

Aus der Klinik für Anästhesiologie mit Schwerpunkt operative  
Intensivmedizin der Medizinischen Fakultät Charité – Universitätsmedizin  
Berlin, Campus Charité Mitte und Campus Virchow – Klinikum

## **DISSERTATION**

### **Der PHQ-4 zur Messung von Depressivität und Angst bei operativen Patienten der Anästhesieambulanz: Untersuchung eines Ultrakurz-Screeninginstrumentes**

zur Erlangung des akademischen Grades  
Doctor medicinae (Dr. med.)

vorgelegt der Medizinischen Fakultät  
Charité – Universitätsmedizin Berlin

von  
Janina Tillinger  
aus Berlin

Datum der Promotion: 04.09.2015

# Inhaltsverzeichnis

Abkürzungsverzeichnis.....	9
1 Einleitung.....	11
1.1 Depressivität und Ängstlichkeit bei operativen Patienten.....	11
1.2 Der PHQ-4 als Ultrakurz-Screeninginstrument für Depressivität und generalisierte Angst.....	13
2 Ziele der Studie .....	17
3 Methodik.....	18
3.1 Studiendesign und Studienzeitraum.....	18
3.2 Patienten .....	18
3.2.1 Ein- und Ausschlusskriterien.....	18
3.2.2 Ablauf des Studieneinschlusses .....	19
3.2.3 Missing-Value-Analyse .....	20
3.3 Messinstrumente .....	22
3.3.1 PHQ-4.....	22
3.3.2 Das BSI als Goldstandard der Screeningverfahren für Depressivität und Angst.....	24
3.3.3 Soziodemographische Charakteristika.....	26
3.3.4 Klinische Charakteristika.....	26
3.4 Statistik.....	27

3.4.1	Reliabilität .....	28
3.4.2	Validität .....	28
3.4.3	ROC-AUC-Analyse .....	31
3.4.4	Analyse der Cutoff-Werte.....	31
4	Ergebnisse .....	33
4.1	Patientencharakteristika .....	33
4.2	Psychometrische Qualität des PHQ-4.....	34
4.2.1	Skalenkennwerte des PHQ-4.....	34
4.2.2	Item-Interkorrelationen des PHQ-4 .....	36
4.2.3	Reliabilität des PHQ-4.....	36
4.2.4	Validität des PHQ-4 .....	36
4.2.5	AUC-Werte des PHQ-2, GAD-2 und des PHQ-4 .....	42
4.2.6	Analyse alternativer Cutoff-Werte des PHQ-2, GAD-2 und PHQ-4.....	44
5	Diskussion .....	47
5.1	Stärken und Schwächen des PHQ-4.....	47
5.1.1	Hauptergebnisse.....	47
5.1.2	Reliabilität .....	48
5.1.3	Konstruktvalidität.....	48
5.1.4	ROC-AUC-Analysen .....	49
5.1.5	Faktorenanalyse .....	49

5.1.6	Sensitivität und Spezifität .....	51
5.2	Methodenkritik .....	53
5.2.1	Studiendesign .....	53
5.2.2	Fehlende Werte .....	55
5.3	Zusammenfassung und Implikationen für Forschung und Klinik .....	55
7	Literaturverzeichnis .....	57
	Lebenslauf .....	67
	Danksagung .....	69
	Eidesstattliche Versicherung .....	70
	Anhang .....	72

## Zusammenfassung

**Einleitung** Depression und Angst sind weitverbreitete Problematiken, die häufig gemeinsam auftreten und mit funktionellen Einschränkungen der Betroffenen einhergehen. Ultrakurz-Screeninginstrumente für Depressivität und Ängstlichkeit haben in ersten Studien bereits eine hohe psychometrische Qualität gezeigt. Es liegen allerdings kaum Studien für den operativen Bereich vor. Die vorliegende Untersuchung geht der Frage nach, in wie weit sich der Selbstauskunftsfragebogen Patient-Health-Questionnaire-4 (PHQ-4) und seine Subskalen Patient Health Questionnaire-2 (PHQ-2) und Generalized Anxiety Disorder Scale-2 (GAD-2) zur Messung von Depressivität, generalisierter Angst und allgemeiner psychischer Belastung bei operativen Patienten in Anästhesieambulanzen eignen.

**Methodik** 3.068 Patienten verschiedener chirurgischer Disziplinen nahmen an der computerbasierten psychologischen Fragebogen-Untersuchung teil. Zur Untersuchung des vier Items umfassenden PHQ-4 und seiner jeweils zwei Items umfassenden Subskalen PHQ-2 und GAD-2 wurden folgende Maße bestimmt: Interne Konsistenz, faktorielle Validität, Konstruktvalidität, Sensitivität, Spezifität, Positiv und Negativ Prädiktiver Wert, Youden-Index und ROC-AUC-Analysen. Als Goldstandard wurden die Skalen des Brief Symptom Inventory (BSI) für Depressivität, Ängstlichkeit, Phobische Angst und Unsicherheit im Sozialkontakt sowie der Gesamtscore Global Severity Index (GSI) verwendet.

**Ergebnisse** Cronbach's Alpha von PHQ-4, PHQ-2 und GAD-2 waren .83, .66 und .78. Die Untersuchung der faktoriellen Validität extrahierte zwei Faktoren, wobei das erste Item den einen Faktor abbildet und die anderen drei Items auf dem anderen Faktor laden. Alle drei Skalen zeigten gute Konstruktvalidität, sowie adäquate diskriminatorische Fähigkeiten mit Areas under the Curve (AUC) zwischen .80 und .86. Die Spezifität des PHQ-4 ( $\geq 6$ ), des PHQ-2 ( $\geq 3$ ) und des GAD-2 ( $\geq 3$ ) befindet sich im sehr hohen Bereich zwischen 89,4% und 94,5% je nach als Goldstandard verwendeter Subskala des BSI, wohin gehend die Sensitivität mit Werten zwischen 46,4% und 61,2% gering ist. Senkt man den Cutoff-Wert des PHQ-4-Gesamtscores auf  $\geq 4$ , zeigen

sich eine Sensitivität und Spezifität von 80,5% und 80,2% bezüglich klinisch relevanter psychischer Belastung gemessen am GSI.

**Schlussfolgerung** Bei einem niedrigeren Cutoff-Wert zeigt der PHQ-4-Gesamtscore eine suffiziente psychometrische Qualität, um klinisch relevante psychische Belastung samt Depressivität und Ängstlichkeit bei operativen Patienten zu detektieren. PHQ-2 und GAD-2 sollten nicht als alleiniges Screeninginstrument für Depressivität und Ängstlichkeit empfohlen werden.

## Abstract

**Objectives** Depression and anxiety are common disorders that are often accompanied by considerable disability and frequently co-occur. Brief screening tools for depression and anxiety have been validated for psychometric quality previously. However, only very few studies exist for the perioperative setting. This study examined the question whether the ultra-short questionnaire PHQ-4 and its subscales Patient Health Questionnaire-2 (PHQ-2) and Generalized Anxiety Disorder Scale-2 (GAD-2) are reliable, valid and precise screening tools for self-reported depression, anxiety and general psychological distress in surgical patients of preoperative anaesthesiological assessment clinics.

**Methods** In sum, 3.068 patients of diverse surgical fields participated in a computer-based psychiatric self-assessment. For the examination of the two-item measures PHQ-2 und GAD-2 and the composite four-item scale PHQ-4 Cronbach's alpha, construct validity, factorial validity, sensitivity, specificity, negative and positive predictive value, Youden index, and ROC-AUC analyses were determined. As criterion measures, we utilized the scores of the Brief Symptom Inventory (BSI) subscales for depression, anxiety, phobic anxiety, interpersonal sensitivity, and the total mean score Global Severity Index (GSI).

**Results** For PHQ-2, GAD-2 and PHQ-4 the Cronbach's alphas were 0.66, 0.78 and 0.83 respectively. The examination of the factorial validity reveals two factors: One consisting of the first item and the second one composed by the three other items. The three scales showed good construct validity and with areas under the curve (AUC) between .80 and .87 they showed adequate accuracy. With their established cut-off points and the respective BSI scales as criterion standards PHQ-2 ( $\geq 3$ ), GAD-2 ( $\geq 3$ ), and PHQ-4 ( $\geq 6$ ) had sensitivities between 46.4% and 61.2% and specificities between 89.4% and 94.5%. For the detection of clinically significant psychological distress according to the GSI the cut-off point for PHQ-4 was lowered to  $\geq 4$  and thus the sensitivity and specificity of the PHQ-4 total scale were 80.5% and 80.2%, respectively.

**Conclusion** This study shows that the PHQ-4 total scale has a sufficient psychometric quality for the detection of self-reported clinically significant psychological distress including depression and/or anxiety in surgical patients at a lowered cut-off point. PHQ-2 and GAD-2 are not recommended as exclusive measures of depression and anxiety in these patients.

## Abkürzungsverzeichnis

ASA	American Society of Anesthesiologists
AUC	Area under the curve
BMI	Body Mass Index
BRIA	Brückenintervention in der Anästhesiologie
BSI	Brief Symptom Inventory
CI	Konfidenzintervall
DSM	Diagnostic and Statistical Manual of Mental Disorders
GAD-2	Generalized Anxiety Disorder-2
GSI	Global Severity Index
ICD-10	International Statistical Classification of Diseases and Related Health Problems-10
M	Mittelwert
NPW	Negativ Prädiktiver Wert
OR	Odds Ratio
PHQ	Patient Health Questionnaire
PPW	Positiv Prädiktiver Wert
Prime-MD	Primary Care Evaluation of Mental Disorders
ROC	Receiver-Operating-Characteristic
SCID	Structured Clinical Interview for DSM-III-R

## Abkürzungsverzeichnis

SCL-90-Revised	Symptom-Checklist-90-Revised
SD	Standardabweichung
SPSS	Statistical Package of Social Science

# 1 Einleitung

## 1.1 Depressivität und Ängstlichkeit bei operativen Patienten

Depression ist eine affektive Störung, die durch gedrückte Stimmung und eine Minderung von Antrieb sowie Aktivität charakterisiert wird<sup>1,2</sup>. Angststörungen umfassen Panikstörungen, Generalisierte Angststörungen, Soziale Phobien, Spezifische Phobien und Posttraumatische Belastungsstörungen<sup>3</sup>. Sowohl bei depressiven als auch bei ängstlichen Patienten<sup>a</sup> ist die Abgrenzung zwischen normalen, also gesunden emotionalen Empfindungen und krankheitsrelevanten, also in Dauer, Intensität und Ausmaß der Beeinträchtigung gesteigerten Reaktionen schwierig<sup>1,2,4</sup>. Depressivität meint folglich einen Zustand aktueller Stimmung, der den Bereich von einer leichten depressiven Verstimmung bis hin zu einer manifesten affektiven Störung umfasst. Mit den Begriffen klinisch relevante Depressivität und Angst werden in der vorliegenden Arbeit Beschwerden beschrieben, die eventuell, aber nicht definitiv, ein Ausmaß erreicht haben, bei dem nach ICD-10 zu kodierende Diagnosen vorliegen. Unter Angst ist sowohl ein gesundes Sorgen, aber auch eine pathologisch übersteigerte Angst zu verstehen. Werden in dieser Arbeit die beiden Begriffe klinisch relevante Depressivität und Angst verwendet, dürfen sie nicht als Diagnosen, sondern nur im Sinne des Vorliegens eines positiven Screeningergebnisses aufgefasst werden.

Depressivität und Angst sind weitverbreitete Problematiken. Betrachtet man nur den Anteil feststehender Diagnosen innerhalb der Europäischen Union, so findet man eine 12-Monats-Prävalenz von 6,9% für Major Depression und von 14% für Angststörungen<sup>b,5</sup>. Auch in der deutschen Bevölkerung gehören Depression und Angst zu den beiden häufigsten psychiatrischen Diagnosen<sup>6</sup>. Beide treten häufig kombiniert auf und gehen mit einer substanziellen funktionellen Einschränkung der Betroffenen einher<sup>5,7-13</sup>.

---

<sup>a</sup> Im Weiteren wird die männliche Form verwendet. Damit sind Frauen und Männer gleichermaßen gemeint außer das Geschlecht ist ausdrücklich spezifiziert.

<sup>b</sup> Hierzu zählten Patienten mit Panikstörungen, Agoraphobie, Sozialer Angst, Generalisierter Angststörung, Spezifischer Phobien, Zwangsstörungen und Posttraumatischer Belastungsstörung.

Unter Patienten verschiedenster medizinischer Fachrichtungen mit somatischen Erkrankungen leidet ein noch größerer Anteil an klinisch relevanter Depressivität und Angst. In diesem Setting sind sogar bis zu viermal mehr Patienten von diesen Problematiken betroffen als in der allgemeinen Bevölkerung. Klinisch relevante Depressivität und Angst tritt hier bei 8 bis 50 Prozent der Patienten mit somatischen Beschwerden auf, je nach Chronizität und Schweregrad der Erkrankung<sup>14-28</sup>.

Bekannt ist, dass Depression sowohl die Lebensqualität senkt als auch die Lebenserwartung verkürzt<sup>15</sup>. Außerdem bestehen Assoziationen mit einer negativen Prognose der medizinischen Erkrankungen, beziehungsweise mit der schlechteren Erholung chirurgischer Patienten<sup>16,19,29-33</sup>, sowie mit einer gesteigerten Rückfallrate von somatischen Erkrankungen<sup>16</sup>. Auch bei operativen Patienten ist die Prävalenz klinisch relevanter Depressivität höher als in der Allgemeinbevölkerung<sup>34</sup>. Studien ergaben, dass klinisch relevante Depressivität mit einer erhöhten Anzahl an perioperativen Komplikationen und erhöhter Morbidität und Mortalität einhergeht, sowie zu einem schlechteren chirurgischen Outcome führt<sup>35-38</sup>. Ferner ergaben Untersuchungen, dass bei depressiven Patienten eine erhöhte Anzahl an Arztbesuchen und längeren krankheitsbedingten Fehlzeiten zu finden ist<sup>9,39,40</sup>. Zu einer verstärkten Ausprägung der genannten Phänomene, die mit klinisch relevanter Depressivität und Ängstlichkeit assoziiert sind, kommt es bei komorbidem Auftreten von Depressivität und Angststörungen<sup>40</sup>. Betroffene Patienten gaben stärkere postoperative Schmerzen an, hatten einen höheren Analgetikabedarf und eine verlängerte Krankenhausverweildauer<sup>41,42</sup>. Ferner scheint Angst einen negativen Einfluss auf die Genesung der Patienten zu haben<sup>43</sup>.

Betrachtet man die operativen Patienten in Anästhesieambulanzen, wird deutlich, dass hier Patienten unterschiedlichster chirurgischer Disziplinen zusammen kommen. Somit wird in den Anästhesieambulanzen eine hohe Anzahl von operativen Patienten versorgt, die in klinisch relevantem Ausmaß unter Depressivität und Angst leiden<sup>44</sup>. Die Prävalenz von Depressivität liegt bei operativen Patientinnen laut einer aktuellen Studie der Arbeitsgruppe, in deren Rahmen auch die vorliegende Arbeit entstanden ist, bei 29,7%<sup>45</sup>.

## 1.2 Der PHQ-4 als Ultrakurz-Screeninginstrument für Depressivität und generalisierte Angst

Untersucht man die Prävalenz von Depressivität und generalisierter Angst in den Anästhesieambulanzen und beschäftigt sich mit den Auswirkungen, die diese Störungen auf die Patienten und damit auch auf das Gesundheitssystem haben, so stellt sich die Frage, auf welche Weise interveniert werden kann. Eine aktuelle Studie hat gezeigt, dass ein großer Anteil der Patienten mit psychischen Beschwerden gerne psychologische Hilfe in Anspruch nehmen würde<sup>44,45</sup>. Es gibt ausreichend Hinweise, dass ein Screening nach Depressivität und Ängstlichkeit bei Patienten mit somatischen Erkrankungen die Basis einer erfolgreichen Therapie darstellt<sup>46-49</sup>.

Wichtig ist hierbei, abgesehen von der klinischen und ökonomischen Effizienz, dass die angewendeten Screeninginstrumente möglichst mit einer geringen zusätzlichen Belastung für den Patienten und das Krankenhaus einhergehen. In diesem Rahmen sind Fragebögen von besonderer Praktikabilität, die vom Personal der Anästhesieambulanzen ohne Hilfe durch einen Spezialisten eingesetzt werden können. Ferner sollten sie wenig zeitaufwendig sein, damit sie sich als anwendbar im Alltag einer Ambulanz erweisen und sich mit geringeren Umständen in den klinischen Ablauf integrieren lassen. Zusammenfassend kann man sagen, dass ein Screeninginstrument gesucht wird, das schnell, einfach durchzuführen und günstig, jedoch auch möglichst aussagekräftig ist.

Ein erprobter Fragebogen im klinischen Alltag ist das 53 Items umfassende BSI (Brief Symptom Inventory)<sup>50-55</sup>. Auf Grund der höheren Praktikabilität bei der Anwendung wäre jedoch ein klinisch effizientes Ultrakurz-Screeninginstrument von großem Nutzen. Ultrakurz-Screener werden typischerweise definiert als Fragebögen, deren Bearbeitung nicht mehr als vier Minuten in Anspruch nimmt und die aus einem bis vier Items bestehen<sup>56</sup>. Hier stehen die Fragebögen der PHQ (Patient Health Questionnaire)-Familie zur Verfügung, welche auch im somatischen Kontext anwendbar sind<sup>57,58</sup>. Eine neue vielversprechende Variante ist der PHQ-4<sup>59</sup>.

Der PHQ-4 ist eine Ultrakurz-Form des Gesundheitsfragebogens für Patienten (Patient Health Questionnaire/PHQ), der zur kombinierten Erfassung der klinisch relevanten

Depressivität und Ängstlichkeit dient. Es handelt sich um einen Selbstauskunftsfragebogen, das heißt er ist von dem Patienten selbst zu beantworten<sup>59</sup>.

Der PHQ-4 ist ein neuer, von Kroenke und Spitzer entwickelter Fragebogen, der 2004 zum ersten Mal Anwendung fand<sup>59</sup>. Er besteht aus vier Items und ist eine Kombination der beiden jeweils zwei Fragen umfassenden Fragebögen PHQ-2 und GAD-2 (Generalized Anxiety Disorder)<sup>59</sup>. Desweiteren existiert eine zusätzliche Frage, die den Grad der subjektiven Beeinträchtigung der Patienten erfasst<sup>60,61</sup>. Der PHQ-2 erfragt die beiden Kernkriterien für Depression (Minderung des Antriebs, gedrückte Stimmung) und ist eine Kurzform des neun Fragen langen PHQ-9. Der GAD-2, als Kurzform des sieben Fragen umfassenden GAD-7, erfragt die zwei Hauptkriterien für generalisierte Angst (innere Spannung, Befürchtungen). Alle vier Fragebögen, das heißt der PHQ-9, PHQ-2, GAD-7 und GAD-2, haben sich als sehr gute Screeninginstrumente im Bereich der US-medizinischen Grundversorgung herausgestellt<sup>57,58,62,63</sup>. Kroenke et al. (2001) konnten anhand einer Stichprobe von 6.000 Patienten verschiedener Kliniken der US-amerikanischen Grundversorgung und geburtshilflicher/gynäkologischer Kliniken zeigen, dass der PHQ-9 eine exzellente Reliabilität, sowie gute Validität aufweist<sup>58</sup>. Außerdem untersuchten Kroenke et al. (2003) die ersten beiden Fragen des PHQ-9, bekannt als PHQ-2, auf ihre Validität. Es ergab sich bei einem Cutoff von  $\geq 3$  auf seiner 0-6-Punkte-Skala eine Sensitivität von 83% und eine Spezifität von 90% hinsichtlich der Diagnose Major Depression, die in dieser Studie mit Hilfe eines als Goldstandard verwendeten strukturierten klinischen Interviews gestellt wurde<sup>57</sup>. Auch der GAD-7 zeigt eine ähnlich gute Reliabilität und Validität<sup>62,63</sup>. Beim GAD-2 beträgt für generalisierte Angststörungen bei einem Cutoff von  $\geq 3$  auf seiner 0-6-Punkte-Skala die Sensitivität 86% und die Spezifität hinsichtlich einer durch ein strukturiertes klinisches Interview diagnostizierten Generalisierten Angststörung 83%<sup>63</sup>.

In einer ersten Studie, veröffentlicht von Kroenke et al. (2009)<sup>59</sup>, wurde der PHQ-4 als Gesamtfragebogen erstmals hinsichtlich seiner Validität untersucht. Hierzu wurden Daten von 2.149 Patienten aus 15 Einrichtungen der US-amerikanischen Grundversorgung analysiert. Es wurde eine explorative Faktorenanalyse durchgeführt, die zwei Faktoren, nämlich Depressivität und Ängstlichkeit, zeigte. Die beiden Items des PHQ-2 bilden hierbei den einen, die Items des GAD-2 den zweiten Faktor<sup>59</sup>. Wenn man die Kürze der Skalen berücksichtigt, sind die internen Konsistenzen mit  $\alpha=.85$  für die

Gesamtskala,  $\alpha=.81$  für die Teilskala Depression und  $\alpha=.82$  für die Unterskala Angst als gut zu bezeichnen<sup>59</sup>. Auch für die Konstruktvalidität der Scores des PHQ-4 konnten Hinweise gefunden werden<sup>49,59</sup>. 2010 zeigten Löwe et al., dass der PHQ-4 auch in der deutschen Allgemeinbevölkerung über gute Screening-Eigenschaften verfügt<sup>49</sup>. Kocalevent et al. (2014) untersuchten den PHQ-4 durch eine Haushaltsbefragung in der kolumbianischen Allgemeinbevölkerung an Hand verschiedener Selbstauskunftsfragebögen und stellten eine gute Reliabilität ( $\alpha=.84$ ) und gute Validität fest<sup>64</sup>.

Außerdem fand der PHQ-4 in verschiedenen klinischen Bereichen Anwendung, ohne dass weitere psychometrische Überprüfungen durchgeführt wurden: Bei Patienten mit chronischen Schmerzen, Migräne und Malignomerkrankungen wurde er verwendet<sup>65-71</sup>. Schmidt et al. (2011) nutzten den PHQ-4 als Maß für Depressivität und Angst bei einer Studie zur Untersuchung von Fragebögen zur Erfassung psychischer Beschwerden in einer bevölkerungsrepräsentativen Stichprobe<sup>72</sup>. Auch außerhalb des klinischen Settings wurden in den letzten Jahren Studien mit dem PHQ-4 durchgeführt. So forschten Zenger et al. (2013) mit Hilfe des PHQ-4 an den Zusammenhängen der Arbeitslosigkeit mit der Stimmung<sup>73</sup> und Nesterko et al. (2014) suchten nach der Prävalenz von Depressivität und Ängstlichkeit in einer Gruppe von Juden sowjetischen Ursprungs in Deutschland<sup>74</sup>. Glaesmer et al. (2014) verwendeten den PHQ-4 bereits als Goldstandard zur Untersuchung eines anderen Screeninginstruments<sup>75</sup>.

Zusammenfassend kann gesagt werden, dass sich der PHQ-4 in ersten Studien als effektives Ultrakurz-Screeninginstrument herausgestellt hat, das sich für die Identifizierung von Personen, die unter Depressivität und Ängstlichkeit leiden, eignet.

Ein sehr gut untersuchter, etablierter und effektiver, aber deutlich längerer Fragebogen sowohl für Depressivität also auch für Ängstlichkeit ist das BSI (Brief Symptom Inventory)<sup>51,53-55,76-79</sup>. Eine Validierung des PHQ-4 anhand dieses Goldstandards bei operativen Patienten blieb jedoch bisher aus. Die vorliegende Studie soll diese Lücke füllen.

Das BSI (deutsche Version: Franke, 2000) ist ein Instrument zur Erfassung subjektiver Beeinträchtigung durch psychische Symptome<sup>52,53</sup> (siehe Anhang 1). Es handelt sich hierbei um eine Kurzform der Symptom Checklist-90 (SCL-R-90), die die

Symptombelastung in der letzten Woche erfragt. Für das BSI wurden diejenigen Items der verschiedenen Subskalen der SCL-R-90 übernommen, die auf dem jeweiligen Faktor am höchsten luden<sup>53</sup>. Sprachlich sind die Fragen leichtverständlich formuliert und es werden psychopathologische Fachausdrücke vermieden, die nicht umgangssprachlich gebräuchlich sind.

Die Items des BSI lassen sich in neun Subskalen, drei globalen Kennwerten und vier Zusatzitems zusammenfassen. In der vorliegenden Arbeit werden die Subskalen Depressivität, Ängstlichkeit, Unsicherheit im Sozialkontakt und Phobische Angst verwendet. Außerdem wird einer der Gesamtscores, nämlich der GSI (Global Severity Index) als ein Maß der allgemeinen psychischen Belastung des Patienten berechnet. Da dieser vor allem durch die Faktoren Angst und Depressivität beeinflusst ist<sup>80</sup>, wird er in dieser Arbeit auch als Indikator des gemeinsamen Auftretens von beidem genutzt. Untersuchungen der internen Konsistenzen ergaben  $\alpha$ -Werte von .72 bis .87 für die Depressions-, .61 bis .81 für die Unsicherheit im Sozialkontakt-, .62 bis .81 für die Angst- und .39 bis .82 für die Phobische Angst-Subskala. Die interne Konsistenz des GSI lag bei  $r=.92$  bis zu  $r=.96$ <sup>52,53,78</sup>. Auch die Validität der Subskalen des BSI wurde hinreichend untersucht: Für die Konstruktvalidität gibt es positive Hinweise, da sowohl divergente als auch konvergente Validität gezeigt werden konnte<sup>53,78</sup>. So korrelieren zum Beispiel das BSI und die SCL-90-R hoch miteinander<sup>81</sup>. Folglich lässt sich festhalten, dass das in dieser Studie als Goldstandard verwendete BSI, ein viel verwendeter, gut akzeptierter Fragebogen ist, der sich zur Messung von allgemeinem psychologischem Stress eignet<sup>50,51,53-55,76-79</sup>.

## **2 Ziele der Studie**

Die vorliegende Untersuchung soll der Frage nachgehen, inwieweit sich der PHQ-4 und seine Subskalen zur Messung von klinisch relevanter Depressivität, Angst und allgemeiner psychischer Belastung bei erwachsenen elektiv operativen Patienten in Anästhesieambulanzen eignen.

Primäres Ziel dieser Studie ist die Untersuchung der psychometrischen Qualität des PHQ-4. Insbesondere werden folgende Kennwerte des PHQ-4 bestimmt: Interne Konsistenz, faktorielle Validität, divergente und konvergente Konstruktvalidität, sowie Sensitivität und Spezifität.

Außerdem sollen die diskriminatorischen Fähigkeiten von PHQ-4, PHQ-2 und GAD-2 und deren Cutoffs näher untersucht werden.

## **3 Methodik**

### **3.1 Studiendesign und Studienzeitraum**

Bei dieser Studie handelt es sich um eine prospektive Beobachtungsstudie. Die Untersuchung des PHQ-4 wurde im Rahmen der BRIA-Studie (Brückenintervention in der Anästhesiologie) an der Klinik für Anästhesiologie mit Schwerpunkt operative Intensivmedizin der Charité-Universitätsmedizin Berlin, Charité Campus Mitte und Campus Virchow-Klinikum durchgeführt. Die BRIA-Studie ist durch die Ethikkommission der Charité genehmigt (EA1/23/2004) und wird durch die Deutsche Forschungsgemeinschaft unterstützt (DFG-GZ KR3836/3-1). In der vorliegenden Untersuchung wurden Patienten der Anästhesieambulanz im Rahmen eines computerbasierten Screeningprogramms befragt. Die Datenerhebung fand von Januar bis Juni 2010 statt. Das Setting der Anästhesieambulanzen hat sich als besonders praktikabel erwiesen, da so Patienten aller chirurgischen Disziplinen eingebunden werden können. Die Befragung umfasste neben dem PHQ-4 und dem BSI soziodemographische Aspekte, klinische Charakteristika und andere psychologische Fragebögen.

### **3.2 Patienten**

#### **3.2.1 Ein- und Ausschlusskriterien**

In die Studie eingeschlossen wurden alle Patienten, die sich zum Zeitpunkt der Patientenrekrutierung präoperativ in den Anästhesieambulanzen der Klinik für Anästhesiologie der Charité Campus Mitte und Campus Virchow-Klinikum vorstellten, nicht anhand der Ausschlusskriterien ausgeschlossen werden mussten, die Einschlusskriterien erfüllten und von denen eine schriftliche Einverständniserklärung vorlag.

Einschlusskriterien waren:

- Patienten mit elektiven operativen Eingriffen
- Vollendung des 18. Lebensjahres
- ausreichende Deutschkenntnisse

- schriftliche Einverständnis zur Studienteilnahme

Ausschlusskriterien waren:

- Notfalloperation
- nicht bereit bzw. in der Lage einen Computer zu benutzen
- Mitarbeiter der Charité
- Verwandte von Mitgliedern der Arbeitsgruppe des Projekts
- Teilnahme an einer anderen klinischen Studie
- Obdachlosigkeit
- mit polizeilicher Begleitung im Krankenhaus
- Alter < 18 Jahre
- mangelnde Sprachkenntnisse
- keine schriftliche Einwilligung

### **3.2.2 Ablauf des Studieneinschlusses**

Von den 7.178 Patienten, die sich in den Anästhesieambulanzen der Charité Campus Mitte und des Campus Virchow-Klinikum zwischen Januar und Juni 2010 vorstellten und hinsichtlich der Einschluss- und Ausschlusskriterien untersucht worden sind, konnten 4.110 aufgrund von Ausschlusskriterien nicht an der Studie teilnehmen, so dass 3.068 Patienten eingeschlossen werden konnten. Von weiteren 216 Patienten, bei denen Werte des PHQ-4 oder der verwendeten BSI-Skalen fehlten, wurde der Datensatz nicht analysiert. Schließlich wurden für die vorliegende Studie die Daten von 2.852 Patienten ausgewertet (siehe Abbildung 1).

Die Ambulanzen waren, abgesehen von gesetzlichen Feiertagen, wochentags zwischen 9.00 und 17.00 Uhr geöffnet. Nach ausführlicher mündlicher und schriftlicher Aufklärung über die Studie und dem Einholen einer schriftlichen Einwilligungserklärung führten die für die Studie geeigneten Patienten das computerbasierte Screening durch. Dies fand während ihrer Wartezeit auf das anästhesiologische Aufklärungsgespräch und die dazu gehörige körperliche Untersuchung vor der Operation statt. Zur Hilfe bei technischen Problemen und für Rückfragen zum Vorgehen war immer ein Studienmitarbeiter anwesend. Alle Fragen konnten per Mausklick beantwortet werden, so dass keine Tastatur bedient werden musste. Die Beantwortung der Befragung dauerte pro Patient ca. 20 min.

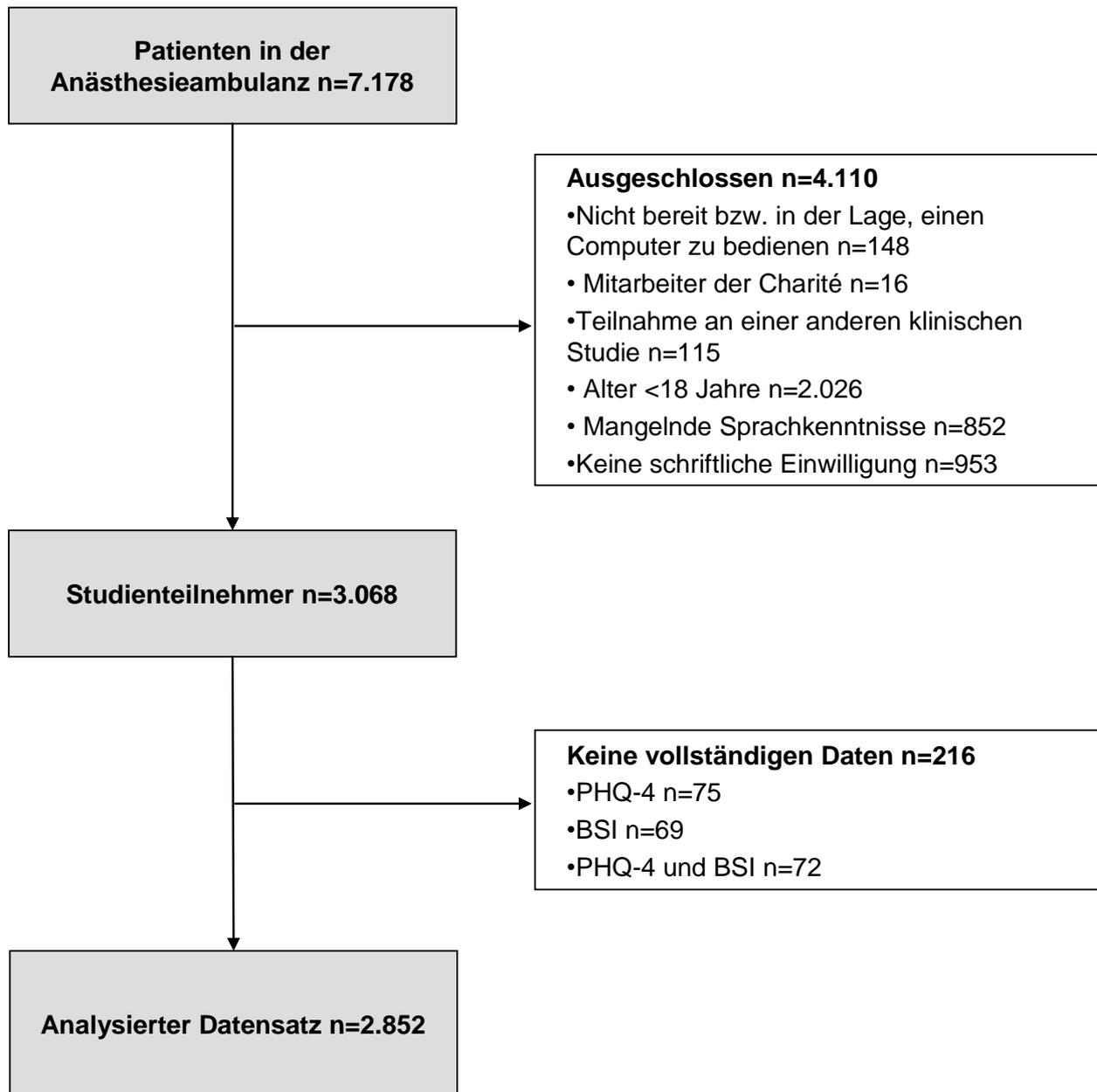


Abbildung 1: Einschlussbaum

### 3.2.3 Missing-Value-Analyse

Von den in die Studie eingeschlossenen 3.068 Patienten gab es bei 216 Patienten (7%) fehlende Werte beim PHQ-4 oder den in dieser Studie verwendeten BSI-Skalen. Im Folgenden werden die 2.852 Patienten mit vollständigen Angaben beim PHQ-4 und den verwendeten BSI-Subskalen und die mit fehlenden Angaben bei jenen Punkten vergleichend untersucht (siehe Tabelle 1). Die univariate Analyse ergab, dass sich die zwei Gruppen signifikant hinsichtlich des Alters und des Bildungsstandes

(Hochschulzugangsberechtigung) unterscheiden. Außerdem unterscheiden sie sich leicht, aber signifikant, hinsichtlich des Geschlechts und des BMI. In allen anderen untersuchten Merkmalshäufigkeiten (BSI-Skala für Depressivität, BSI-Skala für Ängstlichkeit, BSI-Skala für Unsicherheit im Sozialkontakt, BSI-Skala für Phobische Angst, BSI-GSI) unterschieden sich die Gruppen nicht signifikant. Eine multivariate, binäre Regressionsanalyse für dieselben Variablen konnte diese Ergebnisse bestätigen. Nach einem Ausschlussprinzip wurden schrittweise Variablen ausgeschlossen. Zuletzt blieben das Alter (OR: 0,951 [95% CI 0,934-0,968],  $p < 0,000$ ) und die Hochschulzugangsberechtigung (OR: 1,878 [95% CI 1,083-3,257],  $p < 0,025$ ) bestehen.

**Tabelle 1: Vergleich der Patienten mit vollständigen Angaben beim PHQ-4, bei den BSI-Skalen für Depressivität, Ängstlichkeit, Unsicherheit im Sozialkontakt, Phobische Angst und/ oder beim GSI mit den Patienten mit fehlenden Werten beim PHQ-4 und den verwendeten BSI-Subskalen; N=3.068; bei Kardinalskalen: Mittelwert (SD), p-Wert errechnet durch t-Test. Bei Nominalskalen: n(%), p-Werte errechnet durch Chi-Quadrat-Test**

	vollständigen Angaben		fehlende Angaben		p-Werte
	n	M (SD); n(%)	n	M (SD); n(%)	
<b>Alter in Jahren</b>	2.852	47,21 (16,14)	216	54,64(15,55)	<0,001
<b>Hochschulzugangsberechtigung</b>	2.840	1.217 (42,9)	210	61 (29,0)	<0,001
<b>Geschlecht: weiblich</b>	2.852	1.484 (52,0)	216	106 (49,1)	0,041
<b>BMI</b>	2.799	25,86 (5,30)	192	26,64 (4,94)	0,048
<b>BSI-Skala für Depressivität<sup>a</sup></b>	2.852	381 (13,4)	90	12 (13,3)	0,994
<b>BSI-Skala für Ängstlichkeit<sup>b</sup></b>	2.852	316 (11,1)	94	14 (14,9)	0,249
<b>BSI-Skala für Unsicherheit im Sozialkontakt<sup>c</sup></b>	2.852	261 (9,2)	106	10 (9,4)	0,921
<b>BSI-Skala für Phobische Angst<sup>d</sup></b>	2.852	276 (9,7)	104	11 (10,6)	0,761
<b>BSI-GSI<sup>e</sup></b>	2.852	415 (14,6)	83	15 (18,1)	0,371

<sup>a</sup> Klinisch relevante Depressivität nach BSI-Depressivität T-Wert  $\geq 63$

<sup>b</sup> Klinisch relevante Ängstlichkeit nach BSI-Ängstlichkeit T-Wert  $\geq 63$

<sup>c</sup> Klinisch relevante Unsicherheit im Sozialkontakt nach BSI-Unsicherheit im Sozialkontakt T-Wert  $\geq 63$

<sup>d</sup> Klinisch relevante Phobische Angst nach BSI-Phobische Angst T-Wert  $\geq 63$

<sup>e</sup> Klinisch relevante allgemeine psychische Belastung nach BSI-GSI T-Wert  $\geq 63$

### 3.3 Messinstrumente

Im Folgenden werden die für diese Studie verwendeten Messinstrumente näher beschrieben.

#### 3.3.1 PHQ-4

Der PHQ-4 ist ein Selbstauskunftsfragebogen zur Messung von Depressivität und Angst und setzt sich zusammen aus den ersten beiden Items des PHQ-9, bekannt als PHQ-2, und den ersten beiden Fragen des GAD-7, bekannt als GAD-2. Der PHQ-2 erfragt die beiden diagnostischen Kernkriterien für depressive Störungen nach DSM-IV (Minderung des Antriebs, gedrückte Stimmung) und der GAD-2 die Kernkriterien für generalisierte Angststörungen nach DSM-IV (innere Spannung, Befürchtungen)<sup>59,82</sup>. Wie schon in Kapitel 1.2 erwähnt liegen für beide Teile des Fragebogens Studien vor, bei denen sie sich als gute Screeninginstrumente im Bereich der medizinischen Primärversorgung herausgestellt haben<sup>57,63</sup>. Auch der PHQ-4 als Gesamtfragebogen hat in ersten Studien eine gute Reliabilität und Validität bewiesen<sup>49,59,64</sup>.

Der PHQ-4 erfragt, wie oft sich der Patient während der letzten zwei Wochen durch folgende Beschwerden beeinträchtigt gefühlt hat: Interessens- und Freudverlust (Item 1), Niedergeschlagenheit, Schwermut oder Hoffnungslosigkeit (Item 2), Nervosität, Ängstlichkeit oder Anspannung (Item 3) und dem Problem, Sorgen nicht stoppen oder kontrollieren zu können (Item 4). Als Antwortmöglichkeiten stehen „überhaupt nicht“, „an einigen Tagen“, „an mehr als der Hälfte der Tage“ und „beinahe jeden Tag“ zur Verfügung, denen die Zahlenwerte 0 bis 3 zugeordnet sind. Es handelt sich also um eine vierstufige Likert-Skala. Tabelle 2 zeigt die Items des PHQ-4. Bei der Auswertung wird die Summe der Punkte gebildet, so dass für den PHQ-4 Werte von 0 bis 12 möglich sind, das heißt jeweils für den PHQ-2 und GAD-2 Werte von 0 bis 6. Wie es sich in vorherigen Studien bewährt hat, wurde als Cutoff für das Vorhandensein einer klinisch relevanten Depressivität und Ängstlichkeit je ein Punktwert von  $\geq 3$  jeweils im PHQ-2 und im GAD-2 festgelegt<sup>57,63,83</sup>. Diese Werte basieren auf ROC-(Receiver-Operating-Characteristic-) Analysen vorheriger Studien an Patienten der medizinischen Grundversorgung<sup>63,83</sup>. Ursprünglich hatten Kroenke et al. (2009) die Schwere von klinisch relevanter Depression und/oder Angst nach der Punktschwere im PHQ-4 in folgenden Abstufungen gewichtet: normal (0-2), mild (3-5), moderat (6-8), schwer (9-

12)<sup>59</sup>. Als Cutoff-Wert für den PHQ-4 wird in dieser Studie ein Summenwert von  $\geq 6$  Punkten für das Vorliegen einer klinisch relevanten Depression und/ oder Ängstlichkeit verwendet. Löwe et al. (2010) empfehlen diesen aus klinischer Expertise als Cutoff für das Vorliegen einer klinisch relevanten Depression und/ oder Angst, da er dem 96. Perzentilwert entspricht<sup>49</sup>. Kocalevent et al. (2014) erstellten normative Daten für die kolumbianische Bevölkerung. Auch hier entspricht ein Summenwert von 6 im PHQ-4 der 96. Perzentile<sup>64</sup>.

**Tabelle 2: PHQ-4;** Cutoff-Werte: PHQ-2 $\geq 3$ <sup>57,83</sup>, GAD-2 $\geq 3$ <sup>63</sup>, PHQ-4 $\geq 6$ <sup>49</sup>

Wie oft fühlten Sie sich im Verlauf der letzten 2 Wochen durch folgende Beschwerden beeinträchtigt?		Überhaupt nicht	An einzelnen Tagen	An mehr als der Hälfte der Tage	Beinahe jeden Tag
PHQ-2	Wenig Interesse oder Freude an Ihren Tätigkeiten	0	1	2	3
	Niedergeschlagenheit, Schwermut, Hoffnungslosigkeit	0	1	2	3
GAD-2	Nervosität, Ängstlichkeit, Anspannung	0	1	2	3
	Nicht in der Lage sein, Sorgen zu stoppen oder zu kontrollieren	0	1	2	3
Wenn eines oder mehrere dieser Probleme bei Ihnen vorliegen, geben Sie bitte an, wie sehr diese Probleme es Ihnen erschwert haben, Ihre Arbeit zu erledigen, Ihren Haushalt zu regeln oder mit anderen Menschen zu Recht zu kommen.		überhaupt nicht erschwert	etwas erschwert	relativ stark erschwert	sehr stark erschwert

Das fünfte, zusätzliche Item des PHQ-4 ist auch Bestandteil des PHQ-D und dient dazu den Grad der subjektiven Beeinträchtigung und damit der klinischen Relevanz abschätzen zu können. Es geht nicht mit in die Berechnung des PHQ-4-Gesamtscores ein. In dieser Studie bietet es eine Möglichkeit zur Konstruktvalidierung. Es wird erfragt, inwieweit eventuell bestehende Probleme den Alltag der Befragten erschweren. Hier stehen als Antwortmöglichkeiten „überhaupt nicht erschwert“, „etwas erschwert“, „relativ stark erschwert“ und „sehr stark erschwert“ zur Verfügung<sup>60,61</sup>.

Der PHQ-4 wurde bereits wie in Kapitel 1.2 beschrieben hinsichtlich Reliabilität und Validität untersucht: Es zeigte sich nach Kroenke et al. (2009) eine interne Konsistenz der Gesamtskala von  $\alpha=.85$ , der Teilskala Depression von  $\alpha=.81$  und der Subskala von  $\alpha=.82$ <sup>59</sup>. Kocalevent zeigten eine Reliabilität des Gesamtfragebogens PHQ-4 von

$\alpha=.84$ <sup>64</sup>. Löwe et al. (2010) konnten Hinweise für die Validität des PHQ-4 in der deutschen und Kocalevent et al. (2014) in der kolumbianischen Allgemeinbevölkerung liefern<sup>49,64</sup>.

### **3.3.2 Das BSI als Goldstandard der Screeningverfahren für Depressivität und Angst**

Das BSI dient, wie schon in Kapitel 1.2 erläutert, der Erfassung von subjektiver Belastung durch psychische Symptome<sup>52</sup>. Dieses Messinstrument stellt eine Kurzform der SCL-90-R dar und ist, wie der PHQ-4, ein Selbstauskunftsfragebogen. Gefragt wird, inwieweit sich die Untersuchten in der letzten Woche durch verschiedene Symptome, wie zum Beispiel Nervosität und Schwermut, belastet gefühlt haben. Zur Beantwortung steht eine 5-stufige Likert-Skala mit den Antwortmöglichkeiten „überhaupt nicht“, „ein wenig“, „ziemlich“, „stark“ und „sehr stark“ zur Verfügung, denen Zahlenwerte von 0 bis 4 zugeordnet werden.

Das BSI umfasst 53 Items, die zu neun Skalen, drei globalen Kennwerten und vier zusätzlichen Items zusammengefasst werden. Eine Darstellung des gesamten Fragebogens ist dieser Arbeit angehängt (siehe Anhang 1). Die Skalen treffen Aussagen über Somatisierung, Zwanghaftigkeit, Aggressivität/Feindseligkeit, Paranoides Denken, Psychotizismus, Depressivität, Ängstlichkeit, Unsicherheit im Sozialkontakt und Phobische Angst. Zur Validierung des PHQ-4 sind davon folgende Skalen von Bedeutung: Depressivität, Ängstlichkeit, Unsicherheit im Sozialkontakt und Phobische Angst. Die Depressivitäts-Skala (Item 9,16,17,18,35,50) misst Traurigkeit bis hin zu einer schweren, klinisch manifesten Depression (siehe Tabelle 3). Die Items der Skala für Ängstlichkeit (Item 1,12,19,38,45,49) erfragen körperlich spürbare Nervosität bis hin zu tiefer Angst (siehe Tabelle 4). Abgedeckt werden manifeste Angst mit Nervosität, Spannungen, Zittern, Panikattacken und Schreckgefühlen, kognitive Komponenten mit Gefühlen der Besorgnis, sowie somatische Korrelate der Angst. Die Skala „Unsicherheit im Sozialkontakt“ (Item 20,21,22,42) beschreibt leichte soziale Unsicherheit bis hin zum Gefühl völliger persönlicher Unzulänglichkeit (siehe Tabelle 5). Die fünf Fragen der Skala für Phobische Angst (Item 8,28,31,43,47) beschreiben ein leichtes Gefühl von Bedrohung bis hin zur massiven phobischen Angst (siehe Tabelle 6). Die Depressivitäts-Skala dient in der vorliegenden Arbeit als Goldstandard-

Screeningverfahren für Depressivität, die Skalen Ängstlichkeit, Unsicherheit im Sozialkontakt und Phobische Angst als Goldstandards des Angst-Screenings. Des Weiteren wird in dieser Studie der GSI verwendet. Dieser misst die allgemeine psychische Belastung und gibt als einer der drei globalen Kennwerte Auskunft über das Antwortverhalten bei allen 53 Fragen des BSI (siehe Anhang 1)<sup>53</sup>. Zur Berechnung der Skalen des BSI wird die Summe der Itemwerte dividiert durch Anzahl der beantworteten Items. Bei den Skalen für Depressivität, Ängstlichkeit, Unsicherheit im Sozialkontakt und Phobische Angst ist die Auswertung trotz jeweils einem unbeantworteten Item möglich. Die Berechnung des GSI ist bei bis zu 13 fehlenden Werten möglich. Die empfohlenen Cutoff-Werte wurden dem Manual der deutschen BSI-Version entnommen und bei einem T-Wert der Normstichprobe von  $\geq 63$  festgelegt<sup>52,53</sup>.

**Tabelle 3: BSI-Skala für Depressivität<sup>53</sup>**

Wie sehr litten Sie in den letzten Tagen unter ...	überhaupt	ein	ziemlich	stark	sehr
	nicht	wenig			stark
Gedanken, sich das Leben zu nehmen	0	1	2	3	4
Einsamkeitsgefühlen	0	1	2	3	4
Schwermut	0	1	2	3	4
dem Gefühl, sich für nichts zu interessieren	0	1	2	3	4
einem Gefühl der Hoffnungslosigkeit angesichts der Zukunft	0	1	2	3	4
dem Gefühl, wertlos zu sein	0	1	2	3	4

**Tabelle 4: BSI-Skala für Ängstlichkeit<sup>53</sup>**

Wie sehr litten Sie in den letzten Tagen unter ...	überhaupt	ein	ziemlich	stark	sehr
	nicht	wenig			stark
Nervosität oder innerem Zittern	0	1	2	3	4
plötzlichem Erschrecken ohne Grund	0	1	2	3	4
Furchtsamkeit	0	1	2	3	4
dem Gefühl, gespannt oder aufgeregt zu sein	0	1	2	3	4
Schreck- oder Panikanfällen	0	1	2	3	4
so starker Ruhelosigkeit, dass Sie nicht still sitzen können	0	1	2	3	4

**Tabelle 5: BSI-Skala für Unsicherheit im Sozialkontakt<sup>53</sup>**

Wie sehr litten Sie in den letzten Tagen unter ...	überhaupt	ein			sehr
	nicht	wenig	ziemlich	stark	stark
Verletzlichkeit in Gefühlsdingen	0	1	2	3	4
dem Gefühl, dass die Leute unfreundlich sind oder Sie nicht leiden können	0	1	2	3	4
Minderwertigkeitsgefühl gegenüber anderen	0	1	2	3	4
starker Befangenheit im Umgang mit anderen	0	1	2	3	4

**Tabelle 6: BSI-Skala für Phobische Angst<sup>53</sup>**

Wie sehr litten Sie in den letzten Tagen unter ...	überhaupt	ein			sehr
	nicht	wenig	ziemlich	stark	stark
Furcht auf offenen Plätzen oder auf der Straße	0	1	2	3	4
Furcht vor Fahrten in Bus, Straßenbahn, U-Bahn oder Zug	0	1	2	3	4
der Notwendigkeit, bestimmte Dinge, Orte oder Tätigkeiten zu meiden, weil Sie durch diese erschreckt werden	0	1	2	3	4
Abneigung gegenüber Menschenmengen, z.B. beim Einkaufen oder im Kino	0	1	2	3	4
Nervosität, wenn Sie allein gelassen werden	0	1	2	3	4

### 3.3.3 Soziodemographische Charakteristika

Im Rahmen des Screenings wurden mehrere soziodemographische Charakteristika erfragt. Hierzu zählen das Alter in Jahren und drei dichotome Variablen: Das Geschlecht (männlich/weiblich), das Vorhandensein einer Hochschulzugangsberechtigung (ja/nein) und das Zusammenleben in einer festen Partnerschaft (ja/nein).

### 3.3.4 Klinische Charakteristika

Im Folgenden werden die klinischen Charakteristika, die im Rahmen dieser Studie erhoben wurden, beschrieben.

### 3.3.4.1 BMI

Zur klinischen Charakterisierung wurde in der vorliegenden Studie unter anderem der BMI (Body Mass Index) verwendet<sup>84</sup>. Hierzu wurde Körpergröße und Körpergewicht erfragt und daraus der BMI nach folgender Formel errechnet:

$$\text{BMI} = \text{Körpergewicht (kg)} / [\text{Körpergröße (m)}]^2$$

### 3.3.4.2 ASA-Klassifikation

Zur Einschätzung des präoperativen Gesundheitszustands wurde die ASA (American Society of Anesthesiologists)-Klassifikation verwendet. Saklad et al. publizierte 1941 dieses Schema zur Risikoeinschätzung von Patienten vor einer Narkose anhand ihrer systemischen Erkrankungen<sup>85</sup>. 1963 kam es zur Weiterentwicklung der ASA-Klassifikation<sup>86</sup>. Es gibt sechs Stufen: ASA 5 und ASA 6 kommen im Rahmen dieser Studie nicht zur Anwendung, da es sich hierbei um moribunde und hirntote Patienten handelt. Für diese Studie wurde die Einteilung vom jeweiligen behandelnden Anästhesisten vorgenommen. Tabelle 7 stellt die verwendeten Stufen der ASA-Klassifikation da.

---

**Tabelle 7: ASA-Klassifikation**

---

<b>ASA 1</b>	gesunder Patient
<b>ASA 2</b>	Patient mit leichter Allgemeinerkrankung
<b>ASA 3</b>	Patient mit schwerer Allgemeinerkrankung
<b>ASA 4</b>	Patient mit lebensbedrohlicher Erkrankung

---

## 3.4 Statistik

In dieser Studie soll die psychometrische Qualität des PHQ-4 bestimmt werden. Hierzu werden folgende Indikatoren verwendet: Reliabilität (Cronbach's Alpha), faktorielle Validität, divergente und konvergente Validität (durch Korrelation mit konstruktnahen und konstruktfernen Variablen), sowie Sensitivität und Spezifität. Ferner wurde zur Einschätzung der diskriminatorischen Fähigkeiten eine ROC-AUC-Analyse durchgeführt, um die Cutoff-Werte des PHQ-4 und seiner Subskalen zu überprüfen.

Die Daten wurden mit Hilfe der Software SPSS (Statistical Package of Social Science) Version 20.0 (SPSS Inc., Chicago, Illinois 60606, USA) statistisch ausgewertet. Das Signifikanzniveau aller statistischen Tests wurde zweiseitig bei  $p \leq 0,05$  festgelegt. Die 95% Konfidenzintervalle wurden mit dem „confidence interval calculator“<sup>87</sup> oder SPSS berechnet.

Der Umgang mit fehlenden Werten bei einzelnen Items wurde beim BSI dem Test-Manual entnommen (siehe hierzu auch Kapitel 3.3.2)<sup>53</sup>. Beim PHQ-4 wurden keine fehlenden Werte zugelassen.

Die Angabe von Patientencharakteristika erfolgte als Häufigkeiten und Prozente, wenn es sich um nominale Variablen handelte. Ansonsten wurden Median und 25%- sowie 75%-Quartile oder Mittelwerte und Standardabweichung angegeben.

### **3.4.1 Reliabilität**

Die Reliabilität des PHQ-4 wird über das Maß seiner internen Konsistenz, Cronbach's Alpha ermittelt. Cronbach's Alpha bezeichnet das Ausmaß, in dem die Items einer Skala miteinander in Beziehung stehen und ist ein bewährter, robuster Wert. Die Einschätzung der Ergebniswerte richtet sich in dieser Studie nach den Beurteilungsrichtlinien von Fisseni (1997). Hiernach gilt ein Cronbach's Alpha von .80 bis .90 als angemessen<sup>88</sup>.

### **3.4.2 Validität**

#### **3.4.2.1 Faktorielle Validität**

Zur Analyse der faktoriellen Validität des PHQ-4 wird eine explorative Faktorenanalyse nach der Hauptkomponentenmethode (Principal Component Analysis) durchgeführt<sup>89</sup>. Auf diese Weise kann untersucht werden, ob sich viele beobachtbare Variablen (Testitems) auf wenige Variablen (Faktoren) zurückführen lassen. Wenn die manifesten Variablen so deutlich miteinander zusammenhängen, dass sie mehr oder weniger gemeinsame Informationen enthalten, kann man folgern, dass sie den gleichen latenten, dahinterliegenden Faktor abbilden. Sinn dieser Untersuchung ist es, zu überprüfen, inwiefern die Items des PHQ-4 auf gemeinsame Faktoren laden, das heißt miteinander korrelieren, und inwiefern diese Faktoren den beiden Subskalen PHQ-2

und GAD-2 entsprechen. Zur besseren Interpretierbarkeit soll eine Rotation der Faktorenlösung erfolgen. Hierdurch werden hohe Ladungen der Variablen auf die Faktoren weiter verstärkt und niedrige Ladungen weiter abgeschwächt. Somit kann zwischen den Variablen besser differenziert werden. In dieser Studie wird die Rotation nach der geläufigen Varimax-Rotation durchgeführt<sup>90</sup>.

### **3.4.2.2 Konstruktvalidität**

Um einen Nachweis der Konstruktvalidität zu gewährleisten, müssen sowohl die divergente Validität (diskriminante Validität/Diskriminanzvalidität) als auch die konvergente Validität (Konvergenzvalidität) gegeben sein<sup>88</sup>.

Divergente Validität ist dann vorhanden, wenn Messdaten von Parametern, die verschiedene Konstrukte abbilden, nur gering miteinander korrelieren (sofern die Konstrukte auch tatsächlich voneinander unabhängig sind). Konvergente Validität ist gegeben, wenn die Messdaten von Testverfahren, die dasselbe Konstrukt abbilden, hoch miteinander korrelieren. Zur Analyse der divergenten Validität werden Korrelationen mit demographischen Variablen, die im Wesentlichen andere Konstrukte abbilden, verwendet (Alter, Geschlecht, BMI, Hochschulzugangsberechtigung, Zusammenleben in einer festen Partnerschaft). Zur Analyse der konvergenten Validität werden in dieser Studie verschiedene Subskalen des BSI benutzt. Zusätzlich wird noch die subjektive Beeinträchtigung nach dem Zusatzitem des PHQ-4 herangezogen. Zur Beurteilung der Höhe der jeweiligen Korrelationskoeffizienten wurde auf die Empfehlung von Nachtigall und Wirtz (2006) zurückgegriffen. Hiernach sollen sich die konstruktnahen Variablen mindestens im mittleren Bereich von  $r=.50$  bis  $r=.70$  und die konstruktfernen Variablen im Bereich  $r\leq .20$  bis  $r=.50$  befinden<sup>91</sup>.

### **3.4.2.3 Sensitivität, Spezifität, Positiv Prädiktiver Wert, Negativ Prädiktiver Wert**

Wichtige Indikatoren der psychometrischen Qualität eines Screeninginstrumentes sind neben den klassischen drei Testgütekriterien Objektivität, Reliabilität, Validität die Sensitivität und die Spezifität (siehe Abbildung 2). Hiermit kann ermittelt werden, wie exakt die Untersuchungsergebnisse sind<sup>92</sup>. Besonders bei Screeninguntersuchungen spielt dies eine wichtige Rolle.

		Goldstandard		
		positiv	negativ	
Fragebogen	positiv	<b>Richtig positiv</b> (A)	<b>Falsch positiv</b> (B)	PPW = A / A + B
	negativ	<b>Falsch negativ</b> (C)	<b>Richtig negativ</b> (D)	NPW = D / C + D
		Sensitivität = A / A + C	Spezifität = D / B + D	

**Abbildung 2: Vier-Felder-Tafel: Darstellung von Sensitivität, Spezifität, Positiv Prädiktivem Wert (PPW) und Negativ Prädiktivem Wert (NPW)**

Die Sensitivität gibt an, bei welchem Anteil erkrankter Patienten die jeweilige Erkrankung durch die Anwendung des Tests tatsächlich erkannt wird, das heißt ein positives Testresultat auftritt. Sie wird definiert als der Quotient aus richtig positiven Testergebnissen und der Summe aus richtig positiven und falsch negativen Testergebnissen.

$$\text{Sensitivität} = \text{richtig positiv} / \text{falsch negativ} + \text{richtig positiv}$$

Die Spezifität gibt die Wahrscheinlichkeit an, dass tatsächlich Gesunde, die nicht an der betreffenden Erkrankung leiden, im Test auch als gesund erkannt werden. Sie ist definiert als der Quotient aus richtig negativen Testergebnissen und der Summe aus falsch positiven und richtig negativen Testergebnissen; folglich allen Testergebnissen, denen tatsächlich keine Erkrankung zugrunde lag.

$$\text{Spezifität} = \text{richtig negativ} / \text{falsch positiv} + \text{richtig negativ}$$

Bedeutsam in diesem Zusammenhang ist zudem ein Kennwert, der die Wahrscheinlichkeit beziffert, dass ein Patient mit einem positiven Screeningergebnis

auch tatsächlich belastet ist<sup>93</sup>: Der Positiv Prädiktive Wert bezeichnet den Anteil mit positivem Testergebnis, der richtig als belastet gescreent wird.

Positiv Prädiktiver Wert = richtig positiv / falsch positiv + richtig positiv

Ein weiterer Kennwert, der Negativ Prädiktive Wert, bezeichnet den Anteil der Patienten mit negativem Testergebnis, der korrekt als nicht belastet gescreent wird<sup>93</sup>.

Negativ Prädiktiver Wert = richtig negativ / falsch negativ + richtig negativ

### 3.4.3 ROC-AUC-Analyse

Eine weitere Methode, die häufig zur Analyse diskriminatorischer Fähigkeiten von Screeninginstrumenten eingesetzt wird, ist die im Rahmen der Signalentdeckungstheorie<sup>94</sup> entwickelte ROC- (Receiver-Operating-Characteristic-) Kurve<sup>95,96</sup>. Durch die ROC-Analyse wird eine graphische Ausgabe des Verhältnisses von Sensitivität, dargestellt auf der Ordinatenachse, über 1-Spezifität, dargestellt auf der Abszisse, für jeden möglichen Cutoff-Wert des Instruments ausgegeben. Unterschiedliche Punkte auf der Kurve kennzeichnen jeweils eine bestimmte Relation beider Maße, entsprechend verschiedener Cutoff-Punkte. Folgt die ROC-Kurve der Winkelhalbierenden, so beträgt die AUC (Area under the curve)=.50. In diesem Fall kann durch den Test nicht zwischen zwei Gruppen unterschieden werden<sup>95</sup>. Von der deutlichsten Trennung zweier Gruppen kann man bei einer AUC=1 sprechen. Durch die AUC lässt sich die diskriminatorische Qualität eines Instruments bestimmen. Liegt die AUC bei .50, so hat das Instrument keine diskriminatorischen Fähigkeiten, die stärker sind als der Zufall. Bei einer AUC von 1 ist es in der Lage perfekt zwei Gruppen voneinander zu unterscheiden. Zur Interpretation der AUC-Werte werden folgende Kriterien empfohlen: Hohe ( $AUC > .90$ ), mittlere ( $.90 > AUC > .70$ ), kleine ( $.50 < AUC < .70$ ) diskriminatorische Fähigkeiten<sup>97</sup>.

### 3.4.4 Analyse der Cutoff-Werte

Da in der Praxis meist nicht vollständig zwischen belastet und nicht belastet unterschieden werden kann, gilt es einen Kompromiss zwischen Sensitivität und Spezifität zu finden, da nicht beide maximale Werte erreichen können. Es gilt folglich für jeden Anwendungsfall eine Kosten-Nutzen-Abwägung zu treffen und ein

wohlaustariertes Verhältnis zwischen Sensitivität und Spezifität anzustreben<sup>98</sup>. Bei der Abwägung zwischen möglichst hoher Sensitivität und möglichst hoher Spezifität sollten laut Lehr et al. (2008) auch ökonomische und ethische Aspekte berücksichtigt werden<sup>99</sup>.

Löwe et al. (2004) empfehlen für Screeninguntersuchungen, dass die Spezifität mindestens 75% beträgt, und dass die Sensitivität über der Spezifität liegt, damit falsch negative Werte vermieden werden können<sup>100</sup>. Dies ist sinnvoll, da in der klinischen Praxis nach einem positiven Screeningergebnis meist weitere diagnostische Schritte erfolgen, bei einem negativen hingegen wird die Diagnostik nicht weitergeführt<sup>96</sup>.

Eine weitere Möglichkeit das Maß der Güte eines diagnostischen Tests zu beurteilen ist der Youden-Index. Er wird wie folgt berechnet:

$$\text{Youden-Index} = \text{Sensitivität} + \text{Spezifität} - 1$$

Der Youden-Index kann Werte zwischen -1 und +1 annehmen. Ein Cutoff ist dann optimal, wenn der Youden-Index maximal groß ist, denn dann gelingt die Trennung von belasteten und nicht belasteten Patienten am besten<sup>101,102</sup>.

In dieser Studie wird ein Cutoff von  $\geq 3$  Punkten für den PHQ-2 und GAD-2 sowie ein Cutoff von  $\geq 6$  Punkten für den PHQ-4 verwendet<sup>49,57,59,63,83</sup>. Vorherige Studien stufen die Schwere einer klinisch relevanten Depression und/ oder Angst nach folgendem Schema: normal (0-2), mild (3-5), moderat (6-8), schwer (9-12)<sup>59,68</sup>. Diese Studie soll nun die etablierten Cutoff-Werte untersuchen.

## 4 Ergebnisse

### 4.1 Patientencharakteristika

In die vorliegende Untersuchung wurden die Daten von 2.852 operativen Patienten ausgewertet. In Tabelle 8 sind die Charakteristika dieser Personen dargestellt.

**Tabelle 8: Soziodemographische und klinische Charakteristika der untersuchten Stichprobe; N=2.852<sup>+</sup>; Median [25.- 75.Perzentile]; n(%)**

<b>Alter in Jahren</b>	47 [34-60]
<b>Geschlecht: weiblich</b>	1.484 (52)
<b>Mit Partner zusammenlebend</b>	1.772 (63)
<b>Hochschulzugangsberechtigung</b>	1.217 (43)
<b>BMI</b>	25 [22-28]
<b>ASA-Klassifikation<sup>a</sup></b>	
<b>ASA 1</b>	911 (32)
<b>ASA 2</b>	1.545 (55)
<b>ASA 3</b>	361 (13)
<b>ASA 4</b>	7 (0,2)
<b>BSI-Skala für Depressivität<sup>b</sup></b>	381 (13,4)
<b>BSI-Skala für Ängstlichkeit<sup>c</sup></b>	316 (11,1)
<b>BSI-Skala für Unsicherheit im Sozialkontakt<sup>d</sup></b>	261 (9,2)
<b>BSI-Skala für Phobische Angst<sup>e</sup></b>	276 (9,7)
<b>BSI-GSI<sup>f</sup></b>	415 (14,6)

<sup>+</sup>Anzahl variiert je nach Variable wegen fehlender Daten von 2.799 bis 2.852

<sup>a</sup> Präoperativer, körperlicher Zustand nach ASA (American Society of Anesthesiologists)-Klassifikation: ASA 1: gesunde Patienten, ASA 2: Patienten mit leichten Allgemeinerkrankungen, ASA 3: Patienten mit schweren Erkrankungen, ASA 4: Patienten mit lebensbedrohlichen Erkrankungen.

<sup>b</sup> Klinisch relevante Depressivität nach BSI-Depressivität T-Wert  $\geq 63$

<sup>c</sup> Klinisch relevante Ängstlichkeit nach BSI-Ängstlichkeit T-Wert  $\geq 63$

<sup>d</sup> Klinisch relevante Unsicherheit im Sozialkontakt nach BSI-Unsicherheit im Sozialkontakt T-Wert  $\geq 63$

<sup>e</sup> Klinisch relevante Phobische Angst nach BSI-Phobische Angst T-Wert  $\geq 63$

<sup>f</sup> Klinisch relevante allgemeine psychische Belastung nach BSI-GSI T-Wert  $\geq 63$

Das Alter aller Patienten lag im Median bei 47 Jahren mit einem Minimum von 18 Jahren und einem Maximum von 88 Jahren. 1.484 (52%) waren weiblich und 1.368 (48%) männlich. 1.772 (63%) haben mit einem festen Partner zusammengelebt. 1.217 (43%) der Studienteilnehmer haben eine Hochschulzugangsberechtigung. Der BMI lag in der Studienpopulation im Median bei 25. Nach der ASA-Klassifikation gehörten 911

Patienten (32%) zu ASA 1 (gesunde Patienten), 1.545 Patienten (55%) zu ASA 2 (Patienten mit leichten Allgemeinerkrankungen), 361 Patienten (13%) zu ASA 3 (Patienten mit schweren Erkrankungen) und 7 Patienten (0,2%) zu ASA 4 (Patienten mit lebensbedrohlichen Erkrankungen).

Die in dieser Studie als Goldstandard verwendeten Skalen des BSI ergaben folgende Ergebnisse: Die Subskala für Depressivität ergab 381 (13,4%) klinisch relevant belastete Patienten. 316 (11,1%) lagen über dem Cutoff der Skala für Ängstlichkeit, 261 (9,2%) lagen über dem Cutoff der Subskala für Unsicherheit im Sozialkontakt und 276 (9,7%) der Patienten waren von klinisch relevanter phobischer Angst betroffen. Laut des GSI waren 415 (14,6%) psychisch belastet.

## **4.2 Psychometrische Qualität des PHQ-4**

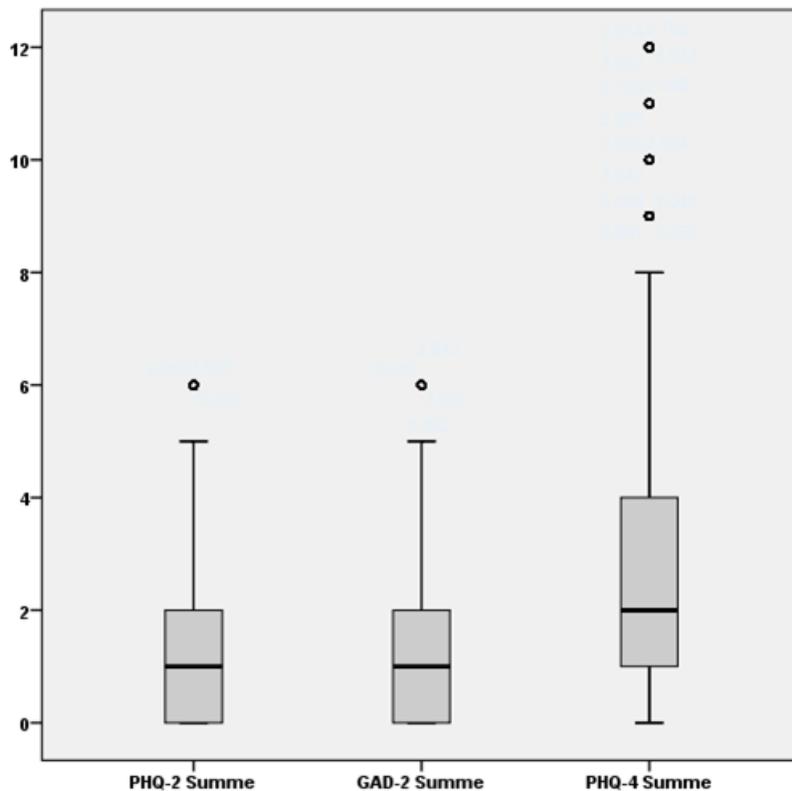
### **4.2.1 Skalenkennwerte des PHQ-4**

Eine detaillierte Darstellung der Skalenkennwerte des PHQ-4 und seiner Subskalen zeigt Tabelle 9. Der Mittelwert beim PHQ-2 lag bei 1,41 (SD=1,40), beim GAD-2 bei 1,22 (SD=1,41) und beim PHQ-4 bei 2,63 (SD=2,58). Der Median des PHQ-2 und des GAD-2 lag jeweils bei 1 [0-2] und der des PHQ-4 bei 2 [1-4]. Alle drei Fragebögen lassen eine rechtsschiefe Verteilung erkennen. Die Kurtosis zeigte eine im Vergleich zur Normalverteilung abgeflachte Verteilung. In Abbildung 3 ist die Verteilung der Summenwerte des PHQ-4 und seiner Subskalen in Boxplots dargestellt. Die Auswertung des PHQ-4 ergab folgende Ergebnisse: Nach der Subskala PHQ-2 waren 496 (17,4%) der untersuchten Stichprobe klinisch relevant depressiv und nach dem GAD-2 380 (13,3%) klinisch relevant ängstlich. 348 (12,2%) der Patienten sind laut PHQ-4 klinisch relevant depressiv und/oder ängstlich. Die Korrelation zwischen dem PHQ-4 und dem PHQ-2, sowie zwischen dem PHQ-4 und dem GAD-2 beträgt jeweils  $r=.92$ . Die Korrelation zwischen PHQ-2 und GAD-2 beträgt  $r=.69$ .

**Tabelle 9: Skalenkennwerte des PHQ-4 und seiner Subskalen in der untersuchten Stichprobe; N=2.852**

	PHQ-2	GAD-2	PHQ-4
<b>Mittelwert (SD)</b>	1,41 (1,40)	1,22 (1,41)	2,63 (2,58)
<b>Median [25.-75.Perzentile]</b>	1 [0-2]	1 [0-2]	2 [1-4]
<b>Modalwert</b>	0	0	0
<b>Schiefe (Standardfehler der Schiefe)</b>	1,10 (0,05)	1,53 (0,05)	1,39 (0,05)
<b>Kurtosis (Standardfehler der Kurtosis)</b>	0,98 (0,92)	2,33 (0,92)	1,92 (0,92)
<b>Spannweite [Minimum- Maximum]</b>	6 [0-6]	6 [0-6]	12 [0-6]
<b>Studienteilnehmer über dem Cutoff <sup>a</sup>; n (%)</b>	496 (17,4)	380 (13,3)	348 (12,2)

<sup>a</sup> Cutoff-Werte: PHQ-2 $\geq$ 3<sup>57,83</sup>, GAD-2 $\geq$ 3<sup>63</sup>, PHQ-4 $\geq$ 6<sup>49</sup>



**Abbildung 3: Boxplot-Diagramm; Verteilung der Gesamtsumme der Punkte bei PHQ-2, GAD-2 und PHQ-4; Box: 25.-75.Perzentile, horizontale Linie: Median, Whiskers: oberes/unteres Quartil  $\pm$  1.5 Box beziehungsweise der niedrigste Wert; Kreise: Ausreißer.**

## 4.2.2 Item-Interkorrelationen des PHQ-4

Betrachtet man die Item-Interkorrelationen des PHQ-4, so finden sich Pearson-Korrelationskoeffizienten zwischen 0,41 und 0,67 (siehe Tabelle 10).

**Tabelle 10: Item-Interkorrelationen des PHQ-4; N=2.852; Korrelation nach Pearson**

	Item 1	Item 2	Item 3	Item 4
Item 1	1			
Item 2	0,50*	1		
Item 3	0,41*	0,64*	1	
Item 4	0,44*	0,67*	0,63*	1

\* Korrelation auf dem Niveau von 0,05 zweiseitig signifikant

Items des PHQ-4: Item 1: Wenig Interesse oder Freude an Ihren Tätigkeiten

Item 2: Niedergeschlagenheit, Schwermut, Hoffnungslosigkeit

Item 3: Nervosität, Ängstlichkeit, Anspannung

Item 4: Nicht in der Lage sein, Sorgen zu stoppen oder zu kontrollieren

## 4.2.3 Reliabilität des PHQ-4

Das Cronbach's Alpha von .66 für den PHQ-2 ist ungenügend. Mit einem Cronbach's Alpha von .78 erweist sich der GAD-2 als im mäßigen Bereich reliabel. Das Cronbach's Alpha für den Gesamtfragebogen PHQ-4 liegt mit .83 im guten Bereich.

## 4.2.4 Validität des PHQ-4

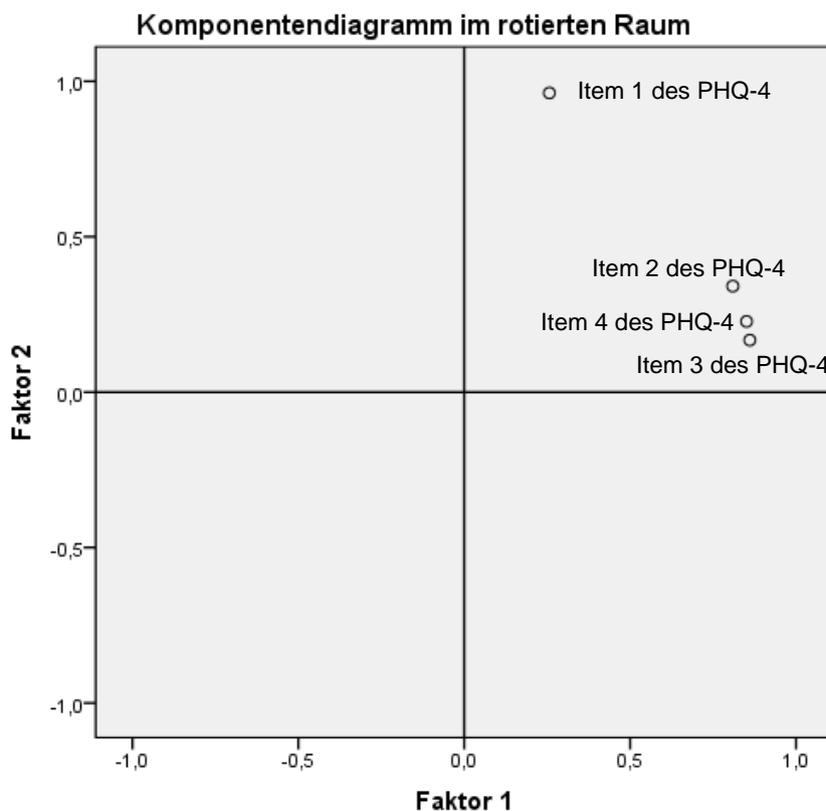
### 4.2.4.1 Faktorielle Validität des PHQ-4

Die faktorielle Struktur des PHQ-4 wurde mittels einer explorativen Faktorenanalyse untersucht. Hierbei wurde eine Hauptkomponentenanalyse als Extraktionsverfahren sowie die Varimax-Rotation als Rotationsmethode eingesetzt. Für den PHQ-4 konnten zwei Faktoren extrahiert werden, die insgesamt 83% der Varianz aufklären. Dabei erklärt der erste Faktor 28% und der zweite Faktor 55% der Gesamtvarianz. Tabelle 11 zeigt die Faktorladungen nach Varimax-Rotation. Item 1 lädt mit einer niedrigen Ladung von 0,257 auf Faktor 1, jedoch mit einer hohen Faktorladung von 0,962 auf Faktor 2. Item 2, 3 und 4 laden mit niedrigen Faktorenladungen zwischen 0,167 und 0,341 auf Faktor 2, jedoch mit hohen Faktorladungen zwischen 0,809 und 0,861 auf Faktor 1. Folglich bildet Item 1 einen Faktor alleine ab, wohingegen sich die restlichen drei Items zu einem weiteren Faktor zusammenfassen lassen. Deutlich wird das in einem Komponentendiagramm im rotierten Raum (Abbildung 4).

**Tabelle 11: rotierte Komponentenmatrix der PHQ-4 Items in der untersuchten Stichprobe (N=2.852), Faktorladungen der PHQ-4 Items**

	Faktor	
	1	2
PHQ-4 Item 1	0,257	0,962
PHQ-4 Item 2	0,809	0,341
PHQ-4 Item 3	0,861	0,167
PHQ-4 Item 4	0,851	0,227

Extraktionsmethode: Hauptkomponentenanalyse. Die Rotation ist in 3 Iterationen konvergiert. Rotationsmethode: Varimax mit Kaiser-Normalisierung



**Abbildung 4: Komponentendiagramm im rotierten Diagramm;** Ordinate: Faktorladungen auf Faktor 2, Abszisse: Faktorenladungen auf Faktor 1; Items des PHQ-4: Item 1: Wenig Interesse oder Freude an Ihren Tätigkeiten; Item 2: Niedergeschlagenheit, Schwermut, Hoffnungslosigkeit; Item 3: Nervosität, Ängstlichkeit, Anspannung, Item 4: Nicht in der Lage sein, Sorgen zu stoppen oder zu kontrollieren

#### 4.2.4.2 Konstruktvalidität des PHQ-4

Die Daten zur Konstruktvalidität des PHQ-4 und seiner Subskalen PHQ-2 und GAD-2 zeigt Tabelle 12. Bei der Analyse der divergenten Validität (Diskriminanzvalidität) zeigten die Variablen Alter, BMI, Hochschulzugangsberechtigung und Zusammenleben

in Partnerschaft bezüglich aller drei Subskalen Pearson-Korrelationskoeffizienten nahe Null. Somit konnte eine sehr geringe Korrelation zwischen PHQ-2, GAD-2 und PHQ-4 und den eben genannten Faktoren gezeigt werden. Die jeweils zur Analyse der konvergenten Validität (Konvergenzvalidität) verwendeten Skalen des BSI korrelierten gut mit PHQ-2, GAD-2 und PHQ-4. Es ergaben sich Werte zwischen  $r=.46$  und  $r=.69$ . Die Korrelation des PHQ-4 und seiner Subskalen mit dem Beeinträchtigungs-Item des PHQ-4 ergab Werte zwischen  $r=.64$  und  $r=.70$ .

**Tabelle 12: Divergente und konvergente Validität von PHQ-2, GAD-2 und PHQ-4; N=2.852;**  
Korrelation nach Pearson

	PHQ-2	GAD-2	PHQ-4
<b>Divergente Validität</b>			
Alter in Jahren	-0,02	-0,07*	0,05*
BMI	0,03	0,01	0,02
Hochschulzugangsberechtigung	0,06*	0,02	-0,04*
Zusammenleben in Partnerschaft	-0,10*	-0,10*	-0,10*
<b>Konvergente Validität</b>			
BSI-Skala für Depressivität	<b>0,61*</b>	0,65*	0,68*
BSI-Skala für Ängstlichkeit	0,50*	<b>0,66*</b>	0,63*
BSI-Skala für Unsicherheit im Sozialkontakt	0,48*	<b>0,55*</b>	0,56*
BSI-Skala für Phobische Angst	0,38*	<b>0,46*</b>	0,46*
BSI-GSI	0,60*	0,67*	<b>0,69*</b>
PHQ-4; Beeinträchtigungs-Item	0,64*	0,64*	0,70*

\* Korrelation auf dem Niveau von 0,05 zweiseitig signifikant  
Fettgedruckt sind untersuchte Korrelationen

#### 4.2.4.3 Sensitivität und Spezifität des PHQ-4

Tabelle 13 und Abbildungen 5 bis 9 zeigen die Analyse der Sensitivität, der Spezifität, des Positiv Prädiktiven Werts und des Negativ Prädiktiven Werts. Die Sensitivität zeigt relativ niedrige Ergebnisse von 61,2% für den PHQ-2 und 46,4% bis 56,3% für den GAD-2 je nach verwendeter Angst-Subskala des BSI als Goldstandard. Die Spezifität der beiden Subskalen des PHQ-4 liegt im sehr guten Bereich mit einem Wert von 89,4% für den PHQ-2 und Werten von 90,2% bis 92,0% für den GAD-2. Für den PHQ-4 als Gesamtfragebogen zeigen sich eine Sensitivität von 51,6% und eine Spezifität von 94,5%. Deutlich erkennbar ist also ein Muster für alle drei Fragebögen, das eine gute

Spezifität, jedoch eine geringe Sensitivität zeigt. Auch beim Positiv und Negativ Prädiktiven Wert zeigt sich ein deutliches Muster. Der Negativ Prädiktive Wert liegt bei allen Fragebögen mit Werten zwischen 92,0% und 93,7% im guten Bereich, wohingegen der Positiv Prädiktive Wert relativ niedrige Ergebnisse zwischen 33,7% und 61,5% zeigt.

**Tabelle 13: Sensitivität, Spezifität, Positiv Prädiktiver Wert (PPW), Negativ Prädiktiver Wert (NPW) von PHQ-2, GAD-2 und PHQ-4 (95% Konfidenzintervall/CI) mit den jeweiligen BSI-Skalen als Goldstandard; N=2.852**

	<b>Sensitivität (95% CI)</b>	<b>Spezifität (95% CI)</b>	<b>PPW (95% CI)</b>	<b>NPW (95% CI)</b>
<b>PHQ-2</b>				
Goldstandard: BSI-Depressivität	61,2% (0.56-0.66)	89,4% (0.88-0.91)	47,0% (0.43-0.51)	93,7% (0.93-0.95)
<b>GAD-2</b>				
Goldstandard: BSI-Ängstlichkeit	56,3% (0.51-0.62)	92,0% (0.91-0.93)	46,8% (0.42-0.52)	94,4% (0.93-0.95)
<b>GAD-2</b>				
Goldstandard: BSI-Unsicherheit im Sozialkontakt	55,6% (0.50-0.62)	90,9% (0.90-0.92)	38,2% (0.33-0.43)	95,3% (0.94-0.96)
<b>GAD-2</b>				
Goldstandard: BSI-Phobische Angst	46,4% (0.41-0.52)	90,2% (0.89-0.91)	33,7% (0.29-0.39)	94,0% (0.93-0.95)
<b>PHQ-4</b>				
Goldstandard: BSI-GSI	51,6% (0.47-0.56)	94,5% (0.94-0.95)	61,5% (0.56-0.67)	92,0% (0.91-0.93)

Cutoff-Werte: PHQ-2 $\geq 3$ <sup>57,83</sup>, GAD-2 $\geq 3$ <sup>63</sup>, PHQ-4 $\geq 6$ <sup>49</sup>

		BSI- Depressivität		
		positiv	negativ	
PHQ-2	positiv	<b>233</b> (richtig positiv)	<b>263</b> (falsch positiv)	PPW = $233/233+263$ = $233/496$ = 47,0%
	negativ	<b>148</b> (falsch negativ)	<b>2.208</b> (richtig negativ)	NPW = $2.208/148+2.208$ = $2.208/2.356$ = 93,7%
		Sensitivität = $233/233+148$ = $233/381$ = 61,2%	Spezifität = $2.208/263+2.208$ = $2.208/2.471$ = 89,4%	N= 2.852

**Abbildung 5: Vier-Felder-Tafel zur Darstellung von Sensitivität, Spezifität, Positiv Prädiktiver Wert (PPW) und Negativ Prädiktiver Wert (NPW) des PHQ-2 mit Goldstandard BSI-Depressivitätsskala**

		BSI- Ängstlichkeit		
		positiv	negativ	
GAD-2	positiv	<b>178</b> (richtig positiv)	<b>202</b> (falsch positiv)	PPW = $178/178+202$ = $178/380$ = 46,8%
	negativ	<b>138</b> (falsch negativ)	<b>2.334</b> (richtig negativ)	NPW = $2.334/138+2.334$ = $2.334/2.472$ = 94,4%
		Sensitivität = $178/178+138$ = $178/316$ = 56,3%	Spezifität = $2.334/202+2.334$ = $2.334/2.536$ = 92,0%	N= 2.852

**Abbildung 6: Vier-Felder-Tafel zur Darstellung von Sensitivität, Spezifität, Positiv Prädiktiver Wert (PPW) und Negativ Prädiktiver Wert (NPW) des GAD-2 mit Goldstandard BSI-Ängstlichkeitsskala**

Ergebnisse

		BSI- Unsicherheit im Sozialkontakt		
		positiv	negativ	
GAD-2	positiv	<b>145</b> (richtig positiv)	<b>235</b> (falsch positiv)	PPW = $145/145+235$ = $145/380$ = 38,2%
	negativ	<b>116</b> (falsch negativ)	<b>2.356</b> (richtig negativ)	NPW = $2.356 /116+2.356$ = $2.356/2.472$ = 95,3%
		Sensitivität = $145/145+116$ = $145/261$ = 55,6%	Spezifität = $2.356 /235+2.356$ = $2.356/2.591$ = 90,9%	N= 2.852

**Abbildung 7: Vier-Felder-Tafel zur Darstellung von Sensitivität, Spezifität, Positiv Prädiktiver Wert (PPW) und Negativ Prädiktiver Wert (NPW) des GAD-2 mit Goldstandard BSI-Unsicherheit im Sozialkontakt**

		BSI- Phobische Angst		
		positiv	negativ	
GAD-2	positiv	<b>128</b> (richtig positiv)	<b>252</b> (falsch positiv)	PPW = $128/128+252$ = $128/380$ = 33,7%
	negativ	<b>148</b> (falsch negativ)	<b>2.324</b> (richtig negativ)	NPW = $2.324/148+2.324$ = $2.324/2.472$ = 94,0%
		Sensitivität = $128/128+148$ = $128/276$ = 46,4%	Spezifität = $2.324/252+2.324$ = $2.324/2.576$ = 90,2%	N= 2.852

**Abbildung 8: Vier-Felder-Tafel zur Darstellung von Sensitivität, Spezifität, Positiv Prädiktiver Wert (PPW) und Negativ Prädiktiver Wert (NPW) des GAD-2 mit Goldstandard BSI-Phobische Angst**

		BSI- GSI		
		positiv	negativ	
PHQ-4	positiv	<b>214</b> (richtig positiv)	<b>134</b> (falsch positiv)	PPW = $214/214+134$ = $214/348$ = 61,5%
	negativ	<b>201</b> (falsch negativ)	<b>2.303</b> (richtig negativ)	NPW = $2.303/201+2.303$ = $2.303/2.504$ = 92,0%
		Sensitivität = $214/214+201$ = $214/415$ = 51,6%	Spezifität = $2.303/134+2.303$ = $2.303/2.437$ = 94,5%	N= 2.852

**Abbildung 9: Vier-Felder-Tafel zur Darstellung von Sensitivität, Spezifität, Positiv Prädiktiver Wert (PPW) und Negativ Prädiktiver Wert (NPW) des PHQ-4 mit Goldstandard BSI-GSI**

#### 4.2.5 AUC-Werte des PHQ-2, GAD-2 und des PHQ-4

Die Größe der Fläche unter der Kurve (Area under the curve) gibt Auskunft über die diskriminatorischen Fähigkeiten. Es ergeben sich für den PHQ-4 und seine beiden Subskalen Werte im mittleren Bereich (Tabelle 14). Die AUC des PHQ-2 beträgt .863 und die des GAD-2 je nach verwendeter Skala des BSI .801 bis .873. Der PHQ-4 zeigt eine annähernd gute AUC von .880. Abbildung 10 zeigt die ROC-Kurven des PHQ-4 und seiner Subskalen, jeweils im Vergleich mit den jeweiligen Skalen des BSI.

**Tabelle 14: Area under the curve des PHQ-2, GAD-2 und PHQ-4; N=2.852**

		AUC
<b>PHQ-2</b>	im Vergleich mit Depressivitätsskala des BSI	.863
<b>GAD-2</b>	im Vergleich mit Ängstlichkeitsskala des BSI	.873
	im Vergleich mit Skala des BSI für Unsicherheit im Sozialkontakt	.833
	im Vergleich mit Skala des BSI für Phobische Angst	.801
<b>PHQ-4</b>	im Vergleich mit dem GSI des BSI	.880

## Ergebnisse

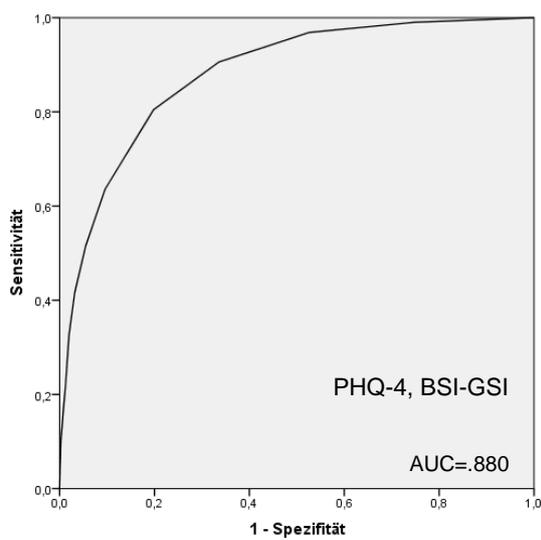
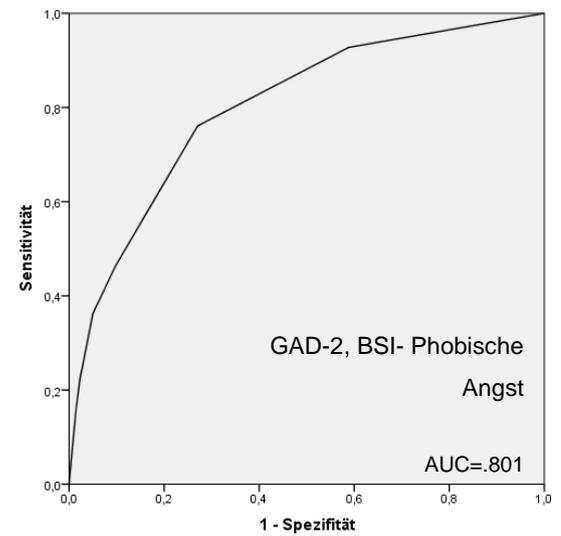
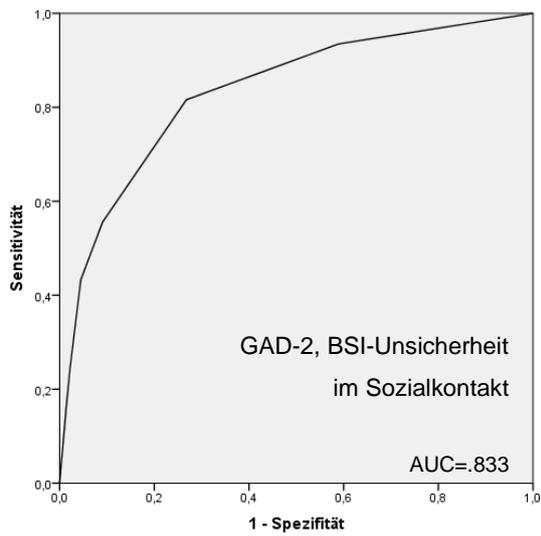
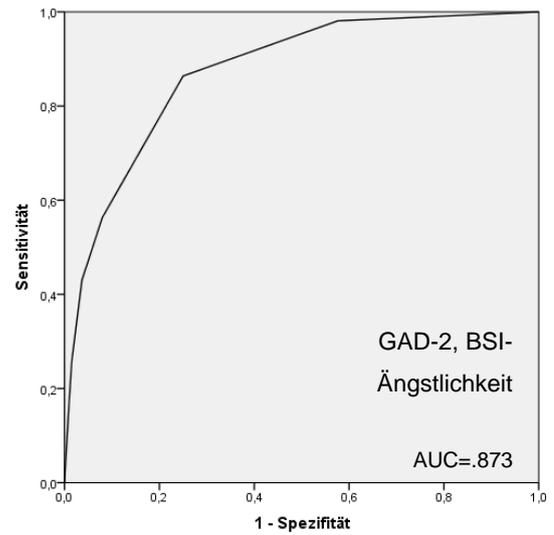
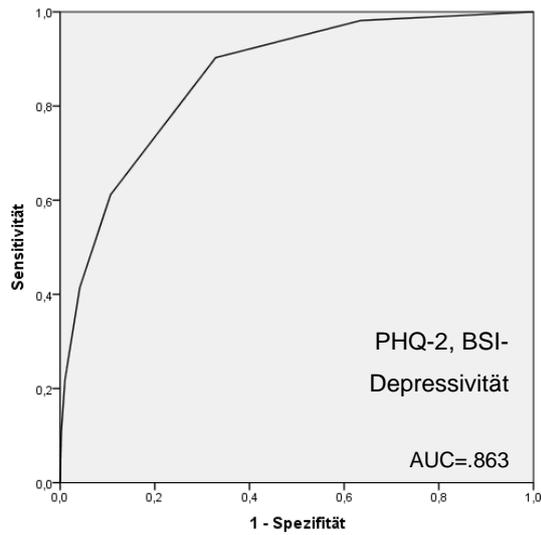


Abbildung 10: ROC-Kurven des PHQ-2, GAD-2 und PHQ-4; N=2.852

#### 4.2.6 Analyse alternativer Cutoff-Werte des PHQ-2, GAD-2 und PHQ-4

Explorativ wurden die Sensitivität und die Spezifität des PHQ-2, GAD-2 und PHQ-4 bei unterschiedlichen Cutoff-Werten untersucht (siehe Tabellen 15-19). Beim PHQ-2 ist bei keinem der möglichen Cutoff-Werte gleichzeitig eine Spezifität von mindestens 75% und eine höhere Sensitivität als Spezifität zu finden. Somit wird das von Löwe et al. (2004) empfohlene Kriterium für einen optimalen Cutoff für keinen der möglichen Cutoff-Werte des PHQ-2 erfüllt. Beim GAD-2 findet man bei einem Cutoff von  $\geq 2$  mit einer Spezifität von 75,0% und einer Sensitivität von 86,4% beim Goldstandard BSI-Ängstlichkeit eine Erfüllung der Kriterien von Löwe et al. (2004). Der Youden-Index ist beim PHQ-2 und GAD-2 mit Maximalwerten zwischen 0,490 und 0,614 jeweils bei einem Cutoff von  $\geq 2$  am höchsten. Die von Löwe et al. (2004) empfohlenen Kriterien werden beim PHQ-4 bei einem Cutoff von  $\geq 4$  erfüllt<sup>100</sup>. Hier findet man eine Spezifität von 80,2% und eine Sensitivität von 80,5%. Die von vorherigen Studien empfohlenen Cutoff-Werte des PHQ-4  $\geq 3$ ,  $\geq 6$  und  $\geq 9$  erfüllen diese Kriterien nicht<sup>49,59</sup>. Verwendet man den Youden-Index zur Bewertung der Cutoff-Werte des PHQ-4, so stellt man fest, dass bei einem Cutoff von  $\geq 4$  der beste Wert erreicht wird (0,61).

**Tabelle 15: Sensitivität und Spezifität des PHQ-2 bei unterschiedlichen Cutoffs anhand des Goldstandards BSI-Depressivität; N=2.852**

Cutoff	Sensitivität	Spezifität	Youden-Index
$\geq 0$	100%	0%	0
$\geq 1$	98,2%	36,5%	0,35
$\geq 2$	90,3%	67,1%	0,57
$\geq 3$	61,2%	89,4%	0,51
$\geq 4$	41,5%	95,8%	0,37
$\geq 5$	21,8%	99,0%	0,21
$\geq 6$	10,8%	99,8%	0,11

**Tabelle 16: Sensitivität und Spezifität des GAD-2 bei unterschiedlichen Cutoffs anhand des Goldstandards BSI-Ängstlichkeit; N=2.852**

Cutoff	Sensitivität	Spezifität	Youden-Index
≥0	100%	0%	0
≥1	98,1%	42,4%	0,41
<b>≥2</b>	<b>86,4%</b>	<b>75,0%</b>	<b>0,61<sup>+</sup></b>
≥3	56,3%	92,0%	0,48
≥4	43,0%	96,3%	0,39
≥5	25,6%	98,5%	0,24
=6	17,4%	99,0%	0,16

<sup>+</sup>Erfüllung der Kriterien für Screeninguntersuchungen nach Löwe et al. (2004): Spezifität mind. 75% und Sensitivität höher als die entsprechende Spezifität<sup>100</sup>

**Tabelle 17: Sensitivität und Spezifität des GAD-2 bei unterschiedlichen Cutoffs anhand des Goldstandards BSI-Unsicherheit im Sozialkontakt; N=2.852**

Cutoff	Sensitivität	Spezifität	Youden-Index
≥0	100%	0%	0
≥1	93,5%	41,1%	0,35
≥2	81,6%	73,2%	0,55
≥3	55,6%	90,9%	0,47
≥4	43,3%	95,5%	0,39
≥5	24,1%	97,8%	0,22
=6	16,5%	98,6%	0,15

**Tabelle 18: Sensitivität und Spezifität des GAD-2 bei unterschiedlichen Cutoffs anhand des Goldstandards BSI-Phobische Angst; N=2.852**

Cutoff	Sensitivität	Spezifität	Youden-Index
≥0	100%	0%	0
≥1	92,8%	41,2%	0,34
≥2	76,1%	72,9%	0,49
≥3	46,4%	90,2%	0,37
≥4	36,2%	95,0%	0,31
≥5	22,1%	97,7%	0,20
=6	15,6%	98,6%	0,14

**Tabelle 19: Sensitivität und Spezifität des PHQ-4 bei unterschiedlichen Cutoffs anhand des Goldstandards BSI-GSI; N=2.852**

Cutoff	Sensitivität	Spezifität	Youden-Index
≥0	100%	0%	0
≥1	99,0%	25,2%	0,24
≥2	96,9%	47,4%	0,57
≥3	90,6%	66,4%	0,57
<b>≥4</b>	<b>80,5%</b>	<b>80,2%</b>	<b>0,61<sup>†</sup></b>
≥5	63,6%	90,4%	0,54
≥6	51,6%	94,5%	0,46
≥7	41,7%	96,8%	0,39
≥8	32,5%	98,0%	0,31
≥9	21,7%	98,8%	0,21
≥10	15,9%	99,3%	0,15
≥11	10,1%	99,7%	0,10
=12	6,7%	99,8%	0,07

<sup>†</sup>Erfüllung der Kriterien für Screeninguntersuchungen nach Löwe et al. (2004): Spezifität mind. 75% und Sensitivität höher als die entsprechende Spezifität<sup>100</sup>

## 5 Diskussion

### 5.1 Stärken und Schwächen des PHQ-4

Im Folgenden werden die Hauptergebnisse der vorliegenden Arbeit zusammengefasst und kritisch diskutiert.

#### 5.1.1 Hauptergebnisse

Die interne Konsistenz des PHQ-4 liegt mit einem Cronbach's Alpha von .83 im guten Bereich. Mit einem Cronbach's Alpha von .66 für den PHQ-2 und .78 für den GAD-2 erweisen sich die beiden Subskalen des PHQ-4 als im mäßigen Bereich reliabel. Die gute Konstruktvalidität des PHQ-4 und seiner Subskalen konnte durch Nachweis der divergenten und konvergenten Validität gezeigt werden. Der PHQ-4 zeigt eine annähernd gute AUC von .863. Bei den Subskalen des PHQ-4 finden sich mäßige bis gute Werte je nach verwendeter Subskala des BSI (zwischen .801 und .873). Der PHQ-4 hat demnach relativ gute diskriminatorische Fähigkeiten. Die Untersuchung der faktoriellen Validität extrahiert zwei Faktoren, wobei das erste Item den einen Faktor abbildet und die anderen drei Items sich zum anderen Faktor zusammenfassen lassen. Die Spezifität befindet sich durchweg, je nach verwendeter Subskala, im sehr hohen Bereich zwischen 89,4% und 94,5%, wohin gehend die Sensitivität mit Werten zwischen 46,4% und 61,2% gering ist. Für den PHQ-4 als Gesamtfragebogen zeigt sich beispielsweise eine Spezifität von 94,5% und eine Sensitivität von 51,6%. Die explorative Untersuchung der Sensitivität und Spezifität des PHQ-4 bei unterschiedlichen Cutoff-Werten zeigt, dass es in der untersuchten Population bei keinem möglichen Cutoff-Wert möglich wäre, gleichzeitig eine optimale Sensitivität und eine optimale Spezifität zu erreichen. Der Youden-Index identifizierte für den PHQ-2 und GAD-2 in dieser Studienpopulation jeweils einen Cutoff von  $\geq 2$  als am besten. Der PHQ-4 hat seinen optimalen Cutoff in dieser Studie bei  $\geq 4$ . Hier findet man den höchsten Youden-Index und eine Spezifität von 80,2% und eine Sensitivität von 80,5%.

### 5.1.2 Reliabilität

Die interne Konsistenz des Gesamtfragebogens PHQ-4 liegt bei einem Cronbach's Alpha von .83. Hiermit konnten die Ergebnisse von Kroenke et al. (2009) und Kocalevent et al. (2014) bestätigt werden, die dem PHQ-4 ebenfalls eine gute interne Konsistenz mit einem Cronbach's Alpha von .85 beziehungsweise .84 nachweisen konnten<sup>59,64</sup>. Mit einem Cronbach's Alpha von .78 erweist sich der GAD-2 als ebenfalls reliabel, wenn auch im mäßigen Bereich. Das Cronbach's Alpha des PHQ-2 ist mit .66 ungenügend. Somit konnte in dieser Stichprobe die gute Reliabilität der Teilfragebögen, die vorherige Studien festgestellt hatten, nicht reproduziert werden. Bisherige Studien hatten sowohl für den PHQ-2 als auch für den GAD-2 ein gutes Maß an interner Konsistenz gezeigt: Kroenke et al. (2009) errechneten ein Cronbach's Alpha von .81 für den PHQ-2 sowie .82 für den GAD-2<sup>59</sup>. Löwe et al. (2010) zeigten ein Cronbach's Alpha von .75 für den PHQ-2 und .82 für den GAD-2<sup>49</sup>. Es zeigt sich in der vorliegenden Studie folglich ein hohes Maß an Zuverlässigkeit beim Gesamtkonstrukt und eine mittelmäßige Zuverlässigkeit bei den Teilfragebögen.

### 5.1.3 Konstruktvalidität

Es konnten positive Hinweise für die Konstruktvalidität des PHQ-4 und seiner Subskalen (Tabelle 12) gefunden werden. Bei der Analyse der divergenten Validität zeigten die Variablen Alter, BMI, Hochschulzugangsberechtigung und Zusammenleben in Partnerschaft bezüglich aller drei Subskalen Pearson-Korrelationskoeffizienten kleiner 0,2. Somit konnte eine sehr geringe Korrelation zwischen den untersuchten Fragebögen und jenen Faktoren, die ein anderes Konstrukt abbilden, gezeigt werden. Außerdem zeigten sich gute Beziehungen zwischen dem PHQ-4, PHQ-2 und GAD-2 und den jeweiligen Skalen des BSI. Die konvergente Validität des PHQ-4 wurde mit Hilfe des GSI untersucht, die des PHQ-2 mit der Depressivitätsskala und für die konvergente Validität des GAD-2 wurden die Skalen für Ängstlichkeit, Unsicherheit im Sozialkontakt und Phobische Angst herangezogen. Wird vorausgesetzt, dass jene Variablen Bereiche desselben Konstrukts abbilden, so zeigen die Pearson-Korrelationskoeffizienten zwischen  $r=.45$  und  $.69$  eine gute konvergente Validität. Die Korrelation des PHQ-4 und seiner Subskalen mit dem Beeinträchtigungs-Item des PHQ-4 ergab Werte zwischen  $r=.64$  und  $r=.70$ , was ebenfalls für eine starke

konvergente Validität spricht. Folglich ist eine Konstruktvalidität beim PHQ-4 und seinen Subskalen PHQ-2 und GAD-2 gegeben. Hiermit konnten die Ergebnisse vorheriger Studien bestätigt werden: Jene untersuchten unter anderem mit Hilfe von Selbstauskunftsfragebögen und soziodemographischen Merkmalen die divergente und konvergente Validität und konnten die Konstruktvalidität des PHQ-4 und oder seiner Subskalen bestätigen<sup>49,57,59,64,83</sup> (siehe Kapitel 1.2).

#### **5.1.4 ROC-AUC-Analysen**

Ein weiteres psychometrisches Qualitätsmerkmal des PHQ-4 findet sich bei der Auswertung der ROC-AUC-Analyse (Tabelle 14). Der PHQ-4 zeigt nach den Kriterien von Swets et al. (1988) eine annähernd gute AUC von .863<sup>97</sup>. Der Fragebogen kann also mäßig gut zwischen zwei Gruppen unterscheiden. Bei den Subskalen des PHQ-4 finden sich mäßige bis gute Werte je nach verwendeter Subskala des BSI (zwischen .801 und .873). In vorherigen Studien konnten bereits gute AUC-Werte für die beiden Subskalen PHQ-2 und GAD-2 festgestellt werden. Kroenke et al. (2003) fanden AUC-Werte zwischen .90 und .93 für den PHQ-2<sup>57</sup>. Für den GAD-2 ergab eine andere Studie AUC-Werte zwischen .80 und .91<sup>63</sup>. In dieser Studie konnten also die vorher beschriebenen guten diskriminatorischen Fähigkeiten von PHQ-2 und GAD-2 bestätigt und eine gute diskriminatorische Fähigkeit des PHQ-4 gezeigt werden.

#### **5.1.5 Faktorenanalyse**

Die Untersuchung der faktoriellen Validität extrahierte zwei Faktoren, wobei das erste Item, die Frage nach der Beeinträchtigung durch wenig Interesse oder Freude an eigenen Tätigkeiten, den einen Faktor abbildet. Die anderen drei Items Beeinträchtigung durch Niedergeschlagenheit, Schwermut, Hoffnungslosigkeit, sowie Nervosität, Ängstlichkeit, Anspannung als auch durch das Gefühl, Sorgen nicht stoppen oder kontrollieren zu können, ergeben den anderen Faktor.

Vorherige Studien konnten zwei Faktoren extrahieren, die jeweils eine Subskala des PHQ-4 abbilden. Kroenke et al. (2009) verwendete eine explorative Faktorenanalyse. Löwe et al. (2010) und Kocalevent et al. (2014) überprüften die Faktorenstruktur des PHQ-4 mit einer konfirmatorischen Faktorenanalyse und bestätigten hierbei die postulierte zweifaktorielle Struktur des PHQ-4<sup>49,59,64</sup>. Die vorliegende Studie wendete

erneut eine explorative Faktorenanalyse an, um unabhängig von vorherigen theoretischen Annahmen die Faktorenstruktur grundlegend zu untersuchen. In der hier untersuchten Stichprobe konnte die postulierte faktorielle Struktur nicht repliziert werden (siehe Tabelle 11).

Die Ergebnisse der vorliegenden Studie zeigen, dass der PHQ-4 nicht zwischen ängstlichen und depressiven Patienten unterscheiden kann. Das Problem der Trennung der beiden Symptombereiche könnte außer methodisch auch inhaltlich bedingt sein, da Depression und Angst häufig als komorbide Beschwerden auftreten<sup>10-12</sup>. Auch bei anderen Selbstauskunftsfragebögen, die Items für die Feststellung von Depressivität und Angst beinhalten, scheint es Schwierigkeiten mit der Trennung der beiden Symptomkomplexe zu geben. Norton et al. (2013) stellten zum Beispiel für die Hospital Anxiety and Depression Scale fest, dass es auch mit diesem Fragebogen nicht möglich ist, gut zwischen klinisch relevant depressiven und ängstlichen Patienten zu differenzieren. Die Untersuchung der faktoriellen Struktur der Hospital Anxiety and Depression Scale ergab, dass durch einen Allgemeinfaktor, wie psychische Belastung, zusätzlich zu den beiden Faktoren Angst und Depressivität ein Großteil der Varianz aufgeklärt werden kann<sup>103</sup>. Für das Beck Depression Inventory II und das Beck Anxiety Inventory konnte eine Faktorenanalyse einen Allgemeinfaktor extrahieren, der ca. 50% der Gesamtvarianz der Items beider Fragebögen aufklärt<sup>104</sup>. Verschiedene Studien untersuchten die Langformen des PHQ-4, PHQ-9 und GAD-7, auf ihre Faktorenstruktur und kamen zu widersprüchlichen Ergebnissen. Gelaye et al. (2013) und Ryan et al. (2013) fanden eine einfaktorielle Struktur von PHQ-9 und GAD-7, wohingegen andere Studien keine Eindimensionalität zeigen konnten<sup>105-110</sup>. In Bezug auf die hier ausgeführte Problematik sollte der PHQ-4 zunächst nur angewendet werden, wenn eine Differenzierung zwischen ängstlichen und depressiven Patienten nicht von Bedeutung ist, sondern die Detektion eines oder beider Merkmale im Vordergrund steht. Durch die in dieser Studie vorliegende Faktorenstruktur des PHQ-4 liegt die Idee, einen PHQ-3 zu entwickeln, nahe. Dadurch könnte eventuell das Problem der mangelnden Differenzierbarkeit gelöst werden. Zu diesem Zweck sollten weitere Studien durchgeführt werden.

### 5.1.6 Sensitivität und Spezifität

Bei der Betrachtung der Sensitivität und Spezifität des PHQ-4 und seiner Subskalen (Tabelle 13) wird ein deutliches Muster erkennbar. Die Spezifität befindet sich durchweg im sehr hohen Bereich zwischen 89,4% und 94,5%, wohingegen die Sensitivität mit Werten zwischen 46,4% und 61,2% geringer ist. Die beiden bisher veröffentlichten Studien zur Untersuchung des PHQ-4 von Kroenke et al. (2009) und Löwe et al. (2010) haben keine Sensitivität und Spezifität des PHQ-4 als Gesamtskala untersucht<sup>49,59</sup>. Fragebögen mit einer hohen Spezifität haben eine geringe Rate an falsch Positiven. Durch seine hohe Spezifität ist der PHQ-4 folglich in der vorliegenden Stichprobe mit diesem Cutoff in der Lage durch eine geringe Rate an falsch Positiven nicht psychisch durch Angst und/oder Depression belastete gut zu erkennen. Er kann also in einem hohen Maß tatsächlich nicht ängstlich und/oder depressive Patienten als solche identifizieren. Die niedrige Sensitivität spricht jedoch dafür, dass der PHQ-4 eine relativ hohe Rate von falsch Negativen hat und folglich ängstliche und/oder depressive Patienten nicht immer als belastet erkennt. Die Tatsache, dass Fragebögen ab einer gewissen Kürze an Qualität einbüßen, wurde bereits von anderen Autoren beschrieben. So untersuchten Mitchell et al. (2007) in einer Metaanalyse verschiedene Ultrakurz-Screeninginstrumente hinsichtlich ihrer Sensitivität und Spezifität und kamen zu dem Ergebnis, dass auf Kosten der Simplizität der Durchführung Abstriche bei der psychometrischen Qualität der Fragebögen gemacht werden. Besonders bezüglich der Sensitivität wurden große Einbußen bei sehr kurzen Fragebögen gefunden<sup>56</sup>.

Vorherige Studien hatten PHQ-2 und GAD-2 einzeln bereits auf ihre Sensitivität und Spezifität hin untersucht und kamen zu folgenden Ergebnissen: Kroenke et al. (2003) zeigten für den PHQ-2 bei einem Cutoff von  $\geq 3$  auf seiner 0-6-Punkte-Skala eine Sensitivität von 83% und eine Spezifität von 90% hinsichtlich der Diagnose Major Depression, die mit Hilfe eines strukturierten klinischen Interviews gestellt wurde<sup>57</sup>. Beim GAD-2 beträgt nach einer Studie von Kroenke et al. (2007) für generalisierte Angststörungen bei einem Cutoff von  $\geq 3$  auf seiner 0-6-Punkte-Skala die Sensitivität 86% und die Spezifität hinsichtlich einer durch ein strukturiertes klinisches Interview diagnostizierte Generalisierten Angststörung 83%<sup>63</sup>. Die sehr positiven Ergebnisse von Kroenke et al. (2003) und Kroenke et al. (2007) konnten in der vorliegenden Studie nicht reproduziert werden. Die Ursache hierfür kann an den unterschiedlichen

Goldstandards liegen. In der vorliegenden Studie wurden Selbstauskunftsfragebögen statt klinischer Interviews verwendet. Kroenke et al. verwendeten das SCID (Structured Clinical Interview for DSM-III-R) und die Prime-MD (Primary Care Evaluation of Mental Disorders)<sup>111,112</sup>. Die Prime-MD besteht zu einem Teil aus Items des PHQ, was mit den positiveren Ergebnissen in Zusammenhang stehen könnte<sup>112</sup>. Im Gegensatz zu den eben genannten Studien wurde in der vorliegenden Arbeit als Goldstandard ein Fragebogen verwendet, was die Ergebnisse in ihrer Vergleichbarkeit an sich einschränkt. Auch die sich unterscheidenden Studienpopulationen könnte dazu beitragen, dass die vorliegende Studie teilweise zu anderen Ergebnissen kommt: Kroenke et al. (2003) und Kroenke et al. (2007) befragten Patienten der US-amerikanischen Grundversorgung und Kliniken für Gynäkologie und Geburtshilfe, die Datenerhebung der vorliegende Studie fand hingegen bei operativen Patienten in Anästhesieambulanzen statt<sup>57,63</sup>.

Die explorative Untersuchung der Sensitivität und Spezifität der Gesamtskala PHQ-4 bei unterschiedlichen Cutoff-Werten zeigt, dass in der untersuchten Population keiner der bisher verwendeten Cutoff-Werte bestätigt werden kann (Tabelle 15-19). Vorherige Studien hatten entweder eine gestufte Einteilung der Schwere von Depression und Angst verwendet (normal 0-2, mild 3-5, moderat 6-8, schwer 9-12) oder einen Cutoff bei  $\geq 6$ , beziehungsweise  $\geq 9$  gesetzt<sup>49,59</sup>. Das von Löwe et al. (2004) empfohlene Kriterium konnte bei keinem der möglichen Cutoff-Werte des PHQ-2 erfüllt werden, da nie eine Spezifität von mindestens 75% und eine höhere Sensitivität als Spezifität erreicht wurde. Beim GAD-2 findet man bei der Verwendung von BSI-Ängstlichkeit als Goldstandard bei einem Cutoff von  $\geq 2$  eine Spezifität von 75,0% und einer Sensitivität von 86,4% und damit eine Erfüllung der Kriterien von Löwe et al. (2004). Der Youden-Index ist beim PHQ-2 und GAD-2 jeweils bei einem Cutoff von  $\geq 2$  am höchsten. Die von Löwe et al. (2004) empfohlenen Kriterien werden bei der Gesamtskala PHQ-4 bei einem Cutoff von  $\geq 4$  erreicht<sup>100</sup>. Bei diesem Cutoff zeigt sich beim PHQ-4 auch ein optimaler Youden-Index (.61), so dass nach den Ergebnissen der vorliegenden Arbeit ein Cutoff von  $\geq 4$  für den PHQ-4 nahegelegt werden kann. Bei diesem Cutoff liegt die Sensitivität bei 80,5% und die Spezifität bei 80,2% (siehe Tabelle 19). Die Cutoff-Werte, die sich in der vorliegenden Studie als die optimalen herausstellten, sind niedriger als die bisher empfohlenen. Da die beiden Subskalen des PHQ-4 ihren empfohlenen Cutoff jeweils bei  $\geq 3$  haben, bedeutet dieses Ergebnis, dass mit den neuen Cutoff-Werten

Patienten mit nur einer Auffälligkeit bei PHQ-2 oder GAD-2 als klinisch relevant depressiv und/ oder ängstlich gescreent werden könnten.

## 5.2 Methodenkritik

### 5.2.1 Studiendesign

Die Ergebnisse der vorliegenden Studie sind nicht uneingeschränkt mit bisher durchgeführten Studien zur Validität des PHQ-4 vergleichbar, da der PHQ-4 zuvor nicht bei operativen Patienten in Anästhesieambulanzen angewendet wurde und sich damit die Populationen unterscheiden. Kroenke et al. (2009), Löwe et al. (2010) und Kocalevent et al. (2014) rekrutierten in ihren Studien zu Validität des PHQ-4 ihre Studienteilnehmer in Kliniken der US-amerikanischen Grundversorgung und durch eine Haushaltsbefragung in Deutschland beziehungsweise Kolumbien<sup>49,59,64</sup>.

Ein Kritikpunkt der vorliegenden Studie könnte ferner sein, dass die Befragung mittels computerbasiertem Fragebogen durchgeführt wurde. Verschiedene Untersuchungen kommen jedoch zu dem Schluss, dass Computer- und Papierfragebögen für Umfragen gleich gut geeignet sind oder die Computerversion sogar überlegen ist<sup>113-115</sup>. Die direkte Eingabe in den Computer durch die Patienten erbrachte den Vorteil, dass dem Personal die häufig sehr zeitraubende Datenübertragung vom Papierfragebogen in den Computer erspart blieb. Somit konnte auch eine potentielle Fehlerquelle ausgeschlossen werden. Die computerbasierte Form hat des Weiteren den Vorteil, dass sie nicht gesondert ausgewertet werden muss, sondern, dass das Programm das Screeningergebnis direkt anzeigt.

Der Einschlussbaum der vorliegenden Studie ist vergleichbar mit dem früherer Arbeiten derselben Arbeitsgruppe, die sich ebenfalls mit der Anwendung von PC-gesteuerten Selbstauskunftsfragebögen bei chirurgischen Patienten in Anästhesieambulanzen beschäftigen<sup>45,116,117</sup>, wodurch eine grundlegende Vergleichbarkeit der Ergebnisse gegeben ist. Bei näherer Betrachtung der Zusammensetzung der Studienpopulation, wird deutlich, dass die Häufigkeit von klinisch relevanter Depressivität und Angst in der vorliegenden Studie vergleichbar ist mit der in vorherigen Arbeiten veröffentlichten Prävalenz von Depressivität und Angst bei chirurgischen Patienten<sup>32,34,36,45,55,118</sup>.

Ein Kritikpunkt an der Methodik der vorliegenden Studie könnte die Wahl des Goldstandards sein, denn es könnte infrage gestellt werden, inwieweit es sinnvoll ist das BSI als Goldstandard bei medizinischen Patienten zu verwenden<sup>119,120</sup>. Obwohl es ein in diesem Bereich häufig genutztes Screeninginstrument ist<sup>50,51,121,122</sup>, zweifeln die Studien von Slaughter et al. (1999) und Tate (1994) die Anwendbarkeit des BSI im medizinischen Setting an<sup>119,120</sup>. Es wurde die Vermutung aufgestellt, dass die hohe Prävalenz von Depressivität beim Erheben der emotionalen Belastung operativer Patienten per Fragebögen auf Items mit somatischen Beschwerden wie Appetitverlust, Schlafprobleme und Libidoverlust zurückgehen könnten. Jene könnten jedoch auch durch die anstehende Operation selbst oder die bestehende physische Ursache zur Operation verursacht werden und weniger Anzeichen von Depressivität sein<sup>123</sup>. Die von Slaughter (1999) und Tate (1994) beschriebene Problematik spielt jedoch in dieser Studie höchstens beim GSI eine Rolle, weil weder der PHQ-4 noch die Subskalen für Depressivität und Ängstlichkeit des BSI somatische Beschwerden erfragen<sup>53,59,119,120</sup>. Außerdem konnten Kerper et al. (2012) für die vorliegende Stichprobe zeigen, dass bei operativen Patienten eine 6-Monatsstabilität der in Selbstauskunftsfragebögen erhobenen psychologische Belastung zu finden ist; es hatte keine signifikanten Veränderungen in Bezug auf die psychische Belastung gegeben. Dies ist ein deutlicher Hinweis dafür, dass die Beschwerden eher nicht durch die anstehende Operation verursacht sind, sondern eher auf klinisch relevante psychische Störungen zurückzuführen sind<sup>124</sup>.

Eine weitere Limitation der vorliegenden Studie ist die Tatsache, dass die Untersuchungen hinsichtlich klinisch relevanter Angst und Depression nicht hinsichtlich feststehender Diagnosen nach ICD-10 oder DSM-V durchgeführt wurden. Als Goldstandard diene ein Fragebogen, kein klinisches strukturiertes Interview. Es gehört jedoch zur psychometrischen Qualität eines Fragebogens, dass er mit anderen Fragebögen übereinstimmt. Auch andere Studien untersuchten Fragebögen anhand anderer Fragebögen<sup>125,126</sup>: So untersuchte zum Beispiel die Studie von De Wit et al. (2007) die Validität des WHO-5 Well-Being Index an Hand der Center for Epidemiologic Studies Depression Scale<sup>126</sup>.

Die Ergebnisse der vorliegenden Studie beruhen auf einer ersten Evaluation des PHQ-4 als eigenständiger Fragebogen an Patienten des operativen Bereichs. Kocalevent et al.

(2014) untersuchten den PHQ-4 als eigenständigen Fragebogen bereits anhand der kolumbianischen Allgemeinbevölkerung<sup>64</sup>. Vorherige Studien verwendeten den PHQ-4 meist im Kontext seiner Langformen PHQ-9, PHQ-8 und/oder GAD-7<sup>49,57,59,63,83</sup>. Folglich sollten weiterführende Studien den PHQ-4 als eigenständiges Screeninginstrument in psychometrischen Studien validieren.

### **5.2.2 Fehlende Werte**

Die Analyse der fehlenden Werte ergab, dass die Daten von 216 Patienten (7%) der eingeschlossenen 2.852 Patienten nicht weiter untersucht werden konnten, weil hier Werte des PHQ-4 oder der in dieser Studie verwendeten BSI-Skalen fehlten. Der Vergleich derjenigen Patienten ohne fehlende Daten mit denen, bei denen jene Werte vorhanden waren, ergab, dass es signifikante Unterschiede bezüglich zwei der untersuchten Variablen gab. Fehlende Werte gab es vermehrt bei älteren Patienten und jenen ohne Hochschulzugangsberechtigung. Folglich sind Ältere und niedriger gebildete Patienten leicht unterrepräsentiert in der untersuchten Studienpopulation.

### **5.3 Zusammenfassung und Implikationen für Forschung und Klinik**

Um festzustellen in wie weit sich der PHQ-4 und seine Subskalen zur Messung von Depressivität und Angst bei operativen Patienten in Anästhesieambulanzen eignen, wurde die psychometrische Qualität des PHQ-4 untersucht. Insbesondere wurden die folgende Kennwerte bestimmt: Interne Konsistenz, faktorielle Validität, divergente und konvergente Konstruktvalidität, sowie Sensitivität und Spezifität.

Verschiedene Untersuchungsergebnisse sprechen für eine gute psychometrische Qualität des PHQ-4. Bei der Reliabilität erbringt die Gesamtskala PHQ-4 gute Ergebnisse. Konstruktvalidität ist sowohl bei PHQ-4 als auch für bei seinen Subskalen gegeben. Auch diskriminatorischen Fähigkeiten sind bei PHQ-2, GAD-2 und PHQ-4 vorhanden.

Es gibt aber auch deutliche Schwachpunkte, wie die Reliabilität der Subskalen PHQ-2 und GAD-2. Beide erscheinen in der vorliegenden Studie nur mäßig reliabel. Die Faktorenstruktur zeigt, dass der PHQ-4 seine Stärke nicht in der Trennung von depressiven und ängstlichen Probanden hat. Dies könnte jedoch inhaltlich begründet

sein, da Depressivität und Ängstlichkeit häufig als komorbide Beschwerden auftreten. Der PHQ-4 sollte also als Screener psychischer Belastung durch einen oder beide Aspekte gesehen werden. Bei einer sehr guten Spezifität mangelt der PHQ-4 in dieser Studie an Sensitivität. Folglich legen die Ergebnisse nahe, dass der PHQ-4, bis weitere Studien vorliegen, nur im Rahmen eines breitgefächerten Screenings angewendet werden sollte.

Im Rahmen weiterer Forschung sollte der hohen Spezifität und niedrigen Sensitivität des PHQ-4 Bedeutung zukommen: Dieses Muster ist insbesondere bei Screeninginstrumenten ein Schwachpunkt. Es sollten weitere Studien durchgeführt werden, die den PHQ-4 hinsichtlich seiner Sensitivität untersuchen. Hierzu könnten andere längere Ängstlichkeits- und Depressivitätsscreener als das BSI als Goldstandard verwendet werden. Bestandteil weiterer Forschung sollte es ferner sein, den PHQ-4 in weiteren Studien durch ein strukturiertes klinisches Interview zur Messung von klinisch relevanter Ängstlichkeit und Depressivität sowie allgemeiner psychischer Belastung im Setting der Anästhesieambulanz zu validieren.

Im Klinikalltag ist zu empfehlen den PHQ-4 vorläufig als Teil eines breitgefächerten Screenings anzuwenden. Das Konzept des BRIA-Projekts an der Charité Berlin hat außer dem verwendeten breitgefächerten Screening eine weitere Lösung gefunden. Es wird jeder Patient, der an dem Screening teilnimmt gefragt, ob der Wunsch besteht mit einem Psychologen zu sprechen. Somit wird denjenigen, die als nicht depressiv und/oder ängstlich gescreent wurden, die Möglichkeit gegeben sich im Erstinterviewer als behandlungsbedürftig psychisch belastet zu zeigen<sup>44</sup>.

## 7 Literaturverzeichnis

1. Hautzinger M. Depression. Göttingen: Hogrefe Verlag; 1998.
2. Hautzinger M, Meyer T. Diagnostik affektiver Störungen In: Stieglitz R-D, Baumann U, Freyberger H (Hrsg.). Psychodiagnostik in der Klinischen Psychologie, Psychiatrie, Psychotherapie. Stuttgart: Georg Thieme Verlag; 2001:S. 418-29.
3. Stieglitz R-D. Diagnostik und Klassifikation in der Psychiatrie. Stuttgart: Kohlhammer; 2008.
4. Fabian E. Die Angst: Geschichte, Psychodynamik, Therapie. Münster: Waxmann Verlag; 2013.
5. Wittchen HU, Jacobi F, Rehm J, et al. The size and burden of mental disorders and other disorders of the brain in Europe 2010. *European Neuropsychopharmacology* 2011;21:655-79.
6. Jacobi F, Wittchen HU, Holting C, et al. Prevalence, co-morbidity and correlates of mental disorders in the general population: results from the German Health Interview and Examination Survey (GHS). *Psychological Medicine* 2004;34:597-611.
7. Löwe B, Spitzer RL, Williams JB, Mussell M, Schellberg D, Kroenke K. Depression, anxiety and somatization in primary care: syndrome overlap and functional impairment. *General Hospital Psychiatry* 2008;30:191-9.
8. Kessler RC, Berglund P, Demler O, et al. The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R). *Journal of the American Medical Association* 2003;289:3095-105.
9. Alonso J, Angermeyer MC, Bernert S, et al. Disability and quality of life impact of mental disorders in Europe: results from the European Study of the Epidemiology of Mental Disorders (ESEMeD) project. *Acta Psychiatrica Scandinavica Supplementum* 2004:38-46.
10. Mineka S, Watson D, Clark LA. Comorbidity of anxiety and unipolar mood disorders. *Annual Review of Psychology* 1998;49:377-412.
11. Schoevers RA, Beekman AT, Deeg DJ, Jonker C, van Tilburg W. Comorbidity and risk-patterns of depression, generalised anxiety disorder and mixed anxiety-depression in later life: results from the AMSTEL study. *International Journal of Geriatric Psychiatry* 2003;18:994-1001.
12. Byers AL, Yaffe K, Covinsky KE, Friedman MB, Bruce ML. High occurrence of mood and anxiety disorders among older adults: The National Comorbidity Survey Replication. *Archives of General Psychiatry* 2010;67:489-96.

13. Haller H, Cramer H, Lauche R, Gass F, Dobos GJ. The prevalence and burden of subthreshold generalized anxiety disorder: a systematic review. *BMC Psychiatry* 2014;14:128.
14. Kujanpää T, Ylisaukko-Oja T, Jokelainen J, et al. Prevalence of anxiety disorders among Finnish primary care high utilizers and validation of Finnish translation of GAD-7 and GAD-2 screening tools. *Scandinavian Journal of Primary Health Care* 2014;32:78-83.
15. Goldberg D. The detection and treatment of depression in the physically ill. *World Psychiatry* 2010;9:16-20.
16. Iosifescu DV. Treating depression in the medically ill. *Psychiatric Clinics of North America* 2007;30:77-90.
17. Moussavi S, Chatterji S, Verdes E, Tandon A, Patel V, Ustun B. Depression, chronic diseases, and decrements in health: results from the World Health Surveys. *Lancet* 2007;370:851-8.
18. Ebmeier KP, Donaghey C, Steele JD. Recent developments and current controversies in depression. *Lancet* 2006;367:153-67.
19. Benton T, Staab J, Evans DL. Medical co-morbidity in depressive disorders. *Annals of Clinical Psychiatry* 2007;19:289-303.
20. Rymaszewska J, Kiejna A. [Depression and anxiety after coronary artery bypass grafting]. *Pol Merkur Lekarski* 2003;15:193-5.
21. Stroobant N, Vingerhoets G. Depression, anxiety, and neuropsychological performance in coronary artery bypass graft patients: a follow-up study. *Psychosomatics* 2008;49:326-31.
22. Hobby JL, Venkatesh R, Motkur P. The effect of psychological disturbance on symptoms, self-reported disability and surgical outcome in carpal tunnel syndrome. *Journal of Bone & Joint Surgery (Br)* 2005;87:196-200.
23. Gallagher R, McKinley S. Anxiety, depression and perceived control in patients having coronary artery bypass grafts. *Journal of Advanced Nursing* 2009;65:2386-96.
24. Caumo W, Schmidt AP, Schneider CN, et al. Risk factors for postoperative anxiety in adults. *Anaesthesia* 2001;56:720-8.
25. Caumo W, Schmidt AP, Schneider CN, et al. Risk factors for preoperative anxiety in adults. *Acta Anaesthesiologica Scandinavia* 2001;45:298-307.
26. Krannich JH, Weyers P, Lueger S, Herzog M, Bohrer T, Elert O. Presence of depression and anxiety before and after coronary artery bypass graft surgery and their relationship to age. *BMC Psychiatry* 2007;7:47.

27. Vingerhoets G. Perioperative anxiety and depression in open-heart surgery. *Psychosomatics* 1998;39:30-7.
28. Johnston M. Anxiety in surgical patients. *Psychological Medicine* 1980;10:145-52.
29. Blumenthal JA, Lett HS, Babyak MA, et al. Depression as a risk factor for mortality after coronary artery bypass surgery. *Lancet* 2003;362:604-9.
30. Evans DL, Charney DS, Lewis L, et al. Mood disorders in the medically ill: scientific review and recommendations. *Biological Psychiatry* 2005;58:175-89.
31. Rosenberger PH, Jokl P, Ickovics J. Psychosocial factors and surgical outcomes: an evidence-based literature review. *Journal of the American Academy of Orthopaedic Surgeons* 2006;14:397-405.
32. Smith PJ, Attix DK, Weldon BC, Greene NH, Monk TG. Executive function and depression as independent risk factors for postoperative delirium. *Anesthesiology* 2009;110:781-7.
33. Kerper L, Spies C, Buspavanich P, et al. Preoperative depression and hospital length of stay in surgical patients. *Minerva Anestesiologica* 2014;80:984-91.
34. Vaerøy H, Juell M, Høivik B. Prevalence of depression among general hospital surgical inpatients. *Nordic Journal of Psychiatry* 2003;57:13-6.
35. Bosch JA, Engeland CG, Cacioppo JT, Marucha PT. Depressive symptoms predict mucosal wound healing. *Psychosomatic Medicine* 2007;69:597-605.
36. Sinikallio S, Aalto T, Airaksinen O, Herno A, Kröger H, Viinamäki H. Depressive burden in the preoperative and early recovery phase predicts poorer surgery outcome among lumbar spinal stenosis patients: a one-year prospective follow-up study. *Spine* 2009;34:2573-8.
37. Borowicz L, Royall R, Grega M, Selnes O, Lyketsos C, McKhann G. Depression and cardiac morbidity 5 years after coronary artery bypass surgery. *Psychosomatics* 2002;43:464-71.
38. Blumenthal JA, Lett HS, Babyak MA, et al. Depression as a risk factor for mortality after coronary artery bypass surgery. *Lancet* 2003;362:604-9.
39. Wittchen HU, Krause P, Hoyer J, et al. Prevalence and correlates of generalized anxiety disorders in primary care. *Fortschritte der Medizin* 2001;119:17-25.
40. Wittchen HU. Generalized anxiety disorder: prevalence, burden, and cost to society. *Depression and Anxiety* 2002;16:162-71.
41. Badner NH, Nielson WR, Munk S, Kwiatkowska C, Gelb AW. Preoperative anxiety: detection and contributing factors. *Canadian Journal of Anaesthesia* 1990;37:444-7.

42. Ray C, Fitzgibbon G. Stress arousal and coping with surgery. *Psychological Medicine* 1981;11:741-6.
43. Wallace LM. Psychological preparation as a method of reducing the stress of surgery. *Journal of Human Stress* 1984;10:62-77.
44. Lange LF, Spies C, Weiß-Gerlach E, et al. Bridging Intervention in Anaesthesiology: First results on treatment need, demand and utilization of an innovative psychotherapy program for surgical patients. *Clinical Health Promotion* 2011;1:4-12.
45. Linnen H, Krampe H, Neumann T, et al. Depression and essential health risk factors in surgical patients in the preoperative anaesthesiological assessment clinic. *European Journal of Anaesthesiology* 2011;28:733-41.
46. US Preventive Services Task Force. Screening for depression in adults: U.S. preventive services task force recommendation statement. *Annals of Internal Medicine* 2009;151:784-92.
47. Katon W, Roy-Byrne P. Anxiety disorders: efficient screening is the first step in improving outcomes. *Annals of Internal Medicine* 2007;146:390-2.
48. Kroenke K, Spitzer RL, Williams JB, Löwe B. The Patient Health Questionnaire Somatic, Anxiety, and Depressive Symptom Scales: a systematic review. *General Hospital Psychiatry* 2010;32:345-59.
49. Löwe B, Wahl I, Rose M, et al. A 4-item measure of depression and anxiety: validation and standardization of the Patient Health Questionnaire-4 (PHQ-4) in the general population. *Journal of Affective Disorders* 2010;122:86-95.
50. Giesinger JM, Kuster MS, Behrend H, Giesinger K. Association of psychological status and patient-reported physical outcome measures in joint arthroplasty: a lack of divergent validity. *Health Qual Life Outcomes* 2013;11:64.
51. Purushotham AD, Upponi S, Klevesath MB, et al. Morbidity after sentinel lymph node biopsy in primary breast cancer: results from a randomized controlled trial. *Journal of Clinical Oncology* 2005;23:4312-21.
52. Derogatis L. Brief Symptom Inventory (BSI): administration, scoring, and procedures manual (3). Minneapolis: National Computer Services; 1993.
53. Franke G. Brief Symptom Inventory von L.R. Derogatis (Kurzform des SCL-90-R)- Deutsche Version. Göttingen: Beltz test; 2000.
54. Davis GE, Yueh B, Walker E, Katon W, Koepsell TD, Weymuller EA. Psychiatric distress amplifies symptoms after surgery for chronic rhinosinusitis. *Otolaryngology-Head and Neck Surgery* 2005;132:189-96.
55. O'Hara MW, Ghoneim MM, Hinrichs JV, Mehta MP, Wright EJ. Psychological consequences of surgery. *Psychosomatic Medicine* 1989;51:356-70.

56. Mitchell AJ, Coyne JC. Do ultra-short screening instruments accurately detect depression in primary care? A pooled analysis and meta-analysis of 22 studies. *British Journal of General Practice* 2007;57:144-51.
57. Kroenke K, Spitzer RL, Williams JB. The Patient Health Questionnaire-2: validity of a two-item depression screener. *Medical Care* 2003;41:1284-92.
58. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *Journal of General Internal Medicine* 2001;16:606-13.
59. Kroenke K, Spitzer RL, Williams JB, Löwe B. An ultra-brief screening scale for anxiety and depression: the PHQ-4. *Psychosomatics* 2009;50:613-21.
60. Arroll B, Goodyear-Smith F, Kerse N, Fishman T, Gunn J. Effect of the addition of a "help" question to two screening questions on specificity for diagnosis of depression in general practice: diagnostic validity study. *British Medical Journal* 2005;331:884.
61. Löwe B. Adding a question that asks whether help is wanted improves sensitivity of a depression screening tool. *Evidence- Based Mental Health* 2006;9:39.
62. Spitzer RL, Kroenke K, Williams JB, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Archives of Internal Medicine* 2006;166:1092-7.
63. Kroenke K, Spitzer RL, Williams JB, Monahan PO, Löwe B. Anxiety disorders in primary care: prevalence, impairment, comorbidity, and detection. *Annals of Internal Medicine* 2007;146:317-25.
64. Kocalevent RD, Finck C, Jimenez-Leal W, Sautier L, Hinz A. Standardization of the Colombian version of the PHQ-4 in the general population. *BMC Psychiatry* 2014;14:205.
65. Polidoro Lima M, Osório FL. Indicators of psychiatric disorders in different oncology specialties: a prevalence study. *Journal of Oncology* 2014;2014:350262.
66. Häuser W, Schmutzer G, Hinz A, Hilbert A, Brähler E. [Prevalence of chronic pain in Germany. A representative survey of the general population]. *Schmerz* 2013;27:46-55.
67. Zenger M, Petrowski K, Ernst J, Götze H, Hinz A. [The protective impact of optimistic expectations in male cancer patients]. *Zeitschrift für Psychosomatische Medizin und Psychotherapie* 2012;58:11-25.
68. Bagley CL, Rendas-Baum R, Maglante GA, et al. Validating Migraine-Specific Quality of Life Questionnaire v2.1 in Episodic and Chronic Migraine. *Headache* 2012;52:409-21.
69. Wang SJ, Wang PJ, Fuh JL, Peng KP, Ng K. Comparisons of disability, quality of life, and resource use between chronic and episodic migraineurs: a clinic-based study in Taiwan. *Cephalalgia* 2013;33:171-81.

70. Lauche R, Häuser W, Jung E, et al. Patient-related predictors of treatment satisfaction of patients with fibromyalgia syndrome: results of a cross-sectional survey. *Clinical and Experimental Rheumatology* 2013;31:34-40.
71. Häuser W, Bohn D, Kühn-Becker H, Erdkönig R, Brähler E, Glaesmer H. Is the association of self-reported childhood maltreatments and adult fibromyalgia syndrome attributable to depression? A case control study. *Clinical and Experimental Rheumatology* 2012;30:59-64.
72. Schmidt S, Petermann F, Beutel ME, Braehler E. Psychological Condition, Ailments and Impairments: Results from a Representative Study. *Zeitschrift für Psychiatrie, Psychologie und Psychotherapie* 2011;59:155-65.
73. Zenger M, Hinz A, Petermann F, Brähler E, Stöbel-Richter Y. [Health and quality of life within the context of unemployment and job worries]. *Psychotherapie, Psychosomatik, Medizinische Psychologie* 2013;63:129-37.
74. Nesterko Y, Seidel N, Brähler E, Glaesmer H. [Depression and anxiety in elderly Jews from former Soviet Union in Germany: the role of discrimination and religiosity]. *Psychiatrische Praxis* 2014;41:76-81.
75. Glaesmer H, Braehler E, Grande G, Hinz A, Petermann F, Romppel M. The German Version of the Hopkins Symptoms Checklist-25 (HSCL-25) --factorial structure, psychometric properties, and population-based norms. *Comprehensive Psychiatry* 2014;55:396-403.
76. Khalil AA, Hall LA, Moser DK, Lennie TA, Frazier SK. The psychometric properties of the brief symptom inventory depression and anxiety subscales in patients with heart failure and with or without renal dysfunction. *Archives of Psychiatric Nursing* 2011;25:419-29.
77. Abu Ruz ME, Lennie TA, Riegel B, McKinley S, Doering LV, Moser DK. Evidence that the brief symptom inventory can be used to measure anxiety quickly and reliably in patients hospitalized for acute myocardial infarction. *Journal of Cardiovascular Nursing* 2010;25:117-23.
78. Geisheim C, Hahlweg K, Fiegenbaum W, Frank M, Schroder B, von Witzleben I. The German version of the Brief Symptom Inventory (BSI): Reliability and validity in a sample of outpatient psychotherapy patients. *Diagnostica* 2002;48:28-36.
79. Cohen L, de Moor C, Amato RJ. The association between treatment-specific optimism and depressive symptomatology in patients enrolled in a Phase I cancer clinical trial. *Cancer* 2001;91:1949-55.
80. Lang AJ, Norman SB, Means-Christensen A, Stein MB. Abbreviated brief symptom inventory for use as an anxiety and depression screening instrument in primary care. *Depression and Anxiety* 2009;26:537-43.

81. Piersma HL, Boes JL, Reaume WM. Unidimensionality of the Brief Symptom Inventory (BSI) in adult and adolescent inpatients. *Journal of Personality Assessment* 1994;63:338-44.
82. American-Psychiatric-Association. *Diagnostical and Statistical Manual of Mental Disorders DSM-IV-TR (4th Edition)*. Washington DC: American Psychiatric Press; 2000.
83. Löwe B, Kroenke K, Gräfe K. Detecting and monitoring depression with a two-item questionnaire (PHQ-2). *Journal of Psychosomatic Research* 2005;58:163-71.
84. Keys A, Fidanza F, Karvonen MJ, Kimura N, Taylor HL. Indices of relative weight and obesity. *Journal of Chronic Diseases* 1972;25:329-43.
85. Saklad M. Grading of Patients for Surgical Procedures. *Anesthesiology* 1941;2:281- 4.
86. Dripps R. New classification of physical status. *Anesthesiology* 1963;24:111.
87. Herbert R. Confidence Interval Calculator. 2013. Zugriff auf <http://www.pedro.org.au/english/downloads/confidence-interval-calculator/> am 25.01.2014.
88. Fisseni H. *Lehrbuch der psychologischen Diagnostik (2.Aufl.)*. Göttingen: Hogrefe Verlag; 1997.
89. Bühner M. *Einführung in die Test- und Fragebogenkonstruktion*. München: Pearson; 2006.
90. Kaiser H. The varimax criterion for analytic rotation in factor analysis. *Psychometrika* 1958;3.
91. Nachtigall C, Wirtz M. *Wahrscheinlichkeitsrechnung und Inferenzstatistik Statistische Methoden für Psychologen Teil 2*. Weinheim: Juventa Verlag; 2006.
92. Altman DG, Bland JM. Diagnostic tests. 1: Sensitivity and specificity. *British Medical Journal* 1994;308:1552.
93. Altman DG, Bland JM. Diagnostic tests 2: Predictive values. *British Medical Journal* 1994;309:102.
94. Lusted LB. Signal detectability and medical decision-making. *Science* 1971;171:1217-9.
95. Altman DG, Bland JM. Diagnostic tests 3: receiver operating characteristic plots. *British Medical Journal* 1994;309:188.
96. Murphy JM, Berwick DM, Weinstein MC, Borus JF, Budman SH, Klerman GL. Performance of screening and diagnostic tests. Application of receiver operating characteristic analysis. *Archives of General Psychiatry* 1987;44:550-5.

97. Swets JA. Measuring the accuracy of diagnostic systems. *Science* 1988;240:1285-93.
98. Stieglitz RD. Screening. *Psychotherapie, Psychosomatik, Medizinische Psychologie* 2007;57:178-89.
99. Lehr D, Hillert A, Schmitz E, Sosnowsky N. Screening depressiver Störungen mittels Allgemeiner Depressions-Skala (ADS-K) und State- Trait Depressions Scales (STDS-T) Eine vergleichende Evaluation von Cut-off-Werten. *Diagnostica* 2008;54:1-10.
100. Löwe B, Spitzer RL, Gräfe K, et al. Comparative validity of three screening questionnaires for DSM-IV depressive disorders and physicians' diagnoses. *Journal of Affective Disorders* 2004;78:131-40.
101. Youden WJ. Index for rating diagnostic tests. *Cancer* 1950;3:32-5.
102. Moosbrugger H, Kelava A. Testtheorie und Fragebogenkonstruktion 2.Auflage. Berlin: Springer Verlag; 2012.
103. Norton S, Cosco T, Doyle F, Done J, Sacker A. The Hospital Anxiety and Depression Scale: a meta confirmatory factor analysis. *Journal of Psychosomatic Research* 2013;74:74-81.
104. Steer RA, Clark DA, Beck AT, Ranieri WF. Common and specific dimensions of self-reported anxiety and depression: the BDI-II versus the BDI-IA. *Behavior Research Therapy* 1999;37:183-90.
105. Ryan TA, Bailey A, Fearon P, King J. Factorial invariance of the Patient Health Questionnaire and Generalized Anxiety Disorder Questionnaire. *British Journal of Clinical Psychology* 2013;52:438-49.
106. Kertz S, Bigda-Peyton J, Bjorgvinsson T. Validity of the Generalized Anxiety Disorder-7 scale in an acute psychiatric sample. *Clinical Psychology & Psychotherapy* 2013;20:456-64.
107. Elhai JD, Contractor AA, Tamburrino M, et al. The factor structure of major depression symptoms: a test of four competing models using the Patient Health Questionnaire-9. *Psychiatry Research* 2012;199:169-73.
108. Forkmann T, Gauggel S, Spangenberg L, Brähler E, Glaesmer H. Dimensional assessment of depressive severity in the elderly general population: psychometric evaluation of the PHQ-9 using Rasch Analysis. *Journal of Affective Disorders* 2013;148:323-30.
109. Gelaye B, Williams MA, Lemma S, et al. Validity of the patient health questionnaire-9 for depression screening and diagnosis in East Africa. *Psychiatry Research* 2013;210:653-61.
110. Chilcot J, Rayner L, Lee W, et al. The factor structure of the PHQ-9 in palliative care. *Journal of Psychosomatic Research* 2013;75:60-4.

111. Spitzer RL, Williams JB, Gibbon M, First MB. The Structured Clinical Interview for DSM-III-R (SCID). I: History, rationale, and description. *Archives of General Psychiatry* 1992;49:624-9.
112. Spitzer RL, Williams JB, Kroenke K, et al. Utility of a new procedure for diagnosing mental disorders in primary care. The PRIME-MD 1000 study. *Journal of the American Medical Association* 1994;272:1749-56.
113. Corman S. Computerized vs pencil and paper collection of network data. *Social Networks* 1990;12:375-84.
114. Robinson R, West R. A comparison of computer and questionnaire methods of history-taking in a genito-urinary clinic. *Psychology & Health* 1992;6:77-84.
115. Murrelle L, Ainsworth B, Bulger J, SC H. Computerized mental health risk appraisal for college students: User acceptability and correlation with standard pencil-and-paper questionnaires. *American Journal of Health Promotion* 1992;7:90-2.
116. Kip MJ, Neumann T, Jugel C, et al. New strategies to detect alcohol use disorders in the preoperative assessment clinic of a German university hospital. *Anesthesiology* 2008;109:171-9.
117. Kleinwächter R, Kork F, Weiss-Gerlach E, et al. Improving the detection of illicit substance use in preoperative anesthesiological assessment. *Minerva Anestesiologica* 2010;76:29-37.
118. Zhong BL, Chen HH, Zhang JF, et al. Prevalence, correlates and recognition of depression among inpatients of general hospitals in Wuhan, China. *General Hospital Psychiatry* 2010;32:268-75.
119. Slaughter J, Johnstone G, Petroski G, Flax J. The usefulness of the Brief Symptom Inventory in the neuropsychological evaluation of traumatic brain injury. *Brain Injury* 1999;13:125-30.
120. Tate DG, Heinrich RK, Maynard F, Buckelew SP. Moderator-variable effect on the Brief Symptom Inventory test-item endorsements of spinal cord injury patients. *Paraplegia* 1994;32:473-9.
121. Franke GH, Heemann U, Kohnle M, Luetkes P, Maehner N, Reimer J. Quality of life in patients before and after kidney transplantation. *Psychology & Health* 2000;14:1037-49.
122. Zabora J, BrintzenhofeSzoc K, Jacobsen P, et al. A new psychosocial screening instrument for use with cancer patients. *Psychosomatics* 2001;42:241-6.
123. Burg JS, Williams R, Burrig RG, Donovan PJ. Psychiatric treatment outcome following traumatic brain injury. *Brain Injury* 2000;14:513-33.

124. Kerper LF, Spies CD, Lößner M, et al. Persistence of psychological distress in surgical patients with interest in psychotherapy: results of a 6-month follow-up. *PloS One* 2012;7:e51167.
125. Stafford L, Judd F, Gibson P, Komiti A, Quinn M, Mann GB. Comparison of the Hospital Anxiety and Depression Scale and the Center for Epidemiological Studies Depression Scale for detecting depression in women with breast or gynecologic cancer. *General Hospital Psychiatry* 2013;36:74-80.
126. de Wit M, Pouwer F, Gemke RJ, Delemarre-van de Waal HA, Snoek FJ. Validation of the WHO-5 Well-Being Index in adolescents with type 1 diabetes. *Diabetes Care* 2007;30:2003-6.

## **Lebenslauf**

Mein Lebenslauf wird aus datenschutzrechtlichen Gründen in der elektronischen Version meiner Arbeit nicht veröffentlicht.



## Danksagung

Ganz besonderer Dank gilt Herrn PD Dipl.-Psych. Dr. Henning Krampe und Frau Dipl.-Psych. Dr. Leonie Kerper, wissenschaftliche Mitarbeiter der Charité, Klinik für Anästhesiologie und operative Intensivmedizin, für die intensive Betreuung und Unterstützung. Sie führten mich mit unendlicher Geduld in das wissenschaftliche Arbeiten ein und standen mir stets zur Seite. Ihr fachliches Wissen, die zeitaufwendigen Rückmeldungen zu meiner Arbeit und ihre motivierenden Worte trugen essentiell zum Erfolg dieser Arbeit bei. Durch die Zusammenarbeit durfte ich viele Erfahrung sammeln, die mir weit über die Dissertation hinaus eine große Hilfe sind und sein werden.

Bedanken möchte ich mich auch bei Frau Prof. Dr. Claudia Spies, Direktorin der Klinik für Anästhesiologie und operative Intensivmedizin der Charité-Universitätsmedizin Berlin, Campus Charité Mitte und Campus Virchow-Klinikum für die Überlassung des Dissertationsthemas und die organisatorische Betreuung.

Bedanken möchte ich mich auch bei allen Mitarbeitern der Anästhesieambulanzen des Campus Charité Mitte und Campus Virchow-Klinikums für die freundliche und unterstützende Zusammenarbeit. Auch den Patienten möchte ich danken, denn durch ihre Bereitwilligkeit, an der Studie teilzunehmen, haben sie diese Arbeit überhaupt erst möglich gemacht.

Für den unentbehrlichen statistischen Rat bedanke ich mich bei Herrn Prof. Dr. Karl Wegscheider, Leiter des Instituts für medizinische Biometrie und Epidemiologie des Universitätsklinikums Hamburg-Eppendorf.

Und was wäre ich ohne meine Familie und Freunde... Danke- an Mama und Papa, die stets an mich glauben und mir perfekte Vorbilder sind. Danke- an Tabea, die von meinem kleinen Schwesterherz zur Unterstützung in jeder Lebenslage wurde. Danke- an Claudi, Friederike, Gaby, Mayra und Max, die ausnahmslos immer für mich da sind. Danke- an den Menschen, der meine gesamte Studienzeit und damit mich geprägt hat... Danke Franzi.

## Eidesstattliche Versicherung

„Ich, Janina Tillinger, versichere an Eides statt durch meine eigenhändige Unterschrift, dass ich die vorgelegte Dissertation mit dem Thema: „Der PHQ-4 zur Messung von Depressivität und Angst bei operativen Patienten der Anästhesieambulanz: Untersuchung eines Ultrakurz-Screeninginstruments“ selbstständig und ohne nicht offengelegte Hilfe Dritter verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel genutzt habe.

Alle Stellen, die wörtlich oder dem Sinne nach auf Publikationen oder Vorträgen anderer Autoren beruhen, sind als solche in korrekter Zitierung (siehe „Uniform Requirements for Manuscripts (URM)“ des ICMJE -[www.icmje.org](http://www.icmje.org)) kenntlich gemacht. Die Abschnitte zu Methodik (insbesondere praktische Arbeiten, Laborbestimmungen, statistische Aufarbeitung) und Resultaten (insbesondere Abbildungen, Graphiken und Tabellen) entsprechen den URM (s.o) und werden von mir verantwortet.

Meine Anteile an den ausgewählten Publikationen entsprechen denen, die in der untenstehenden gemeinsamen Erklärung mit dem/der Betreuer/in, angegeben sind. Sämtliche Publikationen, die aus dieser Dissertation hervorgegangen sind und bei denen ich Autor bin, entsprechen den URM (s.o) und werden von mir verantwortet.

Die Bedeutung dieser eidesstattlichen Versicherung und die strafrechtlichen Folgen einer unwahren eidesstattlichen Versicherung (§156,161 des Strafgesetzbuches) sind mir bekannt und bewusst.“

---

Janina Tillinger

Berlin, 03.02.2015

**Anteilserklärung an der erfolgenden Publikation**

Janina Tillinger hatte folgenden Anteil an der folgenden Publikation:

Publikation 1:

Kerper LF, Spies C, Tillinger J, Wegscheider K, Salz AL, Weiss-Gerlach E, Neumann T, Krampe H: Screening for depression, anxiety, and general psychological distress in pre-operative surgical patients: A psychometric analysis of the Patient Health Questionnaire 4 (PHQ-4). Clinical Health Promotion 2014; 4: 5

Beitrag im Einzelnen:

Substantielle Mitwirkung bei Konzeption und Studiendesign, bei der Rekrutierung der operativen Patienten, Durchführung der Datenerhebung und Datenpflege, Datenanalyse und -interpretation, sowie bei Literaturrecherche, Schreiben des Manuskriptentwurfs und Anfertigung der Publikation in der vorliegenden Form.

---

PD Dipl.-Psych. Dr. Henning Krampe  
Berlin, 03.02.2015

---

Janina Tillinger  
Berlin, 03.02.2015

# Anhang

## Anhang 1: BSI

Wie sehr litten Sie in den letzten Tagen unter ...	überhaupt nicht	ein wenig	ziemlich	stark	sehr stark
1.) Nervosität oder innerem Zittern	0	1	2	3	4
2.) Ohnmachts- und Schwindelgefühle	0	1	2	3	4
3.) der Idee, dass irgendjemand Macht über Ihre Gedanken hat	0	1	2	3	4
4.) dem Gefühl, dass andere an den meisten Ihrer Schwierigkeiten Schuld sind	0	1	2	3	4
5.) Gedächtnisschwierigkeiten	0	1	2	3	4
6.) dem Gefühl, leicht reizbar oder verärgert zu sein	0	1	2	3	4
7.) Herz- oder Brustschmerzen	0	1	2	3	4
8.) Furcht auf offenen Plätzen oder auf der Straße	0	1	2	3	4
9.) Gedanken, sich das Leben zu nehmen	0	1	2	3	4
10.) dem Gefühl, dass man den meisten Menschen nicht trauen kann	0	1	2	3	4
11.) schlechtem Appetit	0	1	2	3	4
12.) plötzlichem Erschrecken ohne Grund	0	1	2	3	4
13.) Gefühlsausbrüchen, denen gegenüber Sie machtlos waren	0	1	2	3	4
14.) Einsamkeitsgefühlen, selbst wenn Sie in Gesellschaft sind	0	1	2	3	4
15.) dem Gefühl, dass es Ihnen schwer fällt, etwas anzufangen	0	1	2	3	4
16.) Einsamkeitsgefühlen	0	1	2	3	4
17.) Schwermut	0	1	2	3	4
18.) dem Gefühl, sich für nichts zu interessieren	0	1	2	3	4
19.) Furchtsamkeit	0	1	2	3	4
20.) Verletzlichkeit in Gefühlsdingen	0	1	2	3	4
21.) dem Gefühl, dass die Leute unfreundlich sind oder Sie nicht leiden können	0	1	2	3	4
22.) Minderwertigkeitsgefühl gegenüber anderen	0	1	2	3	4

## Anhang

23.) Übelkeit oder Magenverstimmung	0	1	2	3	4
24.) dem Gefühl, dass andere Sie beobachten oder über Sie reden	0	1	2	3	4
25.) Einschlafschwierigkeiten	0	1	2	3	4
26.) dem Zwang, wieder und wieder nachzukontrollieren was Sie tun	0	1	2	3	4
27.) Schwierigkeiten, sich zu entscheiden	0	1	2	3	4
28.) Furcht vor Fahrten in Bus, Straßenbahn, U-Bahn oder Zug	0	1	2	3	4
29.) Schwierigkeiten beim Atmen	0	1	2	3	4
30.) Hitzewallungen oder Kälteschauer	0	1	2	3	4
31.) der Notwendigkeit, bestimmte Dinge, Orte oder Tätigkeiten zu meiden, weil Sie durch diese erschreckt werden	0	1	2	3	4
32.) Leere im Kopf	0	1	2	3	4
33.) Taubheit oder Kribbeln in einzelnen Körperteilen	0	1	2	3	4
34.) Dem Gefühl, dass Sie für Ihre Sünden bestraft werden sollten	0	1	2	3	4
35.) Einem Gefühl der Hoffnungslosigkeit angesichts der Zukunft	0	1	2	3	4
36.) Konzentrationsschwierigkeiten	0	1	2	3	4
37.) Schwächegefühl in einzelnen Körperteilen	0	1	2	3	4
38.) dem Gefühl, gespannt oder aufgeregt zu sein	0	1	2	3	4
39.) Gedanken an den Tod und ans Sterben	0	1	2	3	4
40.) dem Drang, jemanden zu schlagen, zu verletzen oder ihm Schmerz zuzufügen	0	1	2	3	4
41.) dem Drang, Dinge zu zerbrechen oder zu zerschmettern	0	1	2	3	4
42.) starker Befangenheit im Umgang mit anderen	0	1	2	3	4
43.) Abneigung gegenüber Menschenmengen, z.B. beim Einkaufen oder im Kino	0	1	2	3	4
44.) dem Eindruck, sich einer anderen Person nie so richtig nahe fühlen zu können	0	1	2	3	4
45.) Schreck- oder Panikanfällen	0	1	2	3	4
46.) der Neigung, immer wieder in Erörterungen und Auseinandersetzungen zu geraten	0	1	2	3	4
47.) Nervosität, wenn Sie allein gelassen werden	0	1	2	3	4
48.) Mangelnder Anerkennung Ihrer Leistung durch andere	0	1	2	3	4

## Anhang

---

49.) So starker Ruhelosigkeit, dass Sie nicht still sitzen können	0	1	2	3	4
50.) dem Gefühl, wertlos zu sein	0	1	2	3	4
51.) dem Gefühl, dass Leute Sie ausnutzen, wenn Sie es zulassen würden	0	1	2	3	4
52.) Schuldgefühlen	0	1	2	3	4
53.) dem Gefühl, dass irgendetwas mit Ihrem Verstand nicht in Ordnung ist	0	1	2	3	4

---