. Als „first-night-effect“ bezeichnet man die Tatsache, dass Patienten gewöhnlich in der ersten Untersuchungsnacht in einem Schlaflabor schlechter schlafen als in den darauffolgenden Nächten. Der Effekt äußert sich in einem unruhigen Schlafverhalten und vermehrten

Bewegungen im Wachzustand. Es ergab sich nur bei zwei Patienten die Möglichkeit in aufeinanderfolgenden Nächten Daten aufzuzeichnen. Die vorhandene Datenmenge reicht für eine Beurteilung der Beeinflussung des Studienergebnisses durch den „first-night-effect“ nicht aus.

Das benutzte Testverfahren bietet keine Möglichkeit der Schlafstadienerkennung. Bewegungsartefakte, die bei alleiniger Betrachtung des Impedanz-Signals wie respiratorische Ereignisse imponieren können, wirken sich negativ auf die Ergebnisse der Studie aus. Das geplante Einsatzgebiet des Untersuchungsgerätes liegt in der häuslichen Umgebung des Patienten. Es ist dort mit einer Abnahme der Bewegungsartefakte zu rechnen.

Weiter besteht die Möglichkeit, Bewegungsartefakte zu identifizieren. In der graphischen Darstellung der Impedanz sind Bewegungsartefakte daran zu erkennen, dass sie, im Gegensatz zu dem durch Atmung produzierten Impedanzsignal, eine höhere, kürzere und steilere Amplitude aufweisen. Diese Eigenschaften könnten bei einer automatischen Auswertung durch einen Computer in einem Algorithmus mit eingebunden werden.

#### (2) Datenerhebung im Schlaflabor

Das Anlegen des LZ-EKG Gerätes erfolgt durch den Untersucher der Studie und wird in der Nacht nicht weiter kontrolliert. Die Messfühler der PSG Untersuchung werden durch das schlafmedizinische Personal angelegt und während der Nacht überwacht. Einige Unterschiede in der Aufzeichnungsqualität der PSG Aufzeichnung gegenüber der Aufzeichnung des Impedanzsignals sind hierdurch zu erklären.

#### (3) Auswertung der Daten

Die Daten der PSG und des LZ-EKG werden durch einen Auswerter beurteilt. Dieser ist durch die anonyme Speicherung der Daten verblindet gegenüber der Identität und den Begleiterkrankungen der untersuchten Versuchspersonen. Die Datenauswertung von PSG und LZ-EKG erfolgt getrennt voneinander und wird durch Übereinstimmungen in den Untersuchungsergebnissen beurteilt (Kriterien vgl. Kap. 3.4). Die Daten werden nur einmal durch einen Auswerter beurteilt. Daher sind keine Aussagen möglich über Diskrepanzen der Ergebnisse, die bei erneuter Auswertung durch denselben Auswerter entstehen können (intrascorer agreement). Auch bestehen keine Erkenntnisse bezüglich der Ergebnisse zwischen zwei verschiedenen Auswertern (interscorer agreement).

#### (4) Visuelle Auswertung des Datenmaterials

Es ist davon auszugehen, dass eine Veränderung der Hardware (z.B. die Verwendung eines kleineren Monitors bei der Auswertung) zu Veränderungen im Ergebnis führt (vgl. Kap. 3.3.3).

### **5.1.3 Einfluss der Körperlage auf die Ergebnisse der Studie**

In der Auswertung der Daten zeigt sich, dass eine Veränderung der Körperlage sowohl in der PSG- als auch in der Impedanz-Aufzeichnung zu einer Veränderung des Signals führen kann. Dies äußert sich zumeist in einer Veränderung der Grundlinie. In seltenen Fällen führt es in der PSG-Aufzeichnung zu einer mangelhaften Ausprägung der Atemamplitude. Dies ist wahrscheinlich dann der Fall, wenn durch die Lageänderung der

dehnungsempfindliche Teil des Gurtes entweder konstant maximal gedehnt oder aber komplett entspannt wird, so dass die Atemexkursionsbewegung nicht mehr auf den Dehnungssensor übertragen werden kann. Da jedoch ein Brust und ein Bauchgurt parallel verwendet wurden, war in der Regel zumindest ein Signal verwertbar.

In der Impedanz zeigt sich ebenso eine Beeinträchtigung des Signals bei Lagewechsel in Form einer Verschiebung der Grundlinie, welche als Veränderung der Grundimpedanz zu werten ist. Ursächlich hierfür sind durch den Lagewechsel bedingte Veränderungen im Gas/Flüssigkeitsverhältnis der durchmessenen Thoraxregion. Auch hier ist im Regelfall nach erfolgter Lageänderung wieder eine gut ausgeprägte Atemexkursionskurve vorhanden. Ausnahmen entstehen bei Elektrodenablösung die durch Lagewechsel bedingt sind.

## 5.2. Diskussion der Studienergebnisse

Die Ergebnisse der Analysen Ereignis gegen Ereignis\* und RDI gegen RDI\* werden getrennt erörtert. Mögliche Ursachen für Unterschiede in den Ergebnissen von PSG und LZ-EKG Untersuchungen werden genannt. Besonders stark abweichende Ergebnisse (Extremwerte), die in der Darstellung nach Bland und Altman [11] (vgl. Abb. 4-9) identifiziert werden können, werden individuell diskutiert.

### 5.2.1 Diskussion der Ereignis / Ereignis\* Analyse

Verschiedene Punkte sind bei der Ereignis / Ereignis\* Analyse zu betrachten:

(1) Unterschiede in der Erkennung von verschiedenen Ereignisarten

Bei der Ereignis / Ereignis\* Analyse lassen sich Unterschiede im Erkennungsvermögen der verschiedenen Ereignisarten feststellen (vgl. Tab. 4-2). In der folgenden Betrachtung werden Gemischte Apnoe wegen ihres zentralen Anteils der Gruppe der Zentralen Apnoe zugeschlagen. Aus der Tabelle ergibt sich die folgende Rangfolge richtig erkannter Ereignisse:

Zentrale Apnoe (97%) > Obstruktive Apnoe (89%) > Hypopnoe (81%)

Von 2154 Fällen Gemischter und Zentraler Apnoe werden durch das Testsystem 97% richtig als Ereignis erkannt. Von 2462 Fällen Obstruktiver Apnoe werden 89% und von 9733 Hypopnoe werden 81% richtig erkannt.

Folgende Mechanismen können zu Impedanzänderungen führen und damit die Rangfolge beeinflussen:

- Intrathorakales Gas- und Flüssigkeitsverhältnis: Veränderung des Flüssigkeitsanteils  
Während der normalen Inspiration entsteht sehr kurzzeitig ein negativer Druck im Thorax, der durch den Einstrom von Luft ausgeglichen wird. Dies führt zu einer Änderung in der Impedanz. Wenn die Atemwege ganz oder teilweise verlegt sind, kann der Druckausgleich durch die Atemluft gar nicht oder nur bedingt stattfinden. Es resultiert ein länger andauernder Unterdruck im Thorax. Dieser negative Druck

fördert den Einstrom von Blut in den Thoraxbereich. Es kommt zu einer Veränderung des Gas- und Flüssigkeitsverhältnisses und damit zu einer Erhöhung der Gesamtimpedanz. In einer graphischen Darstellung würde sich bei einer Obstruktiven Apnoe ein höheres Niveau der Impedanz ergeben. Dadurch verringert sich die relative Amplitude und es bietet sich eine schlechtere Möglichkeit der Diskriminierung. Gleiches gilt in geringerem Ausmaß auch bei einer Hypopnoe.

- Intrathorakales Gas- und Flüssigkeitsverhältnis: Veränderung des Gasanteils  
Bei normaler Atmung bewirkt die Erhöhung des Gasanteils im Thorax die größte Impedanzänderung. Mit dem benutzten System ist es nicht möglich, die Menge der eingeatmeten Luft zu messen. Ob die eingeatmete frische Luft für eine ausreichende Sauerstoffanreicherung des Blutes ausreichend ist, kann nicht gesagt werden.
- Aktive Muskelarbeit  
Die Impedanz der Atemmuskeln ändert sich während der Kontraktion. Bei alleiniger Betrachtung der transthorakalen Impedanz können sogenannte frustrane Atembewegungen, wie sie bei Obstruktiven Apnoen auftreten, zu Signalen führen, die nach den Kriterien der Studie einer suffizienten Atmung entsprechen. Bei Obstruktiven Apnoe und bei Hypopnoe treten Muskelkontraktionen auf, die im Impedanzsignal zu einer schlechteren Erkennbarkeit von Ereignissen führen.
- Gekoppelte Ereignisdefinition  
Die Definition der Hypopnoe in der PSG besteht aus einer Verknüpfung von Atemminderung um mindestens 50% und einer folgenden  $sO_2$  Minderung und/oder einer Arousal Reaktion. In der Studie wird nur die Verknüpfung mit einer  $sO_2$  Minderung als Hypopnoe gewertet, da die EEG Kanäle nicht dargestellt werden. Das Erkennen von Hypopnoe bei einem Monorecordersystem, welches weder eine EEG noch eine  $sO_2$  Messung durchführt, kann nur auf der Auswertung der Atemminderung basieren. Falsch positive Ereignisse\* können entstehen, wenn zwar eine in der PSG sichtbare Atemminderung über 50% vorliegt, die auch in der Impedanzmessung erkannt wird, aber in der PSG keine  $sO_2$  Minderung folgt (oder eine Arousal Reaktion stattfindet). In diesem Falle erkennt das Testsystem zwar die Atemminderung, es kann aber nicht festgestellt werden, ob die zweite Bedingung ( $sO_2$  Abfall oder Arousal) erfüllt wird. Dies kann zu falsch positiven Ereignissen\* führen.

Die genannten Mechanismen beeinflussen vor allem die Messung von Obstruktiven Apnoe und Hypopnoe. Letztere bilden mit 68%, bezogen auf die Gesamtzahl der Ereignisse, den größten Anteil an der Menge von Ereignissen bei SBAS (vgl. Abb. 4-5). Allein durch die Messung der transthorakalen Atemimpedanz ist keine Unterscheidung von Apnoe und Hypopnoe möglich. Dies ist nach Angaben der AASM auch nicht weiter erforderlich, da Apnoe und Hypopnoe nicht nur dieselben Folgen für den Patienten haben, sondern auch die Behandlung der nächtlichen Atmungsstörung in beiden Fällen gleich ist. [1]

## (2) Bewertung der Analyseergebnisse

Die Qualität der Ereigniserkennung ist mit einer Sensitivität von 0,84 (vgl. Kap. 4.2.3) als

gut zu bewerten, was sich auch in der Erkennung von SBAS erkrankten Personen in der RDI / RDI\* Analyse zeigt. Alle erkrankten Personen werden als SBAS krank erkannt. Die Qualität im Erkennen von Zeitabschnitten ohne Ereignis zeigt sich in der Spezifität von 0,95 (vgl. Kap. 4.2.3). Dieser als hoch einzustufende Wert gibt Anlass zu der Vermutung, dass bei einer Messung von Personen ohne SBAS diese auch als solche erkannt werden. Weitere Untersuchungen an einem gemischten Patientenkollektiv aus gesunden und kranken Patienten sind jedoch notwendig, um in der Praxis die Diskriminierungsfähigkeit zwischen diesen beiden Fällen zu bestätigen.

Die Richtigkeit der getroffenen Entscheidungen in der Auswertung der transthorakalen Impedanz wird durch eine hohe Relevanz (positiv prädikativer Wert) von 0,92 (vgl. Kap. 4.2.4) und einer hohen Segreganz (negativ prädikativer Wert) von 0,91 (vgl. Kap. 4.2.4) bestätigt. Diese hohe Übereinstimmung der beiden Systeme ist nicht zufällig. Der Kappa-Koeffizient nach Cohen  $\kappa = 0,81$  (vgl. Kap. 4.2.5) berücksichtigt die Wahrscheinlichkeit von zufällig gleichen Beobachtungen. 81% der getroffenen Entscheidungen gehen demnach über die zufällig zu erwartende Übereinstimmung hinaus. Die AASM sieht einen Kappa Wert von 0,6 als untere Grenze für eine zuverlässige Messung an. [1] Der in der Studie ermittelte Wert ist als hoch anzusehen und unterstreicht die Möglichkeiten der transthorakalen Impedanzmessung zur Erkennung von SBAS Ereignissen.

### 5.2.2 Diskussion der RDI / RDI\* Analyse

Grundsätzlich geht es um die klinische Einsatzfähigkeit des Testsystems. Es stellt sich die Frage, ob die in der Ereignis / Ereignis\* Auswertung beobachteten falschen Ereignisse\* (falsch positive und falsch negative Ereignisse\*) Auswirkungen haben auf die klinische Beurteilung der Patienten im Rahmen der Schweregradeinteilung von SBAS.

Die in der Arbeit errechneten RDIs\* beinhalten sowohl falsch positive als auch falsch negative Ereignisse\*. Als zeitlicher Bezugsrahmen ist die gesamte parallele Aufzeichnungszeit gewählt. Dies entspricht nicht der Definition des RDI, der eigentlich auf die tatsächliche Schlafzeit bezogen ist. Da jedoch im klinischen Alltag die Schlafzeit nur mit einem EEG bestimmbar ist und das LZ-EKG ein solches nicht besitzt, wird auch bei zukünftigen Untersuchungen von Monorecordern die Gesamtaufzeichnungszeit den zeitlichen Rahmen bilden müssen. Wie bereits erwähnt, können bewegungsbedingte Artefakte bei der Impedanzmessung zu falsch positiven Ereignissen\* führen. Bei alleiniger Betrachtung von Ereignissen\*, die in der Schlafzeit auftreten, ist damit zu rechnen, dass die Anzahl falsch negativer Ereignisse\* geringer wird. Bewegungsbedingte Artefakte treten vor allem im wachen Zustand des Patienten auf.

Die Aussagekraft dieser Analyse wird sicher eingeschränkt durch die geringe Patientenzahl ( $n=60$ ) und die Selektion der Studienteilnehmer durch den Aufenthalt im Schlaflabor selbst. Studien mit einer größeren Patientenzahl können wegen der großen Mengen an aufgezeichneten Daten nur in Multicenter Studien durchgeführt werden. Sie bedürfen einer automatischen Auswertung durch ein Computersystem.

Aus der Abb. 4-8 ist ein linearer Zusammenhang zwischen den RDIs und RDIs\* zu erkennen. Der Korrelationskoeffizient nach Pearson ergibt eine Korrelation von  $r = 0,93$  der Wertepaare.

In der Darstellung nach Bland und Altman (vgl. Abb. 4-9) sind die Grenzen der zweifachen Standardabweichung und der Mittelwert der Abweichung bezogen auf alle Messwerte eingezeichnet. Betrachtet man die Ergebnisse\* im RDI Bereich bis 20 E/h und über 20 E/h getrennt, so ergeben sich unter Verwendung der Tab. 7-4 (Anhang) die in Tab. 5-1 errechneten Werte. Man erkennt eine Tendenz zu erhöhten RDIs\* im LZ-EKG bei RDIs zwischen 0 und 20 E/h in der PSG. Bei RDIs über 20 E/h in der PSG fallen die RDIs\* eher niedriger aus. Die Mittelwerte der RDI\* Abweichungen für die Patientengruppen mit einem RDI < 20 E/h und einem RDI > 20 E/h bestätigen diesen Trend. Auch die Standardabweichung wird mit zunehmenden RDI größer.

	RDI ≤ 20 (E/h)	RDI > 20 (E/h)
Mittelwert der Abweichung RDI* zu RDI	0,8	-3,9
Standardabweichung RDI* zu RDI	±3,1	±7,1

**Tab. 5-1:** Mittelwerte der RDI\* Abweichungen

Das geplante Einsatzgebiet des LZ-EKG Recorders ist das ambulante Screening von Patienten auf SBAS. In dieser Hinsicht ist besonders der niedrige Mittelwert der Abweichung und die geringe Standardabweichung im Bereich  $RDI \leq 20$  E/h wichtig. Dieser Bereich dient der Differenzierung zwischen gesunden und mittelschwer erkrankten SBAS Patienten. Bei Patienten mit einem höheren RDI ist die größere Standardabweichung und Abweichung des Mittelwertes tolerabel, da das Untersuchungsziel nicht in der Einteilung des Schweregrades der SBAS Patienten liegt, sondern einzig dem Erkennen von SBAS kranken Patienten dienen soll.

### 5.2.3 Extremwerte der Studie

Zwei Wertepaare liegen außerhalb der zweifachen Standardabweichung (vgl. Abb. 4-9). Diese beiden Extremwerte haben einen großen Einfluss auf die Ergebnisse der Ereignis/ Ereignis\* Analyse und verändern die Ergebnisse zum Negativen. Es ist zu prüfen, ob bei diesen Patienten Fehler in der Aufnahme oder Auswertung begangen wurden, die zu einer Disqualifizierung der Patientenergebnisse führen können. Außerdem ist es notwendig zu ergründen, ob diese schlechten Ergebnisse in einem Zusammenhang mit einem möglichen systematischen Fehler stehen, der Einfluss auf die grundsätzliche Messqualität des Testsystems hat. Es handelt sich bei den Extremwerten um die Messergebnisse der Patienten mit den Identifikationsnummern (ID) 01036 und 01060. Die Aufzeichnungen dieser Patienten wurden ein zweites Mal intensiv überprüft.

#### (1) Extremwert von Patient ID 01036

Der Unterschied RDI\* zu RDI beträgt bei diesem Patienten -30,56 E/h. Bei diesem Patienten wurden 31 E/h nicht durch die transthorakale Atemimpedanz erkannt.

Die Qualität der Aufzeichnung ist gut. Ein Zusammenhang zwischen falsch negativen Ereignissen\* und der Körperlage ist nicht festzustellen. Es fällt jedoch auf, dass dieser Patient die kürzeste  $\emptyset$ Ereigniszeit (13,5 s) und den niedrigsten Wert der Minimalzeit zwischen zwei Ereignissen (27,9 s) hat (vgl. Tab.7-2). Weiterhin fällt auf, dass 95% der gemessenen Ereignisse bei diesem Patienten Hypopnoe sind. Es wird vermutet, dass die Kombination von der vorwiegenden Ereignisart (Hypopnoe) als auch die kurze  $\emptyset$ Ereigniszeit zu diesem schlechten Ergebnis beigetragen haben. Dies kann jedoch nicht mit Sicherheit gesagt werden, da ein Patient (ID 01054) mit gleicher Minimalzeit zwischen zwei Ereignissen (27,9 s) allerdings unterschiedlicher  $\emptyset$ Ereigniszeit (18,2 s) nur eine Abweichung des RDI\* zu RDI Wertes von 0,1E/h hat. Bei diesem Patienten sind nur 73% der gemessenen Ereignisse auf Hypopnoe zurückzuführen. Vielleicht gibt es einen Schwellenwert für das Erkennen von Ereignissen, basierend auf der  $\emptyset$ Ereigniszeit, der im Falle des Patienten (ID 01036) unterschritten wird. Dies ist jedoch nur eine Hypothese, letztendlich bleibt die Ursache für die extreme Abweichung der Ergebnisse ungeklärt.

## (2) Extremwert von Patient ID 01060

Der Unterschied RDI\* zu RDI beträgt bei diesem Patienten -23,43 E/h. Die  $\emptyset$ Ereigniszeit (17,4 s) und die Minimalzeit zwischen zwei Ereignissen (34,6 s) liegen unterhalb der durchschnittlichen Werte. Der Anteil von Hypopnoe an der Gesamtzahl der gemessenen Ereignisse liegt bei 91%. Das oben Gesagte könnte auch hier einen Einfluß auf dieses Ergebnis haben.

Die erneute Betrachtung der Aufzeichnungen dieses Patienten zeigt Mängel in der Aufnahmequalität. Diese können auf Lageänderungen während des Schlafs zurückgeführt werden. So hat sich der Patient um 23:36 Uhr aus der Rückenlage in die Linkslage und um 01:29 Uhr in die Rechtslage bewegt. Innerhalb dieses Zeitraums ist die Qualität der Aufzeichnung schlechter. Nach dem Lagewechsel auf die rechte Seite ist die Qualität wieder besser geworden. Es ist zu vermuten, dass die rechte Elektrode durch die Drehung auf die linke Körperhälfte einen schlechteren Kontakt hatte (eventuell durch Zugspannung am Elektrodenkabel). Durch die Zurückverlagerung hat sich der Kontakt wieder verbessert. Ein Drittel aller falsch negativen Ergebnisse\* bei diesem Patienten traten innerhalb dieses Zeitraums auf. Die Zahl von falsch negativen Ereignissen\* ist in diesem Zeitraum mit 32 E/h stark erhöht gegenüber der Zahl falsch negativer Ereignisse\* in der übrigen Aufzeichnungszeit (22 E/h).

Die dargelegten Gründen erscheinen nicht ausreichend für eine nachträgliche Disqualifizierung der beiden Patienten. Es kann aber auch keine befriedigende Begründung für das Auftreten dieser Extremwerte gegeben werden. Die Daten beider Patienten werden weiter berücksichtigt.

## 5.3 Vergleich der Ergebnisse dieser Studie mit den Resultaten anderer Studien

Geräte für das Erkennen von SBAS werden in vier Kategorien, Level I bis IV, unterteilt (Kap. 2.2.2). [24] Unterschiede liegen in der Anzahl der verwendeten Aufzeichnungskanäle und in der Unterscheidung zwischen stationären und mobilen Geräten.

Zunächst werden Studien betrachtet, in denen eine atemabhängige Impedanz für das Erkennen von Atmungsstörungen benutzt wird. Danach werden die Ergebnisse der



vorliegenden Studie verglichen mit den Ergebnissen von Studien die Monorecorder mit anderen Messverfahren verwenden. Abschließend findet ein Vergleich mit Studien statt, in denen Mehrkanalrecordersysteme zur Anwendung kommen.

### 5.3.1 Vergleich der Studienergebnisse mit atemabhängigen Impedanzstudien

Zu den atemabhängigen Impedanzstudien gehören die Studien von Scharf et al. und Saarelainen et al. [75, 71]. In Tab. 5-2, 5-3 und 5-4 sind die Patientendaten und Ergebnisse dieser Studien und der vorliegenden Studie tabellarisch zusammengefasst. Im folgenden werden die allgemeine Daten, die Ereignis / Ereignis\* Ergebnisse und die RDI / RDI\* Analysen betrachtet.

#### (1) Allgemeine Daten der Studien

Charakteristische Daten der untersuchten Patientengruppen und studienspezifische Eigenheiten sind in Tab. 5-2 aufgeführt.

	Diese Studie	Scharf et al.	Saarelainen et al.
Elektrodenplatzierung	transthorakal durch externe Elektroden	intrathorakal durch Herzschrittmacher	ges. Körper Arm zu Arm
Referenzmethode	PSG	PSG	PSG
Patientenanzahl (n)	60	22	14
Alter (J)	$\bar{\varnothing}58 \pm 11,5$ (33-88)	$\bar{\varnothing}65,6 \pm 13,4$ (k. A.)	$\bar{\varnothing}46 \pm 9,6$ (30-63)
Geschlecht	♂49 : ♀11	♂11 : ♀8	♂12 : ♀2
BMI (kg/m <sup>2</sup> )	$\bar{\varnothing}29,3 \pm 4,4$ (21-39)	$\bar{\varnothing}27,5 \pm 5,8$ (20-41)	$\bar{\varnothing}29,4 \pm 5,3$ (25-47)
$\bar{\varnothing}$ Bezugszeit (h)	$7,6 \pm 0,91$ (TRT)	$6,2 \pm 1,8$ (TST)	$6,9 \pm 1,1$ (TST)
Definition der Apnoe im Impedanzsignal	$t_{\min} = 10$ s Signal < 50 % der Ausgangsamplitude	$t_{\min} = 10$ s Signal < 50 % der Ausgangsamplitude	$t_{\min} = 7$ s Signal < 40 % der Ausgangsamplitude
Definition der Hypopnoe im Impedanzsignal	$t_{\min} = 10$ s Signal < 50 % der Ausgangsamplitude	$t_{\min} = 10$ s Signal < 50 % der Ausgangsamplitude	$t_{\min} = 7$ s Signal 40 - 50 % der Ausgangsamplitude <sup>1</sup>

**Tab. 5-2:** Grunddaten der atemabhängigen Impedanzstudien

Eine Impedanz kann grundsätzlich an jeder beliebigen Körperstelle zwischen zwei

<sup>1</sup> "Hypopnoeas were defined as a period of 50 - 60% decrease in the respiratory amplitude lasting at least 7s and followed by an increase in the respiratory amplitude" [71, S. 110].

Elektroden gemessen werden. Um die Signale der Atmung aufzuzeichnen ist es notwendig, dass die Strecke zwischen den Ableitelektroden zumindest teilweise durch den Thorax führt. Die in den Studien verwendeten Messmethoden der Impedanz unterscheiden sich in der Platzierung der Elektroden. In der hier durchgeführten Studie sind die Elektroden an gegenüberliegenden Seiten der Thoraxwand befestigt. Scharf et al. benutzen eine intrathorakale Elektrode des Schrittmacherimplantats und das Schrittmachergehäuse als extrathorakale Elektrode. Saarelainen et al. befestigen die Elektroden an den distalen oberen Extremitäten. Hauptunterschiede der Messmethoden liegen bei den Strecken und dem Gewebe zwischen den Elektroden.

In allen Studien gilt die PSG als Goldstandard und wird als Referenz benutzt.

Die Patientenzahl (n) ist in den Studien unterschiedlich groß. In der vorliegenden Studie werden die Ergebnisse der gesamten Patientengruppe betrachtet. Die zwei Patienten mit stark abweichenden Ergebnissen verbleiben in der Studie.

Das Durchschnittsalter der Patienten liegt bei Saarelainen et al. relativ niedrig. Dieses könnte darauf zurückzuführen sein, dass Patientengruppen mit Erkrankungen, die vermehrt im Alter auftreten, nicht in die Studie aufgenommen wurden.

Das Verhältnis der Geschlechter in der Studie entspricht in der vorliegenden Studie und bei Saarelainen et al. in etwa der natürlichen Verteilung der Erkrankung. Bei Scharf et al. ist der Frauenanteil erhöht. Dieses kann darauf zurückzuführen sein, dass nur Patienten ausgewählt wurden, die in einem bestimmten Zeitraum einen Herzschrittmacher bekommen hatten.

Der durchschnittliche BMI ist bei allen Studien erhöht. Dies ist bei SBAS Patienten häufig zu erwarten.

Als Bezugszeit ist die in den Studien verwendete Aufzeichnungszeit zu sehen, in denen Testgerät und Goldstandard parallel ausgewertet wurden. Hier bestehen wichtige Unterschiede zwischen den Studien. Bei Scharf et al. und Saarelainen et al. ist die Total Sleep Time (TST) als Bezugszeit gewählt. Dies entspricht der Zeit, in der an Hand der PSG-Auswertung bei Patienten eine Schlafphase vorliegt, die nicht als Wach- oder Bewegungszeit gewertet ist, also den Schlafstadien S1 - S4 oder REM entspricht. Ereignisse, die während Wachstadien und Bewegungsstadien stattfinden, fließen nicht in die Auswertung ein. Das Vorgehen kann zu einer Verbesserung der Ergebnisse führen, da bewegungsbedingte Artefakte im Impedanzsignal gehäuft in den Wachzeiten auftreten. Dieses gehäufte Auftreten kann bei der hier vorliegenden Studie zu einer Erhöhung der Anzahl falsch positiver Ereignisse\* führen. Unter alleiniger Verwendung der Impedanz ist ein sicheres Erkennen des Zustandes Schlaf nicht möglich. In der hier vorgenommenen Studie wird die TRT (Total Recording Time) als Bezugszeit benutzt. Dies entspricht weitgehend den späteren Bedingungen für den Einsatz des Gerätes. Unter Verwendung eines Fragebogens können Wachzeiten des Patienten grob geschätzt werden, sofern der Patient den Zeitraum der Wachphase angeben kann. Diese Angabe ist jedoch nur bedingt verwendbar, da der Zeitpunkt an dem der Patient wieder eingeschlafen ist weder von dem Patienten notiert noch an Hand des aufgezeichneten Signals bestimmt werden kann.

Saarelainen et al. benutzen andere Kriterien für die Bewertung des Impedanzsignals und stützen sich dabei auf die Ergebnisse eigener Voruntersuchungen. Danach führt eine

Apnoe von 10s Dauer in der PSG zu einem ca. 7s dauernden Signalabfall in der Impedanz. In der hier vorliegenden Studie in eine solche Veränderung der zeitlichen Länge eines Ereignisses\* gegenüber eines Ereignisses nicht in diesem Ausmaß festzustellen. Ferner wird eine stärkere Signalminderung als Kriterium für ein Ereignis\* benutzt.

In keiner der Impedanzstudien ist eine Unterscheidung von Hypopnoe und Apnoe anhand des gemessenen Impedanzsignals möglich.

(2) Zusammenstellung der Ergebnisse Ereignis / Ereignis\* in den verschiedenen Studien

Die Ergebnisse des direkten Vergleichs von Ereignissen\* in der Impedanz und Ereignissen in der PSG sind in Tab. 5-3 dargestellt.

	Eigene Studie	Scharf et al.	Saarelainen et al.
Anzahl der Ereignisse in PSG (n)	14389	803	2953
Richtig erkannte Ereignisse* in Impedanz (n)	12142 (84 %)	724 (90 %)	k.A. (77 %)
Falsch positiv* (n)	1126	229	k.A.
Sensitivität	0,84	0,82	k.A.
Spezifität	0,95	0,74	k.A.
positiv prädikativer Wert	0,92	0,76	k.A.
negativ prädikativer Wert	0,91	0,76	k.A.
$\kappa$ -Wert	0,81	0,69	k.A.

**Tab. 5-3:** Ergebnisse der Ereignis / Ereignis\* Analysen

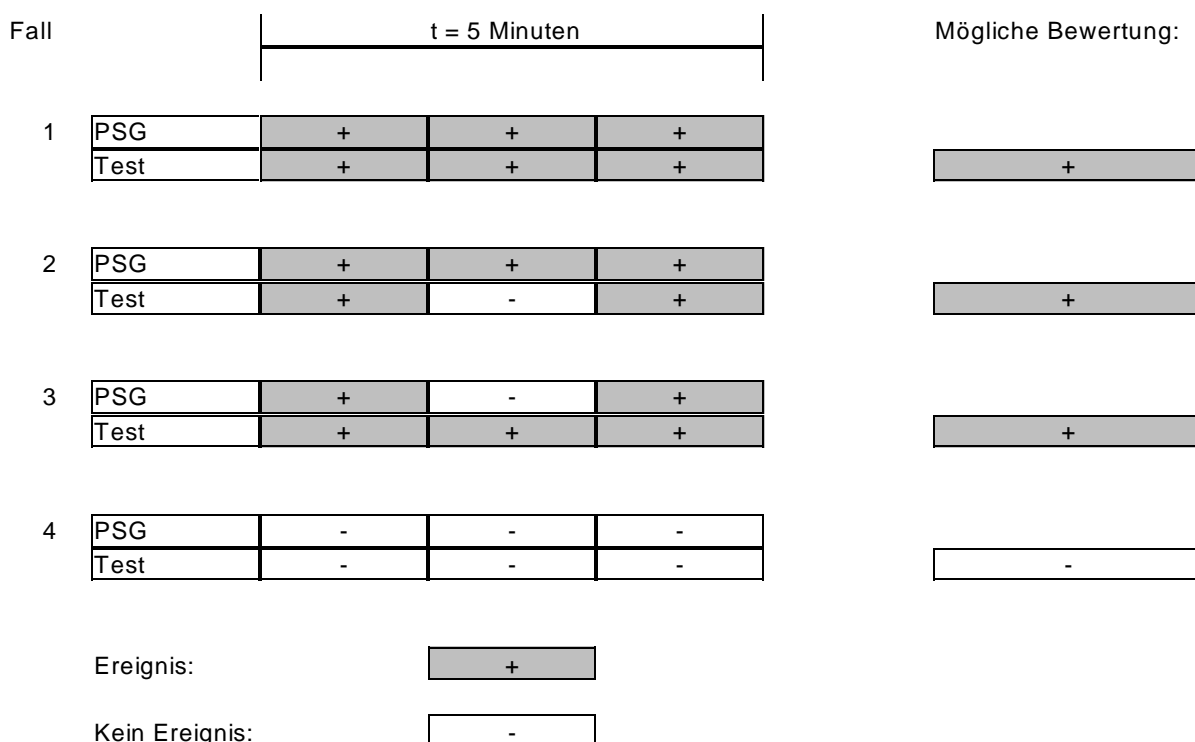
In der Studie von Saarelainen et al. findet man keine weiteren Angaben zu den Ergebnissen der Ereignis / Ereignis\* Analyse. Aus diesem Grunde werden hier nur die Ergebnisse von Scharf et al. zu Vergleichszwecken berücksichtigt.

Es fällt auf, dass die Anzahl (n) der Ereignisse und Ereignisse\* in der vorliegenden Studie sehr viel größer sind als bei Scharf et al. Die Kriterien für richtig erkannte Ereignisse\* für die vorliegende Studie sind in Kap. 3.4 aufgeführt. Scharf et al. wenden ein anderes methodisches Vorgehen an. Die Untersuchungs Nächte werden in Epochen zu 5 Minuten unterteilt und retrospektiv nach dem Auftreten von Ereignissen für beide Messverfahren bewertet. Die Anzahl (n) der Ereignisse in der Tabelle entspricht bei Scharf et al. der Anzahl von 5 Minuten Epochen mit aufgetretenen Ereignissen. Die Anzahl der Ereignisse innerhalb einer 5 Minuten Epoche wird vernachlässigt.

Beide Studien haben vergleichbare Ergebnisse bezüglich der Sensitivität.

In der Studie von Scharf et al. wird jedoch eine deutlich niedrigere Spezifität ermittelt. Dieser Unterschied kann sich möglicherweise aus dem unterschiedlichen methodischen Vorgehen ergeben. Die Spezifität ( $d/(d+b)$ ) ist die Wahrscheinlichkeit, ein negatives Ereignis als solches zu erkennen (vgl. Tab. 3-1). Bezogen auf die Studie handelt es sich um die Wahrscheinlichkeit, einen Zeitabschnitt ohne Ereignis in der PSG als solchen durch die LZ-EKG Aufnahme zu erkennen. In der Formel für die Spezifität wird die Zahl richtig negativer Ereignisse\* der Testmethode (d) bezogen auf die Gesamtzahl wahrer negativer Ereignisse des Goldstandards (d+b). In der Studie von Scharf et al. werden bei 22 untersuchten Patienten 605 richtig negative Ereignisse\* (d) erkannt. Hochgerechnet auf die in dieser Studie untersuchte Patientenzahl (n=60) ergibt dieses theoretisch 1650 richtig negative Ereignisse\* (d). Im Vergleich hierzu haben sich in dieser Studie 22 552 richtig negative Ereignisse\* (Zeitabschnitte) ergeben. Aufgrund der unterschiedlich großen Zahl richtig negativer Ereignisse\* (d) wirken sich gleiche Veränderungen der Zahl falsch positiver Ereignisse\* (b) unterschiedlich stark aus. Die Studie von Scharf et al. ist anfälliger für falsch positive Ereignisse\* (b).

Weitere Unterschiede ergeben sich durch die starre Zeiteinteilung von 5 Minuten bei Scharf et al. In diesem Zeitraum kann mehr als ein Ereignis auftreten. Das kann u.a. zu den in Abb. 5-1 aufgezeigten Bewertungskonstellationen führen.



**Abb. 5-1:** Bewertungskonstellationen bei fixem Zeitrahmen

In der Abb. 5-1 ist eine starre Zeiteinteilung von 5 Minuten mit drei Ereignissen vorgegeben. Vier Fälle (1-4) und deren mögliche Bewertung sind dargestellt. In den Fällen 1 und 4 besteht zwischen PSG und Test vollständige Übereinstimmung. Im Fall 2 wird ein wahres Ereignis in der PSG durch das Testsystem nicht erkannt. Eine mögliche Bewertung des gesamten Zeitabschnittes ist die Bewertung als richtig positiv. Dabei wird ein falsch negatives Ereignis ignoriert. Würde man jedoch den Gesamtzeitraum als falsch negativ bewerten, so würden zwei richtig positive Ereignisse ignoriert. Im Fall 3 zeigt der Test ein positives Ereignis. In der PSG liegt kein Ereignis vor. Eine mögliche Bewertung des gesamten Zeitabschnittes ist die Bewertung als richtig positiv. Dabei wird ein falsch positives Ereignis ignoriert. Wie man leicht erkennt, kommt es durch die starre Zeiteinteilung zu einer ungenauen Bewertung von falsch positiven (b) und falsch negativen (c) Ereignissen.

Ungeachtet der vorliegenden methodischen Unterschiede ist beiden Studien jedoch zu entnehmen, dass das Erkennen von respiratorischen Ereignissen bei SBAS anhand einer Impedanzmessung als sehr zufriedenstellend zu bewerten ist. Dies wird durch die positiven  $\kappa$ -Werte  $>0,6$  bestätigt.

### (3) RDI gegen RDI\* Analyse der Impedanzstudien

Diese Ergebnisse sind nicht durch die oben genannten Probleme der starren Zeiteinteilung beeinträchtigt. Es ist jedoch der unterschiedliche Bezugszeitraum für die Errechnung der RDIs (vgl. Tab. 5 -2) zu bedenken. Ein längerer Bezugszeitraum führt zu einem relativ niedrigeren RDI bei gleicher Anzahl von Ereignissen. Eine Bewertung der reinen Schlafzeit vermindert möglicherweise die Anzahl bewegungsbedingter Artefakte während der Wachphase, die in der Impedanz als Ereignisse\* gewertet werden können. Da aber artefaktbedingte Ereignisse\* in der Impedanz bei der RDI\* Ermittlung als wahre Ereignisse angesehen werden, kann dies bei Nutzung der Gesamtaufzeichnungsdauer als Bezugszeit zu einer Erhöhung der Ereigniszahl\* und zu einer Erhöhung des RDIs\* führen.

	Diese Studie		Scharf et al.		Saarelainen et al. <sup>1</sup>	
RDI in PSG (E/h)	31,4 ± 17,7		24,0 ± 21,0		34,6 ± 31,0	
RDI* in Impedanz (E*/h)	28,9 ± 15,6		22,5 ± 16,0		33,4 ± 29,1	
∅ Differenz RDI* zu RDI	- 2,5		- 1,5		- 1,2	
SD der ∅ Differenz	± 6,6		± 10,6		± 8,4	
Koeffizient nach Pearson	0,93		0,87		0,96	
Sensitivität bei C <sub>x/h</sub> <sup>2</sup>	1,00	C <sub>5/h</sub>	0,94	C <sub>5/h</sub>	0,75	C <sub>5/h</sub>
	0,92	C <sub>15/h</sub>	1,00	C <sub>15/h</sub>	0,89	C <sub>15/h</sub>
	0,84	C <sub>20/h</sub>	1,00	C <sub>20/h</sub>	1,00	C <sub>20/h</sub>
	0,89	C <sub>30/h</sub>	0,78	C <sub>30/h</sub>	1,00	C <sub>30/h</sub>
Spezifität bei C <sub>x/h</sub>	1,00	C <sub>5/h</sub>	0,80	C <sub>5/h</sub>	0,50	C <sub>5/h</sub>
	0,64	C <sub>15/h</sub>	0,89	C <sub>15/h</sub>	0,80	C <sub>15/h</sub>
	0,94	C <sub>20/h</sub>	1,00	C <sub>20/h</sub>	0,71	C <sub>20/h</sub>
	0,97	C <sub>30/h</sub>	0,92	C <sub>30/h</sub>	0,86	C <sub>30/h</sub>

**Tab. 5-4:** Ergebnisse der RDI gegen RDI\* Analysen

Die Studie von Scharf et al. gibt die RDIs und RDIs\* aller Patienten tabellarisch an, wohingegen bei der Studie von Saarelainen et al. diese Werte aus einer Grafik ermittelt werden müssen.

Die Werte der Sensitivität und Spezifität werden für die Daten aller drei Studien errechnet und sind in Tab. 5-4 aufgelistet. Die Cut off Werte (C<sub>x/h</sub>) bei RDIs von 5, 15, und 30 E/h sind von der AASM als Grenzwerte zwischen gesunden Patienten zu Patienten mit milden, moderaten und schweren SBAS Erkrankungen angegeben (vgl. Kap. 4.3.3). Scharf et al. setzen die Grenze von moderaten zu schweren SBAS bei einem RDI von 20 E/h an. Die Frage, welche RDIs letztendlich als Grenzwerte zu sehen sind, ist nicht geklärt. Es handelt sich bei den Aussagen der AASM lediglich um Empfehlungen, die auf der Auswertung verschiedener Studien beruhen. [1]

Die durchschnittlichen RDIs der Studien sind unterschiedlich. Bei Scharf et al. ist der durchschnittliche RDI besonders niedrig. Dies kann darauf beruhen, dass Scharf et al. andere Kriterien für die Teilnahme an der Studie aufstellt haben. Das Patientenkollektiv

<sup>1</sup>Aufgrund fehlender Angaben mussten die RDI\*-Werte aus Grafiken abgelesen werden. Folgeberechnungen basieren auf diesen Daten.

<sup>2</sup>C<sub>x/h</sub> = zugrundeliegender Cut off Wert

stammt aus einer Gruppe von Herzschrittmacherpatienten. Das Vorliegen einer SBAS und damit die große Wahrscheinlichkeit eines erhöhten BMI, ist kein Kriterium für die Teilnahme an der Studie.

Wie man weiter aus Tab. 5-4 erkennt, liegt in allen Studien eine negative mittlere Abweichung ( $\varnothing$ Differenz RDI\* zu RDI) vor. Es bedeutet, dass die Testsysteme weniger Ereignisse\* erkannt haben als im Goldstandard vorhanden sind. Die Abweichung ist in der vorliegenden Studie mit -2,5 E\*/h am Größten. In dieser Studie und auch bei Scharf et al. wird festgestellt, dass bei zunehmenden RDIs die Abweichung der RDIs\* größer wird (vgl. Tab. 5-1). Eine mögliche Ursache für den Unterschied der durchschnittlichen Differenz RDI\* zu RDI kann durch ein Patientenkollektiv mit insgesamt höheren RDIs in der vorliegenden Studie erklärt werden.

In der Studie von Saarelainen et al. wird mit einer kleinen Studiengruppe (n=14) gearbeitet. Dabei hat ein Patient einen RDI von 113 E/h. Rechnet man diesen Patienten heraus, so ergibt sich ein  $\varnothing$  RDI von 28,5 E/h (anstelle von 34,6 E/h).

Die Standardabweichung der durchschnittlichen Differenz ist in der vorliegenden Studie kleiner als in den anderen zwei Studien. Daraus kann man schließen, dass die Werte relativ dicht um die durchschnittliche Differenz streuen.

In allen Studien wird ein hoher Index der Korrelation nach Pearson für den Zusammenhang zwischen RDI und RDI\* ermittelt.

Die Werte der Sensitivität und Spezifität bei verschiedenen Cut off Punkten sind in allen drei Studien gut. In vorliegenden Studie liegt eine niedrige Spezifität von 0,64 bei einem Cut off Wert von 15 E/h vor. Dies bedeutet, dass die Sicherheit mit der eine milde SBAS von einer moderaten SBAS unterschieden werden kann nur bei 64% liegt. Von 11 Patienten, die bei einem Cut off Punkt von 15 E/h in der PSG Untersuchung als „gesund“ eingestuft werden, sind nach den Kriterien der Impedanzauswertung 4 krank. Die durchschnittliche Überbewertung oberhalb des Cut off RDI von 15 E/h liegt bei  $\varnothing$ 1,25 E\*/h für die 4 Patienten. Die tatsächliche Überbewertung dieser Patienten ist somit als gering anzusehen. Das Beispiel zeigt, dass eine Bewertung der Patienten nach RDI Cut off Punkten eine willkürliche Einteilung ist.

### 5.3.2 Vergleich der Studienergebnisse mit weiteren Level IV Geräten

Verschiedene Studien beschäftigen sich mit dem Erkennen von respiratorischen Ereignissen bei SBAS mit Hilfe von Mono- oder Stereorecordersystemen. Allgemein gilt, dass nur ein oder zwei Signale für das Erkennen von Ereignissen benutzt werden. Da grundsätzlich kein EEG aufgezeichnet wird, ist hier, wie auch bei Level III Geräten, keine Schlafstadieneinteilung möglich. Das benutzte LZ-EKG zählt zu der Gruppe der Level IV Geräte, da nur 2 Parameter, nämlich transthorakale Impedanz und EKG, gemessen werden. Streng genommen kann man sogar von einem Monorecordersystem sprechen, da die EKG Aufzeichnung nicht für die Auswertung genutzt wird.

Die meisten Studien nutzen eines der folgenden Signale für das Erkennen von SBAS:

S      Atembewegungen mit Atemgurten

- S Atemfluss
- S Oxymetrie
- S Herzfrequenz

Die verschiedenen Möglichkeiten werden einzeln betrachtet.

(1) Atembewegung mit Atemgurten

Die Atembewegung kann mittels Induktionsplethysmographie oder Piezoelementen in einem Summsignal von Bauch- und Brustgurten gemessen werden. Für das Erkennen von Hypopnoe wird die Sensitivität mit 0,86 und Spezifität mit 0,65 angegeben. [9]

In der hier vorliegenden Studie sind die entsprechenden Werte unabhängig von der zugrundeliegenden Atmungsstörung. Sie betragen für die Sensitivität 0,84 und für die Spezifität 0,95. Bei vergleichbarer Sensitivität ist durch die Messung der transthorakalen Impedanz also eine höhere Spezifität gegeben. Der Anteil an falsch positiven Ereignissen in der transthorakalen Impedanz ist also niedriger.

(2) Atemfluss

Die Messung des Atemflusses durch einen Thermistor wird von der AASM negativ eingeschätzt. [1] Das Signal ist nicht linear zum tatsächlichen Atemfluss, was zu einer generellen Überbewertung des Atemflusses führt. [23] Selbst unter idealen Bedingungen an wachen Versuchspersonen sind die Sensoren ungenau. [9] Auf eine weitergehende Betrachtung wird aus diesem Grund verzichtet.

(3) Oxymetrie

Die Qualität der Ergebnisse bei Messung der Sauerstoffsättigung ist abhängig von dem Sensorhersteller und dem Ort der Messung (Ohr oder Finger). Verschiedene Studien haben anhand von ermittelten RDIs\* die Sensitivität und Spezifität an verschiedenen Cut off Punkten bestimmt (vgl. Tab. 5-5). [19, 14]

	Diese Studie		Douglas et al.		Cooper et al.	
Patientenanzahl (n)	60		200		41	
Sensitivität bei $C_{x/h}$ <sup>1</sup>	1,00	$C_{5/h}$	0,92	$C_{5/h}$	0,6	$C_{5/h}$
	0,92	$C_{15/h}$	0,97	$C_{15/h}$	0,75	$C_{15/h}$
	0,84	$C_{20/h}$	0,99	$C_{20/h}$	k. A.	$C_{20/h}$
Spezifität bei $C_{x/h}$	1,00	$C_{5/h}$	0,67	$C_{5/h}$	0,95	$C_{5/h}$
	0,64	$C_{15/h}$	0,46	$C_{15/h}$	0,86	$C_{15/h}$
	0,97	$C_{20/h}$	0,36	$C_{20/h}$	k. A.	$C_{20/h}$

**Tab. 5-5:** Vergleich mit Oxymeter-Studien

<sup>1</sup> $C_{x/h}$  = zugrundeliegender Cut off Wert



Die Studien der Oxymetrie (Douglas et al. und Cooper et al.) zeigen eine Gegenläufigkeit von Sensitivität und Spezifität. Eine hohe Sensitivität ist verbunden mit einer niedrigen Spezifität und umgekehrt. Eine Bewertung der SBAS nur aufgrund von Veränderungen im Oxymeter Signal ist daher nicht zu empfehlen.

#### (4) Herzfrequenzvariabilität

Die Herzfrequenzvariabilität wird als Veränderung der R-R Zackenzeit gemessen. Die Veränderungen des kardiovaskulären Systems als Antwort auf eine Apnoe oder Hypopnoe ist abhängig von einem funktionierenden, autonomen Nervensystem. Diabetes assoziierte Neuropathien, wie auch die Medikation mit  $\beta$ -Blockern, können zu einer Beeinträchtigung führen. [24] Nicht geeignet ist diese Methode bei Vorliegen von Herzrhythmusstörungen. Der Verdacht einer SBAS wird anhand der Minutenzahl mit zyklischen Herzrhythmusänderungen (CVHR) innerhalb der ersten drei Schlafstunden ermittelt. Stein et al. haben festgestellt, dass ein Minutenanteil von CVHR  $\geq 20\%$  innerhalb der ersten drei Schlafstunden der geeignete Cut off Punkt für die Unterscheidung von SBAS kranken und gesunden Personen ist. Dieser Zeitraum entspricht 36 Minuten. Nach Aussage der Autoren werden 86% der Patienten mit starker OSAHS als krank erkannt. [82] Ein Vergleich der Methoden ist aufgrund des unterschiedlichen methodischen Vorgehens nicht möglich. Die Messmethode scheint vielversprechend zu sein, insbesondere unter dem Aspekt, dass die Untersuchung prinzipiell an jedem digital aufgezeichneten EKG auch retrospektiv durchgeführt werden kann.

Zusammenfassend lässt sich sagen, dass Systeme mit nur einem Messverfahren für das Erkennen von SBAS nicht die Qualität einer Aufzeichnung im PSG Labor erreichen. Bauch- und Brustgurte auf Basis der Induktionspletysmographie scheinen der Qualität der Impedanzsysteme gleichwertig zu sein. Thermistorsensoren haben sich als ungeeignet erwiesen. Die Messung der Sauerstoffsättigung mittels Oxymetrie weist eine schlechte Statistik bezüglich der Diskriminierungsfähigkeit von SBAS kranken zu gesunden Patienten auf. Systeme, die auf der Herzfrequenzvariabilität basieren, erscheinen vielversprechend. Sie sind nicht direkt mit den Ergebnissen der hier vorliegenden Studie zu vergleichen.

### **5.3.3 Vergleichende Betrachtung der Studienergebnisse mit Mehrkanalrecorder-systemen**

Mehrkanalrecorder sind nach Einteilung der ASDA Level III Geräte. [2] Dingli et al. und Bar et al. verwenden unterschiedliche Mehrkanalrecorder. Ihre Ergebnisse werden mit den hier erhaltenen verglichen. In allen Studien gilt die stationäre PSG als Goldstandard. Die von der AASM empfohlenen Richtlinien für das Vorgehen bei einem Vergleich von Testsystemen mit dem Goldstandard und die dabei zu benutzenden statistischen Auswertungsmethoden werden in den Studien zwar genannt, aber nicht konsequent umgesetzt. Daher ist ein Vergleich in vielen Aspekten nur bedingt möglich.

In der Studie von Dingli et al. wird ein „Embletta“ System der Firma Somnologica (Flaga, Reykjavik, Island). Es werden gemessen: Atemfluss mit einer Nasal Canula, Atembewegungen mit Brust- und Bauchgurten (Piezosensoren), Sauerstoffsättigung mittels Pulsoxymetrie und Lageveränderungen durch einen Lagesensor (vgl. Kap. 2.3.5-8). In der Studie gibt es zwei Untersuchungsgruppen. In der ersten Gruppe wird parallel in einem

Schlaflabor die PSG und die Testuntersuchung durchgeführt. Die zweite Gruppe wird in der ersten Nacht zu Hause mit dem Testgerät und in einer darauffolgenden Nacht im Schlaflabor mit PSG untersucht. Nur in der ersten Gruppe findet ein direkt vergleichbarer Test statt. Daher wird nur dieser Studienteil zum Vergleich herangezogen. [18]

In der zweiten Studie von Bar et al. wird ein „Watch PAT100“ der Firma Itamar Medical (Caesarea, Israel) benutzt. Es werden gemessen: Peripherer arterieller Ton, Herzfrequenz (von PAT Signal), Pulsoxymetrie und Aktigraphie. Das Messprinzip für das Erkennen von Apnoe und Hypopnoe beruht auf einer indirekten Messung. Das PAT Signal misst die arteriellen, pulsatilen Volumenänderungen am Finger. Dieses wird durch die  $\alpha$ -adrenerge Innervation der glatten Gefäßmuskulatur des Fingers gesteuert und spiegelt die Aktivität des sympathischen Nervensystems wider. Apnoe und Hypopnoe rufen eine Änderung der sympathischen Aktivität hervor und dadurch eine Veränderung des Fingervolumens. [7] Andere moderne Messmethoden des Sympathikotonus im Schlaf sind die Herzfrequenzvariabilität und die Pulstransitzeit. [58]

#### (1) Allgemeine Daten der Studien

Die zu vergleichenden Daten sind in Tab. 5-6 zusammengefasst. Die Studien zeigen Unterschiede in der Größe der Studienpopulation, im durchschnittlichen Alter, Geschlechterverhältnis und BMI. Die Größe der Studienpopulationen wird in allen Fällen als ausreichend betrachtet. In allen Studien erfolgt die Auswertung der PSG Aufzeichnung nach internationalem Standard. Die Auswertung des Testsystems wird jedoch von den Autoren individuell festgelegt.

In der Studie von Bar et al. sind 33 als gesund einzustufende Testpatienten eingeschlossen, die restlichen Patienten sind aus dem Schlaflabor rekrutiert. Dies ist ein möglicher Grund für das relativ niedrige Durchschnittsalter und die hohe Varianz. Da ein Teil der Testpersonen gesund ist, kann man mit einem niedrigen durchschnittlichen BMI rechnen.

Dingli et al. zeichnen mit ihrem System dieselben Atemsignale auf, die auch bei einer stationären PSG benutzt werden. Allerdings werden andere Kriterien für die Diagnose von Hypopnoe benutzt. Eine  $sO_2$  Messung findet statt, sie wird jedoch nicht in die Auswertung einbezogen.

Bei Bar et al. findet die Auswertung automatisch statt. Die Auswertung basiert auf einem Algorithmus der das PAT Signal, die Herzfrequenz und die Sauerstoffsättigung berücksichtigt.

Keine der Studien liefert Daten zu einer Ereignis gegen Ereignis\* Auswertung. Es werden stets nur ermittelte RDIs und RDIs\* verglichen. Unklar ist, inwieweit ein ermittelter RDI\* beeinflusst ist von falsch positiv oder falsch negativ gemessenen Ereignissen\*.

	Diese Studie	Dingli et al.	Bar et al.
Messprinzip	transthorakale Impedanzmessung	klassische Atemparameter	$\alpha$ -adrenerges Sympathikus Niveau
Referenzmethode	PSG	PSG	PSG
Patientenanzahl (n)	60	39	102
Alter (J)	58 $\pm$ 11,5 (33-88)	46 $\pm$ 9 (k. A.)	41,4 $\pm$ 15,2 (k. A.)
Geschlecht	♂49 : ♀11	♂32 : ♀7	♂78 : ♀24
BMI (kg/m <sup>2</sup> )	29,3 $\pm$ 4,4 (21-39)	32 $\pm$ 6 (k. A.)	26,8 $\pm$ 5,5 (k. A.)
Ø Bezugszeit (h)	7,6 $\pm$ 0,91 (TRT)	6,8 $\pm$ 0,3 (TRT) <sup>1</sup>	k. A.
Definition Apnoe in Testmethode	t <sub>min</sub> = 10s Signal < 50%	t <sub>min</sub> = 10s komplettes Sistieren Atemfluss	k. A. Auswertung basiert auf Algorhythmus aus PAT-Signal, Herzfrequenz und $sO_2$ .
Definition Hypopnoe in Testmethode	t <sub>min</sub> = 10s Signal < 50%	t <sub>min</sub> = 10s Signal < 50% in Gurt-Signalen	

**Tab. 5-6:** Grunddaten der Mehrkanalrecorder Studien

## (2) RDI gegen RDI\* Analysen der Mehrkanalrecordersysteme

In der Tab. 5-7 sind die Werte der RDI / RDI\* Analysen angegeben. In keiner der Vergleichsstudien wird die Sensitivität und Spezifität des Untersuchungsgerätes angegeben. Anhand der Werte für den Korrelationskoeffizient nach Pearson oder den Cohens Kappa Korrelationskoeffizient wird in beiden Studien der Zusammenhang zwischen den erhaltenen Testergebnissen und den wahren PSG Ergebnissen gezeigt.

Die beiden Werte der vorliegenden Studie zeigen eine höhere Korrelation zwischen den Ergebnissen der PSG und der Testmethode. Sofern eine Schlussfolgerung bei dieser geringen Datenlage möglich ist, scheint die hier gewählte Testmethode durchaus auf dem Niveau von Mehrkanalrecordern für das Erkennen von respiratorischen Ereignissen bei SBAS zu liegen.

<sup>1</sup>Es wurden sowohl die TRT als auch die TST-Werte benutzt. Zum besseren Vergleich werden hier nur die TRT-Werte angegeben.

	Diese Studie	Dingli et al.	Bar et al.
RDI in PSG (E/h)	31,4 ± 17,7	35 ± 5,5	k. A.
RDI* in Testmethode (E*/h)	28,9 ± 15,6	k. A.	k. A.
∅ Differenz RDI* zu RDI	- 2,5	3	k. A.
SD der ∅ Differenz	± 6,6	± 9	k. A.
Koeffizient nach Pearson	0,93	k. A.	0,88
Kappa Koeffizient (κ)	0,81	0,62	k. A.

**Tab. 5-7:** Ergebnisse der RDI gegen RDI\* Analysen

#### 5.4 Qualitäten der transthorakalen Impedanz

Verschiedene Aspekte sind bei der Gesamtbetrachtung einer Messmethode zu berücksichtigen. Die Validität der in dieser Studie angewandten Messmethode ist in den vorhergehenden Abschnitten belegt worden. Weitere Qualitätsaspekte eines Messverfahrens sind die Relevanz, Segreganz, Effektivität und Effizienz. Diese vier Faktoren sollen im folgenden erläutert werden.

##### (1) Relevanz und Segreganz

Die Relevanz und Segreganz einer Testmethode sind wichtige Aspekte für die Beurteilung eines Testverfahrens. Sie spiegeln die Sicherheit wider, dass Ereignisse\*, die durch das Testsystem angezeigt werden, auch richtig sind.

Die Relevanz gibt die Wahrscheinlichkeit wider, dass bei einem positiven Testergebnis auch tatsächlich ein Ereignis vorliegt. Sie wird im positiven prädikativen Wert ausgedrückt.

Die Segreganz gibt die Wahrscheinlichkeit wider, dass bei einem negativen Testergebnis auch tatsächlich kein Ereignis vorliegt. Sie wird im negativen prädikativen Wert ausgedrückt.

In der Ereignis gegen Ereignis\* Analyse wird für die Relevanz ein Wert von 0,92 und für die Segreganz ein Wert von 0,91 ermittelt. Diese hohen Werte belegen, dass bei alleiniger Betrachtung des Impedanzsignals man in hohem Maße davon ausgehen kann, dass die gemessenen Ergebnisse richtig sind. Daraus ist zu folgern, dass bei richtig gemessenen Ergebnissen auch die Patienten als krank oder nicht krank erkannt werden. Es ist festzuhalten, dass die Impedanzmessung zur Feststellung von SBAS bezogenen Atmungsstörungen in hohem Maße geeignet ist.

Ob diese Schlussfolgerung auch in der Praxis bestand hat, ist durch weitere Messungen an einem gemischten Kollektiv aus SBAS gesunden und kranken Personen zu überprüfen.

##### (2) Effektivität und Effizienz

Die Effektivität („die richtigen Dinge tun“) und Effizienz („die Dinge richtig tun“) sind korrespondierende Größen und werden gemeinsam betrachtet. Die Effektivität wird in der

Regel als wichtiger eingestuft. Es ist von höherer Priorität Ziele zu erreichen, bevor man den Weg zur Zielfindung optimiert.

In dieser Studie wird die Effektivität beurteilt, mit der man mittels der transthorakalen Atemimpedanz SBAS Erkrankungen nachweisen kann. Die Ergebnisse der Studie zeigen, dass Atmungsstörungen, wie sie bei SABS kranken Personen auftreten, in hohem Maße richtig erkannt werden. Die Effektivität der Messmethode ist also gut.

Die Effizienz einer Messmethode ist ein Maßstab, Ressourcen wirtschaftlich einzusetzen. Sie dient der Einschätzung, ob ein Ziel mit angemessenen Kosten erreicht werden kann. Bei medizinischen Untersuchungen ist es meist schwierig, den Nutzen eines Patienten gegenüber den entstehenden Kosten aufzurechnen. Bei einer Kosten/Nutzen Diskussion handelt es sich um ein ethisches Problem, welches im Rahmen dieser Arbeit nicht diskutiert werden soll. Nimmt man den Nutzen des Patienten, der in einer SBAS Therapie und den damit verbundenen gesundheitlichen Vorteilen liegt, als festes Ziel, so kann man versuchen, den Weg für das Erreichen des Zieles zu optimieren.

Die PSG Untersuchung in einem Schlaflabor ist wegen der stetigen Kontrolle durch eine Person während der Aufzeichnung und der mehrkanaligen Überwachung SBAS relevanter Parameter die sicherste und beste Methode, SBAS zu erkennen. Der große materielle und personelle Aufwand erzeugt hohe Kosten. Daher ist die Untersuchung in einem Schlaflabor der Diagnosesicherung und der Therapieeinstellung vorbehalten. In Deutschland besteht ein relativer Mangel an Schlaflaborplätzen, was zu Wartezeiten von bis zu 4 Monaten führt. Es ist notwendig, dass die Patienten vorab selektiert werden und die Indikation für einen Aufenthalt im Schlaflabor begründet ist. Die Untersuchung in einem Schlaflabor ist aus Kostengründen nicht geeignet für die breite Suche nach SBAS kranken Personen im Sinne eines Screening Programms. Dieses muss außerhalb des Schlaflabors geschehen.

Das Ziel liegt auf der einen Seite in der Kostensenkung und Wartezeitverkürzung im Bereich der stationären Schlafüberwachung (Patienten werden extern selektiert), als auch auf der anderen Seite in der frühzeitigen Erkennung von (symptomlosen) SBAS Erkrankungen. Damit verbindet sich die Möglichkeit der frühzeitigen Therapie für den Patienten. Es ergeben sich indirekte Kostenvorteile durch eine mögliche Minderung von Frühverrentungen, müdigkeitsbedingten Unfällen sowohl am Arbeitsplatz [43, 85] als auch im Straßenverkehr [13, 25, 26, 33] und eine verminderte Inanspruchnahme von medizinischen Leistungen aus dem Gesundheitssystem. [70] Für den Patienten resultiert eine verbesserte Gesundheit mit Verbesserung der Lebensqualität, Senkung von Begleiterkrankungen und ein direkter Gewinn an Lebensjahren.[101]

Dingli et al. ermitteln in einer Studie eine mögliche Senkung der Untersuchungskosten um 42%, sofern Patienten mit einem NLM Gerät vorab untersucht werden und danach entweder direkt eine CPAP Therapie erhalten oder als gesund gelten. Unklare Fälle benötigen jedoch weiterhin eine stationäre PSG Untersuchung.[18]

Die Effektivität der transthorakalen Atemimpedanz wird in dieser Studie festgestellt. Entsprechend kann man der Ansicht von Dingli et al. folgen und auch die mögliche Effizienz des Testverfahrens der vorliegenden Studie feststellen.

Unberücksichtigt bleibt bei dieser Aussage jedoch weiterer volkswirtschaftliche Nutzen. Er setzt sich zusammen aus einer verkürzten Zeitdauer bis zum Therapiebeginn, dadurch

früher gesteigertem Gesundheitsniveau und erhöhter Arbeitsfähigkeit.

## 5.5 Schlussfolgerungen

Die Resultate der durchgeführten Studie lassen sich wie folgt zusammenfassen:

(1) Die Methode der transthorakalen Atemimpedanzmessung kann Ereignisse, wie sie im Rahmen von SBAS als Hypopnoe oder Apnoe auftreten, gut erkennen. Dies wird belegt durch eine ermittelte Sensitivität von 0,84, einer Spezifität von 0,95, einem positiv prädikativen Wert von 0,92 und einem negativ prädikativen Wert von 0,91. Die Ergebnisse beruhen nicht auf einer zufälligen Übereinstimmung, was durch einen hohen Kappa Koeffizient nach Cohen von 0,81 bestätigt wird.

(2) Der RDI gegen RDI\* Vergleich zeigt einen hohen Korrelationsindex nach Pearson von 0,93. Die Untersuchung verschiedener Cut off Werte für die Einteilung von Personen nach SBAS krank / nicht krank bestätigt die bisherigen guten Ergebnisse hinsichtlich der Unterscheidungsfähigkeit der Testmethode.

(3) Der Vergleich mit anderen Studien ist aufgrund fehlender verbindlicher Standards für die Methodik eines Vergleiches von PSG zur Testmethode nur eingeschränkt möglich. Eine Orientierung an dem von der AASM empfohlenen Vorgehen wird in dieser Studie durchgeführt und ist bei kommenden Studien wünschenswert, um eine Vergleichbarkeit von Ergebnissen zu gewährleisten. Innerhalb dieser Methodik ist festzustellen, dass die Impedanzmessung sowohl in dieser Studie als auch bei Saarelainen et al. und Scharf et al. als geeignete Methode zur Erkennung von SBAS angesehen wird. Der Vergleich mit Messgeräten der Kategorie III zeigt, dass die Ergebnisse der Impedanzmessung durchaus auf dem Niveau von Mehrkanalrecordern liegen. Eine Reduktion der Parameter, die für das Erkennen von SBAS gemessen werden müssen, ist durchaus möglich, ohne die Ergebnisqualität zu senken.

(4) Aufgrund der fehlenden Messbarkeit der Schlaftiefe bei Level III und Level IV Geräten ist das Führen eines Schlafprotokolls zu empfehlen. Durch diese Maßnahme können Artefakte, die in der Wachzeit auf Grund von Bewegungen entstehen, als solche identifiziert werden. Dadurch sinkt die Zahl der falsch positiven Ereignisse\*. Unabhängig davon sollten die Patienten einen Fragebogen zur Erfassung von SBAS Symptomen (z.B. die ESS) benutzen, um auch die subjektive Belastung des Patienten durch die Erkrankung zu erfassen.

(5) Die Ergebnisse sind repräsentativ für SBAS kranke Patientengruppen. In einer gemischten Patientengruppe ist eine ausreichende Unterscheidung zwischen SBAS kranken und SBAS gesunden Patienten wahrscheinlich. Um einen abschließenden Beweis dafür zu erbringen, sind weitere Untersuchungen an gemischten Patientengruppen notwendig.

(6) Generell ist die ambulante Untersuchung mit NLM Systemen als positiv einzustufen. Es kann zu einer Früherkennung von SBAS Erkrankungen führen. Daraus resultieren positive Effekte bezüglich der Patientengesundheit und der entstehenden Kosten auf direkte und indirekte Art und Weise.