

# Telemedizin bei psychischen Erkrankungen: Vorstellung und erste Ergebnisse eines telefonbasierten Betreuungsansatzes

## Telemedicine for mental disorders: Presentation and initial results of a telephone-based care approach

Autorinnen/Autoren  
Georg Bornemann 

### Institute

Klinik und Poliklinik für Psychiatrie und Psychotherapie,  
Universitätsklinikum Leipzig, Leipzig, Germany

### Schlüsselwörter

Telemedizin, Psychotherapie, Depression, Angst, Telefon

### Keywords

telemedicine, psychotherapy, depression, anxiety, telephone

eingereicht 21.06.2024

akzeptiert nach Revision 01.10.2024

Artikel online veröffentlicht 2024

### Bibliografie

Fortschr Neurol Psychiatr

DOI 10.1055/a-2439-3965

ISSN 0720-4299

© 2024. Thieme. All rights reserved.

Georg Thieme Verlag KG, Rüdigerstraße 14,  
70469 Stuttgart, Germany

### Korrespondenzadresse

Georg Bornemann

Klinik und Poliklinik für Psychiatrie und Psychotherapie,

Universitätsklinikum Leipzig

Leipzig

Germany

georg\_bornemann@gmx.de



Zusätzliches Material finden Sie unter

<https://doi.org/10.1055/a-2439-3965>

### ZUSAMMENFASSUNG

**Ziel der Studie** Ein bereits in die Regelversorgung einer privaten Krankenversicherung implementiertes und telefonbasiertes Gesundheitsprogramm für Menschen mit depressiven und Angststörungen sowie Stressbelastung soll als Betreuungsansatz ebenso wie erste Evaluationsergebnisse vorgestellt werden.

**Methodik** Grundlage des Programmes ist eine mehrmonatige telefonische Betreuung durch medizinisches Fachpersonal. Schwerpunkte sind die Vermittlung psychoedukativer Inhalte und verhaltenstherapeutischer Techniken inkl. bedarfsgerechter Gewährleistung entlastender Gespräche. Im Programm wurde als Fragebogen die deutsche Version des DASS 21 eingesetzt, der Belastung in den Dimensionen Depression, Angst und Stress sowie eine psychische Gesamtbelastung detektiert. Vorgestellt werden die demografischen Daten sowie die in diesem Rahmen gewonnenen Fragebogenergebnisse aller TN aus den ersten elf Monaten des Programmes, welche zu zwei Zeitpunkten Fragebögen ausgefüllt hatten. Die Veränderungen der DASS 21-(Unter-)Scores nach fünf Monaten im Coaching wurden als primäre Endpunkte festgesetzt. Die Berechnungsgrundlage stellte der t-Test für abhängige Stichproben (Signifikanzniveau  $\alpha = 0,05$ ) dar.

**Ergebnisse** Für 86 von 199 TN lagen zwei ausgefüllte Fragebögen vor (43,2 %). Die Population war im Durchschnitt 49,5 Jahre alt und überwiegend weiblich (61,6 %). Durchschnittlich erstreckte sich ein Coaching über 170 (SD = 58) Tage, 5,6 (1,9) Telefonate und 200 (81) Gesprächsminuten. Zu Beginn des Programmes zeigten sich erhöhte Werte in allen Fragebogenkategorien (Depression: 7,13 [4,99]; Angst: 4,49 [4,07]; Stress: 9,85 [4,16]; Gesamt: 21,47 [11,54]). Nach fünf Monaten konnte eine Reduktion der jeweiligen Belastungen gemessen werden (Depression: -2,14 [3,99], Cohen's  $d = 0,54$ ; Angst: -1,65 [3,36],  $d = 0,49$ ; Stress: -2,24 [4,81],  $d = 0,47$ ; Gesamt: -6,03 [10,70];  $d = 0,56$ ;  $p$  jeweils  $< 0,001$ ). Aufgeschlüsselt nach Belastungsstärke zu Betreuungsbeginn zeigte sich in einigen Unterkategorien eine mit zunehmendem Schweregrad betragsmäßig vergrößerte Score-Reduktion bei mittleren bis großen Effektstärken.

**Schlussfolgerung** Das beschriebene Telefonprogramm ist als begleitender Ansatz bei psychischen Erkrankungen vielversprechend praktikabel. Die dargestellten Ergebnisse lassen wegen des quasi Prä-Post-Studiendesign keine finale Aussage über eine Wirksamkeit des Programmes zu. Jedoch empfiehlt sich auf Grundlage der gezeigten signifikanten Verbesserungen der psychischen Belastung im Verlauf, inklusive relevanter Effektstärken, eine weitere Evaluation dieses Settings unter kontrollierten Bedingungen.

**ABSTRACT**

**Objective** A telephone-based health program for people with depressive and anxiety disorders as well as stress, already implemented in the standard care of a private health insurance company, as well as the initial evaluation results will be presented.

**Methods** The program, based on several months of telephone support for patients from medical professionals, focuses on teaching psychoeducational content, behavioral therapy techniques, and emotional support. In the current program, the German version of the DASS 21 questionnaire was used, which detects distress in the dimensions of depression, anxiety and stress as well as overall psychological stress. The demographic data and the questionnaire results of all participants from the first eleven months of the program, who completed questionnaires at two time points, are presented. The reductions in the DASS 21 (sub)scores after five months of coaching were defined as the primary endpoints. The basis for calculation was the t-test for dependent samples ( $\alpha = 0.05$ ).

**Results** Two completed questionnaires were available from 86 of 199 participants (43.2%). The population was on average

49.5 years old and predominantly female (61.6%). In total, coaching lasted 170 (SD = 58) days and 200 (81) minutes. On average, these minutes were distributed over 5.6 (1.9) calls. At baseline, there were increased scores in all subcategories (depression: 7.13 [4.99]; anxiety: 4.49 [4.07]; stress: 9.85 [4.16]; total: 21.47 [11.54]). After five months, a reduction in each parameter was measured (depression: -2.14 [3.99], Cohen's  $d = 0.54$ ; anxiety: -1.65 [3.36],  $d = 0.49$ ; stress: -2.24 [4.81],  $d = 0.47$ ; total: -6.03 [10.70];  $d = 0.56$ ;  $p < 0.001$  for each). Broken down by severity of stress at the start of care, several subcategories showed an increase in score reduction with increasing severity, with medium to large effect sizes.

**Conclusion** The telephone program described is promising as an accompanying approach to treatment of mental illness. Due to the pre-post study design, the results presented do not allow a final statement to be made about the effectiveness of the program. However, based on the significant improvements in psychological stress shown over the course of the program, including relevant effect sizes, a further evaluation of this setting under controlled conditions is recommended.

## Einleitung

Psychische Störungen gehören zu den häufigsten Entitäten: Innerhalb von zwölf Monaten wird statistisch gesehen eine von drei Personen psychisch krank [1]. Die Lebenszeitprävalenz für Angststörungen wird in der Literatur auf 17 %, diejenige für depressive Störungen auf 11 % bis sogar 20 % geschätzt [2–4]. Obwohl die aktuellen Leitlinien sowohl für Angst- als auch depressive Störungen neben einer Medikation die Implementierung von Psychotherapie beziehungsweise gesprächsbasierten Interventionen propagieren, ist die Versorgungslage in dieser Hinsicht prekär [5, 6]. Wie eine Evaluationsstudie der Bundespsychotherapeutenkammer zeigte, beträgt die durchschnittliche Wartezeit bis zum Beginn einer Behandlung circa 20 Wochen [7]. Eine flächendeckende bedarfsgerechte sowie zeitnahe Behandlung psychischer Erkrankungen gemäß Leitlinienempfehlung ist unter diesen Voraussetzungen kaum zu leisten. In diese Versorgungslücke stoßen zunehmend digitale und telemedizinische Angebote insbesondere im Bereich der psychischen Gesundheit, welche sogar in den aktuellen Leitlinien Erwähnung finden [8] ([5], S. 47 u. 63). Als Telemedizin bezeichnet man den Einsatz internet- beziehungsweise telefonbasierter digitaler sowie audio-visueller Diagnostik- und/oder Behandlungsmethoden, welche den Patient\*innen über eine Distanz angeboten werden. Gegenwärtig wird in diesem Sektor des Gesundheitswesens ein enormes Wachstum beobachtet [9]. Die Chancen derartiger Angebote, wie beispielsweise schnelle Verfügbarkeit und Wohnortunabhängigkeit, liegen auf der Hand und werden in der Fachliteratur diskutiert [10].

Seit März 2022 bietet die Debeka-Krankenversicherung ihren Versicherten das sogenannte Programm „InBalance“ an [11]. Dabei handelt es sich um (video-)telefonische Begleitung für Menschen mit depressiven beziehungsweise Angststörungen durch medizinisches Fachpersonal, welche durch den digitalen Gesundheitsdienstleister widicare GmbH, eine Tochterfirma der Debeka, abgedeckt wird [12].

Das Versorgungsmodell ist als niedrigschwelliges beziehungsweise parallel zu fachärztlicher oder psychotherapeutischer Behandlung laufendes Angebot konzipiert und soll genannte Leitlinienversorgung ergänzen, nicht ersetzen. Die vorliegende Studie stellt das Setting des Programmes InBalance sowie die im laufenden Betrieb gewonnenen Evaluationsergebnisse nach einjähriger Laufzeit vor. Dabei greift sie ein Desiderat auf, indem sie außerhalb eines klassischen Studiensettings ein primär nicht-appbasiertes sowie bereits in die Regelversorgung implementiertes, telemedizinisches Angebot für psychische Gesundheit evaluiert.

## Methoden

### Selektionskriterien

Die Anmeldung für das Programm erfolgte freiwillig. Mögliche Teilnehmende wurden durch die Debeka postalisch angeschrieben. Grundlage für diese Anschreiben waren die untenstehenden, durch entsprechende Fachärzte oder Therapeuten ambulant vergebene Diagnosen, die auf eingereichten Rechnungen der Versicherten durch die Debeka einzusehen waren. Zudem gab es eine einmalige Informationskampagne via Newsletter und Werbung auf der Debeka-Website. Voraussetzungen für eine Coachingteilnahme waren neben dem Versichertenstatus und einem Alter zwischen 18 und 80 Jahren folgende Haupt- oder Nebendiagnosen:

- Leichte oder mittelgradige depressive Episode (F32.0 oder F32.1)
- Rezidivierende depressive Störung, gegenwärtig leichte Episode (F33.0)
- Anhaltende affektive Störungen (F34.-)
- Phobische Störungen (F40.-)
- Andere Angststörungen (F41.-)

Ausschlusskriterien für die Programmteilnahme stellten hingegen weitere F-Diagnosen, demenzielle und maligne Erkrankungen, diverse Leber- und Nierenfunktionsstörungen sowie Gehörlosigkeit dar. Ebenfalls ausgeschlossen wurden Menschen mit diagnostizierten kognitiven Funktionseinschränkungen. Für die Einhaltung der entsprechenden Ein- und Ausschlusskriterien war allein die Debeka mit der Vorselektion verantwortlich. Eine erneute Diagnostik innerhalb des Programmes fand nicht statt.

## Programmstruktur und Inhalte

Wesentliches Merkmal des Gesundheitsprogrammes war der regelmäßige (video-)telefonische Kontakt mit den Teilnehmenden. Betreuungsbeginn stellte ein circa fünfzigminütiges Anamnesegespräch dar. Daran anknüpfend stand ein Zeitkontingent von circa 180 Minuten zur Verfügung. Die drei Zeitstunden dienten im Sinne einer Mischkalkulation als Richtwert und konnten auch bei Notwendigkeit deutlich verlängert beziehungsweise mussten bei weniger Bedarf nicht voll ausgenutzt werden. Die Gesprächsabstände und -längen gestalteten sich nach den Bedürfnissen der Teilnehmenden. So waren gleichwohl Kurzkontakte von wenigen Minuten als auch länger dauernde Gespräche ebenso individuell vereinbar wie Zeitabstände zwischen den Terminen von wenigen Tagen bis hin zu mehreren Wochen. Gesprächsinhalte stellten krankheitsspezifische Psychoedukation und leitliniengestützte verhaltenstherapeutische Ansätze dar. Konkret umfasste das über Gespräche, E-Mails, selbst produzierte Videos, Webinare oder Übungsblätter vermittelte Angebot:

- Übungen zur kognitiven Umstrukturierung (u. a. ABC-Modell nach Albert Ellis, Selbstbeobachtungsprotokolle)
- Achtsamkeitsübungen (z. B. Body-Scan, Gehmeditation, Dankbarkeitsübungen, Genusstaining)
- Anleitungen für Reflexionsübungen (u. a. zu persönlichen Ressourcen)
- Entspannungsübungen (z. B. Atemübungen, Imaginationsübungen, Online-Yoga-Sessions)

Darüber hinaus stand die Exploration individueller Krankheitsfaktoren und krankheitsbegünstigender Verhaltensweisen und entlastende Gespräche im Mittelpunkt. Nicht zuletzt konnten bedarfsgerecht persönliche, krankheitsassoziierte Themen besprochen sowie Hilfestellung bei der Suche nach fachärztlichen oder psychotherapeutischen Behandlungsmöglichkeiten vor Ort gegeben werden, indem auf entsprechende Netzwerke der Debeka verwiesen wurde [13]. Zudem bestand die Möglichkeit des Einsatzes digitaler Angebote wie diverser Entspannungssapps (z. B. 7Mind, SonicTonic) oder psychologischer Online-Programme (Selfapy, Novego), auf die parallel zum Telefon-Coaching durch die Teilnehmenden zurückgegriffen werden konnte. Trotz eines möglichen App-Einsatzes stellten die telefonischen Gespräche weiterhin den Schwerpunkt der Betreuung dar.

Letztere erfolgte in der Regel durch eine Bezugsperson. Das Team der Coachenden bestand aus hinsichtlich Programmstrukturierung und -inhalten eigens geschultem medizinischen Fachpersonal, darunter Ärzte, psychologische Psychotherapeutinnen, Psychologinnen (teilweise in psychotherapeutischer Ausbildung) sowie Sozialarbeiterinnen und Arbeitstherapeuten mit psychosozialen Praxiserfahrungen beziehungsweise entsprechenden Wei-

terbildungen. Die konzeptionelle Leitung oblag dabei einem Arzt sowie einer psychologischen Psychotherapeutin verhaltenstherapeutischer Ausrichtung, welche für die Inhalte und Strukturierung des Programmes verantwortlich zeichneten. Ferner war eine Rückbesprechung der Fälle für die Coaches jederzeit in einer regelmäßig stattfindenden multidisziplinären Fallkonferenz möglich. Konsile bis hin zur Übernahme durch ärztliche oder psychotherapeutische Kolleginnen oder Kollegen stellten eine weitere Säule der Qualitätssicherung dar. Gleichwohl die Teilnehmenden durch die Debeka in das Programm eingesteuert wurden, erfolgte keinerlei Übertragung von Gesundheitsdaten an oder von der Versicherung. Weder waren etwaige Vorerkrankungen für die widicare einsehbar, noch wurden Informationen über vertrauliche Coachinginhalte oder Fragebogenergebnisse an die Debeka übermittelt.

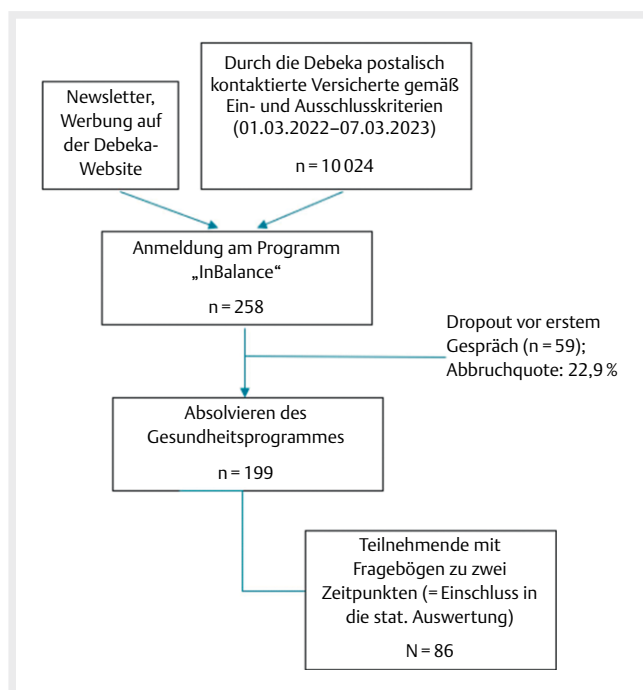
## Datenerhebung und statistische Analyse

Als Messinstrument für die Erfassung der psychischen Belastung der Programmteilnehmenden wurde die deutsche Kurzversion der Depressions-Angst-Stress-Skalen (original: Depression Anxiety Stress Scale), kurz: DASS 21, verwendet [14]. Dieser Fragebogen erfasst die Belastung in den Unterkategorien Depression, Angst und Stress – eben denjenigen Hauptdiagnosen, welche das Gesundheitsprogramm adressiert [15]. Die Unterskalen sind für die genutzte deutsche Version jeweils reliabel mit Werten für Cronbachs Alpha von 0,88 (Depression), 0,76 (Angst) und 0,86 (Stress). Zudem konnte durch die Korrelationstestung mithilfe des Beck-Depressions- bzw. Angst-Inventars für die Depressions- sowie die Angstskala des DASS21 eine ausreichende Konstruktvalidität abgeleitet werden (0,68 für Depression und 0,76 für Angst) [16]. Anhand der Beantwortung von sieben Fragen je Unterkategorie auf einer vierstufigen Likert-Skala von 0 bis 3 wird ein Summenwert zwischen 0 und 21 je Unterkategorie sowie eine Gesamtsumme (0 bis 63) errechnet, wobei die Gesamtsumme als Maß für den Leidensdruck herangezogen werden kann [17]. Die Fragebögen dienten einerseits als Instrument für die anamnestiche Erhebung der psychischen Belastung, andererseits als Grundlage für eine Verlaufsbeurteilung und wurden den Teilnehmenden über eine datengeschützte Plattform für Gesundheitsfragebögen zum Selbstrating zur Verfügung gestellt. Dementsprechend erfolgte das Ausfüllen des DASS 21 zu zwei Zeitpunkten: Erstmalig in Anschluss an das Anamnesegespräch und schließlich fünf Monate nach Programmstart.

Im Zentrum der vorliegenden Arbeit steht die Fragestellung, ob bei Teilnehmenden des telefonbasierten Gesundheitsprogrammes InBalance fünf Monate nach Start des Programmes eine Verringerung der psychischen Belastung erkennbar ist. Als primäre Endpunkte wurden daher die Veränderungen im DASS 21-Gesamtscore sowie in den Unterkategorien Depression, Angst und Stress festgelegt. Deren Verminderung wäre wiederum gleichbedeutend mit einer Reduktion der psychischen Belastung in der jeweiligen Kategorie. Ähnlich wie bei einer einarmigen Interventionsstudie wurden die Mittelwertsunterschiede der Unterkategorien Depression, Angst und Stress sowie des Gesamtscores von Programmstart (t1) zu nach fünf Monaten (t2) durch einen t-Test für abhängige Stichproben analysiert. Zusätzlich zur Betrachtung dieser Mittelwertsdifferenzen über alle Teilnehmenden erfolgte eine gesonderte statistische Analyse für diejenigen Teilnehmenden, welche

► **Tab. 1** Cutoff-Werte für die unterschiedlichen Belastungsgrade der drei Unterkategorien, adaptiert für den DASS 21 ([18], S. 26–28).

	Depression	Angst	Stress
Normal	0–4	0–3	0–7
„Leicht“	5–6	4	8–9
„Mittel“	7–10	5–7	10–12
„Schwer“	11–13	8–9	13–16
„Extrem schwer“	14 und mehr	10 und mehr	17 und mehr



► **Abb. 1** Entwicklung der Inanspruchnahmepopulation.

gemäß Definition der Entwickler des DASS einen initial erhöhten Wert (mindestens leichte Belastung) in den Unterkategorien Depression, Angst und Stress aufwiesen [18] (siehe ► **Tab. 1**). Diese separate Auswertung trägt dem Umstand Rechnung, dass eine Auswertung über alle Teilnehmenden ungenau wäre, da eine Belastung in einer Unterkategorie nicht gleichbedeutend mit einer Punkterhöhung in den übrigen ist. Daraus ergibt sich über alle Teilnehmenden gerechnet eine Verringerung insbesondere des t1-Wertes, wenn auch initial unauffällige beziehungsweise niedrige Werte einfließen. ► **Tab. 1**

Die Berechnungen erfolgten unter Benutzung des Statistik-Programmes GNU pspp, Version 1.6.2-g78a33a. Als Signifikanzniveau der zweiseitigen Tests wurde  $\alpha = 0,05$  gewählt. Um einer bei multiplen Testungen zu erwartenden Alpha-Fehler-Kumulierung vorzubeugen, wurde eine Adjustierung der Signifikanzniveaus mittels Bonferroni-Holm-Korrektur vorgenommen [19]. Die Voraussetzung der Normalverteilung für den t-Test für abhängige Stichproben war in den meisten Fällen gegeben, da das n mit  $> 30$  entweder ausreichend

groß oder für Stichprobengrößen  $< 30$  der Shapiro-Wilk-Test ( $\alpha = 0,05$ ) nicht signifikant geworden war, eine Normalverteilung also angenommen werden konnte (entsprechende rechnerische Nachweise siehe Supplementary 1) [20]. Lediglich die kleine Gruppe von fünf Personen mit „extrem schwerem“ Stress musste aufgrund des signifikanten Shapiro-Wilk-Tests (gleichbedeutend mit mutmaßlicher Verletzung der Normalverteilung) ausgeklammert werden.

## Ergebnisse

### Stichprobe und Programm Kennzahlen

Zwischen Initiierung des Programmes InBalance im März 2022 bis 07. März 2023 wurde 10.024 Versicherten ein Angebotsschreiben unterbreitet. Zwischen 06. April 2022 und 07. März 2023 absolvierten von 258 Anmeldungen 199 das telefonische Gesundheitsprogramm. Für 86 Teilnehmende aus dieser Gruppe lagen zu zwei Zeitpunkten ausgefüllte Fragebögen vor (Quote = 43,2%). ► **Abb. 1**

Von diesen 86 Personen hatten 76 (88,4%) zur zweiten Datenerhebung nach fünf Monaten ihr Programm bereits abgeschlossen, 10 befanden sich noch im Coaching. Das durchschnittliche Alter zu Betreuungsbeginn betrug 49,5 Jahre. Der Großteil (61,6%) war weiblichen Geschlechts. 16 Personen wurden parallel zum Programm ambulant psychotherapeutisch behandelt. 23 befanden sich zu Beginn oder in der Vergangenheit in psychiatrischer beziehungsweise neurologischer fachärztlicher Behandlung. In entsprechender Behandlung befindliche Personen fanden sich in der Auswertung in allen Schweregradeinteilungen wieder (von maximal „normalen“ bis maximal „extrem schweren“ Scores).

Die Laufzeit des Gesundheitsprogrammes betrug für die untersuchte Population im Durchschnitt 170 Tage. Innerhalb dieser Zeit wurden im Mittel 200 Minuten telefoniert, wobei sich diese Zeit auf durchschnittlich 5,6 Telefonate verteilte. Etwas weniger als jede \* r dritte Teilnehmende nutzte im Rahmen des Programmes eine App (0,3 Apps/Teilnahme). Weitere ausführliche demografische Basisdaten finden sich in ► **Tab. 2**.

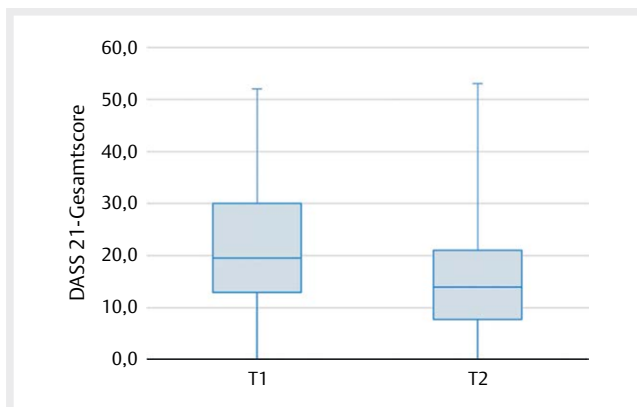
### Fragebogenergebnisse

Legt man die Referenzwerte aus ► **Tab. 1** zugrunde, lässt sich zu Programmbeginn durchschnittlich über alle Teilnehmenden gerechnet eine mittlere Belastung in der Unterkategorie Depression (7,13 [SD = 4,99]), eine leichte bis mittlere Belastung bei Angst (4,49 [4,07]) sowie eine leichte bis mittlere Belastung in der Kategorie Stress (9,85 [4,16]) feststellen. Der durchschnittliche Gesamtscore betrug zu diesem Zeitpunkt 21,47 (11,54). Für alle Unterkategorien zeigte sich nach fünf Monaten im Programm eine signifikante Reduktion der Scores (jeweils  $p < 0,001$ ) um 2,14 (Depression; SD = 3,99), 1,65 (Angst; 3,36) beziehungsweise 2,24 (Stress; 4,81) Punkte. Der Gesamtscore verbesserte sich um 6,03 (10,70;  $p < 0,001$ ) Punkte. Die Effektstärken nach Cohen bewegen sich für diese Veränderungen im mittleren Bereich zwischen 0,47 und 0,56 (siehe ► **Tab. 3**, auch für weitere detaillierte Ergebnisdarstellungen). ► **Abb. 2**

Bei Betrachtung der Ergebnisse für die Unterkategorie Depression, aufgeschlüsselt nach zu Programmbeginn gemessener Belastungsintensität, verändern sich die entsprechenden Kennzahlen. Während für die sehr kleine Gruppe ( $n = 8$ ) der „extrem schwer“ Be-

► **Tab. 2** Übersicht zu den demografischen sowie Telefoniebasisdaten der untersuchten Personengruppe.

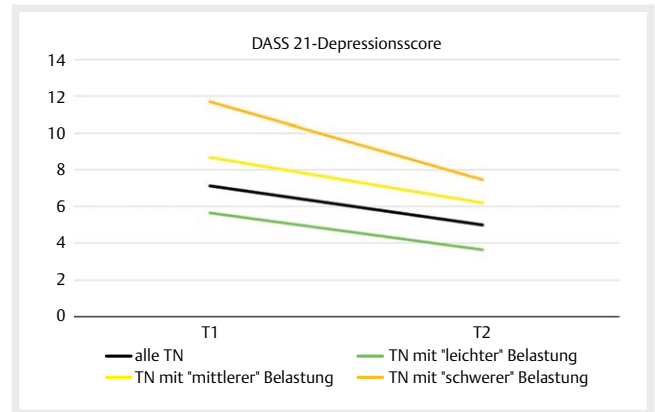
Merkmal	Daten der Teilnehmenden-gruppe (n = 86)
<i>Altersangaben in Jahren</i>	
Mittelwert (Standardabweichung [SD])	49,5 (11,76)
Median	47,5
Minimal- und Maximalalter	25 und 75, Spannweite 50
<i>Geschlechterverteilung, n</i>	
Weiblich	53 (61,6%)
Männlich	33 (38,4%)
<i>(Vor-)Behandlung, n</i>	
Aktuelle ambulante Psychotherapie	16 (18,6%)
Fachärztliche Behandlung aktuell oder in der Vergangenheit	23 (26,7%)
<i>Durchschnittliche Telefoniebasisdaten</i>	
Laufzeit in Tagen (SD); Min. – Max.	170 (58); 41–296
Gesprächsminuten (SD); Min. – Max.	200 (81); 46–355
Anzahl Telefonate (SD); Min. – Max.	5,6 (1,9); 2–10



► **Abb. 2** Verteilung der Gesamtscores des DASS 21 zu t1 und t2 aller Teilnehmenden dargestellt im Boxplot ohne Ausreißer.

lasteten keine Signifikanz gegeben war, stieg die Mittelwertsdifferenz von t1 zu t2 – verstanden als Belastungsreduktion im Verlauf – mit zunehmender initialer Belastung an („leicht“: 2,00 [1,95];  $p = 0,036$ ; „mittel“: 2,48 [2,91];  $p = 0,009$ ; „schwer“: 4,23 [2,80];  $p < 0,001$ ). Die entsprechenden Effektstärken rangieren zwischen 0,85 und 1,51 und somit im Bereich großer Effekte. ► **Abb. 3**

In der Kategorie Angst zeigten sich insgesamt initial weniger Personen belastet nach Definition des Fragebogens. Für die kleinen Gruppen der „leicht“ ( $n = 6$ ) und „schwer“ ( $n = 7$ ) Belasteten ergeben sich in der Folge keine signifikanten Mittelwertsunterschiede. Dem gegenüber stehen deutliche Verbesserungen der größeren Gruppen der „mittleren“ Belastung (2,11 [2,40];  $p = 0,016$ ) und



► **Abb. 3** Veränderung der mittleren DASS 21-Depressionsscores von t1 zu t2 aufgeschlüsselt nach Schwere der Belastung.

„sehr schweren“ Belastung (5,62 [4,96];  $p = 0,028$ ). Die Effektstärken sind mit 0,88 respektive 1,13 als groß einzuordnen. ► **Abb. 4**

Mit Stress „mittel“ und „schwer“ Belastete nach DASS-Definition zeigen ebenfalls signifikante Verbesserungen von t1 zu t2 („mittel“: 3,26 [2,96];  $p < 0,001$ ; „schwer“: 4,39 [5,93];  $p = 0,036$ ) bei mittlerer bis großer Effektstärke. Die Teilnehmenden mit initial „leichter“ Stressbelastung scheinen nicht signifikant zu profitieren. Für die kleine Gruppe ( $n = 5$ ) der „extrem schwer“ Stressbelasteten erfolgte aufgrund der Verletzung der Normalverteilung keine statistische Testung. ► **Abb. 5**, ► **Tab. 3**

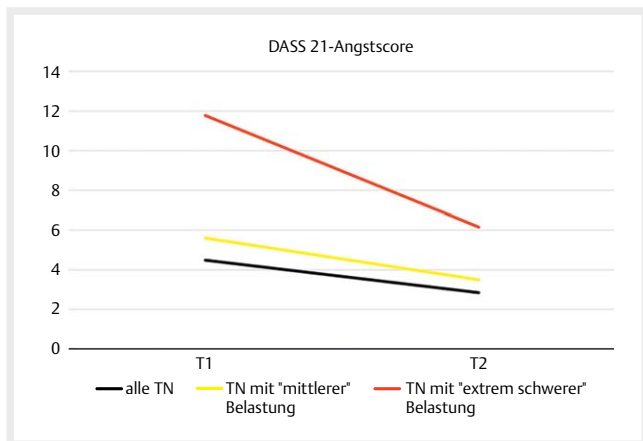
Des Weiteren wurde untersucht, inwiefern sich Teilnehmende mit paralleler psychotherapeutischer beziehungsweise fachärztlicher Behandlung in der initialen Belastungsschwere sowie den Veränderungen zu t2 von der Gruppe ohne Parallel-Behandlung unterschied. Dabei zeigten sich für keine Fragebogenunterskala signifikanten Gruppenunterschiede – weder im initialen Wert zu t1 noch in den Mittelwertsunterschieden zwischen t1 und t2. ► **Tab. 4**

## Diskussion

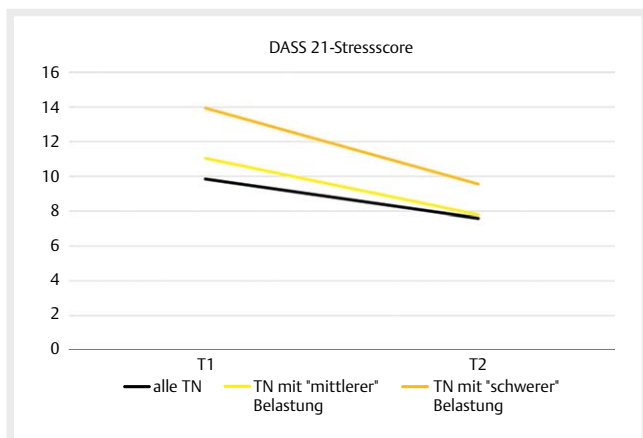
Die vorliegende Arbeit stellt neben dem Modus des Gesundheitsprogrammes erste durch Fragebogen generierte Ergebnisse vor. Im Prä-Post-Vergleich der DASS 21-Scores zeigt sich eine signifikante Verbesserung der psychischen Belastung nach fünfmonatiger Teilnahme am Programm. Diese Verbesserungen lassen sich sowohl für die Gesamtbelastung als auch getrennt für alle Unterkategorien (Depression, Angst, Stress) über alle Teilnehmenden feststellen.

Es ist von Interesse, die durch den Fragebogen vorgegebene Einteilung nach Belastungsschwere in den Kategorien Angst, Stress und Depression in der Auswertung aufzugreifen: Dabei ergeben sich in vier Fällen Gruppen, die bei deutlich eingeschränkter Fallzahl zwischen 5 und 8 entweder keine signifikanten Ergebnisse mehr liefern oder die Testvoraussetzung Normalverteilung verletzen. Die Gruppe der „leicht“ Stressbelasteten scheint trotz vergleichsweise hoher Fallzahl ( $n = 19$ ) nicht zu profitieren. In allen anderen Subgruppen zeigen sich mit zunehmender Belastungsschwere betragsmäßig vergrößerte Score-Reduktionen. Die Effekte fallen mittel bis groß aus und liegen bisweilen bei bemerkenswert gro-





► **Abb. 4** Veränderung der mittleren DASS 21-Angstscores von t1 zu t2 aufgeschlüsselt nach Schwere der Belastung.



► **Abb. 5** Veränderung der mittleren DASS 21-Stressscores von t1 zu t2 aufgeschlüsselt nach Schwere der Belastung.

ßen Werten über 1. Einerseits ist dies mathematisch begründbar, weil Verbesserungen betragsmäßig größer ausfallen können, je größer das Potenzial zur Verbesserung ist. Andererseits sind derartige Effektstärken für Prä-Post-Vergleiche durch die Forschung bereits beobachtet worden [21]. Interessant ist zudem, dass sowohl in initial gemessener Belastung als auch in der erzielten Verbesserung kein signifikanter Unterschied zwischen Teilnehmenden mit und ohne Parallelbehandlung vorzuliegen scheint und es also einen möglichen Hinweis auf eine generelle Nutzbarkeit unabhängig vom Behandlungsstatus gibt.

Aufgrund dessen, dass es sich um ein bereits praktisch implementiertes Telefon-Programm handelt und die Testergebnisse nicht in einem randomisierten und kontrollierten Studiensetting gewonnen wurden, ergeben sich einige Limitationen.

### Limitationen

Eine methodische Schwäche ist in der Akquise der Teilnehmenden zu sehen. Durch die Debeka kann nicht final unterschieden werden kann, wie viele der Teilnehmenden durch die Selektion (und damit mit validierten Diagnosen) und wie viele durch Werbeangebote für

das Programm gewonnen werden konnten. Gleichwohl wird nach interner Statistik der Debeka in der Regel die Mehrzahl durch die Selektion und entsprechende Anschreiben eingesteuert.<sup>1</sup> In jedem Fall galt für alle Teilnahmen der Wunsch nach Unterstützung bei psychischer Belastung – ein Bedarf war demnach als für einen Einschluss in das Coaching ausreichende Voraussetzung gegeben. Die wohl größte Limitation hinsichtlich Validität der vorliegenden Auswertung stellt die fehlende Kontroll-Gruppe dar. Obwohl die dargestellten Ergebnisse mögliche Hinweise für eine positive Wirkung des Programmes sein können, wäre eine Aussage zur tatsächlichen Wirksamkeit verfrüht. Daraus leitet sich ab, dass eine Untersuchung der Intervention unter randomisierten und kontrollierten Bedingungen wünschenswert wäre, möchte man die vorliegenden Ergebnisse bestätigen. Dies wäre insbesondere deshalb wichtig, um die gemessenen Effekte von zeitabhängigen Verbesserungstendenzen abzugrenzen. Denn es ist z. B. bekannt, dass depressive Erkrankungen typischerweise episodisch verlaufen [22]. In diesem Zusammenhang könnten auch mögliche Confounder wie die parallele psychotherapeutische oder fachärztliche Behandlung in der statistischen Betrachtung differenzierter berücksichtigt werden sowie die Validität der eingeschlossenen Diagnosen durch erneute Diagnostik erwogen werden. Eine größere Studienpopulation könnte dabei ebenso die relativ breite Streuung in den Programm-kennzahlen (► **Tab. 2**) ändern, möglicherweise durch geringe Fallzahlen in den Untergruppen nicht erfasste Effekte messen sowie die statistische Power erhöhen. Letztere ist beim Testen kleinerer Probandengruppen beeinträchtigt. Trotz gegebener und bereits dargelegter ausreichender Gütekriterien des DASS 21 ist eine Nutzung gängigerer Fragebögen, wie etwa den Beck-Inventaren, im Kontext weiterführender Forschung zu erwägen.

Die Responderrate von „nur“ 43 % ist am ehesten darauf zurückzuführen, dass Zeitpunkt t2 (fünf Monate nach Start) nicht unbedingt mit dem tatsächlichen Abschluss des Programmes korreliert und ein Ausfüllen der Fragebögen zu zwei Zeitpunkten keine notwendige Voraussetzung für das Absolvieren des Telefonprogrammes war. Die restlichen 57 % durchliefen das Betreuungsangebot dennoch regulär. Es resultiert die Vermutung, dass sich Responder und Non-Responder möglicherweise schon initial unterscheiden, also eine selektive Stichprobe vorliegen könnte. Jedoch konnte bei einer entsprechenden Testung der initialen DASS21-(Unter-)Scores für keine Unterkategorie ein signifikanter Unterschied festgestellt werden (siehe Supplementary 2). Unter kontrollierten Bedingungen und durch Verknüpfung von t2 mit dem Programmende könnte diese Responderquote möglicherweise erhöht werden, gleichwohl die 43 % verglichen mit den anderen Telefonprogrammen des Anbieters (Werte zwischen 20 und 39 %) hoch sind. Eine Erhöhung der Responderquote sowie Fallzahl ist zudem aus Gründen der Repräsentativität anzustreben. Zu beachten ist in diesem Zusammenhang auch, dass die Teilnehmenden des Programmes Privatversicherte sind, deren Nutzungsverhalten neuerer Versorgungsangebote eventuell im Vergleich zur Gesamtbevölkerung alterniert. Dieser Bias ist für weiterführende Untersuchungen zu berücksichtigen.

<sup>1</sup> Nach Debeka-Statistik für das Jahr 2024 waren 78 % der Anmeldungen am Programm infolge der Selektion. Zahlen aus internen Debeka-Statistiken auf Nachfrage.

► **Tab. 3** Ergebnisübersicht über alle Teilnehmenden sowie nach unterschiedlichen Schweregraden. In Klammern stehende Zahlen: zugehörige Standardabweichung (SD). Abweichende Differenzen sind rundungsbedingt.

Kategorie/Merkmal	Alle TN	TN mit „leichter“ Belastung	TN mit „mittlerer“ Belastung	TN mit „schwerer“ Belastung	TN mit „extrem schwerer“ Belastung
<i>Depression, n</i>	86	11	21	13	8
<i>Cutoff</i>	–	≥ 5	≥ 7	≥ 11	≥ 14
T1	7,13 (4,99)	5,64 (0,50)	8,67 (1,06)	11,69 (0,95)	17,50 (2,33)
T2	4,99 (4,50)	3,64 (1,96)	6,19 (2,84)	7,46 (3,10)	10,75 (9,16)
Differenz	2,14 (3,99)	2,00 (1,95)	2,48 (2,91)	4,23 (2,80)	6,75 (8,00)
p-Wert	<0,001	= 0,036	= 0,009	<0,001	= 0,192
Cohen's d	0,54	1,03	0,85	1,51	–
<i>Angst, n</i>	86	6	18	7	13
<i>Cutoff</i>	–	≥ 4	≥ 5	≥ 8	≥ 10
T1	4,49 (4,07)	4 (0,00)	5,61 (0,78)	8,57 (0,53)	11,77 (2,17)
T2	2,84 (3,52)	2,50 (2,26)	3,50 (2,46)	5,71 (4,99)	6,15 (5,67)
Differenz	1,65 (3,36)	1,5 (2,26)	2,11 (2,40)	2,86 (5,01)	5,62 (4,96)
p-Wert	<0,001	= 0,495	= 0,016	= 0,495	= 0,028
Cohen's d	0,49	–	0,88	–	1,13
<i>Stress, n</i>	86	19	19	18	5
<i>Cutoff</i>	–	≥ 8	≥ 10	≥ 13	≥ 17
T1	9,85 (4,16)	8,68 (0,5)	11,05 (0,91)	13,94 (1,11)	
T2	7,60 (4,32)	7,89 (3,02)	7,79 (2,64)	9,56 (5,59)	
Differenz	2,24 (4,81)	0,74 (2,94)	3,26 (2,96)	4,39 (5,93)	
p-Wert	<0,001	= 0,495	<0,001	= 0,036	Normalverteilung verletzt, kein t-Test anwendbar
Cohen's d	0,47	–	1,10	0,74	–
<i>Gesamtscore, n</i>	86				
T1	21,47 (11,54)				
T2	15,43 (10,94)				
Differenz	6,03 (10,70)				
p-Wert	<0,001				
Cohen's d	0,56				

## Schlussfolgerung

Eine abschließende Aussage zu einer Wirksamkeit ist vor dem Hintergrund der dargestellten Limitationen nicht zu treffen. Diese methodischen Einschränkungen folgen vor allem daraus, dass das vorgestellte Programm bereits implementiert ist und sich die Studie auf während der praktischen Anwendung erhobene Evaluationsdaten stützt. Dennoch ergeben sich in der statistischen Auswertung Hinweise auf einen möglichen positiven Effekt dieser Interventionsart als additive Versorgungsstruktur bei psychischen Erkrankungen, denen mithilfe von weiterführenden Studien unter kontrollierten und randomisierten Bedingungen nachgegangen werden sollte. Die Relevanz dieses Ansatzes ist hervorzuheben, da eine (video-)telefonische Betreuung sowohl praktikabel, niedrigschwellig als auch breit verfügbar ist. Vor dem Hintergrund der prekären Versorgungslage für Menschen mit psychischen Erkrankungen ist dies ein wichtiger Punkt. Telefonbasierte Gesundheitspro-

gramme fügen der Palette der alternativen Versorgungsstrukturen eine neue Facette hinzu, beschränken sich Neuerungen auf diesem Gebiet doch zumeist auf Smartphone-Apps oder psychologische Online-Programme, welche automatisiert und zumeist ohne menschlichen Kontakt operieren.

Weiterführend ergibt sich als Ausblick die Frage, ob nicht auch andere Entitäten von einer telemedizinischen beziehungsweise telefonischen Begleitung durch medizinisches Fachpersonal als additive Versorgungsmöglichkeit profitieren könnten. Sowohl ein Einsatz bei anderen psychischen Störungen (zum Beispiel Essstörungen oder Suchterkrankungen) als auch in anderen Fachbereichen, bei denen Gesprächs- und Edukationsbedarf vorhanden sind (etwa Diabetes oder Herz-Kreislauf-Erkrankungen), könnten dahingehend zur Debatte stehen.

► **Tab. 4** Übersicht der durchschnittlichen Fragebogenergebnisse von Teilnehmenden mit Parallelbehandlung verglichen mit Teilnehmenden ohne parallele psychotherapeutische oder ambulant psychiatrische Behandlung. In Klammern stehende Zahlen: zugehörige Standardabweichung (SD). Letzte Spalte: Testung auf Gruppenunterschiede bei gegebener Varianzhomogenität mit T-Test für unabhängige Stichproben.

Kategorie/ Merkmal	Alle TN	TN ohne parallele Behandlung (n = 54)	TN mit psychotherapeutischer und/ oder fachärztlicher Parallel-Behandlung (n = 32)	Signifikanter Gruppenunter- schied?
<i>Depression</i>				
T1	7,13 (4,99)	6,59 (4,86)	8,03 (5,14)	t(84) = 1,30; p = 0,198 kein stat. signifikanter Unterschied
T2	4,99 (4,50)	4,50 (4,15)	5,81 (4,99)	
Differenz	2,14 (3,99)	2,09 (4,31)	2,22 (3,45)	t(84) = 0,14; p = 0,888 kein stat. signifikanter Unterschied
p-Wert	<0,001	= 0,006	= 0,006	
Cohen's d	0,54	0,49	0,64	
<i>Angst</i>				
T1	4,49 (4,07)	4,15 (3,85)	5,06 (4,42)	t(84) = 1,01; p = 0,317 kein stat. signifikanter Unterschied
T2	2,84 (3,52)	2,48 (3,27)	3,44 (3,88)	
Differenz	1,65 (3,36)	1,67 (3,65)	1,63 (2,86)	t(84) = 0,06; p = 0,956 kein stat. signifikanter Unterschied
p-Wert	<0,001	= 0,006	= 0,006	
Cohen's d	0,49	0,46	0,57	
<i>Stress</i>				
T1	9,85 (4,16)	9,63 (4,07)	10,22 (4,35)	t(84) = 0,63; p = 0,528 kein stat. signifikanter Unterschied
T2	7,60 (4,32)	7,00 (4,26)	8,63 (4,30)	
Differenz	2,24 (4,81)	2,63 (4,97)	1,59 (4,52)	t(84) = 0,97; p = 0,337 kein stat. signifikanter Unterschied
p-Wert	<0,001	<0,001	= 0,055	
Cohen's d	0,47	0,53	-	
<i>Gesamtscore</i>				
T1	21,47 (11,54)	20,37 (11,30)	23,31 (11,89)	t(84) = 1,14; p = 0,256 kein stat. signifikanter Unterschied
T2	15,43 (10,94)	13,98 (10,68)	17,88 (11,12)	
Differenz	6,03 (10,70)	6,39 (11,53)	5,44 (9,26)	t(84) = 0,04; p = 0,693 kein stat. signifikanter Unterschied
p-Wert	p < 0,001	<0,001	= 0,006	
Cohen's d	0,56	0,55	0,59	



## Fazit für die Praxis

Niederschwellige telefonbasierte Gesundheitsprogramme können ein praktikabler Betreuungsansatz für Menschen mit depressiven oder Angststörungen sowie Stressbelastung sein.

Ein Einsatz derartiger Angebote, als eigenständiger oder parallel zu anderer Versorgung laufender Ansatz, insbesondere für prekär versorgte Gesundheitsbereiche, sollte diskutiert werden.

Dafür sind weiterführende Wirksamkeitsprüfungen dieses Settings unter kontrollierten und randomisierten Bedingungen vor dem Hintergrund erster positiver Evaluationsergebnisse notwendig und anzustreben.

## Interessenkonflikt

Der Autor, Georg Bornemann, gibt an, nebenberuflich bei der widecare GmbH beschäftigt und in diesem Rahmen mit der Aufgabe der wissenschaftlichen Evaluation der Programme beauftragt zu sein.

## Literatur

- [1] Jachertz N. Psychische Erkrankungen: Hohes Aufkommen, niedrige Behandlungsrate. *Dtsch Arztebl International* 2013; 12/2: 61
- [2] Geiser F, Kleiman A, Albus C et al. Angststörungen. *Internist* 2012; 53: 1289–1295. DOI: 10.1007/s00108-012-3069-7
- [3] Gutiérrez-Rojas L, Porras-Segovia A et al. Prevalence and correlates of major depressive disorder: a systematic review. *Braz J Psychiatry* 2020; 42: 657–672. DOI: 10.1590/1516-4446-2020-0650
- [4] Lim GY, Tam WW, Lu Y et al. Prevalence of Depression in the Community from 30 Countries between 1994 and 2014. *Sci Rep* 2018; 8. DOI: 10.1038/s41598-018-21243-x
- [5] Bundesärztekammer (BÄK), Kassenärztliche Bundesvereinigung (KBV), Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF). Nationale VersorgungsLeitlinie Unipolare Depression – Langfassung, Version 3.1. 2022. DOI: 10.6101/AZQ/000496
- [6] Bandelow B, Aden I, Alpers G et al. Deutsche S3-Leitlinie Behandlung von Angststörungen, Version 2 (2021). Im Internet: [https://register.awmf.org/assets/guidelines/051-028k\\_S3\\_Behandlung-von-Angststoerungen\\_2021-06.pdf](https://register.awmf.org/assets/guidelines/051-028k_S3_Behandlung-von-Angststoerungen_2021-06.pdf); stand: 13.07.2023
- [7] Bundespsychotherapeutenkammer. Ein Jahr nach der Reform der Psychotherapie-Richtlinie. *Wartezeiten* 2018. Im Internet: [https://www.bptk.de/wp-content/uploads/2019/01/20180411\\_bptk\\_studie\\_wartezeiten\\_2018.pdf](https://www.bptk.de/wp-content/uploads/2019/01/20180411_bptk_studie_wartezeiten_2018.pdf); Stand: 13.07.2023
- [8] Sauer mann S, Herzberg J, Burkert S et al. DiGA – A Chance for the German Healthcare System. *J Eur CME* 2021; 11. DOI: 10.1080/21614083.2021.2014047
- [9] Waller M, Stotler C. Telemedicine: a Primer. *Curr Allergy Asthma Rep* 2018; 18. DOI: 10.1007/s11882-018-0808-4
- [10] Hautzinger M, Fuhr K. Kann die Online-Therapie die Psychotherapie sinnvoll ergänzen? *Pro. Nervenarzt* 2018; 89: 94–95. DOI: 10.1007/s00115-017-0379-y
- [11] InBalance – Versorgungsprogramm bei psychischen Belastungen. Im Internet: <https://www.debeka.de/service/progesundheit/psyche/inbalance---versorgungsprogramm-bei-psychischen-belastungen.html>; Stand: 13.07.2023
- [12] widecare GmbH – Über uns. Im Internet: <https://www.widecare.de/ueber-uns>; Stand: 13.07.2023
- [13] Wegweiser Debeka – Terminvermittlung für ambulante Psychotherapie. Im Internet: <https://www.debeka.de/service/progesundheit/psyche/terminvermittlung-ambulante-psychotherapie.html>; Stand: 08.08.2024
- [14] Lovibond PF, Lovibond SH. The structure of negative emotional states: comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. *Behav Res Ther* 1995; 33: 335–343. DOI: 10.1016/0005-7967(94)00075-u
- [15] Henry JD, Crawford JR. The short-form version of the Depression Anxiety Stress Scales (DASS-21): construct validity and normative data in a large non-clinical sample. *Br J Clin Psychol* 2005; 44: 227–239. DOI: 10.1348/014466505X29657
- [16] Nilges P, Essau C. Die Depressions-Angst-Stress-Skalen. *Der DASS – ein Screeningverfahren nicht nur für Schmerzpatienten.* *Schmerz* 2015; 29: 649–657
- [17] Sinclair SJ, Siefert CJ, Slavin-Mulford JM et al. Psychometric evaluation and normative data for the depression, anxiety, and stress scales-21 (DASS-21) in a nonclinical sample of U.S. adults. *Eval Health Prof* 2012; 35: 259–279. DOI: 10.1177/0163278711424282
- [18] Lovibond SH, Lovibond PF. *Manual for the Depression Anxiety Stress Scales.* 2. Aufl. Sydney: Psychology Foundation; 1995
- [19] McLaughlin MJ, Sainani KL, Bonferroni, Holm, and Hochberg corrections: fun names, serious changes to p values. *PM&R* 2014; 6: 544–546. DOI: 10.1016/j.pmrj.2014.04.006.
- [20] Stone ER. t Test, Paired Samples. In: Salkind NJ, Hrsg. *Encyclopedia of research design.* Los Angeles: SAGE; 1560–1565
- [21] Eccles M, Grimshaw J, Campbell M et al. Research designs for studies evaluating the effectiveness of change and improvement strategies. *Qual Saf Health Care* 2003; 12: 47–52. DOI: 10.1136/qhc.12.1.47
- [22] Whiteford HA, Harris MG, McKeon G et al. Estimating remission from untreated major depression: A systematic review and meta-analysis. *Psychol Med* 2013; 43: 1569–1585. DOI: 10.1017/S0033291712001717