

Evaluationsbericht

(gemäß Nr. 14.1 ANBest-IF)



Konsortialführung:	Universitätsklinikum Heidelberg
Förderkennzeichen:	01NVF17034
Akronym:	Mind:Pregnancy
Projekttitel:	Online-basierte Achtsamkeitsintervention zur Stärkung psychischer Stabilität von Schwangeren und Förderung einer physiologischen Geburt
Autoren:	Prof. Dr. Stephanie Wallwiener, Prof. Dr. Wolfgang Greiner, MPH Lena Hasemann, Dr. Mitho Müller
Förderzeitraum:	1. Juli 2018 – 30. Juni 2022

Das dieser Veröffentlichung zugrundeliegende Projekt wurde mit Mitteln des Innovationsausschusses beim Gemeinsamen Bundesausschuss unter dem Förderkennzeichen 01NVF17034 gefördert.

Inhaltsverzeichnis

I. Abkürzungsverzeichnis.....	2
II. Abbildungsverzeichnis.....	4
III. Tabellenverzeichnis.....	4
1. Ziele der Evaluation.....	7
1.1 Wirksamkeitsevaluation.....	7
1.2 Gesundheitsökonomische Evaluation (Universität Bielefeld).....	10
2. Darstellung des Evaluationsdesigns.....	10
2.1 Entwicklung der App „mindmom“.....	10
2.2 Wirksamkeitsevaluation.....	11
2.3 Gesundheitsökonomische Evaluation (Universität Bielefeld).....	20
2.3.1 Fragestellung und Arbeitshypothese.....	20
2.3.2 Datengrundlage und -aufbereitung.....	20
2.4 Statistische Auswertung.....	23
3. Ergebnisse der Evaluation.....	24
3.1 Wirksamkeitsevaluation (LMU).....	24
3.2 Gesundheitsökonomische Evaluation (Universität Bielefeld).....	52
3.2.1 GKV-Perspektive.....	54
3.2.2 Gesellschaftliche Perspektive.....	58
3.2.3 Ergänzende Analysen.....	60
4. Schlussfolgerungen und Empfehlungen des Evaluators.....	64
4.1 Wirksamkeitsevaluation (LMU).....	65
4.2 Gesundheitsökonomische Evaluation (Universität Bielefeld).....	68

Akronym: Mind: Pregnancy
Förderkennzeichen: 01NVF17034

5. Literaturverzeichnis.....	70
6. Anhang	76
7. Anlagen.....	76

I. Abkürzungsverzeichnis

AIC	Akaike Information Criterion
ANOVA	Analysis of Variance
APA	American Psychiatric Association
ATC	Anatomisch-Therapeutisch-Chemischen Klassifikation
AU	Arbeitsunfähigkeit
AUC _G	Area under the Curve with respect to the Ground
BfArM	Bundesinstitut für Arzneimittel und Medizinprodukte
CI	Konfidenzintervall
CLPM	Cross-lagged-panel-model
DDD	Defined Daily Dose
df	Degrees of freedom
DiGA	Digitale Gesundheitsanwendungen
DSM-IV	Diagnostical and Statistical Manual for Mental Disorders – 4 th revision
DSM-5	Diagnostical and Statistical Manual for Mental Disorders – 5 th revision
eCRF	Electronic case report form
eMBI	Electronic mindfulness based intervention
EPDS	Edinburgh Postpartum Depression Scale
FFA-14	Freiburger Fragebogen zur Achtsamkeit - Kurzversion
GKV	Gesetzliche Krankenversicherung
HSA	Hochschulambulanzen
ICER	Inkrementellen Kosten-Nutzen Relation
ICH	International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use
IG	Interventionsgruppe
ITT	Intention to treat
KG	Kontrollgruppe
LOS	Length of stay
M	Mittelwert
Max	Maximum
MCAR	Missing-Completely-At-Random
MFAS	Maternal-Fetal-Attachment-Scale
Min	Minimum
NPC	Ohne persönliches Coaching
PBQ-16	Postpartum-Bonding-Questionnaire – Revision
PC	Mit persönlichem Coaching
PHQ-D	Patient Health Questionnaire - Deutsch
PIA	Psychiatrische Institutsambulanzen

PEQ-G	Patients' Experience Questionnaire - Geburt
PRAQ-R	Pregnancy-Related-Anxiety-Questionnaire – Revision
QALY	Quality adjusted life years
RCT	Randomized controlled trial
SD	Standardabweichung
S.E.	Standardfehler
SGB	Sozialgesetzbuch
STAI	State-Trait-Anxiety-Inventory
TAU	Treatment as usual
WidO	Wissenschaftliches Institut der AOK

II. Abbildungsverzeichnis

Abbildung 1: CONSORT-Flussdiagramm	24
Abbildung 2: FFA-14-Mittel nach Interventionsgruppierung im Studienverlauf	30
Abbildung 3: EPDS-Mittel nach Interventionsgruppierung im Studienverlauf	31
Abbildung 4: STAI-S-Mittel nach Interventionsgruppierung im Studienverlauf	32
Abbildung 5: STAI-T-Mittel nach Interventionsgruppierung im Studienverlauf	32
Abbildung 6: PRAQ-R-Mittel nach Interventionsgruppierung im Studienverlauf	33
Abbildung 7: Flow-Diagramm Studienpopulation der gesundheitsökonomischen Evaluation	53
Abbildung 8: Anteile der mittleren leistungsbereichsspezifischen Kosten an den durchschnittlichen Gesamtkosten pro Studienteilnehmerin der IG und KG (ohne Interventionskosten)	64

III. Tabellenverzeichnis

Tabelle 1: Übersicht über Hypothesen samt fokussierter Messinstrumente	8
Tabelle 2: Ein- und Ausschlusskriterien der Studie	13
Tabelle 3: Anzahl der Gespräche (inkl. Erstgespräche) pro Gruppe	14
Tabelle 4: Studienablauf und Erhebungsinstrumente.....	15
Tabelle 5: Analyseverfahren.....	18
Tabelle 6: Grundlage der Berechnung der Produktivitätsverluste.....	22
Tabelle 7: Übersicht über die Komponenten der neuen Versorgungsform.....	22
Tabelle 8: Patientenstatus - Vergleich Studienzentren	25
Tabelle 9: Deskriptive Statistiken und Tests auf Vergleichbarkeit der Subgruppen	26
Tabelle 10: FFA-14-Mittel nach Interventionsgruppierung im Studienverlauf	29
Tabelle 11: PRAQ-R-Mittel nach Interventionsgruppierung im Studienverlauf.....	33

Tabelle 12: MFAS-Mittel nach Interventionsgruppierung im Studienverlauf.....	36
Tabelle 13: Zellbesetzung der Subgruppen in der Moderationsanalyse für den Faktor "Klinische Diagnose" mit der abhängigen Variable FFA-14	38
Tabelle 14: FFA-14-Mittel nach Interventionsgruppierung im Studienverlauf.....	38
Tabelle 15: Zellbesetzung der Subgruppen in der Moderationsanalyse für den Faktor "Klinische Diagnose" mit der abhängigen Variable EPDS	39
Tabelle 16: EPDS-Mittel nach Interventionsgruppierung im Studienverlauf.....	39
Tabelle 17: Zellbesetzung der Subgruppen in der Moderationsanalyse für den Faktor "Klinische Diagnose" mit der abhängigen Variable STAI.....	40
Tabelle 18: Zellbesetzung der Subgruppen in der Moderationsanalyse für den Faktor "Klinische Diagnose" mit der abhängigen Variable PRAQ-R.....	41
Tabelle 19: PRAQ-Mittel nach Interventionsgruppierung im Studienverlauf.....	41
Tabelle 20: Zellbesetzung der Subgruppen in der Moderationsanalyse für den Faktor "Klinische Diagnose" mit der abhängigen Variable FFA-14	42
Tabelle 21: FFA-14-Mittel nach Interventionsgruppierung und Coaching-Gruppierung im Studienverlauf.....	43
Tabelle 22: Zellbesetzung der Subgruppen in der Moderationsanalyse für den Faktor "Klinische Diagnose" mit der abhängigen Variable FFA-14	43
Tabelle 23: EPDS-Mittel nach Interventionsgruppierung und Coaching-Gruppierung im Studienverlauf	44
Tabelle 24: Zellbesetzung der Subgruppen in der Moderationsanalyse für den Faktor "Klinische Diagnose" mit der abhängigen Variable STAI.....	45
Tabelle 25: STAI-S-Mittel nach Interventionsgruppierung und Coaching-Gruppierung im Studienverlauf	46
Tabelle 26: STAI-T-Mittel nach Interventionsgruppierung und Coaching-Gruppierung im Studienverlauf	47
Tabelle 27: Zellbesetzung der Subgruppen in der Moderationsanalyse für den Faktor "Klinische Diagnose" mit der abhängigen Variable PRAQ-R.....	49
Tabelle 28: FFA-14-Mittel nach Interventionsgruppierung im Studienverlauf.....	49
Tabelle 29: Imputierte FFA-14-Mittel nach Interventionsgruppierung im Studienverlauf.....	50
Tabelle 30: Imputierte PRAQ-R-Mittel nach Interventionsgruppierung im Studienverlauf.....	51
Tabelle 31: Baselinecharakteristika der Studienpopulation	53
Tabelle 32: Inanspruchnahme und Kosten der ambulanten Versorgung von IG und KG	54
Tabelle 33: Inanspruchnahme und Kosten der stationären Versorgung von IG und KG	54
Tabelle 34: Inanspruchnahme und Kosten von Arzneimitteln in IG und KG.....	55
Tabelle 35: Kosten der neuen Versorgungsform pro Versicherte.....	55

Tabelle 36: Übersicht der Inanspruchnahme von Gesundheitsleistungen in IG und KG	56
Tabelle 37: Übersicht der Kosten der Gesundheitsleistungen in IG und KG	57
Tabelle 38: Generalisiertes additives Modell mit inverser Gammaverteilung: Gesamtkosten aus GKV-Perspektive (N=236)	58
Tabelle 39: Arbeitsunfähigkeit und damit verbundene Produktivitätsverluste	59
Tabelle 40: Generalisiertes additives Modell mit inverser Gammaverteilung: Gesamtkosten aus gesellschaftlicher Perspektive (N=236)	59
Tabelle 41: Generalisiertes additives Modell mit negativer Binomialverteilung: Ambulante Arztkontakte (N=236)	60
Tabelle 42: Inanspruchnahme und Kosten der haus- und fachärztlichen Versorgung von IG und KG .	61
Tabelle 43: Generalisiertes additives Modell mit negativer Binomialverteilung: Hausarztkontakte (N=235)	62
Tabelle 44: Erkrankungsspezifische Medikation (ATC N05/N06) (nein/ja) nach Studiengruppe (IG vs. KG)	63

1. Ziele der Evaluation

Zu den häufigsten psychischen Erkrankungen, auch während einer Schwangerschaft, gehören depressive Störungen und Angsterkrankungen. Entsprechend der Veröffentlichungen von Reck et al. aus 2008 und Gawlik et al. aus 2013 sind 10-20 % der Frauen von einer Depression betroffen. Je nach Stichprobenszusammensetzung der Lebenszeitprävalenzen weisen zwischen 13 % und 25 % der Frauen (Reck et al. 2008) Angststörungen auf. Auch kann gesagt werden, dass die Geburtsangst stetig zunimmt – 17 % der werdenden Mütter geben moderate Angst und schwere Angst 6 % vor der Geburt an (Areskog et al. 1981). Die Ängste sind vielfältig und beziehen sich z. B. auf Geburtsverletzungen oder die Gesundheit des Kindes. Aus der aktuellen Datenlage geht hervor, dass 80 % aller Kaiserschnitte aufgrund relativer Indikationen und 10 % auf Angst entfallen. Für die Schwangere, das ungeborene Kind und die Familie stellen affektive Störungen eine große Bürde, für die Kostenträger gesundheitsökonomische Aspekte dar.

Aktuell werden die Schwangeren, auch bei zutreffender Diagnose „psychische Belastung“ nur schwer von Hilfsangeboten im Zusammenhang der Geburt erreicht. Im Mutterpass ist im Rahmen der Regelversorgung der Eintrag „psychische Belastung“ vorgesehen. Dieser ist empirisch nicht unterfüttert. Der Geburtenreport der Techniker Krankenkasse 2017 hat für alle betrachteten Diagnosegruppen gezeigt, dass wenn bei den Schwangeren eine psychische Erkrankung diagnostiziert wurde, häufiger eine Geburt per Sectio vorlag.

Khoury et al. 2013 konnte einen vielversprechenden Ansatz in der Behandlung von Depressions-, Angst- und Stresssymptomen durch achtsamkeitsbasierte- bzw. „Mindfulness-based-Interventionen“ (MBI) zeigen. Die Form der angeleiteten Selbsthilfe ist vergleichbar oder sogar überlegen zur klassischen Psychotherapie, was verschiedenste Studien belegen (Klainin-Yobas et al. 2012).

Hauptziel von Mind:Pregnancy war es, die Depressions- und Angstsymptome bei Schwangeren mit erhöhter Symptomatik signifikant zu reduzieren. Mithilfe einer Web-/ Mobile-App wurde den Schwangeren über einen Zeitraum von acht Wochen, wöchentliche, online-basierte Achtsamkeitsinterventionen im dritten Trimenon zur Verfügung gestellt. Zu den weiteren Zielen zählte die Senkung der Kaiserschnitttrate und die Senkung der peripartalen Depressionen.

1.1 Wirksamkeitsevaluation

Ziel der Wirksamkeitsevaluation war es, die klinische Effektivität einer online-basierten Achtsamkeitsintervention im dritten Trimenon der Schwangerschaft auf die Symptombelastung bei in einem klinischen Screening auffälligen Schwangeren zu untersuchen.

Im dritten Trimenon wurden die schwangeren Frauen gemäß eines der Kernelemente der neuen Versorgungsform auf klinische Auffälligkeit gescreent. Hierzu wurde die *Edinburgh Postnatal Depression Scale* (EPDS; Bergant et al., 1998) verwendet. Frauen, die einen EPDS-Wert > 9 erreichten, wurden als klinisch auffällig betrachtet und ihnen wurde nach einem psychologischem Erstgespräch/Coaching durch die koordinierende interdisziplinäre Versorgungsstelle die Teilnahme an der Interventionsstudie angeboten.

Überprüft werden sollte, ob und wie stark die online-basierte Achtsamkeitsintervention tatsächlich zur Besserung (z.B. niedrigere Prävalenz peripartaler Depression) beiträgt und ob die gemessene Wirkung auf die spezifische Intervention (eMBI) zurückzuführen ist.

Hierbei wurden folgende inhaltliche Hypothesen geprüft: Die Intervention führt bei betroffenen Frauen zu

1. **Sinkender Symptombelastung bzw. einer niedrigeren Prävalenz peripartaler Depressionen** (siehe konfirmatorische Hypothesen 1 – 6 sowie exploratorische Hypothese 17 des Evaluationskonzepts)
2. **Sinkenden Niveaus neuroendokriner Stressmarker** (siehe exploratorische Hypothesen 1 – 3 des Evaluationskonzepts)
3. **Steigender Lebenszufriedenheit** (siehe exploratorische Hypothesen 4 – 6 des Evaluationskonzepts)

4. **Subjektiv positiverem Erleben der Geburt und der Bindung zum Säugling** (siehe exploratorische Hypothesen 7 – 13 des Evaluationskonzepts)
5. **Einer Verringerung der Häufigkeit von primären und sekundären Kaiserschnitten sowie Kinderarztbesuchen** (siehe exploratorische Hypothese 17 des Evaluationskonzepts)

Im Detail sind alle konfirmatorischen und exploratorischen Hypothesen samt jeweils fokussierter Messinstrumente in Tabelle 1 aufgeführt

Tabelle 1: Übersicht über Hypothesen samt fokussierter Messinstrumente

Hypothesenart	Nr.	Hypotheseninhalt	Instrumente
Konfirmatorisch	1	Die Probandinnen der eMBI- weisen verglichen mit Probandinnen der TAU-Gruppe zu den prä- und postpartalen Messzeitpunkten nach Beginn der Intervention (V2 – V7) eine signifikant niedrigere Depressivität und Ängstlichkeit auf.	EPDS & STAI
Konfirmatorisch	2	Bei Probandinnen der eMBI sinken Depressivität und Ängstlichkeit kontinuierlich signifikant zwischen den präpartalen Messzeitpunkten V1 bis V5 und bleiben zu den postpartalen Messzeitpunkten V6 und V7 auf konstant niedrigem Niveau.	EPDS & STAI
Konfirmatorisch	3	Bei Probandinnen der TAU bleiben Depressivität und Ängstlichkeit zu den prä- und postpartalen Messzeitpunkten V1 bis V7 auf konstant hohem Niveau.	EPDS & STAI
Konfirmatorisch	4	Die Probandinnen der eMBI- weisen verglichen mit Probandinnen der TAU-Gruppe zu den präpartalen Messzeitpunkten nach Beginn der Intervention (V2 – V5) eine signifikant niedrigere schwangerschafts- und geburtsspezifische Ängstlichkeit auf.	PRAQ-R
Konfirmatorisch	5	Bei Probandinnen der eMBI sinkt die schwangerschafts- und geburtsspezifische Ängstlichkeit kontinuierlich signifikant zwischen den präpartalen Messzeitpunkten V1 bis V5.	PRAQ-R
Konfirmatorisch	6	Bei Probandinnen der TAU bleibt die schwangerschafts- und geburtsspezifische Ängstlichkeit zu den präpartalen Messzeitpunkten V1 bis V5 auf konstant hohem Niveau.	PRAQ-R
Exploratorisch	1	Die Probandinnen der eMBI- weisen verglichen mit Probandinnen der TAU-Gruppe zu den postpartalen Messzeitpunkten nach Beginn der Intervention (V5 & V7) eine signifikant niedrigere neuroendokrine Stressbelastung auf.	Cortisol (AUC _G)

Exploratorisch	2	Bei Probandinnen der eMBI sinkt die neuroendokrine Stressbelastung kontinuierlich signifikant zwischen den prä- und postpartalen Messzeitpunkten V1, V5 und V7.	Cortisol (AUC _G)
Exploratorisch	3	Bei Probandinnen der TAU bleibt die neuroendokrine Stressbelastung zu den Messzeitpunkten prä- und postpartalen Messzeitpunkten V1, V5 und V7 auf konstant hohem Niveau.	Cortisol (AUC _G)
Exploratorisch	4	Die Probandinnen der eMBI- weisen verglichen mit Probandinnen der TAU-Gruppe zu den präpartalen Messzeitpunkten nach Beginn der Intervention (V3 & V5) eine signifikant höhere gesundheitsbezogene Lebensqualität auf.	EQ-5D
Exploratorisch	5	Bei Probandinnen der eMBI steigt die gesundheitsbezogene Lebensqualität kontinuierlich signifikant zwischen den präpartalen Messzeitpunkten V1, V3 und V5.	EQ-5D
Exploratorisch	6	Bei Probandinnen der TAU bleibt die gesundheitsbezogene Lebensqualität zu den präpartalen Messzeitpunkten V1, V3 und V5 auf konstant niedrigerem Niveau.	EQ-5D
Exploratorisch	7	Die Probandinnen der eMBI- weisen verglichen mit Probandinnen der TAU-Gruppe zu den präpartalen Messzeitpunkten nach Beginn der Intervention (V3 & V5) eine signifikant höheres emotionale Bindung ggü. dem ungeborenen Kind während der Schwangerschaft auf.	MFAS
Exploratorisch	8	Bei Probandinnen der eMBI steigt die emotionale Bindung ggü. dem ungeborenen Kind während der Schwangerschaft kontinuierlich signifikant zwischen den präpartalen Messzeitpunkten V1, V3 und V5.	MFAS
Exploratorisch	9	Bei Probandinnen der TAU bleibt die emotionale Bindung ggü. dem ungeborenen Kind während der Schwangerschaft zu den präpartalen Messzeitpunkten V1, V3 und V5 auf konstant niedrigem Niveau.	MFAS
Exploratorisch	10	Die Probandinnen der eMBI- weisen verglichen mit Probandinnen der TAU-Gruppe zu den postpartalen Messzeitpunkten (V6 & V7) eine signifikant niedrigere Beeinträchtigung der postpartalen Bindung ggü. dem Neugeborenen auf.	PBQ-16
Exploratorisch	11	Bei Probandinnen der eMBI sinkt die Beeinträchtigung der postpartalen Bindung ggü. dem Neugeborenen signifikant zwischen den postpartalen Messzeitpunkten V6 und V7 ab.	PBQ-16
Exploratorisch	12	Bei Probandinnen der TAU bleibt die Beeinträchtigung der postpartalen Bindung ggü. dem Neugeborenen zu den postpartalen Messzeitpunkten V6 und V7 auf konstant hohem Niveau.	PBQ-16

Exploratorisch	13	Die Probandinnen der eMBI- weisen verglichen mit Probandinnen der TAU-Gruppe postpartal zu V6 eine signifikant positivere Patientenerfahrung auf.	PEQ-G
Exploratorisch	14	Die Probandinnen der eMBI- weisen verglichen mit Probandinnen der TAU-Gruppe zu den prä- und postpartalen Messzeitpunkten nach Beginn der Intervention (V5 & V7) eine signifikant höhere Achtsamkeit auf.	FFA-14
Exploratorisch	15	Bei Probandinnen der eMBI steigt die Achtsamkeit kontinuierlich signifikant zwischen den prä- und postpartalen Messzeitpunkten V1, V5 und V7.	FFA-14
Exploratorisch	16	Bei Probandinnen der TAU bleibt die Achtsamkeit zu den prä- und postpartalen Messzeitpunkten V1, V5 und V7 auf konstant niedrigerem Niveau.	FFA-14
Exploratorisch	17	Die Probandinnen der eMBI- weisen verglichen mit Probandinnen der TAU-Gruppe postpartal zu T6 eine signifikant geringere Häufigkeit primärer/sekundärer Kaiserschnitte, postpartal zu V7 eine signifikant geringere Häufigkeit der Kinderarztbesuche sowie zu allen prä- und postpartalen Messzeitpunkten nach Interventionsbeginn (V2 – V7) eine signifikant niedrigere Häufigkeit postpartaler Depressionen auf.	Geburtsmodus, Neonatologie, EPDS Cutoff
<p><i>Notizen.</i> eMBI = electronic mindfulness-based intervention; TAU = Kontrollgruppe; Messzeitpunkte V1: 28. Schwangerschaftswoche (SSW), V2: 30. SSW; V3: 32. SSW; V4: 34. SSW; V5: 36. SSW; V6: 1 Monat pp; V7: 5 Monate pp; AUC_G = Area under the Curve with respect to the ground</p>			

1.2 Gesundheitsökonomische Evaluation (Universität Bielefeld)

Das Ziel der gesundheitsökonomischen Evaluation der randomisiert-kontrollierten Studie (RCT) besteht darin, die Inanspruchnahme von Leistungen der gesetzlichen Krankenkasse (GKV) sowie die dadurch erzeugten Kosten der beiden Studiengruppen (Interventionsgruppe (IG)) und Kontrollgruppe (KG)) abzubilden und zu vergleichen.

2. Darstellung des Evaluationsdesigns

2.1 Entwicklung der App „mindmom“

Die App mindmom wurde den Studienteilnehmerinnen innerhalb der klinischen Laufzeit zur Verfügung gestellt. Sie konnte im App- (iOS) bzw. Google Play-Store (Android) heruntergeladen werden. Das Hosting endete mit der Nachbeobachtungszeit im Juni 2022. Seitdem ist die App nicht mehr im App Store oder bei Google Play erhältlich und kann dem Förderer zur Verfügung gestellt werden. Ein Einsatz in der Versorgung ist jedoch mit Vorbehalt zu sehen und Urheberrechte der einzelnen Elemente müssten vor dem Einsatz außerhalb von Studienbedingungen geklärt werden. Zudem wäre eine Zertifizierung als Medizinprodukt notwendig. Das Codebook ist diesem Bericht beigefügt. Alle Fragebögen, die über die App im Rahmen der Studie abgefragt worden, stellen standardisierte

Fragebögen dar (siehe 3.2) und sind keine Neuentwicklung im Rahmen der Studie. Die Inhalte der App sind ausführlich unter 4.1.1 dargestellt.

2.2 Wirksamkeitsevaluation

Studiendesign

Das Projekt wurde im Rahmen einer prospektiven randomisierten kontrollierten Studie an zwei Zentren umgesetzt. Hierbei wurden klinisch auffällig gescreente Probandinnen, die die spezifische Intervention (Interventionsgruppe; IG) erhielten, verglichen mit klinisch auffällig gescreenten Probandinnen, die ausschließlich eine unspezifische Intervention erhielten (Kontrollgruppe; KG). Die weiteren Ein- und Ausschlusskriterien (neben der klinischen Auffälligkeit) sind in dargestellt. Die Intervention bestand im Zugang und der Teilnahme an einem 8-wöchentlichen Modul. Im Anschluss an jede zweite Sitzung der eMBI (*electronic mindfulness-based intervention*) erfolgte eine Erhebung der psychometrischen Variablen.

Anpassung des Evaluationsplans im Studienverlauf

Entsprechend des deskriptiven Monitorings vom September 2019 war festzustellen, dass die ursprünglich geplanten Fallzahlen zum Abschluss des Versorgungs- und Evaluationsprojekts sowohl hinsichtlich der *Screenings* ($N = 14.686$ bis Q1 2020), der Fälle in den koordinierenden Stellen ($N = 1.469$ bis Q2 2020) als auch bzgl. der Studieneinschlüsse ($N = 1.028$ bis Q2 2020) nicht erreicht werden konnten. Bis zum Q2 2019 wurden $N = 1.100$ *Screenings* durchgeführt (18.7 % gemäß Plan), $N = 241$ Fälle in den koordinierenden Stellen verarbeitet (49.3 % gemäß Plan) und $N = 103$ Fälle in die Evaluation eingeschlossen (32.2 % gemäß Plan).

Die im Vergleich zum Plan niedrigere Zahl an Fällen in den koordinierenden Stellen und Studieneinschlüssen war auf die im Vergleich zum Plan niedrigere Zahl an Screenings und darauf zurückzuführen, dass zwar mit insgesamt ca. 21.5 % weitaus mehr Fälle als auffällig (EPDS > 9) gescreent wurden als erwartet (ca. 7%), allerdings aufgrund von Fallausschlüssen nur ca. 6.8% der Screenings in die Studie eingeschlossen werden konnten. Darüber hinaus gingen mit Dropout-Raten von ca. 26.6% mehr als die erwarteten 20% an Ausschlüssen, Abbrüchen und Dropouts verloren.

Bereits in Version 1.0 des Evaluationskonzepts waren die inhaltlichen konfirmatorischen Hypothesen auf die Wechselwirkung zwischen dem *Between*-(Gruppe) und *Within-Subject-Factor* (Messzeitpunkt) bezogen (die Symptombelastung in den Vergleichsgruppen entwickelt sich über die Messzeitpunkte hinweg unterschiedlich). Die Fallzahlschätzung hingegen wurde jedoch auf kleine Effekte des *Between-Subject-Factor* bezogen, um jene mit hinreichender Sicherheit ausschließen zu können. Eine Fallzahlschätzung bezogen auf die benannte Wechselwirkung führt zu dem Ergebnis, dass die Evaluation der konfirmatorischen Hypothesen mit $N = 196$ Fällen gesichert wäre. Mit einer konservativer angelegten Dropout-Rate von 30% müssten demnach $N = 280$ Falleinschlüsse erfolgen. Mit einer Stichprobe von $N = 196$ können große und mittlere Effekte auf dem *Between-Subject-Factor* noch mit Wahrscheinlichkeit von über 98 % aufgedeckt werden. Lediglich kleine Gruppeneffekte können als Konsequenz der Fallzahlanpassung mit einer Wahrscheinlichkeit von 27%, sie aufzudecken, falls sie wirklich existieren, bei Vorliegen eines statistisch nicht-bedeutsamen *Between-Subject-Factor* nicht ausgeschlossen werden. Dies hatte zur Folge, dass im gegebenen Analysedesign auch kleine Gruppenunterschiede zu einzelnen Messzeitpunkten nicht ausgeschlossen werden können. Da sich die Testung kleiner Effekte des *Between-Subject-Factors* (Unterschied zwischen den über alle Messzeitpunkte gemittelten Gruppenmittelwerten der Evaluationsmaße) nicht auf die konfirmatorischen Hypothesen bezog (Unterschied zwischen den messwiederholten Verläufen der Evaluationsmaße), war die Kernevaluation der Intervention nicht gefährdet. Nach wie vor konnte mit hoher Teststärke geprüft werden, inwieweit mit der Intervention eine Symptomreduktion erzielt werden kann. Darüber hinaus war es möglich, die Daten mittels exploratorischer Mehrebenen-Analysen zu modellieren. Hierbei wurden die messwiederholten Evaluationsmaße (Ebene 1) in den Patientinnen (Ebene 2) genestet. Jene Verfahren haben im Vergleich zu den herkömmlichen

Analysedesigns u.a. den massiven Nachteil, dass Power-Analysen und somit ad-hoc Fallzahlschätzungen meist nicht eindeutig durchzuführen sind (Scherbaum & Ferreter, 2008), was der Grund dafür ist, dass jene Modellierungen nicht ad-hoc für die konfirmatorischen Analysen gewählt wurden. Ergebnisse aus Simulationsstudien (Maas & Hox, 2005) zeigen jedoch auf, dass nur kleine Stichproben auf Ebene 2 (weniger als $N = 50$ Dyaden) zu verzerrten Schätzern der Standardfehler auf Ebene 2 führen. Die exploratorische Anwendung jener Modelle kann – methodeninvariante Ergebnisse vorausgesetzt – somit dazu dienen, die Sicherheit bezüglich des Ausschlusses kleiner Effekte des *Between-Subject-Factors* zu erhöhen.

Basierend auf linearen Extrapolationen der entsprechenden Rekrutierungsraten bis September 2019 und unter der Annahme, dass sich jene ab dem Q3 2019 stabilisierten, waren pro Quartal folgende neu hinzukommende Fallzahlen anzunehmen: ca. $n = 929$ Screenings, ca. $n = 200$ Fälle in den koordinierenden Stellen und ca. $n = 38$ Studieneinschlüsse. Mit einer Ausweitung des Screening-Zeitraums auf das Q3 2020 sowie einer Ausweitung der Fallverarbeitungen in den koordinierenden Stellen und des Studieneinschlusszeitraums auf das Q4 2020 hätten demnach bis zum Q3 2020 ca. 5.743 Fälle gescreent, bis zum Q4 2020 ca. 1.438 Fälle in den koordinierenden Stellen verarbeitet und ca. 328 Fälle in die Evaluation eingeschlossen werden können.

Mit der nun damals angestrebten Zahl an Studieneinschlüssen von $N = 280$ war die Evaluation der konfirmatorischen Hypothesen gesichert. Gemäß der Extrapolation konnte mit einem Erreichen der Fallzahl im Q3 2020 gerechnet werden ($N = 291$ Studieneinschlüsse). Insgesamt wäre dann ein geschätzter Puffer von nur $n = 11$ Studieneinschlüssen jedoch sehr gering ausgefallen, der nicht hinreichend gegen unvorhergesehene negative Schwankungen der Rekrutierungsraten und positive Schwankungen der *Dropout*-Raten absichern hätte können. Dies sei gerade auch deswegen angemerkt, da bei insgesamt damals erst $n = 1$ dokumentierten Studienabschlüssen (Anfang September 2019; Rekrutierungsbeginn Januar 2019; Studiendurchlaufzeit ca. 8 – 9 Monate) die Studien-*Compliance* nur grob geschätzt werden konnte. Sollte ein Erreichen von $N = 280$ Falleinschlüssen bereits vor dem Q4 2020 erfolgt sein, sollten die Rekrutierungsbemühungen folglich dennoch bis zum Ende des Q4 2020 fortgesetzt werden, um die Evaluation der konfirmatorischen Hypothesen gegen zusätzliche listenweisen Dropouts in den statistischen Analysen möglichst abzusichern, die sich skalenspezifisch durch das Fehlen einzelner Itemantworten oder durch post-hoc Fallausschlüsse ergeben konnten und die somit potentiell erst spät im Studienverlauf oder gar erst nach Abschluss der Datenerhebung für jedes Evaluationsmaß und jede konfirmatorische Analyse definitiv bestimmbar waren. $N = 328$ war somit die angestrebte Fallzahl, mit der die Kernevaluation (konfirmatorische Analysen) abgesichert sein sollte.

Darüber hinaus sollte dies auch zur Erhöhung der Sicherheit des Ausschlusses kleiner Gruppeneffekte führen: Mit $N = 328$ Studieneinschlüssen, die extrapoliert bis Q4 2020 erreicht werden konnten, ist ausgehend von einer 30%igen *Dropout*-Rate mit einer Fallzahl von $N = 230$ zu rechnen, was die Sicherheit zum Ausschluss kleiner Gruppeneffekte bereits auf 32% erhöht. Die Studieneinschlüsse konnten erreicht und sogar übertroffen werden, jedoch führten skalenspezifisch fehlende Werte zu einer Minderung der Fallzahlen in den Analysen. Um die konfirmatorischen Analysen gegen Antwort- und Selektionseffekte abzusichern, wurden Multiple Imputationen angewendet (siehe unten).

Finale Fallzahlberechnung

Für die Fallzahlschätzung und *Power*-Analysen wurde *G*Power* (Version 3.1.9.7; Faul et al., 2007; Faul et al., 2009) verwendet. Sie wurde für die konfirmatorischen Hypothesen vorgenommen, welche auf eine Wechselwirkung zwischen dem *Between*- (Interventionsgruppierung) und *Within-Subject-Factor* (Visite) abzielen. Die testschwächsten Analysen waren hierbei jene zur Überprüfung der Hypothesen bzgl. der schwangerschaftsbezogenen Ängste (zweifaktorielle ANOVA mit dem *Between-Subject-Factor* "Interventionsgruppierung" (eMBI vs. TAU) und dem *Within-Subject-Factor* "Visite" (5 Stufen), für welche schließlich die Fallzahlschätzung durchgeführt wurde.

Der Schätzung wurden ein für drei Analysen Bonferroni-adjustierter Fehler 1. Art von $\alpha = .016$ und eine Teststärke von $1-\beta = .90$ zugrunde gelegt. Die statistisch relevante Effektstärke basierte auf den Ergebnissen eines Pilots zu einer *Mindfulness*-basierten Prävention an schwangeren Patientinnen mit Depressionen in der Vorgeschichte (Dimidjian et al., 2016). Basierend auf den in dieser Studie berichteten Mittelwerten, Standardabweichungen und Fallzahlen errechnet sich ein Behandlungseffekt in der Katamnese von Cohen's $d = .34$, der nach Cohen (1988) als kleiner Effekt einzuordnen ist. Somit wird die Fallzahl im beschriebenen ANOVA-Design für kleine Behandlungseffekte ($f = .10$) geschätzt.

Für die Evaluation der eMBI wurden folglich $N = 196$ Fälle (balanciert für zwei Gruppen und zwei Regionen) benötigt. Bei einer angenommenen *Drop-Out-Rate* von 30% wurde eine zu rekrutierende Fallzahl von $N = 280$ festgelegt ($n = 140$ pro Gruppe bzw. pro Region und $n = 70$ pro Gruppe und Region), um die Fallzahl von $N = 196$ zu erreichen.

Tabelle 2: Ein- und Ausschlusskriterien der Studie

Einschlusskriterien	Ausschlusskriterien
GKV-versichert bei einer der teilnehmenden Krankenkassen	Mehrlingsschwangerschaften
vollendetes 18. Lebensjahr	hohes Frühgeburtsrisiko
Internetzugang	Fetale / kongenitale Fehlbildungen
Deutschkenntnisse: fließend	Teilnahme an einer MBI in der jetzigen Schwangerschaft
bei Studieneinschluss gesicherte Schwangerschaft bis max. 28 Schwangerschaftswoche	Suizidalität
EPDS-Score > 9	akute psychotische Episode oder diagnostizierte Psychose
	Akute bipolare affektive Symptomatik
	Akute Substanzkonsumstörung
	<i>Borderline</i> -Persönlichkeitsstörung
	aktuelle traumatische Erlebnisse ohne Bezug zur Schwangerschaft
	Notwendigkeit einer psychiatrischen Akut-Behandlung

Rekrutierung & Studiendauer

Mit der Rekrutierung der Studienteilnehmerinnen wurde am 29.01.2019 begonnen. Schwangere wurden in den Praxen, die sich in den Selektivvertrag eingeschrieben hatten, gescreent, ebenso in den Schwangerenambulanzen der Universitätskliniken Heidelberg und Tübingen. Im Anschluss wurden Schwangere mit einem auffälligen Screening durch die Study Nurse in Heidelberg und Tübingen kontaktiert und zum Erstgespräch eingeladen. Nicht akut behandlungspflichtige Schwangere wurden hier am Ende des Gesprächs durch den Studienarzt in die Evaluationsstudie eingeschlossen und auf die Interventions- oder Kontrollgruppe randomisiert.

Entsprechend des deskriptiven Monitorings der Screening- und Einschlusszahlen vom September 2019 haben wir festgestellt, dass die ursprünglich geplanten Fallzahlen zum Abschluss des Versorgungs- und Evaluationsprojekts sowohl hinsichtlich der Screenings, der Fälle in den koordinierenden Stellen als auch bzgl. der Studieneinschlüsse trotz aller Bemühungen nicht realistisch sind. Daraufhin wurde das Evaluationskonzept von unserem Evaluator Herrn Dr. Mitho Müller angepasst (siehe ausführliche Darstellung der Änderungen weiter oben) und wir haben als Konsortium einen kostenneutralen Antrag auf Laufzeitverlängerung um 6 Monate gestellt, welcher freundlicherweise durch den Projektträger bewilligt wurde. Das Screening fand bis zum 30.09.2020 statt, auffällige Schwangere wurden zu Erstgesprächen in den KOS bis einschließlich Dezember 2020 gesehen und eingeschlossen.

Inanspruchnahme der Erstgespräche

Die Erstgespräche konnten persönlich in den KOS bzw. per Videotelefonie geführt werden. Insgesamt fanden 546 Erstgespräche statt, davon 273 per Video, 144 in der KOS Heidelberg und 129 in der KOS Tübingen. 72 (18 KOS Heidelberg; 54 KOS Tübingen) Schwangere konnten aufgrund z. B. Borderline, Posttraumatische Belastungsstörung, akute Substanzkonsumstörung nicht in die Studie eingeschlossen werden. Daraus resultierend konnten 230 Schwangere in die Interventionsgruppe bzw. 230 Schwangere in die Kontrollgruppe eingeschlossen werden.

Tabelle 3: Anzahl der Gespräche (inkl. Erstgespräche) pro Gruppe

Anzahl Gespräche ^a	TAU	eMBI	Gesamt ^b
1 Erstgespräche	137	102	239
2 Gespräche	27	32	59
3 Gespräche	7	6	13
4 Gespräche	1	2	3
5 Gespräche	2	0	2
Gesamt	174 ^c	142 ^c	316 ^c

Anmerkungen. TAU = Treatment as usual (Kontrollgruppe); eMBI = electronic-mindfulness-based intervention (Interventionsgruppe); a. Mit einzelnen Patientinnen wurden bis zu 5 Gespräche, inklusive Erstgespräch geführt; b. Gesamtzahl an Patientinnen über beide Gruppen hinweg, mit denen die entsprechende Anzahl an Gesprächen geführt wurden; c. Gesamtzahl an Patientinnen in den entsprechenden Gruppen und insgesamt, basierend auf der finalen Fallzahl (siehe Abbildung 1).

Die Differenz aus der Anzahl der Coaching-Gespräche der IG vs. TAU ergibt sich daraus, dass die Gespräche in beiden Gruppen unterschiedlich nachgefragt und wahrgenommen wurden. Teilnehmerinnen der IG hatten den vollständigen Content, d.h. zahlreiche Beiträge, Lehrvideos und Fragebögen innerhalb der App mindmom zur Verfügung. Teilnehmerinnen der TAU standen lediglich die Fragebögen zur Verfügung.

Randomisierung und Verblindung

Neue Patientinnen wurden zunächst im online eCRF-Programm „REDCap“, das auf den Servern des Instituts für Medizinische Informationsverarbeitung Biometrie und Epidemiologie (IBE) der medizinischen Fakultät der LMU gehostet wurde, von durch Dr. Mitho Müller (LMU München, gesundheitswissenschaftliche Evaluation) berechtigten Mitarbeiter:innen der koordinierenden Stellen angelegt. Über eine in „REDCap“ integrierte Randomisierungsliste erhielten die Teilnehmerinnen die für die Evaluator:innen verblindete Zuordnung „0“ bzw. „1“. Diese Randomisierungsliste ($N = 822 + 100\% = 1.644$) wurde vor Erhebungsbeginn von Anton Marx, Mitarbeiter der Lehr- und Forschungseinheit *Psychology of the Learning Sciences* des Departments Psychologie der LMU München nach den Vorgaben von Dr. Mitho Müller, mittels des R-package „blockrand“ (v. 1.3, R v. 3.4.2, RStudio v. 1.1.383) erstellt. Hierbei wurde eine Blockrandomisierung mit zufällig zwischen 2, 4, 6 bzw. 8 Studienteilnehmerinnen variierender Blocklänge vorgenommen. Die Randomisierungsliste wurde nach dokumentierter Erstellung von Anton Marx in „REDCap“ integriert. Die Zuordnung des binären Codes zur Interventions- (*Treatment as usual* + elektronische, mindfullness-basierte Intervention; eMBI) bzw. Kontrollgruppe (*Treatment as usual*; TAU) wurde im Vorfeld der Studie fixiert und ist den Mitarbeiter:innen der koordinierenden Stelle bekannt. Diesen ist die Kommunikation der Zuordnung gegenüber den Evaluator:innen untersagt. Die Entblindung der Evaluatoren vor dem Abschluss der finalen gesundheitswissenschaftlichen Evaluation war weder notwendig noch gestattet. Sie erfolgte nach Studienabschluss der letzten Teilnehmerin.

Datenerhebung inkl. Der primären und sekundären Zielgrößen

Die Erhebung aller evaluationsrelevanter Informationen erfolgte über (1) die App *Mindmom* auf Grundlage von elektronischen Patientenberichten und (2) Eingaben der berechtigten Mitarbeiter:innen der koordinierenden Stellen in den elektronischen Fallbericht (REDCap). Die Datenspeicherung und -übertragung waren verschlüsselt. Die Plausibilitätsprüfung erfolgte, sobald die Daten in das System eingegeben wurden. Die geschützte Dateneingabe wurde durch individuelle Zugangsdaten für jeden Studienmitarbeiter gewährleistet. Alle Patientinnen erhielten pseudonymisierte Studien-IDs durch die koordinierenden Stellen. Die Schlüsseldatei lag nur den berechtigten Studienmitarbeitern und App-Entwicklern und verschlüsselt vor. Die Evaluator:innen erhielten die evaluationsrelevanten Informationen ohne Schlüsseldatei oder anderen Informationen, die eine Identifikation ermöglichte. Für die gesundheitswissenschaftliche Evaluation werden lediglich die pseudonymisierten Daten verwendet. Über das Studienpseudonym wurden darüber hinaus die Datensätze aus der APP- und der eCRF- Datenbank zusammengeführt. Für die Erfassung der evaluationsrelevanten Endpunkte wurden ausschließlich standardisierte, validierte Erhebungsinstrumente verwendet.

Edinburgh Postpartum Depression Scale (EPDS)

Die EPDS (Bergant et al., 1998) wird zur Erfassung depressiver Symptome im Zeitraum der zurückliegenden 7 Tage im Prä- und Postpartalzeitraum verwendet und besteht aus 10 Items, die auf einer Skala von 0 bis 3 Punkten beantwortet werden. In der Summe sind somit 0 bis 30 Punkte möglich, wobei „0“ absolute Symptombefreiheit und „30“ maximale Symptombelastung bedeutet. Die Skala weist bei einem Schwellenwert von > 9 eine Sensitivität von 0.96, eine Spezifität von 1.00 sowie einen positiven Vorhersagewert von 1.00 für die Diagnose einer klinisch signifikanten Depression auf. In unserer Stichprobe erreicht die Skala eine gute interne Konsistenz, deren Punktschätzer zwischen *Cronbach's* $\alpha = .83$ (Visite 2) und $.89$ (Visite 6) variieren.

Tabelle 4: Studienablauf und Erhebungsinstrumente

	Zeitraum	Zu erhebende Daten	Erhebungsinstrument
V1	28. SSW	Psychiatrische Diagnosestellung Depressivität Angst Geburtsangst Präpartales <i>Bonding</i> Lebensqualität Achtsamkeit Cortisol	PHQ-D EPDS STAI PRAQ-R MFAS EQ-5D FFA-14 Morgenprofil
V2	30. SSW	Depressivität Angst Geburtsangst	EPDS STAI PRAQ-R
V3	32. SSW	Depressivität Angst Geburtsangst Präpartales <i>Bonding</i> Lebensqualität	EPDS STAI PRAQ-R MFAS EQ-5D
V4	34. SSW	Depressivität Angst Geburtsangst	EPDS STAI PRAQ-R
V5	36. SSW	Depressivität Angst Geburtsangst Präpartales <i>Bonding</i> Lebensqualität Achtsamkeit Cortisol	EPDS STAI PRAQ-R MFAS EQ-5D FFA-14 Morgenprofil

V6	1 Monat pp	Depressivität Angst Geburtserleben Postpartales <i>Bonding</i>	EPDS STAI PEQ-G PBQ-16
V7	5 Monate pp	Depressivität Angst Postpartales <i>Bonding</i> Achtsamkeit Cortisol	EPDS STAI PBQ-16 FFA-14 Morgenprofil

Gesundheitsfragebogen für Patienten (PHQ-D)

Der PHQ-D (Löwe et al., 2002) ist ein Selbstauskunftsfragebogen, welcher dem Screening und der Fallidentifikation sowie der Erhebung des Schweregrades von psychischen Störungen dient. Für die Studienzwecke wurde der Bogen als Interview durchgeführt. Die Fragen zu den einzelnen Störungsbildern wurden bei der Entwicklung des Instrumentes aus den diagnostischen Kriterien des DSM-IV (Saß et al., 2003) abgeleitet. Es besteht jedoch eine hohe Übereinstimmung zu den konzipierten Kriterien für des DSM-5 (Falkei & Wittchen, 2015): Die *American Psychiatric Association* (APA) empfiehlt die Skalen des PHQ im Konzept des DSM-5 für die Schweregradmessung. Die Kriteriumsvalidität der deutschen Version des PHQ-D wird als hervorragend beurteilt (Gräfe et al., 2004): Die interne Konsistenz für die kontinuierlichen Skalen beträgt Cronbach's $\alpha = 0.88$ für das Depressionsmodul und $\alpha = .79$ für das Somatisierungsmodul (Gräfe et al., 2004).

State-Trait-Anxiety-Inventory (STAI)

Der STAI-Fragebogen (Laux et al., 1981) besteht aus zwei Dimensionen: Der *State*-Dimension (STAI-S; Zustandsangst) und der *Trait*-Dimension (STAI-T; Eigenschaftsangst) á 20 *Items*, die jeweils auf einer Skala von 1 bis 4 Punkten bewertet werden. Der Summenscore variiert somit zwischen 20 und 80 Punkten, wobei „20“ absolute Symptombefreiheit und „80“ maximale Symptombelastung bedeuten. In unserer Stichprobe erreicht die Skala eine sehr gute interne Konsistenz, deren Punktschätzer für den STAI-S zwischen Cronbach's $\alpha = .93$ (Visite 1) und $.96$ (Visite 6) und für den STAI-T zwischen Cronbach's $\alpha = .92$ (Visite 1) und $.95$ (Visite 6) variieren.

Pregnancy-Related-Anxiety-Questionnaire (PRAQ-R)

Die deutsche Übersetzung der Kurzversion des PRAQ-Fragebogens (Dubber et al., 2015) besteht aus 10 *Items*, bei denen schwangerschaftsbezogene Ängste auf einer Skala von 1 bis 5 Punkten evaluiert werden. Der Summenscore variiert somit zwischen 10 und 50 Punkten, wobei „10“ absolute Angstfreiheit und „50“ maximale schwangerschaftsbezogene Angst bedeuten. In unserer Stichprobe erreicht die Skala eine gute interne Konsistenz, deren Punktschätzer zwischen Cronbach's $\alpha = .81$ (Visite 1) und $.86$ (Visite 5) variieren.

Freiburger Fragebogen zur Achtsamkeit (FFA-14)

Der FFA ist ein Selbstbericht, der der Erfassung verschiedener Aspekte von Achtsamkeit dient. Die *Items* werden auf einer Skala von 1 („fast nie“) bis 4 („fast immer“) bewertet. Das Instrument konnte durch die Entwickler auf eine einfaktorielle 14-Item-Kurzversion (FFA-14) reduziert werden, die eine gute Konstruktvalidität aufweist und veränderungssensitiv ist (Walach et al., 2006). Daraus resultiert eine theoretische Variationsbreite von 14 („minimale Achtsamkeit“) bis 56 („maximale Achtsamkeit“). Die diskriminante und konvergente Validität des FFA-14 kann als gesichert betrachtet werden (Heidenreich et al., 2006). In unserer Stichprobe erreicht die Skala eine gute bis sehr gute interne Konsistenz, deren Punktschätzer zwischen Cronbach's $\alpha = .89$ (Visite 1) und $.92$ (Visite 7) variieren.

Cortisol-Morgenprofil

Die Cortisol-Aufwachprofile (*Awakening-Response*; Stalder et al., 2016) wurden als Indikatoren stressbedingter Veränderung der Aktivität der Hypothalamus-Hypophysen-Nebennierenrinden-

Achsen (Diego et al., 2006; Field et al., 2006; Kirschbaum & Hellhammer, 2007) mittels Speichelproben (Kirschbaum, 1991) erfasst.

Die Patientinnen wurden instruiert, an einem Werktag eigenen Speichel direkt nach dem Aufwachen sowie 30, 45 und 60 Minuten danach zu entnehmen. Die Analyse der Speichelproben wurde unter Anleitung von Prof. C. Kirschbaum im Labor der Abteilung für Biologische Psychologie an der Technischen Universität Dresden durchgeführt. Die Speichelproben wurden bis zur Analyse im Labor bei -20°C gelagert (Haltbarkeit über 6 Monate). Nach dem Auftauen wurden die Salivetten mit 3.000 rpm fünf Minuten zentrifugiert, was zu einem Überstand mit geringer Viskosität führt. Zur Bestimmung der Speichelcortisolmenge wurde ein hoch sensitiver kommerzieller Immunoassay mit Chemilumineszenz-Detektion benutzt (CLIA, IBL-International, Hamburg). Die Nachweisgrenze des Assay liegt bei $0.005\ \mu\text{g}/\text{dl}$. Die Intra-Assay-Varianzen ($n = 20$, $20\ \mu\text{l}$ Saliva) liegen bei 3.1% für $0.125\ \mu\text{g}/\text{dl}$, 3.0% für $0.656\ \mu\text{g}/\text{dl}$ und 3.4% für $2.917\ \mu\text{g}/\text{dl}$.

Als Maß des totalen Cortisol-Outputs wurde gemäß aktueller Leitlinien (Stalder et al., 2016) die *Area under the Curve with respect to the Ground* (AUC_G ; Pruessner et al., 2003) verwendet. Dieses Maß ist als das Integral der aus den Messungen resultierenden Cortisol-Kurve und berücksichtigt somit die Zeitabstände zwischen den Messungen im Gegensatz zu statistischen Tests für messwiederholte Daten.

EQ-5D

Bei dem EQ-5D (Greiner, 2012) handelt es sich um ein generisches Indexinstrument zur Erfassung der gesundheitsbezogenen Lebensqualität. Der Fragebogen umfasst 5 Dimensionen mit je 3 Antwortstufen sowie eine visuelle Analogskala (VAS). Verschiedene Studien zeigen, dass der EQ-5D sowohl im Kontext der Depression im Allgemeinen (Eriksson et al., 2017; Hirani et al., 2014) als auch im Rahmen der Depression bei Schwangeren (Turkstra et al., 2015) ein elaboriertes Instrument darstellt. In unserer Stichprobe erreicht die Gesamtskala eine befriedigende interne Konsistenz, deren Punktschätzer zwischen *Cronbach's* $\alpha = .71$ (Visite 1) und $.80$ (Visite 5) variieren.

Maternal-Fetal-Attachment-Scale (MFAS)

Die deutsche Übersetzung des MFAS-Fragebogens (Dubber et al., 2015) erfasst die emotionale Bindung der Mutter gegenüber dem ungeborenen Kind während der Schwangerschaft (*präpartales Bonding*) mit 24 *Items*, die auf einer Skala von 1 bis 7 Punkten bewertet werden. Der Summenscore variiert somit zwischen 24 und 168 Punkten, wobei „24“ ein absolut minimales *Bonding* und „168“ ein maximales *Bonding* indizieren. In unserer Stichprobe erreicht die Skala eine gute interne Konsistenz, deren Punktschätzer zwischen *Cronbach's* $\alpha = .84$ (Visite 1) und $.85$ (Visite 5) variieren.

Postpartum-Bonding-Questionnaire (PBQ-16)

Die deutsche Kurzversion des PBQ-Fragebogens (Reck et al., 2006) erfasst Beeinträchtigungen in der emotionalen Bindung der Mutter gegenüber dem Kind (*postpartales Bonding*) mit 16 *Items*, die auf einer Skala von 0 bis 5 bewertet werden. Der Summenscore reicht somit von 0 bis 80 Punkten, wobei „0“ keine Beeinträchtigung und „80“ maximale Beeinträchtigung bedeuten. In unserer Stichprobe erreicht die Skala eine sehr gute interne Konsistenz mit Punktschätzer von *Cronbach's* $\alpha = .91$ (Visite 1) und $.92$ (Visite 2).

Patients Experience Questionnaire Geburt (PEQ-G)

Der PEQ-G dient dazu Patienten nach einer stationären Entbindung zu den Erfahrungen zu befragen, die diese während ihres Krankenhausaufenthaltes gemacht haben. Sie berücksichtigt die spezielle Situation und richtet sich an den besonderen Informationsbedarf von werdenden Müttern. Der Kurzfragebogen deckt mit insgesamt 16 Kernfragen mehrere versorgungsrelevante Themenbereiche ab. Die *Items* werden auf einer sechsstufigen Antwortskala mit 1 als bestmögliche Bewertung (= 100 %) und 6 als schlechtestmögliche Bewertung (= 0%) bewertet. Der Score ergibt sich aus der Addition der einzelnen Prozentwerte dividiert durch die Gesamtanzahl der Einzelantworten. Bisher liegen keine Ergebnisse bzgl. der Validität der *Items* vor. In unserer Stichprobe erreicht die Skala eine befriedigende bis sehr gute interne Konsistenz mit Punktschätzer von *Cronbach's* $\alpha = .71$ (Zufriedenheit mit Organisation und Service) und $.95$ (Zufriedenheit mit Betreuung durch Hebammen).

Auswertungsstrategie

Alle statistischen Analysen wurden mit IBM® SPSS® (Version 29.0.0.0.(241)) durchgeführt. Tabelle 4 gibt einen Überblick über die spezifischen Verfahren zu den jeweiligen Hypothesen. Der Missing-Completely-At-Random-Test (MCAR-Test, Little, 1988) wurde zur Analyse fehlender Werte aufgrund von *Drop-outs* und *Missings* verwendet, um auszuschließen, dass Fallausschlüsse die Ergebnisse verzerren können. Darüber hinaus wurden die Gruppen hinsichtlich soziodemografischer und medizinischer Drittvariablen mittels *t*-, *U*- und χ^2 -Tests auf Vergleichbarkeit geprüft. Wenn Unterschiede signifikant waren, wurden die entsprechenden Variablen auf Assoziationen mit den Ergebnisvariablen untersucht und ggf. als Kovariaten in den Modellen kontrolliert.

Die kritischen Fehler 1. Art für die 3 Tests zur Überprüfung der konfirmatorischen Hypothesen wurden auf $\alpha_{\text{lokal}} = .05 / 3 = .016$ Bonferroni-adjustiert, um einen globalen Fehler 1. Art von $\alpha_{\text{global}} = .05$ nicht zu überschreiten. Die vorausgehenden, deskriptiven und exploratorischen Analysen erfolgen ohne Adjustierung des kritischen Fehlers 1. Art auf einem konventionellen Niveau von $\alpha_{\text{lokal}} = .05$. Wegen des dadurch kumulierten globalen Fehlers 1. Art haben die exploratorischen Analysen somit hypothesengenerierenden Charakter, da diese so mit einer höheren Wahrscheinlichkeit Zufallsbefunde umfassen können.

Alle empirischen Fehler 1. Art (*p*-Werte) werden nicht-adjustiert berichtet, im Fall der konfirmatorischen Analysen mit dem adjustierten kritischen Fehler 1. Art und im Fall der exploratorischen Analysen mit dem nicht-adjustierten kritischen Fehler 1. Art verglichen. Als post-hoc Test wurde *Dunn's Multiple Comparison Procedure* angewandt (Dunn, 1961). Mithilfe dieses Verfahrens lässt sich unter Festlegung des globalen Fehlerniveaus von $\alpha_{\text{global}} = .05$, gegebenen Freiheitsgraden, gegebener Fehlervarianz und gegebener Zelhäufigkeiten hypothesengeleitet eine kritische Differenz (Ψ_{Dunn}) bestimmen, die überschritten werden muss, um einen Mittelwertsunterschied als signifikant zu erachten. Bei nicht-hypothesenrelevanten signifikanten Effekten wird auf die Durchführung eines post-hoc-Tests verzichtet. Bei solchen signifikanten Effekten erfolgt daher die Interpretation anhand des Unterschiedes zwischen dem größten und dem kleinsten Mittelwert, die für den signifikanten Effekt maßgeblich sind.

Effektstärken der metrischen Analysen werden als partielles η^2 berechnet. Dies ist ein stichprobenbasiertes Maß der aufgeklärten Varianz. Nach Cohen (1988) sind diesbezüglich .01 kleine, .06 mittlere und .14 große Effekte. Effektstärken für χ^2 -Tests werden als $w^2 = \frac{\chi^2}{N}$ bestimmt. Nach Cohen (1988) sind diesbezüglich .01 kleine, .09 mittlere und .25 große Effekte. Für die vorausgehenden, deskriptiven Analysen werden keine Effektstärken berechnet.

Tabelle 5: Analyseverfahren

Konfirmatorische Hypothese(n)	Verfahren	Anzahl der Tests	Teststatistik	Effekt
1 - 3	2 (Gruppe) x 7 (Messzeitpunkt) - ANOVA mit Messwiederholung auf dem Faktor „Messzeitpunkt“	2	F-Verteilung	Gruppe x Messzeitpunkt-Wechselwirkung
4 - 6	2 (Gruppe) x 5 (Messzeitpunkt) - ANOVA mit Messwiederholung auf dem Faktor „Messzeitpunkt“	1	F-Verteilung	Gruppe x Messzeitpunkt-Wechselwirkung
Exploratorische Hypothese(n)	Verfahren	Anzahl der Tests	Teststatistik	Effekt
1 - 9 & 14 - 16	2 (Gruppe) x 3 (Messzeitpunkt) - ANOVA mit Messwiederholung auf dem Faktor „Messzeitpunkt“	9	F-Verteilung	Gruppe x Messzeitpunkt-Wechselwirkung

10 - 12	2 (Gruppe) x 2 (Messzeitpunkt) - ANOVA mit Messwiederholung auf dem Faktor „Messzeitpunkt“	1	F-Verteilung	Gruppe x Messzeitpunkt- Wechselwirkung
13	ANOVA ohne Messwiederholung	5	F-Verteilung	Gruppen-Haupteffekt
17	4-Felder χ^2 -Test	3	χ^2 -Verteilung	Gruppen-Haupteffekt

Darüber hinaus wurden exploratorische Moderationsanalysen zur Aufklärung der Modifizierbarkeit der gefundenen Zusammenhänge durchgeführt. Diese wurden im Rahmen der exploratorischen Analysen von Interaktionseffekten zwischen Drittvariablen und der Interventionsgruppierungsvariable im messwiederholten ANOVA-Design analysiert.

Ferner wurde bereits ein Teil der Daten mittels exploratorischer Mehrebenen-Analysen modelliert. Hierbei wurden die messwiederholten Evaluationsmaße (Ebene 1) der konfirmatorischen Modelle inklusive des *Manipulation Checks* in den Patientinnen (Ebene 2) genestet. Ergebnisse aus Simulationsstudien (Maas & Hox, 2005) zeigen auf, dass nur kleine Stichproben auf Ebene 2 (weniger als $N = 50$ Dyaden) zu verzerrten Schätzern der Standardfehler auf Ebene 2 führen. Aufgrund einer mangelnden Datenbasis konnten à priori keine Annahmen zum Kovarianztyp der messwiederholten Daten getroffen werden, weshalb für all diese Analysen von einer unstrukturierten Varianz-Covarianz-Matrix ausgegangen wurde. In den Mehrebenenanalysen wurden nur solche Effekte modelliert, die auch Bestandteil der konfirmatorischen Analysen und des *Manipulation Checks* waren. Die nicht-messwiederholten Variablen wurden in den Modellen als Prädiktoren mit festem Format behandelt. Die exploratorische Anwendung jener Modelle kann – methodeninvariante Ergebnisse vorausgesetzt – somit dazu dienen, die Sicherheit bezüglich des Ausschlusses kleiner Effekte des *Between-Subject-Factors* zu erhöhen.

Multiple Imputationen

Um fehlende Werte zuverlässig zu schätzen und die konfirmatorischen Analysen somit gegen Antwort- und Selektionseffekte abzusichern, wurden Multiple Imputationen (Enders, 2010) verwendet. Vorab wurde eine mit fehlenden Werten verbundene Systematik mittels des *MCAR-Tests* nach Little (1988) untersucht (siehe oben). Anschließend wurden gemäß der Empfehlung von Enders (2010) $n = 20$ Datensätze geschätzt.

In das Imputationsmodell gingen alle Variablen des Analysemodells ein. Diese Liste umfasst die Gruppierungsvariablen als Prädiktoren, sowie die Kovariaten und alle abhängigen Variablen aus den konfirmatorischen Analysen und dem *Manipulation Check* als Prädiktoren und zu prädizierende Variablen. Hierbei wurden die skalenspezifischen Restriktionen (skalenspezifische Minima und Maxima, siehe 2.1.: Beschreibung der Instrumente) als Nebenbedingungen eingegeben. Bei der Anzahl der Kinder im Haushalt wurde das Maximum willkürlich auf 10 festgelegt, um zumindest der Größenordnung zentraleuropäischer Haushalte gerecht zu werden. Für den automatisierten Schätzprozess wurden maximal 500 Fallziehungen sowie maximal 20 Parameterziehungen zugelassen. Die visuelle Analyse des Iterationsprozesses in jedem der 20 vervollständigten Datensätze ließ keine systematische über zufällige Variation hinausgehende Streuung der Mittelwerte und Standardabweichungen der geschätzten Werte erkennen. Für die Imputationsanalysen wurden die oben aufgeführten konfirmatorischen Modelle und der *Manipulation Check* in jedem der 20 vervollständigten Datensätze durchgeführt und anschließend die Resultate aus allen Datensätzen inklusive des Originaldatensatzes *gepooled* (einfaches arithmetisches Mittel).

2.3 Gesundheitsökonomische Evaluation (Universität Bielefeld)

Das Ziel der gesundheitsökonomischen Evaluation der randomisiert-kontrollierten Studie (RCT) besteht darin, die Inanspruchnahme von Leistungen der gesetzlichen Krankenkasse (GKV) sowie die dadurch erzeugten Kosten der beiden Studiengruppen (Interventionsgruppe (IG)) und Kontrollgruppe (KG)) abzubilden und zu vergleichen.

2.3.1 Fragestellung und Arbeitshypothese

Die zentrale Hypothese der gesundheitsökonomischen Evaluation lautet: Die (direkten) Behandlungskosten fallen in der IG geringer aus, als in der KG. Der Empfehlung des Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (2020) folgend, wurde als primäre Perspektive die Sicht der GKV eingenommen, wobei die folgenden Leistungsbereiche einbezogen wurden: Ambulante Leistungen, stationäre Leistungen, Arzneimittel, Leistungen in Hochschulambulanzen (HSA), psychiatrischen Institutsambulanzen (PIA) und ambulante Operationen, Hebammenleistungen, Heil- und Hilfsmittel, Haushaltshilfe sowie die Kosten der Intervention selbst. Im Falle von Mehrkosten in der Interventionsgruppe sollte zusätzlich eine Kosten-Effektivitätsanalyse, durch die Beurteilung der inkrementellen Kosten-Nutzen Relation (ICER), vorgenommen werden. Da das Ergebnis des Gruppenvergleichs der Gesamtkosten keinen eindeutigen Rückschluss auf Mehrkosten in der IG zulässt, wurde keine Kosten-Effektivitätsanalyse durchgeführt. Im Sinne einer Ausweitung der gesundheitsökonomischen Evaluation auf die gesellschaftliche Perspektive, wurden die Produktivitätsverluste (indirekte Kosten) ermittelt und Selbstzahlerleistungen (OTCs und Dienstleistungen) der Teilnehmerinnen erhoben.

Neben der zentralen Betrachtung der Gesamtkosten, wurden ergänzende aus gesundheitsökonomischer Sicht relevante Analysen durchgeführt. Da zwei Drittel der Patienten mit der Diagnose Depression durch den Hausarzt versorgt werden (Gerste & Roick, 2014), erfolgte eine Differenzierung der Inanspruchnahme ambulanter Arztkontakte nach Arztgruppen (Haus- und relevante Fachärzte), welche zusätzlich im Zeitverlauf betrachtet wurde. Zudem wurde das Vorliegen einer erkrankungsspezifischen Medikation (ATC-Gruppen N05, N06) betrachtet. Um mögliche Treiber der Gesamtkosten der Studienpopulation zu identifizieren, wurden darüber hinaus Kostentreiberanalysen durchgeführt.

Abweichung vom ursprünglich geplanten Vorgehen:

Laut Evaluationskonzept sollten die folgenden Leistungsarten in die gesundheitsökonomische Evaluation einbezogen werden: ambulante ärztliche Leistungen, ambulant verordnete Arzneimittel, Krankenhaus, medizinische Rehabilitation inkl. Mutter-Kind-Kuren, Pflegeleistungen, Heil- und Hilfsmittel, Hebammenleistungen sowie Haushaltshilfe. In Abstimmung mit den beteiligten Medizinerinnen wurden diese zu Projektbeginn geringfügig angepasst. Da Pflegeleistungen im Rahmen der Pflegeversicherung für die vorliegende Population nicht relevant sind, wurden diese ausgeschlossen. Da davon auszugehen ist, dass im Beobachtungszeitraum rund um die Entbindung keine Rehabilitationsleistungen in Anspruch genommen werden, wurden auch diese exkludiert. Als für die Studienpopulation potentiell relevant wurden dagegen Leistungen aus den Bereichen HSA/ PIA/ ambulante Operationen sowie Mutterschaftsgeld ergänzt, wobei Letzteres als Transferzahlung separat ausgewiesen wurde.

2.3.2 Datengrundlage und -aufbereitung

Die Datengrundlage der Analyse bilden insbesondere die GKV-Routinedaten (Sekundärdaten), welche durch die relevanten im Rahmen des Projektes erhobenen Primärdaten ergänzt wurden. Die Zusammenführung erfolgte über das in der koordinierenden Stelle (Vertrauensstelle) vergebene Pseudonym. Die Zuordnungsliste der Identifikationsschlüssel der Primär- und Sekundärdaten wurde von der Vertrauensstelle verwaltet, sodass ein Rückschluss auf die personenbezogenen Daten im Rahmen der Evaluation ausgeschlossen war. Nach der Vereinheitlichung der Formate der GKV-Routinedaten erfolgte die Prüfung hinsichtlich der Vollständigkeit und Plausibilität. Dies umfasste neben der Identifizierung potenziell fehlender und doppelter Datenzeilen auch unplausible Werte, woraus sich ggf. eine Korrektur und erneute Datenlieferung ergab. An die Prüfung schloss sich die Aufbereitung der für die Evaluation relevanten Variablen an. Hierfür wurde zunächst eine einheitliche

Abgrenzung des individuellen Beobachtungszeitraumes von 40 Wochen (ab 28. SSW) vorgenommen. Für den Studienbeginn wurde demnach das Einschlussdatum um die Differenz zwischen der aktuellen Schwangerschaftswoche (SSW) und der 28. SSW verschoben. Wenn dadurch der Beobachtungsbeginn nach dem 30.09.2020 lag, führte dies zum Ausschluss der Teilnehmerinnen, da die Beobachtungszeit somit nicht vollständig, d.h. bis zum neuen Beobachtungsende, in den GKV-Routinedaten abgebildet war.

Als ambulante Kennzahlen wurden die Anzahl der Arztkontakte, die damit verbundenen Kosten pro Teilnehmerin sowie die durchschnittlichen Fallkosten im Beobachtungszeitraum untersucht. Aufgrund der Besonderheiten der Leistungsabrechnung im ambulanten Sektor, wurde die Abgrenzung des Beobachtungszeitraumes angepasst, sodass alle ambulanten Fälle im Quartal des Beobachtungsbeginns sowie in den drei darauffolgenden Quartalen eingeschlossen wurden. Für die Differenzierung von Haus- und relevanten Facharztkontakten wurde die Arztgruppenzuordnung (Kassenärztliche Bundesvereinigung [KBV], 2021) zugrunde gelegt, wobei neben der hausärztlichen Versorgung (Gruppen 1, 2 und 3) Besuche bei Fachärzten bzw. Therapeuten aus den Bereichen der Neurologie, Psychiatrie, Psychosomatik und Psychotherapie (Gruppen 51, 53, 58, 60, 61, 68) sowie Gynäkologie und Geburtshilfe (Gruppen 15, 16, 17, 18) fokussiert wurden. Die stationären Abrechnungsdaten bilden die Anzahl der Krankenhausfälle, die Aufenthaltsdauer, die Kosten sowie die Kosten pro Fall ab, wobei Letztere sich aus der Gruppe der Teilnehmerinnen mit mindestens einem stationären Aufenthalt ergeben.

Die Arzneimitteldaten wurden mithilfe der Stammdatei des GKV-Arzneimittelindex beim Wissenschaftliches Institut der AOK (Wissenschaftliches Institut der AOK [WIdO], 2022) ergänzt. Die Schlüsseltabelle ermöglicht eine Zuordnung der amtlichen Anatomisch-Therapeutisch-Chemischen Klassifikation (ATC-Codierung) sowie der definierten Tagesdosis (Defined Daily Dose, DDD) zu den dokumentierten Pharmazentralnummern. Während Letztere als technische Maß- und Vergleichseinheit für die verordnete Arzneimittelmenge dienen, konnten mithilfe des ATC-Codes Verschreibungen spezifischer Wirkstoffgruppen identifiziert werden. Als erkrankungsassoziierte Medikation wurden Psycholeptika (N05) und Psychoanaleptika (N06) betrachtet, wobei aufgrund der insgesamt geringen Inanspruchnahme von Arzneimitteln eine stark aggregierte Auswertung (binär) erfolgte.

Für alle weiteren relevanten Leistungsbereiche (HSA/ PIA/ ambulantes Operieren; Heilmittel, Hilfsmittel und Hebammenleistungen) wurden jeweils die Menge der Leistungsanspruchnahme anhand der Fälle, Verordnungen oder Tage sowie die Kosten berichtet. Aus den Kosten der einzelnen Leistungsbereiche ergeben sich schließlich die Gesamtkosten aus GKV-Perspektive. Das Mutterschaftsgeld wurde als Transferzahlung der GKV separat ausgewiesen.

Als relevante indirekte Kosten wurden die Produktivitätsverluste einbezogen. Für die Berechnung wurde, um eine Überschätzung der entstandenen Kosten zu vermeiden, der Friktionskostenansatz gewählt (Greiner & Damm, 2012). Dementsprechend sind für die monetäre Bewertung nur diejenigen Kosten berücksichtigt worden, die durch Krankheit entstehen, bis die Arbeitsstelle nachbesetzt wird. Um diese Zeitspanne zu erfassen, wurde die aktuelle durchschnittliche Vakanzzeit der deutschen Arbeitsmarktstatistik von 99 Tagen angesetzt (WISTA, 2018). Die Grundlage der Berechnung bilden die Daten zur Arbeitsunfähigkeit (AU) der Versicherten. Den Empfehlungen für gesundheitsökonomische Evaluationen folgend, wurden Durchschnittswerte auf der Basis der amtlichen Einkommensstatistik verwendet (Greiner & Damm, 2012).

durchschnittl. Prod. verlust

$$= \text{Arbeitsunfähigkeitstage} \times \frac{\text{Arbeitnehmerentgelt in Deutschland pro Jahr}}{(\text{Arbeitnehmer} \times 365 \text{ Tage})}$$

Das durchschnittliche Entgelt pro Arbeitnehmer pro Tag wurde mithilfe der Erwerbstätigenrechnung des Statistischen Bundesamtes (2022a) und der volkswirtschaftlichen Gesamtrechnung (Destatis, 2022b) für die Jahre der Studienlaufzeit ermittelt (Tabelle 5). Auf dieser Grundlage ließen sich die Produktivitätsverluste der Studienteilnehmerinnen anhand der individuellen Anzahl der in den Routinedaten dokumentierten AU-Tage berechnen.

Tabelle 6: Grundlage der Berechnung der Produktivitätsverluste

Jahr	Anzahl Arbeitnehmer (1000 Personen)	Arbeitnehmerentgelt (Mrd. €)	Entgelt pro Arbeitnehmer	Entgelt pro Arbeitnehmer pro Tag
2019	41117	1855,52	45.127,81 €	123,64 €
2020	40860	1852,10	45.327,95 €	124,19 €
2021	40991	1915,23	46.723,18 €	128,01 €

Um die Kosten der Intervention zu ermitteln, wurden zunächst die einzelnen Komponenten der neuen Versorgungsform identifiziert. Dabei ist zu berücksichtigen, dass lediglich die Kosten der online-basierten Achtsamkeitsintervention (d) gesondert in den Vergleich der Gesamtkosten einfließen, da diese ausschließlich für die IG anfielen. Im Sinne einer umfassenden Betrachtung sollten jedoch die potenziellen Kosten der gesamten Versorgungsform im Hinblick auf den Einsatz in der Regelversorgung ermittelt werden. Hierfür wurde die Expertise der Konsortialführung und der am Projekt beteiligten Kassenärztlichen Vereinigung (KV) hinzugezogen. Die Ermittlung der potenziellen Kosten der Komponenten a bis c erfolgte durch die Recherche vergleichbarer Leistungen aus dem einheitlichen Bewertungsmaßstab (EBM). Für die Komponente d wurden die zu Projektbeginn kalkulierten Kosten der online-basierten Achtsamkeitsintervention, inkl. Berücksichtigung möglicher Abweichungen (+/-30%) zugrunde gelegt.

Tabelle 7: Übersicht über die Komponenten der neuen Versorgungsform

Komponenten der neuen Versorgungsform
a) Sektorenübergreifendes Screening auf Depression und Angst beim niedergelassenen Gynäkologen oder in einer der beteiligten Kliniken
b) Koordination des Erstgesprächs, Einführung in die online-basierte Achtsamkeitsintervention und weitere Betreuung der Schwangeren
c) Psychologische Erstgespräch zur affektiven Diagnosestellung und ein Herausfiltern der akut gefährdeten Patientinnen
d) Online-basierte Achtsamkeitsintervention (Dauer der Anwendung: 8 Wochen)

Neben den Sekundärdaten wurden die für die Evaluation relevanten Primärdaten einbezogen. Diese umfassten u.a. projektbezogene Merkmale (Studiengruppe, Studiocenter), soziodemografische Eigenschaften (Alter, Bildung, Einkommen, Anzahl der im Haushalt lebenden Kinder) sowie weitere potenzielle Kontrollvariablen (Stressniveau, Depressivität, physische und psychische Vorerkrankung bei Studieneinschluss). Im Falle fehlender Werte können entsprechende Analysen auf einer reduzierten Stichprobe basieren. Die für die gesundheitsökonomische Evaluation aus gesellschaftlicher Perspektive vorgesehene Erhebung der Selbstzahlerleistungen erfolgte mittels eines für das Projekt entwickelten papierbasierten Kostenwochenbuches (Anlage 1). Dies orientiert sich an der Grundidee validierter Instrumente zur Erhebung von Gesundheitsausgaben (z. B. FIMA (Seidl et al., 2015)), bezieht sich jedoch lediglich auf die Kostenfaktoren nicht-verschreibungspflichtiger, homöopathischer und naturheilkundlicher Arzneimittel sowie alternativer diagnostischer Verfahren und Heilmethoden. Das Instrument wurde in einem Pre-Test erprobt und angepasst. Aufgrund der sehr geringen Rücklauf- bzw. Ausfüllquote und -qualität, sind nur sehr stark aggregierte Auswertungen möglich. Demnach wurde lediglich die Anzahl verschiedenartiger Präparate (z. B. Femibion und Vitamin D) bzw. Leistungen (z. B. zusätzlicher Ultraschall und Stillberatung) betrachtet, da die Angaben zur Anzahl der Packungen bzw. Dienstleistungen nicht einheitlich auswertbar waren. Bei den OTC Präparaten wurden Nahrungsergänzungsmittel, Wirkstoffe und homöopathische Mittel differenziert. Sonstige Angaben der Teilnehmenden (z. B. Kosmetika) wurden ausgeschlossen. Bei den Dienstleistungen wurde zwischen Diagnostik, Kursen, Kursen für Mutter und Kind, Beratungsleistungen und Homöopathie/Heilpraktiken unterschieden. Sonstige Angaben zu nicht-gesundheitsassoziierten Leistungen blieben unberücksichtigt. Eine Auswertung der Kosten zu Lasten der Versicherten war aufgrund der unzureichenden Datenqualität nicht möglich.

Abweichung vom ursprünglich geplanten Vorgehen:

Aufgrund der eingeschränkten Datenqualität wurde der Leistungsbereich der Haushaltshilfe lediglich deskriptiv ausgewertet und nicht in den Gesamtkosten berücksichtigt. Auf die Auswertung der Krankengeldzahlungen wurde gänzlich verzichtet. Da die Mutterschaftsgeldzahlungen nicht von allen beteiligten Krankenkassen geliefert werden konnten und diese vom Versicherungsstatus abhängig sind, basieren die berichteten Ergebnisse auf einer reduzierten Stichprobe.

Aufgrund des geringen Rücklaufs und der unzureichenden Datenqualität konnten die Selbstzahlerleistungen nicht in die Analyse der Gesamtkosten aus gesellschaftlicher Perspektive einbezogen werden.

2.4 Statistische Auswertung

Im Rahmen der gesundheitsökonomischen Evaluation wurden deskriptive und inferenzstatistische Auswertungsmethoden angewendet. An die Untersuchung der relevanten Baselinecharakteristika der Studienpopulation schloss sich die Analyse der vorab definierten Outcomes der gesundheitsökonomischen Evaluation an. Den Auswertungen, welche mithilfe der Statistiksoftware R Version 4.2.1 (R Core Team, 2020) durchgeführt wurden, lag ein Signifikanzniveau von 5 % zugrunde. Als geeignete inferenzstatistische Methoden für unabhängige Stichproben wurden t-Tests, Wilcoxon-Mann-Whitney-U-Tests und Chi²-Tests (`t.test()`, `wilcox.test()`, `chisq.test()`, Package stats 4.2.1 (R Core Team, 2020)) eingesetzt. Zur Vorbereitung multivariater Analysen sowie im Kontext der Kostentreiberanalysen wurden weitere Analysemethoden (z. B. `cor.test()` Package stats 4.2.1 (R Core Team, 2020)) genutzt. Darüber hinaus wurde die Verteilung fehlender Werte innerhalb der berücksichtigten Primärdaten mittels Missing-completely-at-random-test nach Little (1988) geprüft (`na.test()`, Package misty 0.4.6 (Takuya Yanagida, 2022)). Die Untersuchung der zentralen Hypothesen der gesundheitsökonomischen Evaluation hinsichtlich der Gesamtkosten und der ambulanten (Haus-)Arztkontakte erfolgte schließlich mittels geeigneter Regressionsmodelle. Dabei wurde auf additive Modelle für Lage-, Skalen- und Formparameter (Generalized Additive Model for Location, Scale and Shape (GAMLSS)) zurückgegriffen, welche den Ansatz der verallgemeinerten linearen Modelle zugrunde legen. Charakteristisch sind demnach die Lockerung der Normalverteilungsannahme für die abhängige Variable und die Möglichkeiten der Annahme alternativer Verteilungen (Rigby & Stasinopoulos, 2005). Eine Übersicht der relevanten Outcomes und Methoden für den Gruppenvergleich finden sich im Anhang.

Anhand des Akaike-Informationskriteriums (Akaike Information Criterion – AIC) und der Residuenplots wurden für die Analyse der Gesamtkosten aus GKV- und gesellschaftlicher Perspektive generalisierte additive Modelle mit inverser Gammaverteilung (`gamlss()`, Package `gamlss` 5.4-3 (Rigby & Stasinopoulos, 2005)) als geeignet identifiziert, welche der Schiefe der Verteilung der erhobenen Daten gerecht wird. Für die Analyse der ambulanten Arztkontakte und der Hausarztkontakte wurden generalisierte additive Modelle mit negativer Binomialverteilung (`gamlss()`, Package `gamlss` 5.4-3 (Rigby & Stasinopoulos, 2005)) eingesetzt, welche sich besonders für die Analyse von Zähldaten eignen. Für die Entwicklung der Regressionsmodelle wurden zunächst die potenziell relevanten unabhängigen Variablen und Interaktionsterme theoriegeleitet bzw. anhand vorab durchgeführter Analysen ausgewählt. Basierend auf dieser Auswahl wurde das beste Modell anhand des AIC mittels „backward selection“ (`stepAIC()`, Package `gamlss` 5.4-3 (Rigby & Stasinopoulos, 2005)) identifiziert (Fahrmeir et al., 2013). Für multiples Testen wurde mittels Bonferroni-Korrektur kontrolliert. Für die Modellauswahl und die Bewertung der Modellgüte wurden die Dichtekurven der analysierten Variablen mit den Dichtekurven der Modellschätzung verglichen. Zudem wurden weitere diagnostische Verfahren zur Bewertung der Modellgüte der Regressionsanalysen (z. B. Residuenplots (Stasinopoulos et al., 2017)) eingesetzt. Die Interpretation der Ergebnisse erfolgte anhand der exponierten Regressionskoeffizienten (`exp()`, Package base 4.2.1 (R Core Team, 2020)), welche den Faktor angeben, um den sich der Erwartungswert der abhängigen Variablen ändert, wenn sich lediglich die interessierende erklärende Variable um eine Einheit erhöht. Zudem wurden Konfidenzintervalle (`confint()`, Package stats 4.2.1 (R Core Team, 2020)) auf Basis der profile-log Likelihood-Methode berechnet. Die GKV-Daten-Auswertung erfolgte nach den Grundsätzen guter Sekundärdatenanalyse. Um die Stabilität der Ergebnisse zu prüfen, wurden Sensitivitätsanalysen durchgeführt. Dabei wurde zum einen möglichen Abweichungen der Höhe der Interventionskosten Rechnung getragen und zum anderen die Studienpopulation erweitert. Für die Regressionsanalysen wurde jeweils das

Ausgangsmodell analog zur Hauptanalyse gewählt und eine erneute automatische Modellselektion vorgenommen.

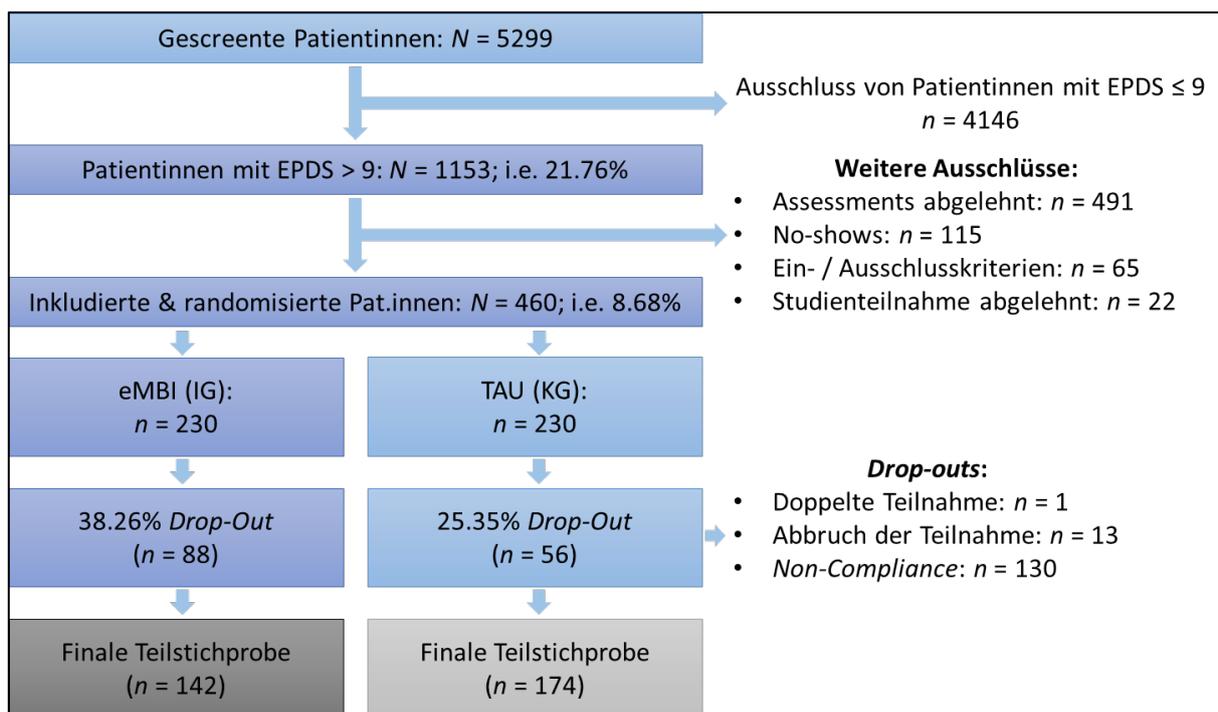
3. Ergebnisse der Evaluation

3.1 Wirksamkeitsevaluation (LMU)

Fallzahl

Die geplante Fallzahl konnten aufgrund gesteigerter Rekrutierungsbemühungen übertroffen werden. Die Rekrutierung wurde nicht unmittelbar mit Einschluss der geplanten Fallzahl gestoppt, um die Evaluation der konfirmatorischen Hypothesen unter der hohen Durchlaufzeit gegen zusätzliche *Dropouts* in den statistischen Analysen abzusichern, die sich skalenspezifisch durch das Fehlen einzelner Werte oder durch Fallausschlüsse nach Erhebungsende ergeben können und die somit potentiell erst im späteren Studienverlauf oder gar erst nach Abschluss der Datenerhebung für jedes Evaluationsmaß und jede konfirmatorische Analyse definitiv bestimmbar sind. Die finale Stichprobe kann dem CONSORT-Flussdiagramm (Abbildung 1) entnommen werden. Von den 547 Frauen, die in den koordinierenden Zentren untersucht wurden, wurden $N = 460$ (84.10%) in die Studie eingeschlossen.

Abbildung 1: CONSORT-Flussdiagramm



13 Teilnehmerinnen haben die Studie abgebrochen, 130 wurden von der Studie, aufgrund Non-Compliance, ausgeschlossen. Für einen Studienausschluss musste einer der folgenden beiden Gründe vorliegen:

1. Kein vollständiges Ausfüllen der 3 für die konfirmatorischen Hypothesen relevanten Fragebögen (EPDS, STAI, PRAQ-R).
2. >2 Module der eMBI wurden nicht absolviert.

Diese beiden Kriterien führten zu einem hohem Ausschluss an Studienteilnehmerinnen, welche jedoch ohnehin nicht für die Auswertung der konfirmatorischen Hypothesen von Nutzen gewesen wären.

Tabelle 8: Patientenstatus - Vergleich Studienzentren

Patientenstatus	HD	TÜ	Gesamt
Abgeschlossen	137	180	317
Abgebrochen	5	8	13
Ausgeschlossen	57	73	130
Gesamt	199	261	460

Aufgrund skalenspezifischer fehlender Werte kann die finale Fallzahl der einzelnen Analysen abweichen. Die spezifischen Auswirkungen auf die jeweiligen Teststärken werden direkt im Anschluss der jeweiligen confirmatorischen Analysen bewertet. Die Teststärken geben an, wie hoch die Wahrscheinlichkeit war, einen Effekt einer bestimmten Größe zu finden, wenn er denn wirklich existiert. Zur Interpretation statistisch nicht bedeutsamer Unterschiede würde es hierzu Werte über $1-\beta \geq .80$ benötigen. Allgemein war die Teststärke in all diesen Analysen für große ($f = 0.40$) Inner- und Zwischensubjekteffekte, sowie für mittlere ($f = 0.25$) Innersubjekteffekt nahezu $1-\beta = 1.0$.

Tabelle 9: Deskriptive Statistiken und Tests auf Vergleichbarkeit der Subgruppen

	Gesamt	TAU	eMBI	t (p)		Gesamt	TAU	eMBI	t (p)
Alter der Patientin (Jahre) ^a <i>M (SD; n)</i>	32.6 (4.6; 316)	32.8 (4.6; 174)	32.3 (4.7; 142)	0.95 (.35)	Geburtsgewicht (g) ^b <i>M (SD; n)</i>	3370.4 (475.9; 308)	3411.8 (438.6; 170)	3319.4 (515.3; 138)	1.70 (.09)
Gestationsalter bei Studien- einschluss (Wochen) ^c <i>M (SD; n)</i>	21.2 (4.3; 309)	21.4 (4.2; 169)	21.0 (4.3; 140)	0.78 (.43)	Größe bei Geburt (cm) ^d <i>M (SD; n)</i>	51.5 (2.6; 302)	51.7 (2.5; 166)	51.2 (2.6; 136)	1.52 (.13)
Gestationsalter bei Geburt (Wochen) ^e <i>M (SD; n)</i>	39.2 (1.7; 292)	39.2 (1.7; 165)	39.2 (1.6; 127)	-0.14 (.89)	PHQ-Stressbewertung (Punkte) ^f <i>M (SD; n)</i>	6.8 (3.4; 315)	7.2 (3.7; 173)	6.3 (3.0; 142)	2.28 (.02)
Bildung der Patientin Häufigkeiten (%)	Gesamt ^g	TAU ^g	eMBI ^h	U (p)	Nettohaushaltseinkommen Häufigkeiten (%)	Gesamt ⁱ	TAU ⁱ	eMBI ⁱ	U (p)
Abitur	149 (51.4)	97 (58.1)	52 (42.3)		< 1.500 €	70 (24.8)	36 (21.8)	34 (29.1)	
Fachhochschulreife	48 (16.6)	29 (17.4)	19 (15.4)		1.500 - 2.999 €	128 (45.4)	76 (46.1)	52 (44.4)	
Mittlere Reife	78 (26.9)	32 (19.2)	46 (37.4)	8396.5 (< .01)	3.000 - 4.999 €	62 (22.0)	39 (23.6)	23 (19.7)	8744.5 (.15)
Volks-/Hauptschule	14 (4.8)	8 (4.8)	6 (4.9)		5.000 - 8.000 €	21 (7.4)	13 (7.9)	8 (6.8)	
Kein Abschluss	1 (0.3)	1 (0.6)	0 (0.0)		> 8.000 €	1 (0.4)	1 (0.6)	0 (0.0)	
Umfang der Beschäftigung Häufigkeiten (%)	General ^k	TAU ^k	eMBI ^k	U (p)	Familienstand Häufigkeiten (%)	General	TAU	eMBI	χ ² (p)
Beschäftigungsverbot	113 (40.9)	68 (42.5)	45 (38.8)		Verheiratet	192 (65.8)	113 (67.3)	79 (63.7)	
Arbeitslos	35 (12.7)	16 (10.0)	19 (16.4)	9150.0 (.84)	Partnerschaft	94 (32.2)	53 (31.5)	41 (33.1)	5.35 (.34)
Teilzeit	60 (21.7)	33 (20.6)	27 (23.3)		Ledig	5 (1.7)	1 (0.6)	4 (3.2)	

Akronym: Mind:Prengancy
 Förderkennzeichen: 01NVF17034

Vollzeit	68 (24.6)	43 (26.9)	25 (21.6)		Geschieden	1 (0.3)	1 (0.6)	0 (0.0)			
Herkunftsland Häufigkeiten (%)	General	TAU	eMBI	χ^2 (p)	Derzeitige Häufigkeiten (%)	Psychotherapie		General	TAU	eMBI	χ^2 (p)
Deutschland	249 (85.3)	140 (83.3)	109 (87.9)	19.17 (.17)	Falsch	55 (17.5)	30 (17.3)	25 (17.7)	0.01 (.99)		
Andere	43 (14.7)	28 (16.7)	15 (12.1)		Wahr	259 (82.5)	143 (82.7)	116 (82.3)			
Gravidität Häufigkeiten (%)	General ^l	TAU ^l	eMBI ^l	U (p)	Parität Häufigkeiten (%)	General ^m	TAU ^m	eMBI ^m	U (p)		
1. Schwangerschaft	143 (45.3)	82 (47.1)	61 (43.0)	11822.0 (.48)	1. Geburt	177 (56.0)	102 (58.6)	75 (52.8)	11675.5 (.35)		
2. Schwangerschaft	85 (26.9)	44 (25.3)	41 (28.9)		2. Geburt	103 (32.6)	53 (30.5)	50 (35.2)			
3. Schwangerschaft	47 (14.9)	29 (16.7)	18 (12.7)		3. Geburt	30 (9.5)	15 (8.6)	15 (10.6)			
≥ 4. Schwangerschaft	41 (13.0)	19 (10.9)	22 (15.5)		≥ 4. Geburt	6 (1.9)	4 (2.3)	2 (1.4)			
Anzahl der Kinder zu Hause Häufigkeiten (%)	General ^m	TAU ^m	eMBI ^m	U (p)	Geburtsmodus Häufigkeiten (%)	General	TAU	eMBI	χ^2 (p)		
Kein Kind	167 (57.4)	105 (62.5)	62 (50.4)	9075.5 ($< .05$)	Spontangeburt	179 (57.9)	103 (60.6)	76 (54.7)	5.35 (.34)		
1 Kind	96 (33.0)	49 (29.2)	47 (38.2)		Primäre C-Sectio	42 (13.6)	23 (13.5)	19 (13.7)			
2 Kinder	25 (8.6)	12 (7.1)	13 (10.6)		Sekundäre C-Sectio	64 (20.7)	30 (17.6)	34 (24.5)			
≥ 3 Kinder	3 (1.0)	2 (1.2)	1 (0.8)		Vaginal-operativ	24 (7.8)	14 (8.2)	10 (7.2)			
Geschlecht des Indexkindes Häufigkeiten (%)	General	TAU	eMBI	χ^2 (p)	APGAR Werte Häufigkeiten (%) (10 Min.)	General ⁿ	TAU ⁿ	eMBI ⁿ	U (p)		

Akronym: Mind: Pregnancy
 Förderkennzeichen: 01NVF17034

					10	279 (92.7)	154 (92.2)	125 (93.3)	
Weiblich	126 (40.6)	72 (42.1)	54 (38.8)		9	16 (5.3)	8 (4.8)	8 (6.0)	
				0.34 (.64)	8	5 (1.7)	4 (2.4)	1 (0.7)	11053.0 (.69)
Männlich	184 (59.4)	99 (57.9)	85 (61.2)		7	1 (0.3)	1 (0.6)	0 (0.0)	

Notizen. t = t -Wert; p = empirischer α -Fehler; M = arithmetisches Mittel; SD = Standardabweichung; U = statistischer Kennwert des U -Tests; χ^2 = statistischer Kennwert des χ^2 -Tests; a. Min = 20.0; Max = 45.0; b. Min = 1,790.0; Max = 4,550.0; c. Min = 12.4; Max = 29.1; d. Min = 41.0; Max = 60.0; e. Min = 30.0; Max = 41.9; f. Min = 0.0; Max = 20.0; g. Median = Abitur; h. median = Fachhochschulreife; i. Median = 1.500 – 2.999 €; k. Median = arbeitslos; l. Median = 2; m. Median = 0; n. Median = 10.

Konfirmatorische Analysen

Voruntersuchung und deskriptive Ergebnisse

In den MCAR-Test wurden soziodemographische, schwangerschafts-, geburts- und kindbezogene, medizinische und psychologische Daten einbezogen. Dieser war nicht signifikant ($\chi^2 = 25\ 104.74$, $df = 25\ 073$, $p = .442$). Es kann somit (mit hoher Sicherheit) davon ausgegangen werden, dass listenweise Fallausschlüsse keine Auswirkungen auf die Ergebnisse haben.

Tabelle 5 enthält alle deskriptiven Charakteristiken zur Beschreibung der Stichprobe, sowie die Tests auf Vergleichbarkeit bzgl. zentraler konfundierender Variablen. Die eMBI und TAU unterschieden sich signifikant hinsichtlich der PHQ-Stressbewertung, der Anzahl der Kinder im Haushalt und hinsichtlich des Bildungsniveaus (siehe Tabelle 5). Die eMBI wies einen niedrigeren Stresswert auf, hatte mehr Kinder sowie einen niedrigeren Bildungsabschluss als die TAU.

Diese potenziell konfundierenden Variablen wurden mit den Endpunkten korreliert: Die PHQ-Stressbewertung war mit jedem Endpunkt zu jeder Visite signifikant assoziiert (Korrelations-Range $r = [.201;.479]$; $p < .001$; Fallzahl-Range $n = [213;301]$). Die Anzahl an Kindern im Haushalt waren signifikant assoziiert mit dem FFA-14, dem STAI-S und -T, als auch mit dem PRAQ-R (Korrelations-Range $\rho = [-.241;.173]$; Signifikanz-Range $p = [< .001; .048]$; Fallzahl-Range $n = [271;290]$). Das Bildungsniveau war signifikant assoziiert mit der EPDS, als auch mit dem STAI-S und -T (Korrelations-Range $\rho = [-.171;-.122]$; Signifikanz-Range $p = [< .017; .041]$, Fallzahl-Range $n = [198;278]$). Somit stellen diese Variablen für die respektiven Endpunkte bedeutsame konfundierende Variablen dar, die somit in den entsprechenden Modellen als Kovariaten kontrolliert wurden.

Im Folgenden werden nun die Ergebnisse des *Manipulation Checks* und der konfirmatorischen Analysen vorgestellt. Die konfirmatorischen Hypothesen zielen darauf ab, dass sich die Mittelwertsverläufe über die Messzeitpunkte hinweg zwischen den Gruppen unterscheiden müssten. Dies würde sich in signifikanten Wechselwirkungen zwischen den Faktoren „Gruppe“ und „Visite“ äußern, weshalb diese Effekte hier betont dargestellt werden.

Manipulation-Check (FFA-14; exploratorische Analysen 14 - 16)

Hierzu wurde eine 2 (Interventionsgruppierung) x 3 (Visite) - ANOVA mit der PHQ-Stressbewertung und der Anzahl der Kinder im Haushalt als Kovariaten verwendet. Die Annahme der Sphärizität wurde signifikant verletzt ($p < .001$) und daher Huynh-Feldt-korrigiert ($\epsilon = .925$).

Es gab einen signifikanten Haupteffekt des Faktors „Visite“ ($F(1.850, 344.164) = 6.543$, $p = .002$, $\eta^2 = .034$, $n = 190$), der auf einen Anstieg der FFA-14-Werte zwischen Visite 1 ($M = 33.0$, $S.E. = 0.5$) und Visite 7 ($M = 35.6$, $S.E. = 0.6$) zurückzuführen ist.

Tabelle 10: FFA-14-Mittel nach Interventionsgruppierung im Studienverlauf

Gruppe	Visite	M	S.E.	95% CI untere Grenze	95% CI obere Grenze
TAU (n = 116)	V1	33.3	0.7	31.9	34.6
	V5	34.0	0.6	32.7	35.2
	V7	35.0	0.7	33.6	36.5
eMBI (n = 74)	V1	32.8	0.9	31.1	34.5
	V5	35.8	0.8	34.2	37.4
	V7	36.2	0.9	34.4	38.0

Notizen. M = Arithmetisches Mittel; S.E. = Standardfehler; CI = Konfidenzintervall; TAU = Treatment as usual; eMBI = electronic mindfulness-based intervention; n = finale Fallzahl; Werte für folgende Kovariaten in dem Modell adjustiert: PHQ-Stressbewertung, Anzahl der Kinder im Haushalt

Der Interaktionseffekt zwischen der Interventionsgruppierung (TAU $n = 116$ vs. eMBI $n = 74$) und dem Faktor „Visite“ war statistisch bedeutsam ($F(1.850, 344.164) = 3.351$, $p = .040$, $\eta^2 = .018$, $n = 190$), was darauf hindeutet, dass die FFA-14-Werte nur in der eMBI zwischen Visite 1 und Visite 7 signifikant anstiegen (Tabelle 9). *Dunn's Post-hoc-Test* ($\Psi_{Dunn} = 1.9$) zeigte zusätzlich einen signifikanten Mittelwertsunterschied zwischen Visite 1 und Visite 5 in der eMBI auf (Tabelle 9, Abbildung 2).

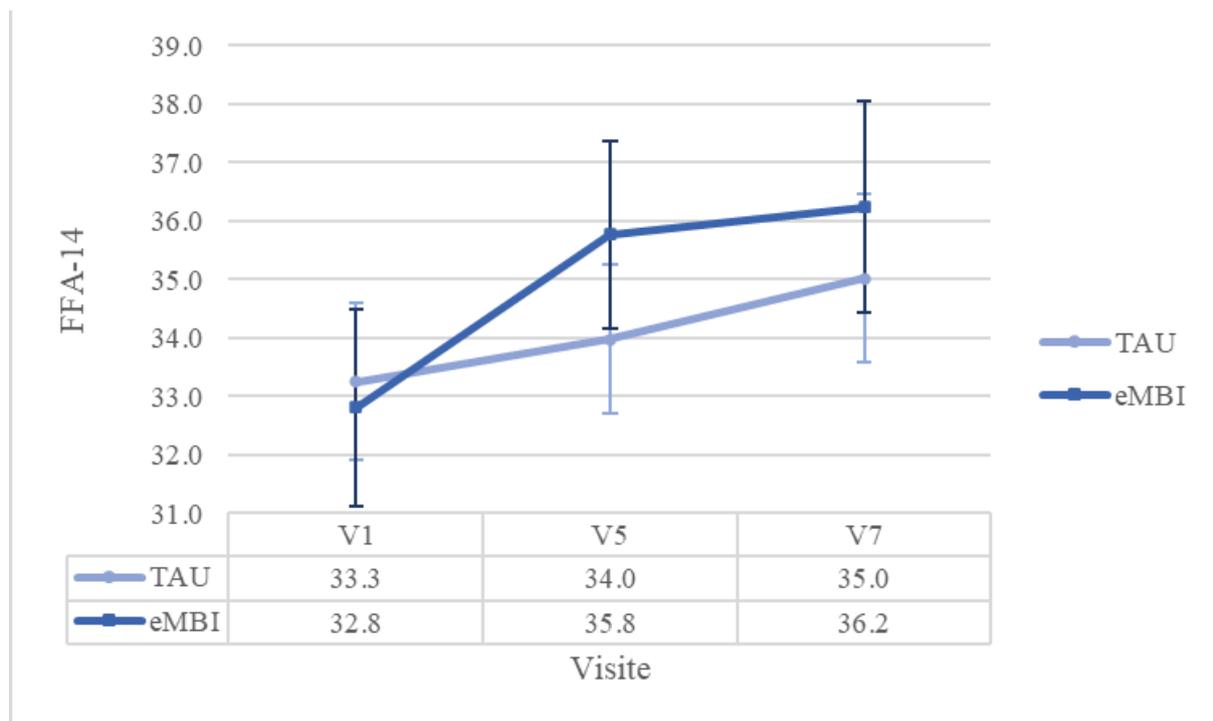


Abbildung 2: FFA-14-Mittel nach Interventionsgruppierung im Studienverlauf

Außerdem wurde ein signifikanter Haupteffekt der PHQ-Stressbewertung ($F(1, 186) = 12.417, p < .001, \eta^2 = .06, n = 190$) festgestellt, der auf einen negativen Zusammenhang zwischen der Stressevaluation und den FFA-14-Werten zurückzuführen ist. Weitere signifikante Haupt- ($p \geq .287, \eta^2 \leq .006, n = 190$) oder Interaktionseffekt ($p \geq .339, \eta^2 \leq .006, n = 190$) wurden nicht aufgedeckt.

Mittlere ($f = 0.25$) Zwischensubjekt-Effekte konnten mit einer Wahrscheinlichkeit von $1-\beta = .940$ und kleine Innersubjekt-Effekte ($f = 0.10$) mit einer Wahrscheinlichkeit von $1-\beta = .852$ aufgedeckt werden. Nur kleine Zwischensubjekt-Effekte konnten mit einer Teststärke von $1-\beta = .205$ im Falle nicht signifikanter Effekte nicht hinreichend ausgeschlossen werden.

Zusammenfassend lässt sich schlussfolgern, dass die Achtsamkeitsintervention im statistischen Mittel tatsächlich die Achtsamkeit erhöht. Je höher allerdings die Stresseinschätzung bewertet wird, desto geringer sind die Werte der Achtsamkeit (unabhängig von der Gruppenzugehörigkeit).

Hauptuntersuchungen

Depressivität (EPDS; konfirmatorische Analysen 1 - 3)

Es wurde eine 2 (Interventionsgruppierung) x 7 (Visite) ANOVA mit der PHQ-Stressbewertung sowie dem Bildungsniveau als Kovariaten verwendet. Die Sphärizitätsannahme wurde signifikant verletzt ($p < .001$) und Huynh-Feldt-korrigiert ($\epsilon = .781$).

Es gab keinen signifikanten Interaktionseffekt zwischen der Interventionsgruppierung (TAU $n = 94$ vs. eMBI $n = 59$) und dem Faktor „Visite“ ($F(4.684, 697.908) = 1.972, p = .086, \eta^2 = .013, n = 153$). Für eine visuelle Betrachtung des Interaktionseffekts siehe Abbildung 3. Für die PHQ-Stressbewertung ($F(1,149) = 36.861, p < .001, \eta^2 = 0.198, n = 153$) und für das Bildungsniveau ($F(1, 149) = 7.231, p = .008, \eta^2 = .046, n = 153$) konnten signifikante Haupteffekte aufgedeckt werden, die auf die signifikanten Assoziationen dieser Maße mit der EPDS zurückzuführen sind.

Darüber hinaus konnten keine weiteren signifikanten Haupt- ($p = .022, \eta^2 = .035, n = 153$) oder Interaktionseffekte ($p \geq 0.544, \eta^2 \leq .005, n = 153$) aufgedeckt werden. Die Teststärke beträgt $1 - \beta = .988$ für mittlere ($f = .25$) Zwischensubjekteffekte. Kleine Inner- und Zwischensubjekt-Effekte ($f = 0.10$) konnten mit einer Teststärke von nur $1-\beta = 0.653$ bzw. $1-\beta = 0.293$ im Falle nicht signifikanter Effekte nicht hinreichend ausgeschlossen werden.

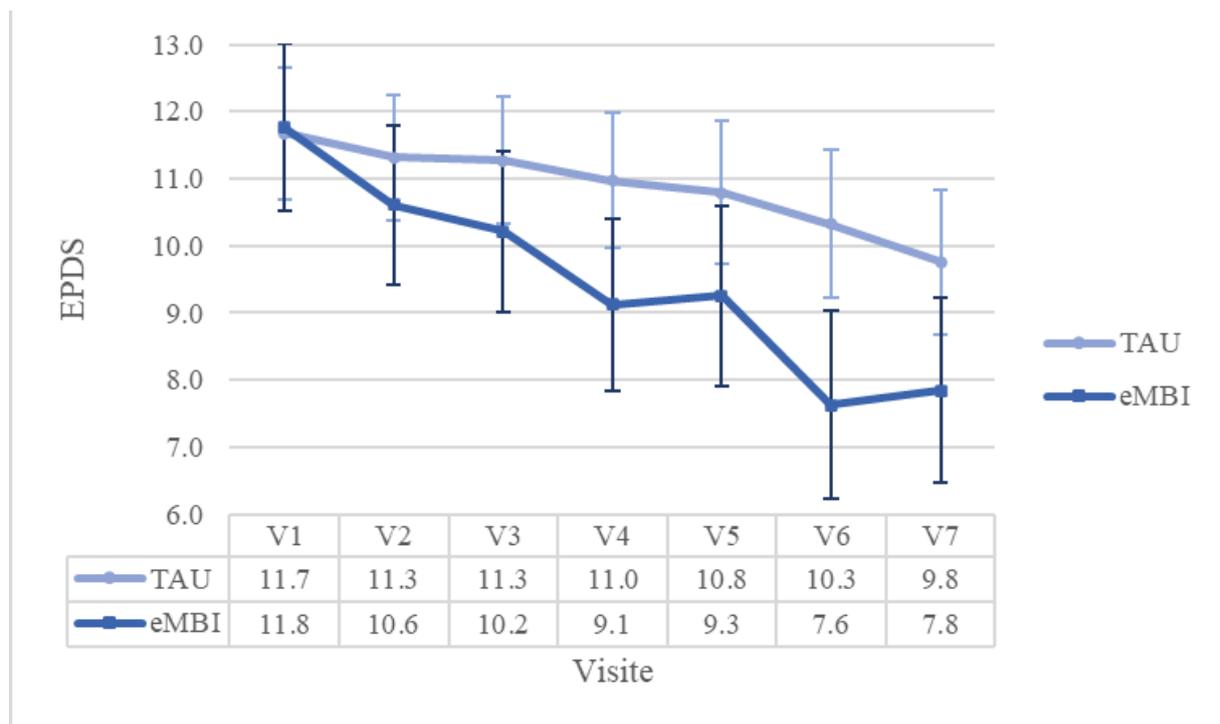


Abbildung 3: EPDS-Mittel nach Interventionsgruppierung im Studienverlauf

Somit lässt sich festhalten, dass die Intervention nicht zu einer Abnahme der depressiven Symptomatik führt; zumindest nicht gemäß den verwendeten statistischen Modellen und Messinstrumente. Je höher allerdings die Stressbewertung und je niedriger das Bildungsniveau sind, desto höher wurden die depressiven Symptome eingeschätzt – unabhängig von der Gruppenzugehörigkeit.

Angstsymptome (STAI; konfirmatorische Analysen 1 - 3)

Es wurde eine 2 (Interventionsgruppierung) x 7 (Visite) MANOVA mit den beiden Subskalen STAI-T und STAI-S als abhängige Variablen, sowie der PHQ-Stressbewertung, der Anzahl der Kinder im Haushalt, und dem Bildungsniveau als Kovariaten verwendet.

Es gab keinen signifikanten Interaktionseffekt zwischen der Interventionsgruppierung (TAU $n = 91$ vs. eMBI $n = 56$) und dem Faktor „Visite“ ($F(12, 131) = 1.314, p = .218, \eta^2 = .107, n = 147$). Für eine visuelle Betrachtung der univariaten Interaktionseffekte siehe Abbildung 4. Es ergab sich ein signifikanter Haupteffekt der PHQ-Stressbewertung ($F(2, 141) = 28.562, p < .001, \eta^2 = .288, n = 147$), der auf den signifikanten positiven Zusammenhang des Maßes mit den STAI-Werten zurückzuführen ist.

Die Ergebnisse waren darüber hinaus negativ für alle anderen Haupt- ($p \geq .122, \eta^2 \leq .029, n = 147$) und Interaktionseffekte ($p \geq .246, \eta^2 \leq .104, n = 147$). Die Teststärke betrug $1 - \beta = .986$ für mittlere Zwischensubjekteffekte ($f = 0.25$). Nur kleine Inner- und Zwischensubjekt-Effekte ($f = 0.10$) konnten mit $1 - \beta = .702$ und $1 - \beta = .287$ im Falle non-signifikanter Ergebnisse mit nicht hinreichender Wahrscheinlichkeit ausgeschlossen werden.

Es lässt sich festhalten, dass die Intervention nicht zu einer Abnahme der Angstsymptomatik führte; zumindest nicht gemäß den verwendeten statistischen Modellen und Messinstrumenten. Je höher allerdings die Stressbewertung ausfiel, desto höher wurden die Angstsymptome eingeschätzt – unabhängig von der Gruppenzugehörigkeit.

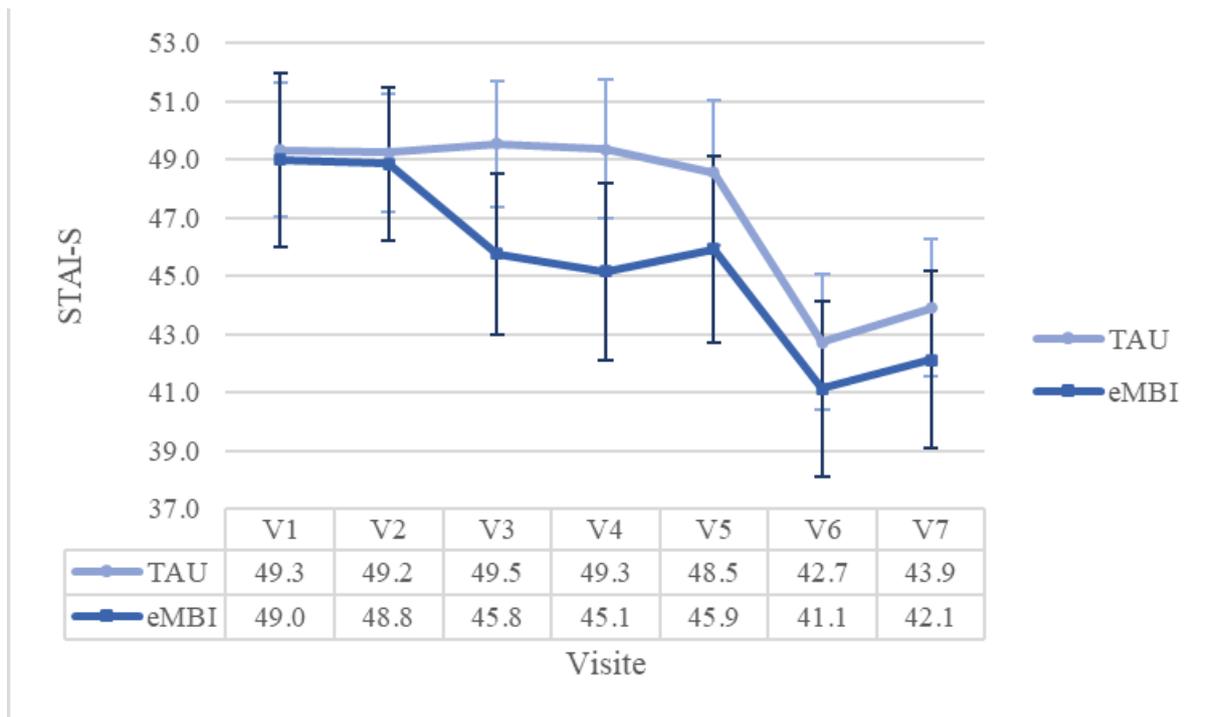


Abbildung 4: STAI-S-Mittel nach Interventionsgruppierung im Studienverlauf

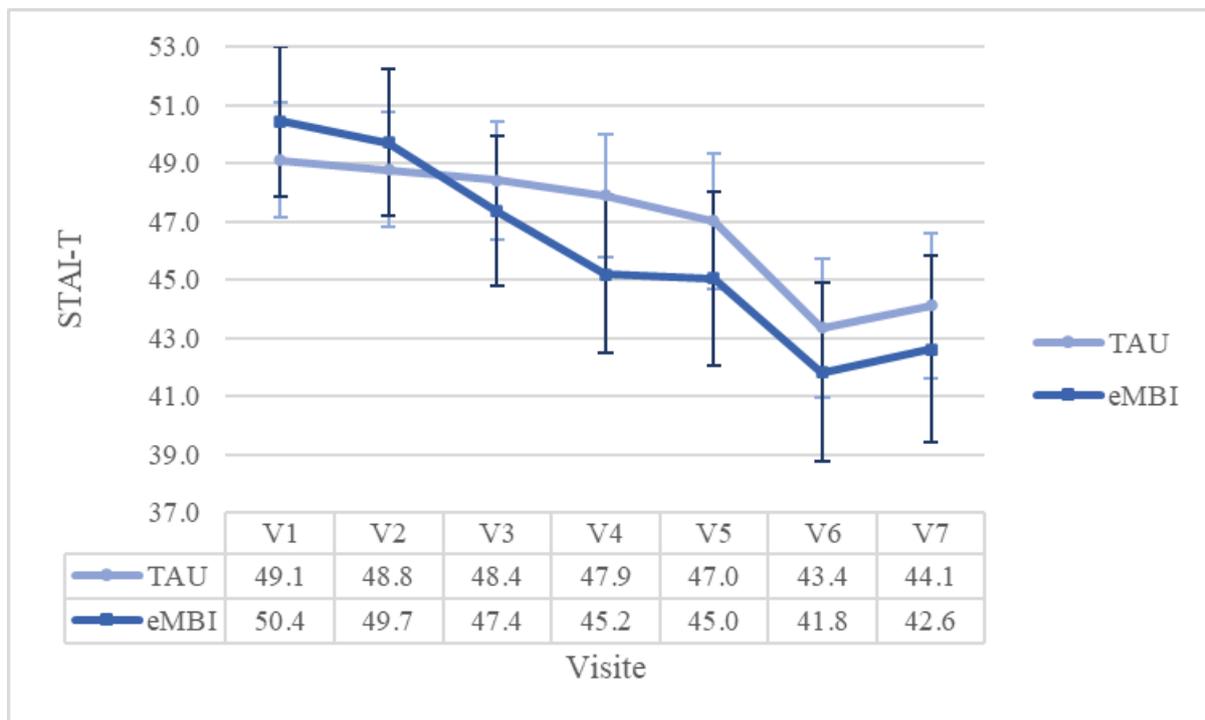


Abbildung 5: STAI-T-Mittel nach Interventionsgruppierung im Studienverlauf

Schwangerschaftsbezogene Ängste (PRAQ-R; konfirmatorische Analysen 4 – 6)

Es wurde eine 2 (Interventionsgruppierung) x 5 (Visite) ANOVA mit der PHQ-Stressbewertung und der Anzahl der Kinder im Haushalt als Kovariaten angewendet. Die Sphärizitätsannahme wurde signifikant verletzt ($p < .001$) und Huynh-Feldt-korrigiert ($\epsilon = .802$).

Es konnte ein signifikanter Interaktionseffekt zwischen der Interventionsgruppierung (TAU $n = 148$ vs. eMBI $n = 87$) und dem Faktor „Visite“ ($F(3.208, 741.004) = 3.558, p = .012, \eta^2 = .015, n = 235$) aufgedeckt werden, was auf sinkende PRAQ-R-Werte in der eMBI zwischen Visite 2 und Visite 4 hindeutete

(Tabelle 10). *Dunn's post-hoc-Test* ($\Psi_{Dunn} = 1.4$) ergab keine weiteren signifikanten Unterschiede (Tabelle 10, Abbildung 6).

Tabelle 11: PRAQ-R-Mittel nach Interventionsgruppierung im Studienverlauf

Gruppe	Visite	M	S.E.	95% CI untere Grenze	95% CI obere Grenze
TAU (n = 148)	V1	25.6	0.6	24.5	26.8
	V2	25.6	0.6	24.4	26.8
	V3	25.5	0.6	24.3	26.7
	V4	25.9	0.6	24.6	27.1
	V5	26.0	0.7	24.7	27.4
eMBI (n = 87)	V1	26.2	0.8	24.7	27.8
	V2	26.5	0.8	25.0	28.1
	V3	26.2	0.8	24.5	27.8
	V4	25.1	0.8	23.4	26.7
	V5	25.2	0.9	23.4	27.0

Notizen. M = Arithmetisches Mittel, S.E. = Standardfehler; CI = Konfidenzintervall; TAU = Treatment as usual; eMBI = electronic mindfulness-based intervention; n = finale Fallzahl; Werte für folgende Kovariaten in dem Modell adjustiert: PHQ-Stressbewertung, Anzahl der Kinder im Haushalt

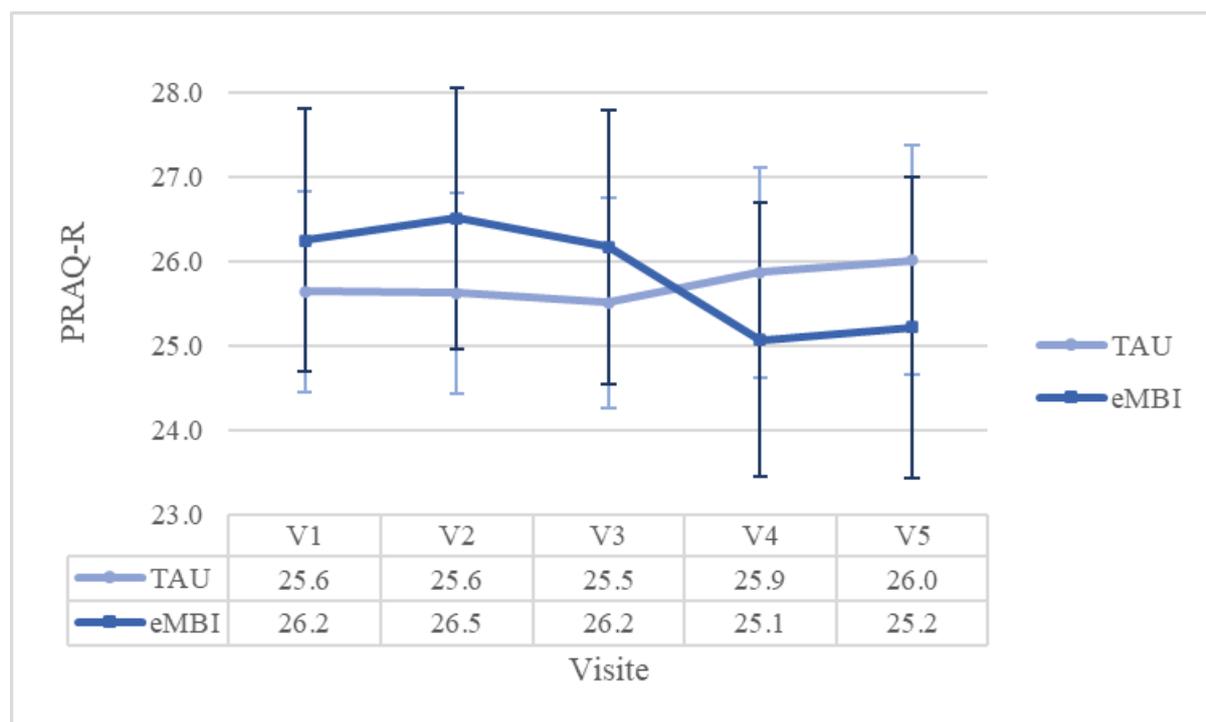


Abbildung 6: PRAQ-R-Mittel nach Interventionsgruppierung im Studienverlauf

Darüber hinaus konnte ein signifikanter Haupteffekt der PHQ-Stressbewertung ($F(1,231) = 18.277, p < .001, \eta^2 = .073, n = 235$) sowie der Anzahl der Kinder im Haushalt ($F(1,231) = 12.438, p < .001, \eta^2 = .051, n = 235$) aufgedeckt werden, welcher sich auf den signifikanten Zusammenhang zwischen diesen Variablen und dem PRAQ-R bezieht.

Es wurden keine weiteren signifikanten Haupt- ($p \geq .832, \eta^2 \leq .001, n = 235$) oder Interaktionseffekte ($p \geq .263, \eta^2 \leq .006, n = 235$) aufgedeckt. Die Teststärke betrug annähernd $1 - \beta = 1.0$ selbst für kleine ($f = 0.10$) Innersubjekt-Effekte sowie $1 - \beta = 0.967$ für mittlere ($f = 0.25$) Zwischensubjekt-Effekte. Nur kleine Zwischensubjekt-Effekte ($f = 0.10$) konnten mit einer Teststärke von nur $1 - \beta = 0.239$ im Falle nicht signifikanter Effekte nicht hinreichend ausgeschlossen werden.

Rate an postpartalen Depressionen (EPDS; Teil der exploratorischen Analysen 17)

Für die Überprüfung der klinischen Auffälligkeit im Rahmen der Depressivität wurden im Rahmen der exploratorischen Analysen (17) 4-Felder- χ^2 -Tests verwendet. Die Raten an Patientinnen, die oberhalb des *Cutoffs* (> 9) auf der EPDS scores zu Visite 1 ($\chi^2 = 0.134, p = .714$), Visite 2 ($\chi^2 = 0.862, p = .353$), Visite 3 ($\chi^2 = 3.579, p = .058$), Visite 4 ($\chi^2 = 3.416, p = .065$), V5 ($\chi^2 = 0.500, p = .480$) und Visite 7 ($\chi^2 = 3.574, p = .059$) scheint nicht im Zusammenhang mit der Intervention zu stehen. Zu V6 hingegen ist die Rate an klinisch auffälligen Frauen in der IG (27.96%) signifikant geringer ($\chi^2 = 7.299, p = .007$) als in der KG (45.97%).

Es handelt sich um messwiederholte nominale Daten. Für zwei Zellen wäre hier die Anwendung eines McNamar-Tests möglich. Dieser gäbe Auskunft über die Veränderung der nominalen Variable innerhalb einer Gruppe. Bei 7 Messzeitpunkten wird die Anzahl von 2 Zellen aber deutlich überschritten, was ebenfalls zu einer wiederholten Anwendung des Tests und somit zu einer Fehlerkumulierung führen würde.

Nun zielt darüber hinaus die Hypothese auf Unterschiede zwischen den Gruppen ab. Mit der Anwendung von 7 χ^2 Tests ist die Fehlerkumulierung im Vergleich zum Vorgehen bei der Anwendung des McNamar-Tests deutlich reduziert.

Auch ist das Problem der Fehlerkumulierung für die nicht-adjustierten kritischen Fehler 1. Art in den exploratorischen Analysen im Methodenteil adressiert. Es handelt sich bei diesen Ergebnissen um exploratorische Analysen, die ausschließlich aufgrund der inhaltlichen passung an dieser Stelle des Berichts eingeordnet wurden. Es wird an entsprechender Stelle nun ausdrücklich auf den exploratorischen Charakter der Analysen hingewiesen. Die Anwendung dieser Analysestrategie war auch bereits im Evaluationskonzept beschrieben. Die Anwendung parametrischer Verfahren ist bei dem Datenniveau invalide. Daher erfolgt keine Anpassung der statistischen Analyse an dieser Stelle es sei denn, die Gutachter:innen haben einen alternativen Vorschlag.

Zusammenfassend lässt sich festhalten, dass die Intervention im statistischen Mittel zur Abnahme der schwangerschaftsbezogenen Ängste führte. Zusätzlich standen höhere Stressbewertungen und eine geringere Anzahl der Kinder im Haushalt im Zusammenhang zu stärkeren Ängsten – unabhängig von der Gruppenzugehörigkeit.

Exploratorische Analysen

Die potenziell konfundierenden Variablen (siehe oben) wurden mit den sekundären Outcomes korreliert: Die PHQ-Stressbewertung war mit der EQ-5D, der MFAS, dem PBQ-16 und dem PEQ-G assoziiert (Korrelations-Range $r = [-.294; .243]$; Signifikanz-Range $p = [< .001; .030]$; Fallzahl-Range $n = [191; 300]$). Das Bildungsniveau war signifikant assoziiert mit dem EQ-5D zu Visite 3 ($p = .157; p = .019; n = 271$) und der MFAS zu Visite 1 ($p = -.129; p = .035; n = 266$). Die Anzahl der Kinder im Haushalt waren signifikant assoziiert mit der EQ-5D, der MFAS, dem PBQ-16 und dem PEQ-G (Korrelations-Range $p = [-.233; .252]$; Signifikanz-Range $p = [< .001; .019]$). Somit stellen diese Variablen für die respektiven Outcomes bedeutsame konfundierende Variablen dar, die somit in den entsprechenden Modellen kontrolliert wurden.

Cortisol (AUC_G; exploratorische Analysen 1 – 3)

Insgesamt haben 214 Probandinnen Cortisolproben zurückgesendet, was einer Rücklaufquote von 46,52% entspricht. Viele Cortisolproben waren nicht verwertbar. Da die Cortisolausschüttung einer zirkadianen Rhythmik unterliegt mit einem morgendlichen Peak, sollte die erste Probe beim aufwachen morgens erfolgen. Diese Uhrzeit wurde von der Probandin dann notiert. Leider wurde hier oftmals eine wesentlich spätere Uhrzeit in der Vormittags- bzw. Mittagszeit notiert, so dass dieser Erstmesswert nicht mehr als reliabel einzuschätzen war.

Es wurde eine 2 (Interventionsgruppierung) x 3 (Visite) ANOVA verwendet. Die Sphäritätsannahme wurde nicht signifikant verletzt ($p = .279$) und somit nicht korrigiert ($\epsilon = 1.0$). Es gab keinen signifikanten Interaktionseffekt zwischen der Interventionsgruppierung (TAU $n = 119$ vs. eMBI $n = 95$) und dem Faktor „Visite“ ($F(2, 424) = 1.665, p = .190, \eta^2 = .008, n = 214$). Der Haupteffekt „Interventionsgruppierung“ (TAU $n = 119$ vs. eMBI $n = 95$) war ebenso nicht signifikant ($F(1, 212) = 0.023, p = .879, \eta^2 < .001, n = 214$). Allerdings war der Haupteffekt „Visite“ statistisch bedeutsam ($F(2,$

424) = 110.609, $p < .001$, $\eta^2 = .343$, $n = 214$), was auf ein Absinken der Werte zwischen Visite 1 ($M = 986.64$ ng/ml, $S.E. = 25.17$) und Visite 7 ($M = 607.30$ ng/ml, $S.E. = 17.18$) zurückzuführen ist. Somit lässt sich festhalten, dass die Intervention nicht zu einer Abnahme des *Cortisoloutputs* führt; zumindest nicht gemäß den verwendeten Modellen und Messinstrumenten.

Gesundheitszustand (EQ-5D exploratorische Analysen 4 - 6)

Es wurde eine 2 (Interventionsgruppierung) x 3 (Visite) ANOVA mit der PHQ-Stressbewertung, dem Bildungsniveau und der Anzahl der Kinder im Haushalt als Kovariaten verwendet. Die Sphäritätsannahme wurde signifikant verletzt ($p < .001$) und somit Huynh-Feldt-korrigiert ($\epsilon = .961$).

Es gab keinen signifikanten Interaktionseffekt zwischen der Interventionsgruppierung (TAU $n = 137$ vs. eMBI $n = 92$) und dem Faktor „Visite“ ($F(1.923, 430.667) = 0.617$, $p = .093$, $\eta^2 = .011$). Allerdings war der Haupteffekt „Visite“ statistisch bedeutsam ($F(1.923, 430.667) = 5.119$, $p = .007$, $\eta^2 = .022$), was auf eine Abnahme der Werte zwischen Visite 1 ($M = 0.72$, $S.E. = 0.01$) und Visite 7 ($M = 0.66$, $S.E. = 0.01$) zurückzuführen ist.

Darüber hinaus konnten für die Kovariaten der PHQ-Stressbewertung ($F(1, 224) = 12.144$, $p < .001$, $\eta^2 = .051$, $n = 229$) und des Bildungsniveaus ($F(1, 224) = 4.703$, $p = .031$, $\eta^2 = .021$, $n = 229$) signifikante Haupteffekte aufgedeckt werden, die auf die signifikanten Assoziation der Variablen mit dem *Outcome* zurückzuführen sind. Des Weiteren existierte ein signifikanter Interaktionseffekt zwischen dem Faktor „Visite“ und dem Bildungsniveau ($F(1.923, 430.667) = 5.311$, $p = .006$, $\eta^2 = 0.02$, $n = 229$), was darauf zurückzuführen ist, dass die Assoziation zum Outcome zwischen den Messzeitpunkten variiert. Darüber hinaus konnten keine weiteren signifikanten Haupt- ($p \geq .594$, $\eta^2 \leq .001$, $n = 229$) oder Interaktionseffekte ($p \geq 0.308$, $\eta^2 \leq .005$, $n = 229$) aufgedeckt werden.

Somit lässt sich festhalten, dass die Intervention nicht zu einem Anstieg des Gesundheitszustandes führt; zumindest nicht gemäß den verwendeten Modellen und Messinstrumenten. Je höher allerdings die Stressbewertung und je niedriger das Bildungsniveau ist, desto niedriger wurde der Gesundheitszustand eingeschätzt – unabhängig von der Gruppenzugehörigkeit.

Präpartales Bonding (MFAS, exploratorische Analysen 7 - 9)

Es wurde eine 2 (Interventionsgruppierung) x 3 (Visite) ANOVA mit der PHQ-Stressbewertung, dem Bildungsniveau und der Anzahl der Kinder im Haushalt als Kovariaten verwendet. Die Sphäritätsannahme wurde nicht signifikant verletzt ($p = .329$) und somit nicht korrigiert ($\epsilon = 1.0$).

Es gab einen signifikanten Interaktionseffekt zwischen der Interventionsgruppierung (TAU $n = 124$ vs. eMBI $n = 78$) und dem Faktor „Visite“ ($F(2, 394) = 4.858$, $p = .024$, $\eta^2 = .008$, $n = 202$), der auf den Anstieg der Werte zwischen Visite 1 und Visite 5 in der TAU zurückzuführen ist. *Dunn's post-hoc-Test* ($\Psi_{Dunn} = 3.5$) ergab zudem signifikante Unterschiede zwischen diesen Werten und denen aus Visite 3 in der KG (Tabelle 11).

Zusätzlich konnten für die PHQ-Stressbewertung ($F(1, 197) = 6.169$, $p = .014$, $\eta^2 = .030$, $n = 202$) und für die Anzahl der Kinder im Haushalt ($F(1, 197) = 5.959$, $p = .016$, $\eta^2 = .029$, $n = 202$) signifikante Haupteffekte aufgedeckt werden, die auf die signifikanten Assoziationen dieser Maße mit dem *Outcome* zurückzuführen sind. Darüber hinaus konnten keine weiteren signifikanten Haupt- ($p \geq .155$, $\eta^2 \leq .009$, $n = 202$) oder Interaktionseffekte ($p \geq 0.444$, $\eta^2 \leq .004$, $n = 202$) aufgedeckt werden.

Tabelle 12: MFAS-Mittel nach Interventionsgruppierung im Studienverlauf

Gruppe	Visite	M	S.E.	95% CI untere Grenze	95% CI obere Grenze
TAU (n = 124)	V1	116.1	1.9	112.3	119.8
	V3	121.2	1.8	117.7	124.8
	V5	125.9	1.8	122.3	129.5
eMBI (n = 78)	V1	119.4	2.4	114.7	124.2
	V3	121.6	2.3	117.1	126.1
	V5	123.8	2.3	119.2	128.3

Notizen. M = Arithmetisches Mittel, S.E. = Standardfehler; CI = Konfidenzintervall; TAU = Treatment as usual; eMBI = electronic mindfulness-based intervention; n = finale Fallzahl; Werte für folgende Kovariaten in dem Modell adjustiert: PHQ-Stressbewertung, Bildungsniveau, Anzahl der Kinder im Haushalt

Somit lässt sich festhalten, dass die Intervention nicht zu einem Anstieg des Präpartalen Bondings führt; zumindest nicht gemäß den verwendeten Modellen und Messinstrumenten. Allerdings stiegen die Werte in der Kontrollgruppe. Je höher darüber hinaus die Stressbewertung und die Anzahl der Kinder im Haushalt sind, desto niedriger wurde das präpartale Bonding eingeschätzt – unabhängig von der Gruppenzugehörigkeit.

Postpartale Beeinträchtigung des Bondings (PBQ-16, exploratorische Analysen 10 – 12)

Es wurde eine 2 (Interventionsgruppierung) x 2 (Visite) ANOVA mit der PHQ-Stressbewertung, dem Bildungsniveau und der Anzahl der Kinder im Haushalt als Kovariaten verwendet. Die Sphärizitätsannahme kann bei nur zwei Stufen auf dem Innersubjektfaktor nicht verletzt werden; die Freiheitsgrade wurden somit nicht korrigiert ($\epsilon = 1.0$).

Es gab keinen signifikanten Interaktionseffekt zwischen der Interventionsgruppierung (TAU $n = 94$ vs. eMBI $n = 63$) und dem Faktor „Visite“ ($F(1, 153) = .047, p = .829, \eta^2 < .001, n = 157$). Allerdings konnte für die PHQ-Stressbewertung ($F(1, 153) = 4.486, p = .036, \eta^2 = .028, n = 157$) ein signifikanter Haupteffekt aufgedeckt werden, der auf die signifikante Assoziation dieses Maßes mit dem Outcome zurückzuführen ist. Des Weiteren existierte ein signifikanter Interaktionseffekt zwischen dem Faktor „Visite“ und der Anzahl der Kinder im Haushalt ($F(1, 153) = 6.227, p = .014, \eta^2 = .039, n = 157$), was darauf zurückzuführen ist, dass die Assoziation zum Outcome zwischen den Messzeitpunkten variiert. Darüber hinaus konnten keine weiteren signifikanten Haupt- ($p \geq .282, \eta^2 \leq .008, n = 157$) oder Interaktionseffekte ($p \geq .072, \eta^2 \leq .021, n = 157$) aufgedeckt werden.

Somit lässt sich festhalten, dass die Intervention nicht zu einem Absinken der postpartalen Beeinträchtigung des Bondings führt; zumindest nicht gemäß den verwendeten Modellen und Messinstrumenten. Je höher aber die Stressbewertung und die Anzahl der Kinder im Haushalt ist, desto höher wurde das präpartale Bonding eingeschätzt – unabhängig von der Gruppenzugehörigkeit.

Patientenerfahrung (PEQ-G; exploratorische Analyse 13)

Für die 5 Skalen wurde je eine ANOVA mit dem Zwischensubjektfaktor „Interventionsgruppierung“ (TAU Fallzahl-Range $n = [110;118]$ vs. eMBI Fallzahl-Range $n = [72;88]$) und den jeweils relevanten Kovariaten verwendet. Insgesamt zeigten sich auf keiner der Skalen Effekte der Intervention ($p \geq .051, \eta^2 \leq .021$). Hingegen waren die Effekte der Kovariaten in den Modellen stets statistisch bedeutsam. So hängt die PHQ Stressbewertung mit der Bewertung der ärztlichen Versorgung ($F(1, 180) = 4.817, p = .029, \eta^2 = .026, n = 199$) und der Betreuung durch die Hebammen ($F(1, 188) = 4.330, p = .039, \eta^2 = .023, n = 210$) zusammen, während die Anzahl der Kinder mit der Bewertung der ärztlichen Versorgung ($F(1, 180) = 3.972, p = .048, \eta^2 = .022, n = 199$), der Betreuung durch die Hebammen ($F(1, 188) = 14.460, p < .001, \eta^2 = .071, n = 210$), der Betreuung durch die Pflege ($F(1, 186) = 6.541, p = .011, \eta^2 = .034, n = 204$) und der Weiterempfehlung ($F(1, 198) = 5.867, p = .016, \eta^2 = .029, n = 218$) zusammenhängt.

Somit lässt sich festhalten, dass die Intervention die positive Bewertung der Patientenerfahrung nicht erhöht; zumindest nicht gemäß den verwendeten Modellen und Messinstrumenten. Je höher allerdings die Stressbewertung, desto niedriger fallen die Bewertungen der ärztlichen Versorgung und der Betreuung durch die Hebammen aus. Je höher die Anzahl der Kinder im Haushalt, desto höher

werden die ärztliche Versorgung, die Betreuung durch die Hebammen und die Pflege und Weiterempfehlung bewertet – all dies unabhängig von der Gruppenzugehörigkeit.

Sectio-Rate, Neonatologie und klinische Auffälligkeit (exploratorische Analysen 17)

Für die Überprüfung der exploratorischen Hypothesen 17 wurden 4-Felder- χ^2 -Tests verwendet. Die Sectio-Rate ($\chi^2 = 1.640, p = .200, \text{TAU } n = 170$ vs. eMBI $n = 139, w^2 = .005$) und die Rate an Vorstellungen beim Kinderarzt nach Entbindung ($\chi^2 = 2.631, p = .105, \text{TAU } n = 167$ vs. eMBI $n = 134, w^2 = .008$) scheinen nicht im Zusammenhang zur Intervention zu stehen. Ebenso verhält es sich mit den Raten an Patientinnen, die oberhalb des *Cutoffs* (> 9) auf der EPDS scoren zu Visite 1 ($\chi^2 = 0.134, p = .714, \text{TAU } n = 168$ vs. eMBI $n = 123, w^2 < .001$), Visite 2 ($\chi^2 = 0.862, p = .353, \text{TAU } n = 168$ vs. eMBI $n = 126, w^2 = .003$), Visite 3 ($\chi^2 = 3.579, p = .058, \text{TAU } n = 170$ vs. eMBI $n = 132, w^2 = .012$), Visite 4 ($\chi^2 = 3.416, p = .065, \text{TAU } n = 168$ vs. eMBI $n = 129, w^2 = .012$), Visite 5 ($\chi^2 = 0.500, p = .480, \text{TAU } n = 168$ vs. eMBI $n = 126, w^2 = .002$) und Visite 7 ($\chi^2 = 3.574, p = .059, \text{TAU } n = 136$ vs. eMBI $n = 111, w^2 = .014$). Zu Visite 6 hingegen ist die Rate an klinisch auffälligen Frauen in der eMBI- ($n = 26, 27.96\%$) signifikant geringer ($\chi^2 = 7.299, p = .007, \text{TAU } n = 124$ vs. eMBI $n = 93, w^2 = .033$) als in der TAU-Gruppe ($n = 57, 45.97\%$).

Somit lässt sich zusammenfassend festhalten, dass die Intervention in keinem Zusammenhang zur Sectio-Rate oder der Rate der Vorstellungen beim Kinderarzt nach Entbindung steht, dass die Intervention allerdings im Zusammenhang zu einer niedrigeren Rate an klinisch auffälligen Frauen zum ersten postpartalen Messzeitpunkt zu stehen scheint.

Moderationsanalysen

Moderationsanalysen sollen Aufschluss darüber geben, ob die Stufen eines weiteren Faktors einen Einfluss auf die Stärke von Effekten haben. Für die Moderationsanalysen wurden die oben aufgeführten konfirmatorischen Modelle und der *Manipulation Check* um die jeweiligen Moderatorvariablen als Zwischensubjektfaktoren erweitert. Wenn ein Moderator Auswirkungen auf einen Effekt zeigt, sollte sich dies in signifikanten Interaktionseffekten von anderen Faktoren oder Wechselwirkungen mit diesem Faktor äußern. Im Folgenden werden nur die Haupt- und Interaktionseffekte, die sich aus dieser Erweiterung ergeben sowie die Zieleffekte im Rahmen der eMBI-Intervention, berichtet. Auf die Effekte von Kovariaten wird inhaltlich nicht mehr näher eingegangen.

Moderation durch klinische Diagnosen

Depressive Störungen wurden durch die klinische Diagnostik im Rahmen des Erstgesprächs mit den ICD-10-Codes F32-F33, Angststörungen mit den ICD-10-Codes F40-F41, somatoforme/dissoziative Störungen mit den ICD-10-Codes F44-F45, F48.1 und F68.0 sowie akute Stressreaktionen mit den ICD-10-Codes F43 und F62 identifiziert. Die Frauen mit und ohne einer solchen definierten klinischen Diagnose wurden jeweils zu einer Gruppe zusammengefasst (mit Diagnose $n = 126$ vs. ohne Diagnose $n = 190$). Die beiden Gruppen (mit vs. ohne Diagnose) unterschieden sich in keiner zu den primären Analysen zusätzlichen und wesentlichen Drittvariable ($p \geq .071, w \leq .010, n = 316$). Darüber hinaus wurde die PHQ-Stressbewertung als Kovariate nicht in Betracht gezogen, da eine starke inhaltliche Nähe zur Konstruktion der diagnostischen Gruppierung besteht.

Manipulation Check (FFA-14)

Hierzu wurde eine 2 (Interventionsgruppierung) x 2 (Diagnostische Gruppierung) x 3 (Visite) - ANOVA mit der Anzahl der Kinder im Haushalt als Kovariate verwendet. Die Annahme der Sphärizität wurde signifikant verletzt ($p < .001$) und daher Huynh-Feldt-korrigiert ($\epsilon = .921$). Die Zellbesetzung der Subgruppen ist in Tabelle 12 dargestellt.

Tabelle 13: Zellbesetzung der Subgruppen in der Moderationsanalyse für den Faktor "Klinische Diagnose" mit der abhängigen Variable FFA-14

	Ohne Diagnose	Mit Diagnose	Gesamt
TAU	69	47	116
eMBI	47	27	74
Gesamt	116	74	190

Notizen. TAU = Treatment as usual; eMBI = electronic mindfulness-based intervention

Es ergab sich ein signifikanter Haupteffekt „Visite“ ($F(1.842, 340.831) = 14.864, p < .001, \eta^2 = .074, n = 190$), der auf einen Anstieg der Achtsamkeit von Visite 1 ($M = 32.5, S.E. = 0.5$) nach Visite 7 ($M = 35.2, S.E. = 0.6$) zurückzuführen ist. Als weiterer signifikanter Faktor stellte sich die diagnostische Gruppierung (ohne Diagnose $n = 116$ vs. mit Diagnose $n = 74$) heraus ($F(1,185) = 16.242, p < .001, \eta^2 = .081, n = 190$), welcher auf den mittleren Unterschied der Werte zwischen den Gruppen mit ($M = 32.1, S.E. = 0.8$) und ohne Diagnose ($M = 36.1, S.E. = 0.6$) zurückführen lässt.

Diese Haupteffekte wurden um eine signifikante Wechselwirkung zwischen der Interventionsgruppierung (TAU $n = 116$ vs. eMBI $n = 74$) und dem Faktor „Visite“ ($F(1.842, 340.831) = 4.934, p = .009, \eta^2 = .026, n = 190$) ergänzt, der auf einen Anstieg der Werte zwischen Visite 1 und Visite 7 in der eMBI zurückzuführen ist (Tabelle 13). *Dunn's Multiple Comparison Procedure* ($\Psi_{Dunn} = 2.3$) wies zudem auf einen Anstieg der Werte zwischen Visite 1 und Visite 5 in der eMBI hin (Tabelle 13). Weitere Mittelwerte unterschieden sich nicht signifikant. Auch konnten keine weiteren Haupt- ($p \geq 158, \eta^2 \leq .011, n = 190$) oder Interaktionseffekte ($p \geq .074, \eta^2 \leq .014, n = 190$) aufgedeckt werden.

Tabelle 14: FFA-14-Mittel nach Interventionsgruppierung im Studienverlauf

Gruppe	Visite	M	S.E.	95% CI untere Grenze	95% CI obere Grenze
TAU ($n = 116$)	V1	32.7	0.7	31.4	34.0
	V5	33.4	0.7	32.2	34.7
	V7	34.3	0.7	32.8	35.7
eMBI ($n = 74$)	V1	32.4	0.9	30.7	34.1
	V5	35.8	0.8	34.2	37.5
	V7	36.1	0.9	34.3	38.0

Notizen. M = Arithmetisches Mittel, S.E. = Standardfehler; CI = Konfidenzintervall; TAU = Treatment as usual; eMBI = electronic mindfulness-based intervention; n = finale Fallzahl; Werte für folgende Kovariaten in dem Modell adjustiert: Anzahl der Kinder im Haushalt

Zusammenfassend lässt sich festhalten, dass die Achtsamkeitsintervention im statistischen Mittel tatsächlich die Achtsamkeit erhöhte. Lag eine klinische Diagnose vor, war die Achtsamkeit im statistischen Mittel geringer (unabhängig von der Zugehörigkeit zur Interventionsgruppierung).

Depressivität (EPDS)

Es wurde eine 2 (Interventionsgruppierung) x 2 (Diagnostische Gruppierung) x 7 (Visite) - ANOVA mit dem Bildungsniveau als Kovariate verwendet. Die Annahme der Sphärizität wurde signifikant verletzt ($p < .001$) und daher Huynh-Feldt-korrigiert ($\epsilon = .784$). Die Zellbesetzung der Subgruppen ist in Tabelle 14 dargestellt.

Tabelle 15: Zellbesetzung der Subgruppen in der Moderationsanalyse für den Faktor "Klinische Diagnose" mit der abhängigen Variable EPDS

	Ohne Diagnose	Mit Diagnose	Gesamt
TAU	55	39	94
eMBI	34	25	59
Gesamt	89	64	153

Notizen. TAU = Treatment as usual; eMBI = electronic mindfulness-based intervention

Es zeigte sich eine signifikante Wechselwirkung zwischen der Interventionsgruppierung (TAU $n = 94$ vs. eMBI $n = 59$) und dem Faktor „Visite“ ($F(4.703, 696.006) = 2.747, p = .021, \eta^2 = .018, n = 153$). Mittels *Dunn's Multiple Comparison Procedure* ($\Psi_{Dunn} = 2.2$) konnten signifikante Gruppenunterschiede zu den Visiten 4, 5, 6 und 7 aufgedeckt werden (Tabelle 15). Darüber hinaus sanken die Werte im Mittel zwischen Visite 1 und Visite 7 in der eMBI, nicht aber in der TAU (Tabelle 15).

Tabelle 16: EPDS-Mittel nach Interventionsgruppierung im Studienverlauf

Gruppe	Visite	<i>M</i>	<i>S.E.</i>	95% <i>CI</i> untere Grenze	95% <i>CI</i> obere Grenze
TAU ($n = 94$)	V1	12.1	0.5	11.1	13.1
	V2	11.8	0.5	10.8	12.7
	V3	11.8	0.5	10.8	12.8
	V4	11.3	0.5	10.3	12.4
	V5	11.3	0.5	10.2	12.3
	V6	11.0	0.6	9.8	12.1
	V7	10.3	0.6	9.2	11.4
eMBI ($n = 59$)	V1	11.7	0.6	10.5	13.0
	V2	10.4	0.6	9.2	11.6
	V3	9.9	0.6	8.6	11.2
	V4	8.9	0.7	7.5	10.2
	V5	8.9	0.7	7.5	10.3
	V6	7.2	0.7	5.7	8.6
	V7	7.6	0.7	6.2	9.0

Notizen. *M* = Arithmetisches Mittel, *S.E.* = Standardfehler; *CI* = Konfidenzintervall; TAU = Treatment as usual; eMBI = electronic mindfulness-based intervention; *n* = finale Fallzahl; Werte für folgende Kovariaten in dem Modell adjustiert: Bildungsniveau

Des Weiteren waren die Faktoren „Interventionsgruppierung“ ($F(1, 148) = 11.372, p < .001, \eta^2 = .071$; TAU $n = 94$ vs. eMBI $n = 59$) und „Diagnostische Gruppierung“ ($F(1, 148) = 11.071, p = .001, \eta^2 = .070$; ohne Diagnose $n = 89$ vs. mit Diagnose $n = 64$) statistisch bedeutsam, was auf die höheren Mittelwerte der TAU ($M = 11.4, S.E. = 0.4$) und der Gruppe mit klinischen Diagnosen ($M = 11.3, S.E. = 0.5$), im Vergleich zur eMBI ($M = 9.2, S.E. = 0.5$) bzw. zur Gruppe ohne klinische Diagnosen ($M = 9.3, S.E. = 0.4$) zurückzuführen ist. Neben einem signifikanten Haupteffekt der Kovariaten ($F(1, 148) = 10.874, p = .001, \eta^2 = .068, n = 153$) gab es keine weiteren statistisch bedeutsamen Effekte ($p \geq .063, \eta^2 \leq .023, n = 153$).

Durch die Hinzunahme des Moderators „Klinische Diagnose“ und der damit verbundenen Bindung zusätzlicher Residualvarianz, scheint die Intervention im statistischen Mittel zu einer Abnahme der depressiven Symptomatik zu führen.

Angstsymptome (STAI)

Es wurde eine 2 (Interventionsgruppierung) x 2 (Diagnostische Gruppierung) x 7 (Visite) - MANOVA mit der Anzahl der Kinder im Haushalt und dem Bildungsniveau als Kovariaten sowie den beiden Subskalen STAI-S und STAI-T als abhängige Variablen verwendet. Die Zellbesetzung der Subgruppen ist in Tabelle 16 dargestellt.

Tabelle 17: Zellbesetzung der Subgruppen in der Moderationsanalyse für den Faktor "Klinische Diagnose" mit der abhängigen Variable STAI

	Ohne Diagnose	Mit Diagnose	Gesamt
TAU	53	38	91
eMBI	33	23	56
Gesamt	86	61	147

Notizen. TAU = Treatment as usual; eMBI = electronic mindfulness-based intervention

Zunächst konnten multivariate Effekte der Interventionsgruppierung ($F(2, 140) = 4.365, p = .014, \eta^2 = .059$, TAU $n = 91$ vs. eMBI $n = 56$), der diagnostischen Gruppierung ($F(2, 140) = 7.587, p < .001, \eta^2 = .098$, ohne Diagnose $n = 86$ vs. mit Diagnose $n = 61$) sowie der dreifachen Wechselwirkung aus den Faktoren „Visite“, „Interventionsgruppierung“ und „Diagnostische Gruppierung“ ($F(12, 130) = 1.840, p = .048, \eta^2 = .145, n = 147$) aufgedeckt werden. Diese Effekte wurden anschließend mit zwei post-hoc 2 (Interventionsgruppierung) x 2 (Diagnostische Gruppierung) x 7 (Visite) - ANOVAs mit der Anzahl der Kinder im Haushalt und dem Bildungsniveau als Kovariaten überprüft. Weitere Effekte waren statistisch nicht bedeutsam ($p \geq .060, \eta^2 \leq .131, n = 147$).

Situationsbezogene Angst (STAI-S)

Die Annahme der Sphärizität wurde signifikant verletzt ($p < .001$) und daher Huynh-Feldt-korrigiert ($\epsilon = .788$). Die Haupteffekte der Interventionsgruppierung ($F(1, 141) = 7.434, p = .007, \eta^2 = .050$, TAU $n = 91$ vs. eMBI $n = 56$) und der diagnostischen Gruppierung ($F(1, 141) = 6.615, p = .011, \eta^2 = .045$, ohne Diagnose $n = 86$ vs. mit Diagnose $n = 61$) waren weiterhin signifikant. Diese Effekte sind auf die höheren Mittelwerte der TAU ($M = 48.7, S.E. = 0.9$) und der Gruppe mit klinischen Diagnosen ($M = 48.5; S.E. = 1.1$), im Vergleich zur eMBI ($M = 44.5, S.E. = 1.2$) bzw. zur Gruppe ohne klinische Diagnosen ($M = 44.7, S.E. = 1.0$) zurückzuführen. Die dreifache Wechselwirkung war in der post-hoc-Analyse nicht mehr statistisch bedeutsam ($F(4.729, 666.749) = 1.814, p = .112, \eta^2 = .013, n = 147$).

Merkmalsbezogene Angst (STAI-T)

Die Annahme der Sphärizität wurde signifikant verletzt ($p < .001$) und daher Greenhouse-Geisser-korrigiert ($\epsilon = .682$). Der Haupteffekt der Interventionsgruppierung ($F(1, 141) = 3.573, p = .061, \eta^2 = .025$, TAU $n = 91$ vs. eMBI $n = 56$) war in der post-hoc-Analyse nicht mehr statistisch bedeutsam, der Effekt der diagnostischen Gruppierung ($F(1,141) = 13.210, p < .001, \eta^2 = .086$, ohne Diagnose $n = 86$ vs. mit Diagnose $n = 61$) hingegen schon. Dieser Effekt ist auf den höheren Mittelwert der Gruppe mit klinischen Diagnosen ($M = 49.7; S.E. = 1.2$) im Vergleich zur Gruppe ohne klinische Diagnosen ($M = 43.9, S.E. = 1.0$) zurückzuführen. Die dreifache Wechselwirkung war in der post-hoc-Analyse nicht mehr statistisch bedeutsam ($F(4.089,576.566) = 0.863, p = .488, \eta^2 = .006, n = 147$).

Zusammenfassend kann also auch durch die Zunahme des Moderators „Klinische Diagnose“ und der damit verbundenen Bindung zusätzlicher Residualvarianz nicht davon ausgegangen werden, dass die Intervention zur Senkung der generellen Angstsymptome führt.

Schwangerschaftsbezogene Ängste (PRAQ-R)

Es wurde eine 2 (Interventionsgruppierung) x 2 (Diagnostische Gruppierung) x 5 (Visite) - ANOVA mit der Anzahl der Kinder im Haushalt als Kovariate verwendet. Die Annahme der Sphärizität wurde signifikant verletzt ($p < .001$) und daher Huynh-Feldt-korrigiert ($\epsilon = .803$). Die Zellbesetzung der Subgruppen ist in Tabelle 17 dargestellt.

Tabelle 18: Zellbesetzung der Subgruppen in der Moderationsanalyse für den Faktor "Klinische Diagnose" mit der abhängigen Variable PRAQ-R

	Ohne Diagnose	Mit Diagnose	Gesamt
TAU	92	57	149
eMBI	53	34	87
Gesamt	145	91	236

Notizen. TAU = Treatment as usual; eMBI = electronic mindfulness-based intervention

Es konnte eine signifikante Wechselwirkung zwischen den Faktoren „Visite“ und „Interventionsgruppierung“ (TAU $n = 149$ vs. eMBI $n = 87$) aufgedeckt werden ($F(3.211, 741.630) = 3.702, p = .010, \eta^2 = .016, n = 236$), welcher auf einen Mittelwertsunterschied zwischen Visite 2 und Visite 4 in der eMBI zurückzuführen ist (Tabelle 18). Mittels *Dunn's Multiple Comparison Procedure* ($\Psi_{Dunn} = 1.5$) konnte ein signifikanter Gruppenunterschied zwischen der TAU und der eMBI zu Visite 4 aufgedeckt werden (Tabelle 18). Weitere Mittelwertsunterschiede waren statistisch nicht bedeutsam (Tabelle 18).

Tabelle 19: PRAQ-Mittel nach Interventionsgruppierung im Studienverlauf

Gruppe	Visite	<i>M</i>	<i>S.E.</i>	95% <i>CI</i> untere Grenze	95% <i>CI</i> obere Grenze
TAU ($n = 149$)	V1	26.2	0.6	25.0	27.5
	V2	26.3	0.6	25.0	27.5
	V3	26.0	0.7	24.7	27.3
	V4	26.4	0.7	25.1	27.7
	V5	26.5	0.7	25.1	27.9
eMBI ($n = 87$)	V1	26.1	0.8	24.5	27.8
	V2	26.5	0.8	24.9	28.1
	V3	26.1	0.9	24.4	27.8
	V4	24.8	0.9	23.2	26.5
	V5	25.1	0.9	23.3	26.9

Notizen. *M* = Arithmetisches Mittel, *S.E.* = Standardfehler; *CI* = Konfidenzintervall; TAU = Treatment as usual; eMBI = electronic mindfulness-based intervention; *n* = finale Fallzahl; Werte für folgende Kovariaten in dem Modell adjustiert: Anzahl der Kinder im Haushalt

Neben einem signifikanten Haupteffekt der Kovariaten ($F(1, 231) = 9.395, p = .002, \eta^2 = .039, n = 236$) konnte ein signifikanter Haupteffekt der diagnostischen Gruppierung aufgedeckt werden ($F(1, 231) = 7.006, p = .009, \eta^2 = .029, n = 236$). Dieser ist auf den Mittelwertsunterschied zwischen der Gruppe mit ($M = 27.3, S.E. = 0.7$) und der Gruppe ohne Diagnose ($M = 24.7, S.E. = 0.6$) zurückzuführen. Darüber hinaus gab es keine weiteren statistisch bedeutsamen Effekte ($p \geq .063, \eta^2 \leq .010, n = 236$).

Durch die Hinzunahme des Moderators „Klinische Diagnose“ und der damit verbundenen Bindung zusätzlicher Residualvarianz, scheint die Intervention im statistischen Mittel weiterhin zu einer Abnahme der schwangerschaftsbezogenen Ängste zu führen.

Moderation durch persönliches Coaching

Während der Analyse auf Parallelität der Effekte zwischen den beiden Studienzentren sind statistisch bedeutsame Unterschiede bezüglich der Wirksamkeit aufgefallen. Hierbei trat die Intervention am Studienzentrum in Heidelberg als wirksamer hervor. Während der Ursachensuche fielen Unterschiede in der Behandlung der Patientinnen je nach Studienstandort auf: Während die Patientinnen in Heidelberg bei Bedarf persönliche Beratungsgespräche durch das Studienpersonal erhielten, war dies in Tübingen nicht der Fall. Somit tat sich ein weiterer Studienfaktor - der des „persönliche Coachings“ - auf. Persönliches Coaching wurde durch die Tatsache identifiziert, dass mehr als ein persönliches Gespräch zum Studieneinschluss zwischen den Studienteilnehmerinnen und den Mitarbeiterinnen in den koordinierenden Zentren stattgefunden hat. Da die Gruppierungen der einzelnen Häufigkeitskategorien klein und die Verteilung stark rechtsschief (alle Quartile der Verteilung fielen auf den Wert 2) ausfielen, wurden die Frauen ohne zusätzliches Coaching ($n = 239$) und Frauen mit mehr als einem Gespräch ($n = 77$) jeweils zu einer Gruppe zusammengefasst.

Die beiden Gruppen unterschieden sich in keiner zu den primären Analysen zusätzlichen und wesentlichen Drittvariable ($p \geq .073$, $\eta^2 \leq .015$, $n = 239$). Somit wurden die konfundierenden Variablen aus der konfirmatorischen Hauptuntersuchung in den entsprechenden Modellen als Kovariaten kontrolliert. Da die Effekte durch diesen Faktor die Zentrums-Effekte deckend aufklären, berichten wir nur die Effekte dieses zur Intervention scheinbar zusätzlich wirksamen Faktors.

Manipulation Check (FFA-14)

Hierzu wurde eine 2 (Interventionsgruppierung) x 2 (Coaching-Gruppierung) x 3 (Visite) - ANOVA mit der PHQ-Stressbewertung und der Anzahl der Kinder im Haushalt als Kovariaten verwendet. Die Annahme der Sphärizität wurde signifikant verletzt ($p < .001$) und daher Huynh-Feldt-korrigiert ($\epsilon = .936$). Die Zellbesetzung der Subgruppen ist in Tabelle 19 dargestellt.

Tabelle 20: Zellbesetzung der Subgruppen in der Moderationsanalyse für den Faktor "Klinische Diagnose" mit der abhängigen Variable FFA-14

	NPC	PC	Gesamt
TAU	90	26	116
eMBI	52	22	74
Gesamt	142	48	190

Notizen. TAU = Treatment as usual; eMBI = electronic mindfulness-based intervention; NPC = Ohne persönliches Coaching; PC = Mit persönlichem Coaching

Es ergab sich ein signifikanter Haupteffekt „Visite“ ($F(1.873, 344.619) = 9.679$, $p < .001$, $\eta^2 = .050$, $n = 190$), der auf einen Anstieg der Achtsamkeit von Visite 1 ($M = 32.5$, $S.E. = 0.6$) nach Visite 7 ($M = 35.8$, $S.E. = 0.7$) zurückzuführen ist. Darüber hinaus war die Wechselwirkung zwischen „Interventionsgruppierung“ (TAU $n = 116$ vs. eMBI $n = 74$) und „Visite“ statistisch signifikant ($F(1.873, 344.619) = 4.560$, $p = .013$, $\eta^2 = .024$, $n = 190$). Mittels *Dunn's Multiple Comparison Procedure* ($\Psi_{Dunn} = 1.9$) konnte ein signifikanter Anstieg der Werte in der eMBI, allerdings auch in der TAU festgestellt werden (Tabelle 20). Jedoch wurde die Achtsamkeit der eMBI zu Visite 5 als höher bewertet (Tabelle 20).

Des Weiteren war die Wechselwirkung zwischen den Faktoren „Coaching-Gruppierung“ (NPC $n = 142$ vs. PC $n = 48$) und „Visite“ statistisch bedeutsam ($F(1.873, 344.619) = 4.585$, $p = .012$, $\eta^2 = .024$, $n = 190$), die auf einen Anstieg der Werte zwischen Visite 1 und Visite 7 in der Gruppe mit Coaching zurückzuführen ist (Tabelle 20). Mittels *Dunn's Multiple Comparison Procedure* ($\Psi_{Dunn} = 2.4$) konnte darüber hinaus ein signifikanter Anstieg der Werte bereits zwischen Visite 1 und Visite 5 in der Gruppe mit Coaching abgesichert werden (siehe Tabelle 20). Im Vergleich dazu stiegen die Werte in der Gruppe ohne Coaching nicht statistisch bedeutsam an (Tabelle 20). Weitere signifikante Mittelwertsunterschiede konnten nicht festgestellt werden (Tabelle 20). Ferner war der Haupteffekt der PHQ-Stressbewertung statistisch bedeutsam ($F(1, 184) = 12.642$, $p < 0.001$, $\eta^2 = 0.064$, $n = 190$). Es konnte keine weiteren signifikanten Effekte aufgedeckt werden ($p \geq .089$, $\eta^2 \leq .013$, $n = 190$).

Tabelle 21: FFA-14-Mittel nach Interventionsgruppierung und Coaching-Gruppierung im Studienverlauf

Gruppe	Visite	M	S.E.	95% CI untere Grenze	95% CI obere Grenze
TAU (n = 116)	V1	32.4	0.8	30.8	34.0
	V5	32.9	0.8	31.4	34.4
	V7	35.0	0.9	33.3	36.7
eMBI (n = 74)	V1	32.5	0.9	30.7	34.4
	V5	36.1	0.9	34.4	37.8
	V7	36.5	1.0	34.6	38.5
NPC (n = 142)	V1	33.6	0.6	32.3	34.8
	V5	35.0	0.6	33.9	36.2
	V7	35.4	0.7	34.1	36.8
PC (n = 48)	V1	31.3	1.0	29.3	33.4
	V5	34.0	1.0	32.0	35.9
	V7	36.1	1.1	33.9	38.4

Notizen. M = Arithmetisches Mittel, S.E. = Standardfehler; CI = Konfidenzintervall; TAU = Treatment as usual; eMBI = electronic mindfulness-based intervention; NPC = Ohne persönliches Coaching; PC = Mit persönlichem Coaching; n = finale Fallzahl; Werte für folgende Kovariaten in dem Modell adjustiert: PHQ-Stressbewertung, Anzahl der Kinder im Haushalt

Es lässt sich zusammenfassen, dass die Achtsamkeitsintervention durch die Hinzunahme des Moderators „persönliches Coaching“ und der damit verbundenen Bindung zusätzlicher Residualvarianz im statistischen Mittel weiterhin die Achtsamkeit erhöhte. Davon unabhängig stieg die Achtsamkeit im Verlauf der Studie v.a. in der Gruppe mit mehreren persönlichen Gesprächen an.

Die q-1-fachen Wechselwirkungen einer q-faktoriellen ANOVA beziehen sich immer auf die Gesamtstichprobe. In der betreffenden Wechselwirkung sind somit sowohl die eMBI als auch die TAU-Patientinnen enthalten. Der Interaktionsterm untersucht das Zusammenspiel der beiden Faktoren „Coaching“ und „Visite“ in der Gesamtstichprobe.

Depressivität (EPDS)

Es wurde eine 2 (Interventionsgruppierung) x 2 (Coaching-Gruppierung) x 7 (Visite) - ANOVA mit der PHQ-Stressbewertung und dem Bildungsniveau als Kovariaten verwendet. Die Annahme der Sphärizität wurde signifikant verletzt ($p < .001$) und daher Huynh-Feldt-korrigiert ($\epsilon = .793$). Die Zellbesetzung der Subgruppen ist in Tabelle 21 dargestellt.

Tabelle 22: Zellbesetzung der Subgruppen in der Moderationsanalyse für den Faktor "Klinische Diagnose" mit der abhängigen Variable FFA-14

	NPC	PC	Gesamt
TAU	73	21	94
eMBI	41	18	59
Gesamt	114	39	153

Notizen. TAU = Treatment as usual; eMBI = electronic mindfulness-based intervention; NPC = Ohne persönliches Coaching; PC = Mit persönlichem Coaching

Es existierte eine statistisch bedeutsamer Haupteffekt der Interventionsgruppierung ($F(1, 147) = 7.845$, $p = .006$, $\eta^2 = .051$, TAU $n = 94$ vs. eMBI $n = 59$), der auf einen bedeutsamen Mittelwertsunterschied zwischen der eMBI ($M = 9.3$, $S.E. = 0.5$) und der TAU ($M = 11.2$, $S.E. = 0.4$) zurückzuführen ist. Des Weiteren war die Wechselwirkung zwischen den Faktoren „Coaching-Gruppierung“ (NPC $n = 114$ vs. PC $n = 39$) und „Visite“ statistisch bedeutsam ($F(4.758, 699.423) = 2.441$; $p = .036$, $\eta^2 = .016$; $n = 153$), was auf das Absinken zwischen Visite 1 und Visite 7 in der Gruppe mit persönlichem Coaching

zurückzuführen ist (Tabelle 22). Mittels *Dunn's Multiple Comparison Procedure* ($\Psi_{Dunn} = 2.6$) konnten keine weiteren signifikanten Mittelwertsunterschiede aufgedeckt werden (Tabelle 22) .

Tabelle 23: EPDS-Mittel nach Interventionsgruppierung und Coaching-Gruppierung im Studienverlauf

Coaching-Gruppierung	Interventionsgruppierung	Visite	M	S.E.	95% CI untere Grenze	95% CI obere Grenze
NPC (n = 114)	TAU (n = 73)	V1	11.3	0.6	10.2	12.4
		V2	10.9	0.5	9.9	12.0
		V3	10.8	0.5	9.7	11.8
		V4	10.4	0.6	9.3	11.5
		V5	10.6	0.6	9.4	11.8
		V6	10.4	0.6	9.1	11.6
		V7	9.9	0.6	8.7	11.1
	eMBI (n = 41)	V1	11.3	0.8	9.8	12.8
		V2	10.6	0.7	9.1	12.0
		V3	10.4	0.7	8.9	11.8
		V4	9.9	0.8	8.4	11.4
		V5	10.6	0.8	9.0	12.2
		V6	7.9	0.9	6.3	9.6
		V7	7.7	0.8	6.0	9.4
	Total	V1	11.3	0.5	10.4	12.2
		V2	10.7	0.4	9.9	11.6
		V3	10.6	0.4	9.7	11.4
		V4	10.2	0.5	9.2	11.1
		V5	10.6	0.5	9.6	11.6
		V6	9.2	0.5	8.1	10.2
		V7	8.8	0.5	7.8	9.8
PC (n = 39)	TAU (n = 21)	V1	13.0	1.0	10.9	15.0
		V2	12.7	1.0	10.7	14.6
		V3	13.0	1.0	11.0	15.0
		V4	13.0	1.1	10.9	15.1
		V5	11.5	1.1	9.3	13.6
		V6	10.1	1.2	7.8	12.4
		V7	9.2	1.2	6.9	11.5
	eMBI (n = 18)	V1	12.7	1.1	10.5	15.0
		V2	10.7	1.1	8.6	12.8
		V3	9.8	1.1	7.7	12.0
		V4	7.3	1.1	5.1	9.6
		V5	6.3	1.2	4.0	8.6
		V6	6.9	1.3	4.4	9.4
		V7	8.2	1.2	5.7	10.6
	Total	V1	12.9	0.8	11.3	14.4
		V2	11.7	0.7	10.2	13.1
		V3	11.4	0.7	10.0	12.9
		V4	10.2	0.8	8.6	11.7
		V5	8.9	0.8	7.3	10.5
		V6	8.5	0.9	6.8	10.2
		V7	8.7	0.8	7.0	10.4

Notizen. *M* = Arithmetisches Mittel, *S.E.* = Standardfehler; *CI* = Konfidenzintervall; TAU = Treatment as usual; eMBI = electronic mindfulness-based intervention; NPC = Ohne persönliches Coaching; PC = Mit persönlichem Coaching; *n* = finale Fallzahl; Werte für folgende Kovariaten in dem Modell adjustiert: PHQ-Stressbewertung, Bildungsniveau

Darüber hinaus war die dreifache Wechselwirkung zwischen den Faktoren „Interventionsgruppierung“, „Coaching-Gruppierung“ und „Visite“ statistisch bedeutsam ($F(4.758, 699.423) = 3.033, p = .012, \eta^2 = .020, n = 153$; Zellbesetzung siehe Tabelle 22). Mittels *Dunn's Multiple Comparison Procedure* ($\Psi_{Dunn} = 4.2$) wurden Gruppenunterschiede zwischen der eMBI und der TAU zu Visite 4 und Visite 5 in der Gruppe mit *Coaching* aufgedeckt (Tabelle 22). Darüber hinaus sanken die Werte in der eMBI der Gruppe mit *Coaching* zwischen Visite 1 und Visite 7 statistisch bedeutsam (Tabelle 22). Weitere bedeutsame Mittelwertsunterschiede konnten nicht ermittelt werden (Tabelle 22).

Ferner war der Haupteffekt der PHQ-Stressbewertung ($F(1,147) = 36.571, p < .001, \eta^2 = .199, n = 153$) und der Haupteffekt des Bildungsniveaus ($F(1,184) = 7.542, p = .007, \eta^2 = .049; n = 153$) statistisch bedeutsam. Es konnte keine weiteren signifikanten Effekte aufgedeckt werden ($p \geq .071, \eta^2 \leq .017, n = 153$).

Die Ergebnisse weisen darauf hin, dass durch die Hinzunahme des Moderators „persönliches *Coaching*“ und die damit verbundene zusätzliche Aufklärung von Residualvarianz, die Intervention im statistischen Mittel zu einer Abnahme der depressiven Symptomatik führt. Dies allerdings in Abhängigkeit davon, ob ein persönliches *Coaching* stattgefunden hat, oder nicht. Auch scheint das *Coaching* an sich bereits eine symptomreduzierende Wirkung zu entfalten.

Angstsymptome (STAI)

Es wurde eine 2 (Interventionsgruppierung) x 2 (Coaching-Gruppierung) x 7 (Visite) - MANOVA mit der PHQ-Stressbewertung, der Anzahl der Kinder im Haushalt und dem Bildungsniveau als Kovariaten sowie den Subskalen STAI-S und STAI-T als abhängige Variablen verwendet. Die Zellbesetzung der Subgruppen ist in Tabelle 23 dargestellt.

Tabelle 24: Zellbesetzung der Subgruppen in der Moderationsanalyse für den Faktor "Klinische Diagnose" mit der abhängigen Variable STAI

	NPC	PC	Gesamt
TAU	70	21	91
eMBI	39	17	56
Gesamt	109	38	147

Notizen. TAU = Treatment as usual; eMBI = electronic mindfulness-based intervention; NPC = Ohne persönliches Coaching; PC = Mit persönlichem Coaching

Zunächst konnten multivariate Effekte der PHQ-Stressbewertung ($F(2, 139) = 28.088, p < .001, \eta^2 = .288, n = 147$), der Interventionsgruppierung ($F(2, 139) = 3.251, p = .042, \eta^2 = .045, TAU n = 91$ vs. eMBI $n = 56$), der zweifachen Wechselwirkung aus den Faktoren „Interventionsgruppierung“ (TAU $n = 91$ vs. eMBI $n = 56$) und „Visite“ ($F(12, 129) = 2.280, p = .012, \eta^2 = .175, n = 147$) sowie der dreifachen Wechselwirkung aus den Faktoren „Visite“, „Interventionsgruppierung“ und „Coaching-Gruppierung“ ($F(12, 129) = 2.361, p = .009, \eta^2 = .180, n = 147$, Zellbesetzung siehe Tabelle 23) aufgedeckt werden. Diese Effekte wurden anschließend mit zwei post-hoc 2 (Interventionsgruppierung) x 2 (Coaching-Gruppierung) x 7 (Visite) - ANOVAs mit der PHQ-Stressbewertung, der Anzahl der Kinder im Haushalt und dem Bildungsniveau als Kovariaten überprüft. Weitere Effekte waren statistisch nicht bedeutsam ($p \geq .231, \eta^2 \leq .107, n = 147$).

Situationsbezogene Angst (STAI-S)

Die Annahme der Sphärizität wurde signifikant verletzt ($p < .001$) und daher Huynh-Feldt-korrigiert ($\epsilon = .817$). Die Haupteffekte der PHQ-Stressbewertung ($F(1, 140) = 43.198, p < .001, \eta^2 = .236, n = 147$) und der Interventionsgruppierung ($F(1, 140) = 3.914, p < .050, \eta^2 = .027, TAU n = 91$ vs. eMBI $n = 56$)

waren weiterhin signifikant. Letzterer Effekt ist auf die höheren Mittelwerte der TAU ($M = 48.1$; $S.E. = 1.0$) im Vergleich zur eMBI ($M = 45.0$, $S.E. = 1.2$) zurückzuführen.

Tabelle 25: STAI-S-Mittel nach Interventionsgruppierung und Coaching-Gruppierung im Studienverlauf

Interventions- gruppierung	Coaching- Gruppierung	Visite	<i>M</i>	<i>S.E.</i>	95% <i>CI</i> untere Grenze	95% <i>CI</i> obere Grenze
TAU (<i>n</i> = 91)	NPC (<i>n</i> = 70)	V1	48.6	1.3	46.0	51.3
		V2	48.4	1.2	46.1	50.7
		V3	48.7	1.2	46.3	51.1
		V4	48.0	1.3	45.4	50.6
		V5	48.0	1.4	45.2	50.8
		V6	42.8	1.4	40.1	45.4
		V7	44.5	1.4	41.9	47.2
	PC (<i>n</i> = 21)	V1	51.4	2.4	46.6	56.2
		V2	52.3	2.1	48.1	56.5
		V3	52.7	2.2	48.3	57.1
		V4	54.0	2.4	49.2	58.8
		V5	50.2	2.6	45.2	55.3
		V6	42.5	2.5	37.6	47.3
		V7	41.7	2.5	36.9	46.6
	Total	V1	50.0	1.4	47.2	52.8
		V2	50.3	1.2	47.9	52.8
		V3	50.7	1.3	48.2	53.2
		V4	51.0	1.4	48.3	53.8
		V5	49.1	1.5	46.2	52.0
		V6	42.6	1.4	39.8	45.4
		V7	43.1	1.4	40.3	45.9
eMBI (<i>n</i> = 56)	NPC (<i>n</i> = 39)	V1	48.7	1.8	45.1	52.3
		V2	48.4	1.6	45.2	51.5
		V3	47.3	1.7	44.0	50.5
		V4	46.9	1.8	43.3	50.5
		V5	48.3	1.9	44.5	52.1
		V6	41.6	1.8	37.9	45.2
		V7	41.2	1.8	37.6	44.8
	PC (<i>n</i> = 17)	V1	49.9	2.7	44.6	55.2
		V2	49.7	2.3	45.0	54.3
		V3	41.8	2.5	37.0	46.7
		V4	40.9	2.7	35.6	46.2
		V5	40.6	2.8	35.0	46.2
		V6	40.3	2.7	34.9	45.7
		V7	44.4	2.7	39.0	49.8
	Total	V1	49.3	1.6	46.1	52.6
		V2	49.0	1.4	46.2	51.8
		V3	44.5	1.5	41.6	47.5
		V4	43.9	1.6	40.7	47.1
		V5	44.4	1.7	41.0	47.8
		V6	41.0	1.7	37.7	44.2
		V7	42.8	1.7	39.5	46.1

Notizen. *M* = Arithmetisches Mittel, *S.E.* = Standardfehler; *CI* = Konfidenzintervall; TAU = Treatment as usual; eMBI = electronic mindfulness-based intervention; NPC = Ohne persönliches Coaching; PC = Mit persönlichem Coaching; *n* = finale Fallzahl; Werte für folgende Kovariaten in dem Modell adjustiert: PHQ-Stressbewertung, Bildungsniveau, Anzahl der Kinder im Haushalt

Tabelle 26: STAI-T-Mittel nach Interventionsgruppierung und Coaching-Gruppierung im Studienverlauf

Interventions- gruppierung	Coaching- Gruppierung	Visite	<i>M</i>	<i>S.E.</i>	95% <i>CI</i> untere Grenze	95% <i>CI</i> obere Grenze
TAU (<i>n</i> = 91)	NPC (<i>n</i> = 70)	V1	48.5	1.1	46.3	50.8
		V2	48.2	1.1	46.0	50.5
		V3	47.7	1.2	45.4	50.0
		V4	46.8	1.2	44.5	49.2
		V5	46.5	1.3	43.9	49.1
		V6	43.7	1.4	41.0	46.5
		V7	44.6	1.4	41.8	47.4
	PC (<i>n</i> = 21)	V1	50.9	2.1	46.8	55.1
		V2	50.4	2.1	46.4	54.5
		V3	50.8	2.1	46.7	55.0
		V4	51.3	2.2	47.0	55.6
		V5	48.7	2.4	44.0	53.5
		V6	41.7	2.5	36.8	46.7
		V7	41.9	2.6	36.8	47.1
	Total	V1	49.7	1.2	47.4	52.1
		V2	49.3	1.2	47.0	51.7
		V3	49.3	1.2	46.9	51.6
		V4	49.1	1.3	46.6	51.6
		V5	47.6	1.4	44.9	50.3
		V6	42.7	1.4	39.9	45.6
		V7	43.3	1.5	40.3	46.2
eMBI (<i>n</i> = 56)	NPC (<i>n</i> = 39)	V1	50.4	1.6	47.3	53.5
		V2	49.1	1.5	46.1	52.2
		V3	47.8	1.6	44.7	50.9
		V4	46.5	1.6	43.3	49.7
		V5	47.5	1.8	44.0	51.0
		V6	42.3	1.9	38.6	45.9
		V7	42.0	1.9	38.2	45.9
	PC (<i>n</i> = 17)	V1	50.7	2.3	46.1	55.2
		V2	51.1	2.3	46.6	55.7
		V3	46.3	2.3	41.7	50.9
		V4	42.3	2.4	37.5	47.0
		V5	39.5	2.6	34.3	44.7
		V6	41.2	2.8	35.7	46.7
		V7	44.6	2.9	38.9	50.3
	Total	V1	50.5	1.4	47.7	53.3
		V2	50.1	1.4	47.4	52.9
		V3	47.1	1.4	44.3	49.9
		V4	44.4	1.5	41.5	47.3
		V5	43.5	1.6	40.3	46.7
		V6	41.7	1.7	38.4	45.1
		V7	43.3	1.8	39.8	46.8

Notizen. *M* = Arithmetisches Mittel, *S.E.* = Standardfehler; *CI* = Konfidenzintervall; TAU = Treatment as usual; eMBI = electronic mindfulness-based intervention; NPC = Ohne persönliches Coaching; PC = Mit persönlichem Coaching; *n* = finale Fallzahl; Werte für folgende Kovariaten in dem Modell adjustiert: PHQ-Stressbewertung, Bildungsniveau, Anzahl der Kinder im Haushalt

Auch die zweifache Wechselwirkung zwischen den Faktoren „Interventionsgruppierung“ (TAU *n* = 91 vs. eMBI *n* = 56) und „Visite“ war weiterhin statistisch bedeutsam ($F(4.902, 686.330) = 3.078, p = .010, \eta^2 = .022, n = 147$). Gemäß *Dunn's Multiple Comparison Procedure* ($\Psi_{Dunn} = 4.6$) konnte zwar in beiden Gruppen eine signifikante Abnahme der Symptomatik zwischen Visite 1 und Visite 7 verzeichnet werden, jedoch lagen die Werte der Interventionsgruppe zu Visite 3, 4 und 5 statistisch bedeutsam unterhalb der Werte der Kontrollgruppe (Tabelle 25). Weitere Mittelwertsunterschiede waren nicht signifikant (Tabelle 25).

Die dreifache Wechselwirkung war in der post-hoc-Analyse ebenfalls weiterhin statistisch bedeutsam ($F(4.902, 686.330) = 4.422, p < .001, \eta^2 = .031, n = 147$, Zellbesetzung siehe Tabelle 26). Gemäß *Dunn's Multiple Comparison Procedure* ($\Psi_{Dunn} = 9.1$) ergaben sich innerhalb des Arms mit *Coaching* in der eMBI signifikant niedrigere Werte zu Visite 3, 4 und 5 im Vergleich zur TAU (siehe Tabelle 14). Darüber hinaus sanken die Werte statistisch bedeutsam in der eMBI des Arms mit *Coaching* zwischen Visite 1 und Visite 7 ab (Tabelle 24). Weitere statistisch bedeutsame Unterschiede konnten nicht ermittelt werden (Tabelle 24).

Merkmalsbezogene Angst (STAI-T)

Die Annahme der Sphärizität wurde signifikant verletzt ($p < .001$) und daher Greenhouse-Geisser-korrigiert ($\epsilon = .691$). Der Haupteffekte der PHQ-Stressbewertung ($F(1, 140) = 56.559, p < .001, \eta^2 = .288, n = 147$) war weiterhin signifikant. Der der Interventionsgruppierung ($F(1, 140) = 0.798, p = .373, \eta^2 = .006, \text{TAU } n = 91 \text{ vs. eMBI } n = 56$) hingegen nicht.

Die zweifache Wechselwirkung zwischen den Faktoren „Interventionsgruppierung“ (TAU *n* = 91 vs. eMBI *n* = 56) und „Visite“ war weiterhin statistisch bedeutsam ($F(4.146, 580.459) = 2.923, p = .019, \eta^2 = .020$). Gemäß *Dunn's Multiple Comparison Procedure* ($\Psi_{Dunn} = 4.2$) konnte zwar in beiden Gruppen eine Abnahme der Symptomatik zwischen Visite 1 und Visite 7 verzeichnet werden, jedoch lagen die Werte der eMBI zu Visite 4 unterhalb der Werte der TAU (Tabelle 25). Weitere Mittelwertsunterschiede waren statistisch nicht bedeutsam (Tabelle 25).

Auch die dreifache Wechselwirkung war in der post-hoc-Analyse weiterhin signifikant ($F(4.146, 580.459) = 4.689, p < .001, \eta^2 = .032, n = 147$; Zellbesetzung siehe Tabelle 26). Gemäß *Dunn's Multiple Comparison Procedure* ($\Psi_{Dunn} = 8.2$) ergaben sich innerhalb des Arms mit *Coaching* in der eMBI signifikant niedrigere Werte zu Visite 4 und Visite 5 im Vergleich zur TAU. Darüber hinaus sanken die Werte statistisch bedeutsam in der TAU des Arms mit *Coaching* und in der eMBI des Arms ohne *Coaching* zwischen Visite 1 und Visite 7 ab. Weitere signifikante Unterschiede konnten nicht ermittelt werden (Tabelle 25).

Zusammenfassend konnte gezeigt werden, dass die Intervention durch die Zunahme des Moderators „persönliches Coaching“ und der damit verbundenen Aufklärung zusätzlicher Residualvarianz im statistischen Mittel zu einer Abnahme der Angst-Symptomatik führen kann, dies v.a. abhängig davon, ob ein Coaching stattgefunden hat, oder nicht. Hierbei sind die Effekte bzgl. der merkmalsbezogenen Angst allerdings weniger eindeutig.

Schwangerschaftsbezogene Ängste (PRAQ-R)

Es wurde eine 2 (Interventionsgruppierung) x 2 (*Coaching*-Gruppierung) x 5 (Visite) - ANOVA mit der PHQ-Stressbewertung und der Anzahl der Kinder im Haushalt als Kovariaten verwendet. Die Annahme der Sphärizität wurde signifikant verletzt ($p < .001$) und daher Huynh-Feldt-korrigiert ($\epsilon = .809$). Die Zellbesetzung der Subgruppen ist in Tabelle 26 dargestellt.

Neben signifikanten Effekten der Kovariaten (PHQ-Stressbewertung: $F(1,229) = 18.862, p < .001, \eta^2 = .076, n = 235$; Anzahl der Kinder im Haushalt: $F(1, 229) = 12.537, p < .001, \eta^2 = .052, n = 235$) konnten keine weiteren statistisch bedeutsamen Effekte aufgedeckt werden ($p \geq .102, \eta^2 \leq .012, n = 235$).

Somit ist durch die Hinzunahme des Moderators „persönliches Coaching“ und der damit verbundenen zusätzlichen Aufklärung von Residualvarianz die Intervention im Hinblick auf schwangerschaftsbezogene Ängste als nicht wirksam zu beurteilen.

Tabelle 27: Zellbesetzung der Subgruppen in der Moderationsanalyse für den Faktor "Klinische Diagnose" mit der abhängigen Variable PRAQ-R

	NPC	PC	Gesamt
TAU	115	33	148
eMBI	64	23	87
Gesamt	179	56	235

Notizen. TAU = Treatment as usual; eMBI = electronic mindfulness-based intervention; NPC = Ohne persönliches Coaching; PC = Mit persönlichem Coaching

Mehrebenenanalysen

Im Folgenden werden nur die Zieleffekte im Rahmen der Evaluation der eMBI, berichtet. Auf die Effekte von Kovariaten wird inhaltlich nicht mehr näher eingegangen. Die Fallzahlen n werden listenweise angegeben.

Manipulation Check (FFA-14)

Es konnte ein signifikanter Haupteffekt des Faktors „Visite“ aufgedeckt werden ($F(2, 237.120) = 3.037, p < .050, \eta^2 = .025, n = 190$), der auf den mittleren Anstieg der Werte zwischen Visite 1 ($M = 33.5, S.E. = 0.4$) und Visite 7 ($M = 36.0, S.E. = 0.5$) zurückzuführen ist. Darüber hinaus war die Wechselwirkung zwischen der Interventionsgruppierung (TAU $n = 116$ vs. eMBI $n = 74$) und dem Faktor „Visite“ statistisch bedeutsam ($F(2, 238.701) = 7.095, p = .001, \eta^2 = .056, n = 190$). Zur Interpretation dieses Effekts wurde die Methode von Cumming (2009) verwendet. Mit dieser Methode wird der statistisch bedeutsame Unterschied zwischen zwei Punktschätzern dadurch bestimmt, wenn die korrespondierenden Konfidenzintervalle nicht mehr als 50% überlappen. Gemäß dieser Methode zeigt sich ein statistisch bedeutsamer Gruppenunterschied zwischen der eMBI und der TAU zu Visite 5 sowie ein signifikanter Anstieg der mittleren Werte zwischen Visite 1 und Visite 7 in der eMBI (Tabelle 27). Es konnten keine weiteren statistisch bedeutsamen Mittelwertsunterschiede aufgedeckt werden.

Tabelle 28: FFA-14-Mittel nach Interventionsgruppierung im Studienverlauf

Visite	Gruppe	n	M	$S.E.$	95% CI untere Grenze	95% CI obere Grenze
v1 ($n = 190$)	TAU	116	33.5	0.5	32.4	34.5
	eMBI	74	33.5	0.6	32.3	34.8
v5 ($n = 190$)	TAU	116	34.0	0.5	33.0	35.0
	eMBI	74	36.5	0.6	35.3	37.7
v7 ($n = 190$)	TAU	116	35.2	0.6	33.9	36.4
	eMBI	74	36.8	0.8	35.3	38.2

Notizen. M = Arithmetisches Mittel, $S.E.$ = Standardfehler; CI = Konfidenzintervall; TAU = Treatment as usual; eMBI = electronic mindfulness-based intervention; n = finale Fallzahl; Werte für folgende Kovariaten in dem Modell adjustiert: PHQ-Stressbewertung, Anzahl der Kinder im Haushalt

Daneben existierte nur noch ein weiterer Haupteffekt der Kovariaten der PHQ-Stressbewertung ($F(1, 289.287) = 16.697, p < .001, \eta^2 = .055, n = 190$). Weitere statistisch bedeutsame Effekte konnten nicht aufgedeckt werden ($p \geq .052, \eta^2 \leq .025, n = 190$).

Depressivität (EPDS)

Es konnte ein signifikanter Haupteffekt der Interventionsgruppierung (TAU $n = 94$ vs. eMBI $n = 59$) aufgedeckt werden ($F(1, 281.449) = 4.066, p = .045, \eta^2 = .014, n = 153$), der auf die im Mittel erhöhten Werte der TAU ($M = 10.5, S.E. = 0.3$) im Vergleich zur eMBI ($M = 9.6, S.E. = 0.3$) zurückzuführen ist.

Daneben existierten nur noch Haupteffekte der Kovariaten der PHQ-Stressbewertung ($F(1, 279.713) = 62.892, p < .001, \eta^2 = .184, n = 153$) und des Bildungsniveaus ($F(1, 281.962) = 4.512, p = .035, \eta^2 = .016, n = 153$). Weitere statistisch bedeutsame Effekte konnten nicht aufgedeckt werden ($p \geq .168, \eta^2 \leq .034, n = 153$).

Situationsbezogene Angst (STAI-S)

Es existierte lediglich ein Haupteffekt der Kovariaten der PHQ-Stressbewertung ($F(1, 276.614) = 65.969, p < .001, \eta^2 = .193, n = 147$). Weitere statistisch bedeutsame Effekte konnten nicht aufgedeckt werden ($p \geq .065, \eta^2 \leq .044, n = 147$).

Merkmalsbezogene Angst (STAI-T)

Es existierte lediglich ein Haupteffekt der Kovariaten der PHQ-Stressbewertung ($F(1, 280.032) = 81.215, p < .001, \eta^2 = .225, n = 147$) sowie ein Interaktionseffekt zwischen der Visite und der PHQ-Stressbewertung ($F(6, 246.184) = 2.894, p = .010, \eta^2 = .012, n = 147$). Weitere statistisch bedeutsame Effekte konnten nicht aufgedeckt werden ($p \geq .234, \eta^2 \leq .029, n = 147$).

Schwangerschaftsbezogene Ängste (PRAQ-R)

Es existierten lediglich Haupteffekte der Kovariaten der PHQ-Stressbewertung ($F(1, 286.109) = 28.849, p < .001, \eta^2 = .092, n = 235$) sowie der Anzahl der Kinder im Haushalt ($F(1, 286.692) = 15.583, p = .010, \eta^2 = .052, n = 235$). Weitere statistisch bedeutsame Effekte konnten nicht aufgedeckt werden ($p \geq .237, \eta^2 \leq .020, n = 235$).

Zusammenfassend konnten die Mehrebenenanalysen keine über die primären Analysen hinausgehende Wirksamkeit der Intervention nachweisen.

Multiple Imputationen

Im Folgenden werden nur die *gepoolten* Ergebnisse der Zieleffekte im Rahmen der Evaluation der eMBI berichtet. Auf die Effekte von Kovariaten wird inhaltlich nicht mehr näher eingegangen.

Manipulation Check (FFA-14)

Hierzu wurde eine 2 (Interventionsgruppierung) x 3 (Visite) - ANOVA mit der PHQ-Stressbewertung und der Anzahl der Kinder im Haushalt als Kovariaten verwendet. Die Annahme der Sphärizität wurde signifikant verletzt ($p < .001$) und daher Huynh-Feldt-korrigiert ($\epsilon = .940$).

Es konnte ein statistisch bedeutsamer Interaktionseffekt zwischen der Interventionsgruppierung (TAU $n = 174$ vs. eMBI $n = 142$) und dem Faktor „Visite“ aufgedeckt werden ($F(1.881, 575.636) = 5.470, p = .012, \eta^2 = .018, n = 316$). *Dunn's Post-hoc-Test* ($\Psi_{Dunn} = 1.3$) deutet auf signifikante Anstiege der Achtsamkeit sowohl in der eMBI als auch in der TAU hin. Jedoch lagen die Werte der eMBI im statistischen Mittel zu Visite 5 und 7 signifikant über denen der TAU (Tabelle 28). Darüber hinaus existierte nur noch ein signifikanter Haupteffekt der PHQ-Stressbewertung ($F(1, 306) = 17.419, p < .001, \eta^2 = .054, n = 316$). Weitere statistisch bedeutsame Effekte konnten nicht aufgedeckt werden ($p \geq .066, \eta^2 \leq .013, n = 316$).

Tabelle 29: Imputierte FFA-14-Mittel nach Interventionsgruppierung im Studienverlauf

Gruppe	Visite	M	S.E.	95% CI untere Grenze	95% CI obere Grenze
TAU ($n = 174$)	V1	33.5	0.5	32.4	34.5
	V5	34.0	0.5	33.0	35.0
	V7	35.1	0.6	33.8	36.3
eMBI ($n = 142$)	V1	33.7	0.6	32.5	34.8
	V5	36.4	0.6	35.3	37.6
	V7	36.5	0.7	35.1	37.8

Notizen. M = Arithmetisches Mittel, S.E. = Standardfehler; CI = Konfidenzintervall; TAU = Treatment as usual; eMBI = electronic mindfulness-based intervention; n = finale Fallzahl; Werte für folgende Kovariaten in dem Modell adjustiert: PHQ-Stressbewertung, Anzahl der Kinder im Haushalt

Depressivität (EPDS)

Es wurde eine 2 (Interventionsgruppierung) x 7 (Visite) ANOVA mit der PHQ-Stressbewertung sowie dem Bildungsniveau als Kovariaten verwendet. Die Sphäritätsannahme wurde signifikant verletzt ($p < .001$) und Huynh-Feldt-korrigiert ($\epsilon = .821$).

Es gab keinen signifikanten Interaktionseffekt zwischen der Interventionsgruppierung (TAU $n = 174$ vs. eMBI $n = 142$) und dem Faktor „Visite“ ($F(4.924, 1499.815) = 2.030, p = .106, \eta^2 = .007, n = 316$). Allerdings konnte ein statistisch bedeutsamer Haupteffekt der Gruppierungsvariable (TAU $n = 174$ vs. eMBI $n = 142$) aufgedeckt werden ($F(1, 304.238) = 4.251, p = .047, \eta^2 = .014$), der auf den in der TAU ($M = 10.5, S.E. = 0.3$) im Vergleich zur eMBI ($M = 9.7, S.E. = 0.3$) erhöhten Gesamtmittelwert zurückzuführen ist.

Darüber hinaus konnten nur noch statistisch bedeutsame Haupteffekte der Kovariaten aufgedeckt werden (PHQ-Stressbewertung: $F(1, 304.238) = 70.161, p < .001, \eta^2 = 0.188, n = 316$; Bildungsniveau: $F(1, 304.238) = 6.088, p = .018, \eta^2 = .020, n = 316$). Darüber hinaus existierten keine weiteren signifikanten Effekte ($p \geq 0.273, \eta^2 \leq .005, n = 316$).

Angstsymptome (STAI)

Es wurde eine 2 (Interventionsgruppierung) x 7 (Visite) MANOVA über die Subskalen STAI-S und STAI-T als abhängige Variablen, mit der PHQ-Stressbewertung, der Anzahl der Kinder im Haushalt sowie dem Bildungsniveau als Kovariaten verwendet. Es ergab sich lediglich ein signifikanter Haupteffekt der PHQ-Stressbewertung ($F(2, 301.952) = 46.344, p < .001, \eta^2 = .236, n = 316$), sowie eine statistisch bedeutsame Wechselwirkung zwischen der Stressbewertung und dem Faktor „Visite“ ($F(12, 291.952) = 2.405, p = .020, \eta^2 = .090, n = 316$). Aufgrund dessen, dass diese Effekte im Hinblick auf die Zielhypothesen keine inhaltliche Bedeutung besitzen, wurde auf eine post-hoc Analyse dieser Effekte verzichtet. Die Ergebnisse waren negativ für alle anderen Haupt- ($p \geq .065, \eta^2 \leq .074, n = 316$) und Interaktionseffekte ($p \geq .059; \eta^2 \leq .083; n = 316$).

Schwangerschaftsbezogene Ängste (PRAQ-R)

Es wurde eine 2 (Interventionsgruppierung) x 5 (Visite) ANOVA mit der PHQ-Stressbewertung und der Anzahl der Kinder im Haushalt als Kovariaten angewendet. Die Sphäritätsannahme wurde signifikant verletzt ($p < .001$) und Huynh-Feldt-korrigiert ($\epsilon = .823$).

Es konnte ein signifikanter Interaktionseffekt zwischen der Interventionsgruppierung (TAU $n = 174$ vs. eMBI $n = 142$) und dem Faktor „Visite“ ($F(3.293, 1015.189) = 3.271, p = .026, \eta^2 = .011, n = 316$) aufgedeckt werden. *Dunn's post-hoc-Test* ($\Psi_{Dunn} = 1.1$) zeigte signifikant höhere mittlere Werte der TAU im Vergleich zur eMBI zu den Visiten 4 und 5 auf (Tabelle 29). Es ergaben sich keine weiteren statistisch bedeutsamen Unterschiede, allerdings stiegen die Mittelwerte in der TAU zwischen Visite 1 und 5 deskriptiv an, während sie in der eMBI deskriptiv absanken (Tabelle 29).

Darüber hinaus konnten signifikante Haupteffekt der Kovariaten aufgedeckt werden (PHQ-Stressbewertung: $F(1, 308.143) = 27.791, p < .001, \eta^2 = .083, n = 316$; Anzahl der Kinder im Haushalt: $F(1, 308.143) = 15.355, p < .001, \eta^2 = .047, n = 316$). Es wurden keine weiteren signifikanten Haupt- ($p \geq .483, \eta^2 \leq .003, n = 316$) oder Interaktionseffekte ($p \geq .104, \eta^2 \leq .007, n = 316$) aufgedeckt.

Tabelle 30: Imputierte PRAQ-R-Mittel nach Interventionsgruppierung im Studienverlauf

Gruppe	Visite	M	S.E.	95% CI untere Grenze	95% CI obere Grenze
TAU ($n = 174$)	V1	25.8	0.5	24.7	26.8
	V2	25.8	0.6	24.8	26.9
	V3	25.6	0.6	24.5	26.7
	V4	25.8	0.6	24.7	27.0
	V5	26.1	0.6	24.9	27.3
eMBI ($n = 142$)	V1	25.8	0.6	24.6	27.0
	V2	25.7	0.6	24.5	26.9

Akronym: Mind:Pregnancy
Förderkennzeichen: 01NVF17034

V3	25.4	0.6	24.1	26.6
V4	24.6	0.6	23.4	25.8
V5	24.7	0.7	23.4	26.1

Notizen. M = Arithmetisches Mittel, S.E. = Standardfehler; CI = Konfidenzintervall; TAU = Treatment as usual; eMBI = electronic mindfulness-based intervention; n = finale Fallzahl; Werte für folgende Kovariaten in dem Modell adjustiert: PHQ-Stressbewertung, Anzahl der Kinder im Haushalt

Zusammenfassend konnten die Multiplen Imputationen keine über die primären Analysen hinausgehende Wirksamkeit der Intervention nachweisen.

Geplante, aber nicht durchgeführte Analysen

In Übereinstimmung mit der ICH-Richtlinie E9 (ICH, 1998) sollten die primären Analysen gemäß dem *Intention-To-Treat-Principle* (ITT) durchgeführt werden. Da jedoch bei nur einer Probandin ein Protokollverstoß (doppelte Teilnahme) festgestellt werden konnte, wurden nach Rücksprache mit der Konsortialleitung alle Analysen unter Ausschluss der doppelten Teilnahme durchgeführt (*Per-Protocol*).

Des Weiteren sollten gefundene Zusammenhänge einer explorativen Kausalprüfung unterzogen werden, was mittels des *Cross-lagged-panel-Designs* der Studie (Outcomes & Störvariablen werden zu mehreren Messzeitpunkten erhoben) mithilfe der Verwendung von *Cross-lagged-panel-models* (CLPMs) hätte erfolgen können. Jedoch ist die Verwendung von CLPMs gerade bei genesteten Designs (wie z.B. dem vorliegenden messwiederholten Design) in jüngster Zeit in die Kritik geraten, da die mit CLPMs untersuchten Effekte massiv überschätzt werden können (Lucas, 2022). Darüber hinaus scheinen zum derzeitigen Stand der über das Evaluationskonzept hinausgehenden exploratorischen Analysen keine in der Studie erfassten messwiederholten *Confounder* zu existieren, die aufgrund mehr oder weniger zeitstabilen Zusammenhängen zu den primären Outcomes einer Kontrolle durch CLPMs bedürften. Aus diesen Gründen wurde von der Umsetzung der CLPMs abgesehen.

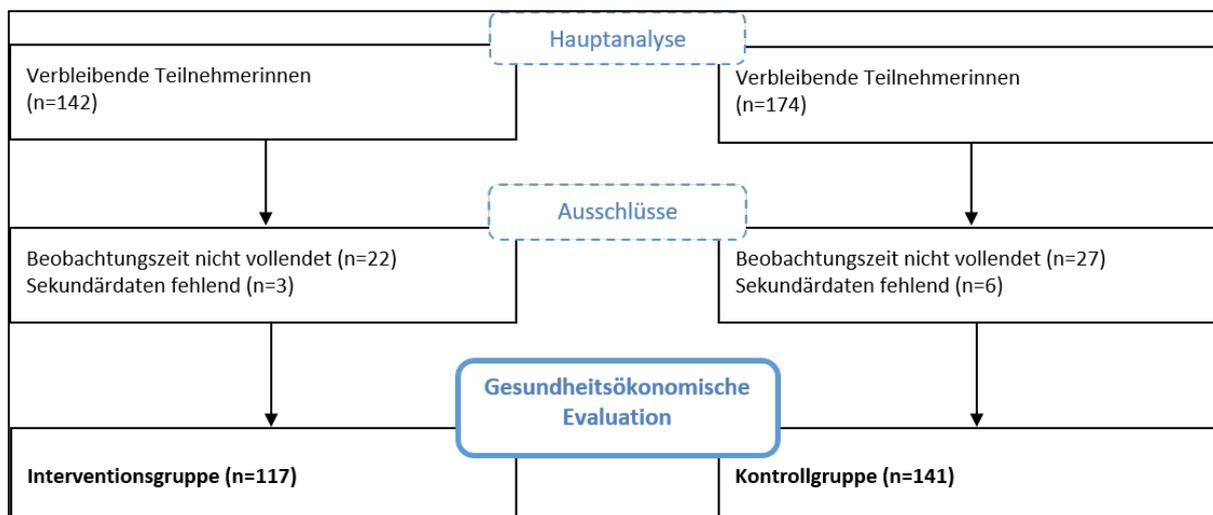
Darüber hinaus waren exploratorische Pfadanalysen auf Mediationsprozesse zur Aufklärung der vermittelnden Mechanismen der gefundenen Zusammenhänge geplant. Nachzeitigem Stand der exploratorischen Analysen wurden bisher keine Konstellation von Variablen ermittelt, die auf medierende Zusammenhänge hinweisen, weshalb weiterführende Mediationsanalysen nicht umgesetzt wurden.

Darüber hinaus waren exploratorische Moderationsanalysen zur Aufklärung der Modifizierbarkeit der gefundenen Zusammenhänge geplant. Zur Analyse diese sollten ursprünglich Pfadanalysen zur Anwendung kommen, um hierbei Mediationsanalysen kombinieren zu können. Da letztere nicht umgesetzt wurden, jedoch bereits im Rahmen der exploratorischen Analysen von Interaktionseffekte zwischen Drittvariablen und der Gruppierungsvariable moderative Effekte aufgedeckt werden konnten (siehe oben), wurde von der weiteren Verwendung der Pfadmodelle abgesehen - auch nicht zuletzt aus ähnlichen Problemen, die bei der Anwendung von CLPM zum Tragen kommen. Die Moderationsanalysen im messwiederholten ANOVA-Design sind weiter oben berichtet.

3.2 Gesundheitsökonomische Evaluation (Universität Bielefeld)

Da der definierte Beobachtungszeitraum von 40 Wochen nicht von allen in die Wirksamkeitsevaluation eingeschlossenen Teilnehmerinnen (N = 316) bis zum 30.06.2021 vollendet wurde (n = 49) bzw. die GKV-Routinedaten nicht für alle Teilnehmerinnen (vollständig) vorlagen (n = 9), basieren die folgenden Analysen auf einer reduzierten Studienpopulation (N = 258).

Abbildung 7: Flow-Diagramm Studienpopulation der gesundheitsökonomischen Evaluation



Der Beobachtungsbeginn beider Studiengruppen lag zwischen März 2019 und September 2020. Es besteht kein statistisch signifikanter Unterschied zwischen der IG und der KG hinsichtlich des Beobachtungsbeginns sowie in Bezug auf die Überschneidung der Beobachtungszeit mit der COVID 19-Pandemie ab März 2020 stellt die Baselinecharakteristika der Studienpopulation der gesundheitsökonomischen Evaluation (N = 258) dar. Die Teilnehmerinnen sind im Mittel 32 Jahre alt und haben mehrheitlich ein Haushaltseinkommen (netto) von über 1.500 €. Die Studiengruppen unterscheiden sich im Bildungsgrad signifikant (W = 8168; p = 0,005*). Der Anteil der Teilnehmerinnen, in deren Haushalt bisher kein Kind lebt, ist in der KG (60,74 %) deutlich höher als in der IG (48,51 %). Hier besteht allerdings kein statistisch signifikanter Unterschied. Hinsichtlich des Stressniveaus weist die KG einen signifikant höheren Skalenwert (7,26) auf, als die IG (6,30) (t = 2,269; p = 0,024*). Die Verteilung der Studiengruppen auf die beiden Center sowie die durchschnittliche Anzahl der geführten Gespräche unterscheiden sich nicht signifikant. Die Untersuchung fehlender Werte mittels "Missing-completely-at-random-test" nach Little (1988) zeigte keine Auffälligkeiten (X² = 35,77; p = 0,573) und legitimiert somit den Ausschluss unvollständiger Fälle in den Regressionsanalysen.

Tabelle 31: Baselinecharakteristika der Studienpopulation

	Intervention (n = 117)	Kontrolle (n = 141)	Teststatistik	p-Wert
Alter (n = 258)	32,33 (SD = 4,51)	32,35 (SD = 4,14)	t = 0,026	0,979
Bildung (n = 236)				
kein Abschluss	0 (0,00 %)	0 (0,00 %)	W = 8168	0,005*
Volks-/Hauptschule	6 (5,88 %)	7 (5,22 %)		
Mittlere Reife	38 (37,25 %)	25 (18,66 %)		
Fachhochschulreife	14 (13,73 %)	22 (16,42 %)		
Abitur	44 (43,14 %)	80 (59,70 %)		
Haushaltseinkommen (netto) (n = 227)				
< 1.500 €	27 (28,42 %)	32 (24,24 %)	W = 6799	0,248
1.500 - 2.999 €	45 (47,37 %)	58 (43,94 %)		
3.000 - 4.999 €	17 (17,89 %)	32 (24,24 %)		
5.000 - 8.000 €	6 (6,32 %)	10 (7,58 %)		
Anzahl der im Haushalt lebenden Kinder (n = 236)				
0	49 (48,51 %)	82 (60,74 %)	W = 5952,5	0,060
1	40 (39,60 %)	42 (31,11 %)		
2	11 (10,89 %)	10 (7,41 %)		
3	1 (0,99 %)	1 (0,74 %)		
Stress (PHQ) (n = 258)	6,30 (SD = 2,81)	7,26 (SD = 3,77)	t = 2,269	0,024*

Depressivität (EPDS) (n = 258)	14,14 (SD = 3,49)	14,01 (SD = 3,40)	t = -0,285	0,776
Psychische Vorerkrankung (n = 258)				
Nein	68 (58,12 %)	81 (57,45 %)	X ² = 0,000	1,000
Ja	49 (41,88 %)	60 (42,55 %)		
Physische Vorerkrankung (n = 258)				
Nein	54 (46,15 %)	60 (42,55 %)	X ² = 0,206	0,650
Ja	63 (53,85 %)	81 (57,45 %)		
Studiencenter (n = 258)				
Heidelberg	57 (48,72 %)	59 (41,84 %)	X ² = 0,959	0,327
Tübingen	60 (51,28 %)	82 (58,16 %)		
Anzahl Gespräche (n = 258)	1,38 (0,61)	1,35 (0,74)	t = -0,334	0,739
<i>Mittelwert (SD = Standardabweichung), *signifikant (p<0,05)</i>				

3.2.1 GKV-Perspektive

Ambulante Leistungen

Alle in die Analysen eingeschlossenen Teilnehmerinnen nahmen im Beobachtungszeitraum mindestens eine ambulante Leistung in Anspruch. Im Durchschnitt hatte die IG 30,74 und die KG 31,30 Arztkontakte. Die Kosten der ambulante-ärztlichen Versorgung lagen bei durchschnittlich 1.475,06 € (IG) bzw. 1.491,34 € (KG). Der Wilcoxon-Mann-Whitney-U-Test zeigt keine signifikanten Unterschiede der Verteilung der Arztkontakte (W = 8376; p = 0,831) in IG und KG sowie der ambulanten Kosten (W = 8265; p = 0,978). Eine differenzierte Analyse der ambulanten Arztkontakte ist in den erweiterten Analysen in Kapitel 3.2.3 zu finden.

Tabelle 32: Inanspruchnahme und Kosten der ambulanten Versorgung von IG und KG

	Intervention (n =117)	Kontrolle (n =141)	Teststatistik	p-Wert
Ambulante Versorgung	117 (100,00 %)	141 (100,00 %)		
Arztkontakte	30,74 (SD = 15,73)	31,30 (SD = 16,70)	W = 8376	0,831
Kosten (€)	1.475,06 (SD = 1.077,06)	1.491,34 (SD = 1.065,37)	W = 8265	0,978
Fallkosten (€) (n = 258)	98,14 (SD = 44,40)	102,83 (SD = 48,21)		
<i>Mittelwert (SD = Standardabweichung), *signifikant (p<0,05)</i>				

Stationäre Leistungen

Etwa 97 % der Teilnehmerinnen wurden mindestens einmal im Beobachtungszeitraum stationär behandelt (Tabelle 32). In den meisten Fällen handelte es sich hierbei um die Entbindung bzw. mit der Schwangerschaft assoziierte Krankenhausbehandlungen. Die durchschnittliche Anzahl der Fälle betrug 1,17 in der IG und 1,20 in der KG. Analog zur mittleren Verweildauer (LOS) (IG: 5,30 Tage; KG: 6,45 Tage), lagen die stationären Kosten der KG (3.509,82 €) etwas über der IG (3.363,67 €). Die Signifikanztests zeigen keinen Unterschied der Verteilungen von LOS (W = 8528,5; p = 0,634) und stationären Kosten (W = 7791; p = 0,443). Die Kosten pro Fall, welche sich auf die stationär behandelten Teilnehmerinnen (n = 252) beziehen, waren in der IG (2.962,99 €) etwas höher als in der KG (2.910,04 €).

Tabelle 33: Inanspruchnahme und Kosten der stationären Versorgung von IG und KG

	Intervention (n = 117)	Kontrolle (n = 141)	Teststatistik	p-Wert
Stationäre Versorgung	114 (97,44 %)	138 (97,87 %)		
Fälle	1,17 (SD = 0,56)	1,20 (SD = 0,62)		
Length of stay (LOS)	5,30 (SD = 3,75)	6,45 (SD = 8,95)	W = 8528,5	0,634
Kosten (€)	3.363,67 (SD = 1.643,36)	3.509,82 (SD = 2.692,00)	W = 7791	0,443

Fallkosten (€) (n = 252)	2.962,99 (SD = 982,29)	2.910,04 (SD = 1.158,86)		
Mittelwert (SD = Standardabweichung), *signifikant (p<0,05)				

Arzneimittel

Insgesamt wurde etwa 77 % der Teilnehmerinnen im Beobachtungszeitraum mindestens ein Medikament verschrieben (Tabelle 33). Die Anzahl der Arzneimittelverordnungen liegt bei durchschnittlich 3,12 (IG) bzw. 3,30 (KG). Die mittleren Arzneimittelkosten waren in der IG (125,56 €) geringer als in der KG (190,92 €), wobei Letztere eine vergleichsweise hohe Standardabweichung (641,98 €) aufweist. Dagegen lag die Anzahl der DDDs in der IG (174,95) etwas über der KG (159,07). Die unkontrollierten Tests zeigen keinen signifikanten Gruppenunterschied hinsichtlich der DDDs (W = 8174; p = 0,900) bzw. der Arzneimittelkosten (W = 8630,5; p = 0,520).

Tabelle 34: Inanspruchnahme und Kosten von Arzneimitteln in IG und KG

	Intervention (n =117)	Kontrolle (n =141)	Teststatistik	p-Wert
Arzneimittel	91 (77,78 %)	108 (76,60 %)		
Verordnungen	3,12 (SD = 3,06)	3,30 (SD = 3,50)		
DDD	174,95 (SD = 259,08)	159,07 (SD = 228,00)	W = 8174	0,900
Kosten (€)	125,56 (SD = 210,85)	190,92 (SD = 641,98)	W = 8630,5	0,520
Kosten pro Verordnung (€) (n = 199)	32,56 (SD = 35,93)	47,77 (SD = 91,02)		
Mittelwert (SD = Standardabweichung), *signifikant (p<0,05)				

Kosten der neuen Versorgungsform

Als vergleichbare Leistung zur Ermittlung der potenziellen Kosten des Screenings auf affektive Symptome (Tabelle 34) wurde die GOP 35600 ausgewählt. Diese umfasst die Durchführung von Screenings bzw. die Anwendung standardisierter Testverfahren/Fragebögen und ist im Kapitel IV des EBM („Arztgruppenübergreifende spezielle Gebührenordnungspositionen“) zu finden. Die Ziffer ist mit 3,83 € (34 Punkten) bewertet. Da für die koordinierenden Leistungen (b), welche unter anderem die Vermittlung des Erstgesprächs und die Einführung in die online-basierte Intervention umfassten, keine vergleichbare GOP aus einem passenden Versorgungsbereich gefunden werden konnte, wurde auf die GOP 37105 aus dem Bereich der „Versorgung gemäß Anlage 27 und 30 zum Bundesmantelvertrag-Ärzte (BMV-Ä) und der Vereinbarung nach § 132g Abs. 3 SGB V, Zuletzt geändert durch Art. 8 G v. 2022 I 969“ zurückgegriffen. Diese bildet neben der Koordination von diagnostischen, therapeutischen und rehabilitativen Maßnahmen und der pflegerischen Versorgung in der stationären Pflegeeinrichtung auch die Steuerung des multiprofessionellen Behandlungsprozesses ab und wird mit 30,98 € vergütet. Hinzu kommt die GOP 01470, welche zunächst befristet bis Ende 2022 als Zusatzpauschale für das Ausstellen einer Erstverordnung für eine DiGA, abrechnungsfähig ist. Für diese Leistung werden von der GKV 2,03 € erstattet. Für das psychologische Erstgespräch (c) wurde die „Probatorische Sitzung“ (GOP 35150) aus den GOPs gemäß der Psychotherapie-Richtlinie gewählt. Die Leistung bezeichnet eine Einzelbehandlung mit einer Dauer von mindestens 50 Minuten, umfasst unter anderem eine differentialdiagnostische Abklärung und wird mit 79,88 € vergütet. Für den Einsatz der Achtsamkeitsintervention (d) im Rahmen des Projektes wurden zu Projektbeginn Kosten in Höhe von 216,86 € pro Teilnehmerin kalkuliert. Insgesamt ergeben sich somit hypothetische Kosten in Höhe von 333,58 € pro Versicherte. In den Vergleich der Gesamtkosten der IG und KG fließen lediglich die Kosten der online-basierten Achtsamkeitsintervention (d) in Höhe von 216,86 € gesondert ein, wobei in Sensitivitätsanalysen höhere Kosten (+ 30 %) bzw. geringere Kosten (-30 %) angenommen werden.

Tabelle 35: Kosten der neuen Versorgungsform pro Versicherte

Komponenten der neuen Versorgungsform	Kosten pro Versicherte	Studiengruppe
a) Screening	3,83 €	IG + KG
b) Koordination in KOS	33,01 €	

c) Psychologisches Erstgespräch	79,88 €	
d) Online-basierte Achtsamkeitsintervention	216,86 €	IG
SUMME	333,58 €	

Gesamtkosten aus GKV-Perspektive

Neben den drei zentralen Leistungsbereichen der ambulanten und stationären Behandlung sowie Arzneimittelversorgung, wurden in die Berechnung der Gesamtkosten Leistungen aus den Bereichen HSA/ PIA/ ambulantes Operieren, Hilfsmittel, Heilmittel und Hebammenleistungen sowie die Interventionskosten einbezogen.

Tabelle 36 geben eine Übersicht über die Inanspruchnahme und die damit verbundenen Kosten zu Lasten der GKV. 36,75 % der IG und 36,17 % der KG nahmen Leistungen aus dem Bereich HSA/ PIA/ ambulantes Operieren in Anspruch und verursachten damit Kosten in Höhe von durchschnittlich 110,99 € (IG) bzw. 108,15 € (KG) pro Person. Die Inanspruchnahme und insbesondere die Kosten von Hilfsmitteln (z. B. Milchpumpen, Produkte zur Blutzuckermessung und Insulininjektion) fielen in der IG (1,32; 65,65 €) etwas geringer aus als in der KG (1,70; 121,77 €). Auch die durchschnittliche Anzahl der Heilmittelleistungen (v. a. Krankengymnastik und manuelle Therapie) und die mittleren Kosten lagen in der IG (0,69; 49,00 €) unterhalb der KG (0,89; 73,26 €). Bei den Hebammenleistungen ist zu berücksichtigen, dass die Mengenangabe aufgrund der teils unterschiedlichen Dokumentationsweise nur bedingt interpretierbar ist. Die Kosten der Hebammenversorgung fielen in den beiden Studiengruppen sehr ähnlich aus (IG: 1.156,71 €; KG: 1. 196,95 €). Für die Inanspruchnahme und die Kosten keiner dieser Leistungsbereiche zeigen sich im unkontrollierten Test signifikante Unterschiede zwischen der IG und der KG.

Bedingt durch die Datenqualität bzw. -verfügbarkeit sind die Ergebnisse zur Inanspruchnahme von Haushaltshilfe nur sehr eingeschränkt interpretierbar und gehen nicht in die Berechnung der Gesamtkosten ein. Das Mutterschaftsgeld nach § 24i SGB V, Zuletzt geändert durch Art. 8 G v. 2022 I 969 wird an Mitglieder der GKV für die Zeit der gesetzlich vorgegebenen Mutterschutzfristen sowie für den Tag der Entbindung gezahlt. Innerhalb der betrachteten Teilstichprobe (N = 104) erfolgte die Mutterschaftsgeldzahlung über durchschnittlich 101,53 (IG) bzw. 102,27 (KG) Tage. Die damit für die GKV verbundenen Kosten lagen bei 1.337,75 € pro Versicherte in der IG und 1.335,85 € in der KG, wobei die Ergebnisse aufgrund der Datengrundlage nur eingeschränkt interpretierbar sind.

Die durchschnittlichen Gesamtkosten der IG liegen bei 6.563,49 € (SD = 2.474,05 €) und damit knapp 129,00 € unterhalb der KG (6.692,21 €; SD = 3.608,22 €). Der inferenzstatistische Vergleich mittels Wilcoxon-Mann-Whitney-U-Test, zeigt auch hier keinen signifikanten Unterschied zwischen den beiden Studiengruppen (W = 7301; p = 0,112).

Tabelle 36: Übersicht der Inanspruchnahme von Gesundheitsleistungen in IG und KG

Leistungsinanspruchnahme	Intervention (n =117)	Kontrolle (n =141)	Teststatistik	p-Wert
Ambulante Versorgung Arztkontakte	30,74 (SD = 15,73)	31,30 (SD = 16,70)	W = 8376	0,831
Stationäre Versorgung LOS	5,30 (SD = 3,75)	6,45 (SD = 8,95)	W = 8528,5	0,634
Arzneimittel DDD	174,95 (SD = 259,08)	159,07 (SD = 228,00)	W = 8174	0,900
HSA/ PIA/ ambulantes Operieren Fälle	0,59 (SD = 0,98)	0,59 (SD = 1,04)	W = 8227	0,966
Hilfsmittel Positionsnummern	1,32 (SD = 1,96)	1,70 (SD = 2,43)	W = 8723	0,390
Heilmittel Positionsnummern	0,69 (SD = 1,53)	0,89 (SD = 2,55)	W = 8047	0,642
Hebammenleistungen Tage	201,49 (SD = 114,05)	225,35 (SD = 105,23)	W = 9222	0,103
[Haushaltshilfe	0,27 (SD = 0,82)	0,35 (SD = 1,27)		

Positionsnummern]				
[Mutterschaftsgeld (N = 104) Tage]	101,53 (SD = 5,06)	102,41 (SD = 7,53)		
Online-basierte Achtsamkeitsintervention	1	NA		
<i>Mittelwert (SD = Standardabweichung), *signifikant (p<0,05)</i>				

Tabelle 37: Übersicht der Kosten der Gesundheitsleistungen in IG und KG

Kosten	Intervention (n = 117)	Kontrolle (n = 141)	Teststatistik	p-Wert
Ambulant (€)	1.475,06 (SD = 1.077,06)	1.491,34 (SD = 1.065,37)	W = 8265	0,978
Stationär (€)	3.363,67 (SD = 1.643,36)	3.509,82 (SD = 2.692,00)	W = 7791	0,443
Arzneimittel (€)	125,56 (SD = 210,85)	190,92 (SD = 641,98)	W = 8630,5	0,520
HSA/ PIA/ ambulantes Operieren (€)	110,99 (SD = 237,91)	108,15 (SD = 207,83)	W = 8243,5	0,992
Hilfsmittel (€)	65,65 (SD = 152,81)	121,77 (SD = 495,83)	W = 8847,5	0,280
Heilmittel (€)	49,00 (SD = 99,75)	73,26 (SD = 256,69)	W = 8051,5	0,649
Hebammenleistungen (€)	1.156,71 (SD = 834,98)	1.196,95 (SD=1.040,53)	W = 8258	0,987
[Haushaltshilfe Kosten (€)]	229,81 (SD = 823,66)	463,31 (SD = 1.826,94)		
[Mutterschaftsgeld (N = 104) Kosten (€)]	1.337,75 (SD = 69,61)	1.335,85 (SD = 107,10)		
Online-basierte Achtsamkeitsintervention (€)	216,86 (+/- 30 %)	NA		
SUMME (€)	6.563,49 (SD = 2.474,05)	6.692,21 (SD = 3.608,22)	W = 7301	0,112
<i>Mittelwert (SD = Standardabweichung), *signifikant (p<0,05)</i>				

Ergänzend zum unkontrollierten Test, stellt

Tabelle 37 das generalisierte additive Modell der Gesamtkosten aus GKV-Perspektive dar. Aufgrund des Ausschlusses unvollständiger Datensätze, bildet die Analyse 236 Studienteilnehmerinnen ab. Das Ausgangsmodell lautet:

$$\begin{aligned} \text{Gesamtkosten_direkt}_i = & \beta_0 + \beta_1 * \text{Studiengruppe}_i + \beta_2 * \text{Center}_i + \beta_3 * \text{Anzahl_Kinder}_i + \beta_4 * \text{Stressniveau}_i \\ & + \beta_5 * \text{Depressivität}_i + \beta_6 * \text{psychische_Vorerkrankung}_i + \beta_7 * \text{physische_Vorerkrankung}_i \\ & + \beta_8 * \text{Studiengruppe}_i * \text{Center}_i + \beta_9 * \text{Studiengruppe}_i * \text{physische_Vorerkrankung}_i \\ & + \beta_{10} * \text{Center}_i * \text{physische_Vorerkrankung}_i + \beta_{11} * \text{Studiengruppe}_i * \text{Center}_i * \text{physische_Vorerkrankung}_i + \\ & \epsilon_i \end{aligned}$$

Daraus ergab sich per Modellselektion der folgende finale Regressionsterm:

$$\begin{aligned} \text{Gesamtkosten_direkt}_i = & \beta_0 + \beta_1 * \text{Studiengruppe}_i + \beta_2 * \text{Center}_i + \beta_3 * \text{Anzahl_Kinder}_i + \beta_4 * \text{Stressniveau}_i + \\ & \beta_5 * \text{psychische_Vorerkrankung}_i + \beta_6 * \text{physische_Vorerkrankung}_i + \beta_7 * \text{Studiengruppe}_i * \text{Center}_i + \\ & \beta_8 * \text{Center}_i * \text{physische_Vorerkrankung}_i + \epsilon_i \end{aligned}$$

Der exponierte Regressionskoeffizient der Studiengruppe zeigt sehr geringfügig ($\text{Exp}(\beta) = 1,009$) höhere Kosten für die IG im Vergleich zur KG an, wobei der Einfluss nicht signifikant ist ($p = 0,889$). Aus dem Interaktionsterm aus Studiengruppe und -center ergeben sich zudem nicht signifikant höhere Kosten für die IG im Studiencenter Tübingen ($\text{Exp}(\beta) = 1,167$, $p = 0,075$). Ein statistisch signifikanter Einfluss der Intervention auf die direkten Gesamtkosten ist damit nicht festzustellen. Signifikante Prädiktoren stellen dagegen das Studiencenter, die Anzahl der im Haushalt lebenden Kinder, das Stressniveau zu Studienbeginn (PHQ-Stress-Skala) sowie eine physische Vorerkrankung dar. Dem Vergleich der Dichtekurven zufolge deckt das Modell die Mitte der beobachteten Werte gut ab, weist

jedoch Abweichungen im unteren und oberen Randbereich auf. Die Residuen-Plots sowie der QQ-Plot zeigen keine weiteren Auffälligkeiten.

Tabelle 38: Generalisiertes additives Modell mit inverser Gammaverteilung: Gesamtkosten aus GKV-Perspektive (N=236)

	Exp(β)	95 %-KI	p-Wert
Intercept	5.544,354	4784,735; 6424,569	<0,001*
Studiengruppe (IG)	1,009	0,891; 1,142	0,889
Center (TU)	0,862	0,744; 0,999	0,049*
Anz. der im Haushalt lebenden Kinder (1)	0,782	0,713; 0,858	<0,001*
Anz. der im Haushalt lebenden Kinder (2)	0,841	0,724; 0,977	0,025*
Anz. der im Haushalt lebenden Kinder (3)	0,721	0,458; 1,135	0,159
Stress (PHQ)	1,018	1,004; 1,031	0,010*
Physische Vorerkrankung (Ja)	0,842	0,743; 0,954	0,007*
Psychische Vorerkrankung (Ja)	1,074	0,983; 1,172	0,115
Studiengruppe (IG) : Center (TU)	1,167	0,985; 1,383	0,075
Center (TU) : Physische Vorerkrankung (Ja)	1,145	0,966; 1,357	0,119
*signifikant ($p < 0,05$) Globale Devianz: 4259,411 AIC: 4283,411 SBC: 4324,977 N = 236 Freiheitsgrade: 12 Residuale Freiheitsgrade: 224			

Zur Prüfung der Stabilität der Ergebnisse der zentralen Analyse wurden Sensitivitätsanalysen durchgeführt, welche mögliche Abweichungen der Annahmen über die Interventionskosten sowie die Ausschlüsse aufgrund des unvollständigen Beobachtungszeitraumes adressieren. Unter Annahme 30 % höherer (281,92 €) und 30 % niedrigerer (151,80 €) Interventionskosten, ändert sich das Ergebnis des deskriptiven und testbasierten Gruppenvergleichs nicht. Demnach weist die IG im Vergleich zur KG weiterhin nicht signifikant geringere Gesamtkosten auf. Für den modellbasierten Vergleich wurde jeweils ein erneuter Auswahlprozess („backward selection“) durchgeführt, wobei sich keine (wesentlichen) Änderungen der selektierten unabhängigen Variablen und der Ergebnisse ergaben. Lediglich der exponierte Regressionskoeffizient der Studiengruppe liegt im Falle geringerer Interventionskosten unter 1,00 ($\text{Exp}(\beta) = 0,997$, $p = 0,965$). Im Falle höherer Interventionskosten bleibt die Richtung des Effektes dagegen bestehen ($\text{Exp}(\beta) = 1,021$; $p = 0,748$). Beide Ergebnisse sind nicht signifikant. Dies gilt in beiden Szenarien auch für den Interaktionsterm aus Studiengruppe und -center. Im Rahmen einer weiteren Sensitivitätsanalyse wurde der Umfang der Ausschlüsse aus der gesundheitsökonomischen Evaluation reduziert, indem Teilnehmerinnen eingeschlossen wurden, deren Beobachtungszeitraum in den Routinedaten lediglich um einen Monat verkürzt war. Damit gingen in das Gesamtkostenmodell 250 Teilnehmerinnen ein ($N=236+14$). Wenngleich der exponierte Regressionskoeffizient der Studiengruppe hier ebenfalls unter 1,00 liegt ($\text{Exp}(\beta) = 0,996$, $p = 0,942$), bleibt das Ergebnis nicht signifikant. Der Einfluss des Interaktionsterms aus Studiengruppe und -center ändert sich auch hier nur geringfügig ($\text{Exp}(\beta) = 1,170$, $p = 0,061$).

3.2.2 Gesellschaftliche Perspektive

Produktivitätsverluste

Für die sekundäre Analyse aus gesellschaftlicher Perspektive sind, neben den Kosten aus Perspektive der GKV, die Produktivitätsverluste (Tabelle 38) relevant, welche auf den AU-Daten und dem durchschnittlichen Entgelt pro Arbeitnehmer pro Tag basieren (Kapitel 0). Sowohl die AU-Tage (IG: 1,37; KG: 1,38; $W = 7833$; $p = 0,211$) als auch die Produktivitätsverluste (IG: 169,40 €; KG: 170,55

€; $W = 7835$; $p = 0,213$) liegen in den Studiengruppen auf sehr ähnlichem Niveau und unterscheiden sich nicht signifikant. Hierbei ist zu berücksichtigen, dass die Interpretierbarkeit aufgrund der teils uneinheitlichen Dokumentation der AU in den Routinedaten eingeschränkt ist.

Tabelle 39: Arbeitsunfähigkeit und damit verbundene Produktivitätsverluste

	Intervention (n = 117)	Kontrolle (n = 141)	Teststatistik	p-Wert
AU	17 (14,53 %)	13 (9,22 %)		
AU-Tage	1,37 (SD = 4,47)	1,38 (SD = 6,07)	$W = 7833$	0,211
Kosten (€)	169,40 (SD = 553,92)	170,55 (SD = 751,67)	$W = 7835$	0,213
<i>Mittelwert (SD = Standardabweichung), *signifikant ($p < 0,05$)</i>				

Gesamtkosten aus gesellschaftlicher Perspektive

Die durchschnittlichen Gesamtkosten aus gesellschaftlicher Perspektive berücksichtigen neben den oben dargestellten direkten Gesamtkosten auch die indirekten Kosten (Produktivitätsverluste). Im Mittel liegen die Gesamtkosten aus gesellschaftlicher Perspektive der IG (6.732,89€; SD = 2.646,74 €) unter der KG (6.862,76 €; SD = 3.801,57 €). Dem Wilcoxon-Mann-Whitney-U-Test zufolge ist der Unterschied nicht signifikant ($W = 7212$; $p = 0,082$). Das Ausgangs- sowie das finale Modell gleichen den oben dargestellten Regressionstermen der Analyse aus GKV-Perspektive. Auch die Ergebnisse der Analysen beider Perspektiven unterscheiden sich nicht wesentlich voneinander. Demnach sind ergibt sich weder für die Studiengruppe ($\text{Exp}(\beta) = 1,006$; $p = 0,923$) noch für den Interaktionsterm aus Studiengruppe und -center ($\text{Exp}(\beta) = 1,188$; $p = 0,056$) ein signifikanter Einfluss auf die Gesamtkosten aus gesellschaftlicher Perspektive (Tabelle 39). Analog zum Modell der direkten Gesamtkosten sind die Abweichungen der Modellschätzung in den Randbereichen zu berücksichtigen.

Tabelle 40: Generalisiertes additives Modell mit inverser Gammaverteilung: Gesamtkosten aus gesellschaftlicher Perspektive (N=236)

	Exp(β)	95 %-KI	p-Wert
Intercept	5.527,035	4745,52; 6437,255	<0,001*
Studiengruppe (IG)	1,006	0,885; 1,145	0,923
Center (TU)	0,869	0,746; 1,012	0,072
Anz. der im Haushalt lebenden Kinder (1)	0,792	0,72; 0,871	<0,001*
Anz. der im Haushalt lebenden Kinder (2)	0,847	0,726; 0,989	0,037*
Anz. der im Haushalt lebenden Kinder (3)	0,716	0,447; 1,145	0,164
Stress (PHQ)	1,018	1,004; 1,032	0,011*
Physische Vorerkrankung (Ja)	0,836	0,734; 0,951	0,007*
Psychische Vorerkrankung (Ja)	1,078	0,984; 1,18	0,109
Studiengruppe (IG) : Center (TU)	1,188	0,997; 1,416	0,056
Center (TU) : Physische Vorerkrankung (Ja)	1,151	0,965; 1,372	0,119
*signifikant ($p < 0,05$) Globale Devianz: 4285,295 AIC: 4309,295 SBC: 4350,861 N = 236 Freiheitsgrade: 12 Residuale Freiheitsgrade: 224			

Selbstzahlerleistungen

Aus gesellschaftlicher bzw. Versichertenperspektive sind auch die Kosten für Gesundheitsleistungen, welche die Versicherten selbst übernommen haben, relevant. Da der Rücklauf und die Datenqualität der Kostenwochenbücher sehr eingeschränkt waren, konnte diese Kostenkomponente nicht in die Analyse der Gesamtkosten einfließen. Der Rücklauf der Kostenwochenbücher lag bei insgesamt 50,39 %, wovon jedoch mehr als die Hälfte leer, unvollständig oder unplausibel und damit nicht auswertbar war. In die aggregierten Analysen fließen somit lediglich 20,54 % der Gesamtpopulation ein. Die Ergebnisse sind demnach nicht aussagekräftig. Die IG nahm durchschnittlich 2,80 und die KG 3,12 unterschiedliche OTCs (z. B. Folsäure, Vitamin D) in Anspruch. Am häufigsten wurden in beiden Gruppen Nahrungsergänzungsmittel und rezeptfreie Arzneimittel angegeben. Die Anzahl verschiedener selbst zu finanzierenden Gesundheitsdienstleistungen lag bei durchschnittlich 1,78 (IG) bzw. 1,60 (KG). Darunter fallen vor allem Kurse und Beratungsleistungen (z. B. spezielle Geburts-/Stillvorbereitung), diagnostische Leistungen sowie Leistungen verschiedener Anbieter von Homöopathie/ Heilpraktik.

3.2.3 Ergänzende Analysen

Ambulante Versorgung

Im Hinblick auf die Hypothesen der gesundheitsökonomischen Evaluation wurde eine detaillierte Untersuchung der ambulanten Versorgung vorgenommen. In Tabelle 40 sind die Ergebnisse der Auswertung der Anzahl der ambulanten Arztkontakte mittels eines generalisierten additiven Modells mit negativer Binomialverteilung dargestellt. Aus dem Ausgangsmodell

$$\text{Anzahl_Arztkontakte}_i = \beta_0 + \beta_1 * \text{Studiengruppe}_i + \beta_2 * \text{Anzahl_Kinder}_i + \beta_3 * \text{Stressniveau}_i + \beta_4 * \text{Depressivität}_i + \beta_5 * \text{psychische_Vorerkrankung}_i + \beta_6 * \text{physische_Vorerkrankung}_i + \beta_7 * \text{Studiengruppe}_i * \text{psychische_Vorerkrankung}_i + \varepsilon_i$$

ergab sich mittels Modellselektion der folgende finale Regressionsterm:

$$\text{Anzahl_Arztkontakte}_i = \beta_0 + \beta_1 * \text{Studiengruppe}_i + \beta_2 * \text{Anzahl_Kinder}_i + \beta_3 * \text{Stressniveau}_i + \beta_4 * \text{psychische_Vorerkrankung}_i + \beta_5 * \text{physische_Vorerkrankung}_i + \beta_6 * \text{Studiengruppe}_i * \text{psychische_Vorerkrankung}_i + \varepsilon_i$$

Die Analyse liefert ein signifikantes Ergebnis des Interaktionsterms aus Studiengruppe und psychischer Vorerkrankung, welches auch unter Berücksichtigung der Bonferroni-Korrektur ($p=0,05/8$) bestehen bleibt (Tabelle 40). Aus den exponierten Regressionskoeffizienten ergibt sich, dass Teilnehmerinnen der IG mit einer psychischen Vorerkrankung im Vergleich zur KG ohne psychische Vorerkrankung mehr (+27,3 %) Arztkontakte aufweisen. Eine Aussage über den Einfluss der Studiengruppe allein ist nicht möglich. Weitere relevante Einflussvariablen sind die Anzahl der im Haushalt lebenden Kinder, das Stressniveau sowie eine physische Vorerkrankung. Anhand des Vergleichs der Dichtekurven ist erkennbar, dass die Modellschätzung die beobachteten Werte nicht vollständig abdeckt und tendenziell überschätzt.

Tabelle 41: Generalisiertes additives Modell mit negativer Binomialverteilung: Ambulante Arztkontakte (N=236)

	Exp(β)	95 %-KI	p-Wert
Intercept	31,237	26,517; 36,797	<0,001*
Studiengruppe (IG)	0,859	0,737; 1	0,052
Anz. der im Haushalt lebenden Kinder (1)	0,836	0,738; 0,948	0,006*
Anz. der im Haushalt lebenden Kinder (2)	0,966	0,788; 1,186	0,744
Anz. der im Haushalt lebenden Kinder (3)	0,479	0,248; 0,928	0,030*
Stress (PHQ)	1,021	1,003; 1,04	0,025*
Physische Vorerkrankung (Ja)	0,799	0,713; 0,895	<0,001*
Psychische Vorerkrankung (Ja)	1,053	0,898; 1,234	0,524
Studiengruppe (IG) : Psychische Vorerkrankung (Ja)	1,407	1,116; 1,775	0,004*

*signifikant ($p < 0,05$)
 Globale Devianz: 1863,292
 AIC: 1883,292
 SBC: 1917,93
 N = 236
 Freiheitsgrade: 10
 Residuale Freiheitsgrade: 226

Für eine differenziertere Analyse der ambulanten Leistungen wurde die hausärztliche Versorgung sowie die Inanspruchnahme bestimmter Fachärzte untersucht. Dabei ist zu berücksichtigen, dass die Summe der dargestellten Mengen und Kosten (Tabelle 41), aufgrund nicht eindeutiger Zuordnung einzelner Fälle zu einer Facharztgruppe, geringfügig von der Gesamtinanspruchnahme abweicht. Im betrachteten Zeitraum nahmen 87,18 % der IG- und 85,11 % der KG-Teilnehmerinnen die hausärztliche Versorgung in Anspruch. Der geringfügige Unterschied der durchschnittlichen Anzahl der Hausarztkontakte zwischen IG (5,71) und KG bei (5,18) ist nicht signifikant ($p = 0,480$). Der Anteil der Teilnehmerinnen, die im Beobachtungszeitraum mindestens einmal einen Facharzt der Neurologie, Psychiatrie, Psychosomatik oder einen psychologischen Psychotherapeuten besuchten, liegt in der IG mit 30,77 % über dem Anteil in der KG (24,11 %). Die Anteile weichen deutlich von den Ergebnissen zur derzeitigen Psychotherapie auf Grundlage der Primärdaten (Tabelle 41) ab. Dies ist zum einen darauf zurückzuführen, dass die Baseline-Charakteristika auf den Selbstangaben der Teilnehmerinnen zu Studienbeginn basieren. Die Ergebnisse in Tabelle 41 basieren dagegen auf den GKV-Routinedaten und beziehen sich auf den Beobachtungszeitraum zwischen dem Quartal des Beobachtungsbeginns und den folgenden drei Quartalen. Da es sich hier um die Phase kurz vor und einige Monate nach der Geburt handelt, erscheint die geringe Inanspruchnahme der Fachärzte und Therapeuten plausibel. Zudem können die oben genannten nicht eindeutigen Fachgruppenzuordnungen zu einer gewissen Unterschätzung der fachärztlichen Inanspruchnahme geführt haben. Im Mittel hatte die KG (3,48) etwas mehr Kontakte als die IG (3,01). Der Signifikanztest zeigt auch hier keinen Unterschied auf ($p = 0,432$). Erwartungsgemäß nahmen alle analysierten Studienteilnehmerinnen eine ambulante gynäkologische bzw. geburtshilfliche Leistung in Anspruch. Die durchschnittliche Anzahl dieser Facharztkontakte lag bei 13,22 (IG) bzw. 13,28 (KG). Sonstige Fachärzte wurden ebenfalls etwas häufiger von der KG besucht (IG: 8,54; KG: 9,51).

Relativ betrachtet, macht das Gebiet der Frauenheilkunde und Geburtshilfe den größten Anteil der durchschnittlichen ambulanten Inanspruchnahme (IG: 43 %; KG: 43 %) sowie der ambulanten Kosten pro Studienteilnehmerin (IG: 49 % KG: 48 % €) aus. Hinsichtlich der Versorgung durch Fachärzte der Psychiatrie usw. fällt auf, dass die durchschnittliche Anzahl der Arztkontakte den geringsten Anteil (IG: 10 %; KG: 11 %) ausmacht, die damit verbundenen durchschnittlichen Kosten dagegen den zweitgrößten Anteil (IG: 19 %; KG: 22 %) an den durchschnittlichen ambulanten Kosten haben. Dies spiegelt sich auch in den Fallkosten (IG: 300,59 €; KG: 448,08 €) dieses Bereiches der Facharztversorgung wider.

Tabelle 42: Inanspruchnahme und Kosten der haus- und fachärztlichen Versorgung von IG und KG

	Intervention (n = 117)	Kontrolle (n = 141)	Test-statistik	p-Wert
Hausärztliche Versorgung	102 (87,18 %)	120 (85,11 %)		
Arztkontakte	5,71 (SD = 5,46)	5,18 (SD = 5,10)	W = 7829	0,480
Kosten (€)	185,24 (SD = 184,82)	163,86 (SD = 192,07)	W = 7590	0,269
Fallkosten (€) (n = 223)	51,71 (SD = 23,86)	50,35 (SD = 23,76)		
Fachärztliche Versorgung: Neurologie, Psychiatrie, Psychotherapie, Psychosomatische Medizin	36 (30,77 %)	34 (24,11 %)		
Arztkontakte	3,01 (SD = 7,12)	3,48 (SD = 8,50)	W = 7881	0,432
Kosten (€)	284,35 (SD = 712,21)	322,48 (SD = 812,31)	W = 7805,5	0,343
Fallkosten (€) (n = 63)	300,59 (SD = 289,27)	448,08 (SD = 294,71)		

Fachärztliche Versorgung: Frauenheilkunde und Geburtshilfe	117 (100,00 %)	141 (100,00 %)		
Arztkontakte	13,22 (SD = 4,71)	13,28 (SD = 5,15)	W = 8104	0,808
Kosten (€)	728,58 (SD = 282,51)	718,52 (SD = 219,04)	W = 8450	0,736
Fallkosten (€) (n = 258)	162,16 (SD = 64,95)	160,41 (SD = 57,14)		
Fachärztliche Versorgung: Sonstige	117 (100,00 %)	140 (99,29 %)		
Arztkontakte	8,67 (SD = 6,78)	9,26 (SD = 7,24)	W = 8656	0,493
Kosten (€)	276,87 (SD = 412,70)	286,49 (SD = 448,84)	W = 7817,5	0,470
Fallkosten (€) (n = 257)	39,69 (SD = 42,48)	43,34 (SD = 51,88)		
<i>Mittelwert (SD = Standardabweichung), *signifikant (p<0,05)</i>				

Für die modellbasierte Analyse der Hausarztkontakte ergab sich aus dem Ausgangsmodell

$$\begin{aligned} \text{Anzahl_Hausarztkontakte}_i = & \beta_0 + \beta_1 * \text{Studiengruppe}_i + \beta_2 * \text{Center}_i + \beta_3 * \text{Anzahl_Kinder}_i + \beta_4 * \text{Bildung} \\ & + \beta_5 * \text{Stressniveau}_i + \beta_6 * \text{Depressivität}_i + \beta_7 * \text{psychische_Vorerkrankung}_i + \\ & \beta_8 * \text{physische_Vorerkrankung}_i \\ & + \beta_9 * \text{Studiengruppe}_i * \text{Center}_i + \beta_{10} * \text{Studiengruppe}_i * \text{physische_Vorerkrankung}_i \\ & + \beta_{11} * \text{Center}_i * \text{physische_Vorerkrankung}_i + \beta_{12} * \text{Studiengruppe}_i * \text{Center}_i * \text{physische_Vorerkrankung}_i + \\ & \epsilon_i \end{aligned}$$

der folgende Term

$$\text{Anzahl_Hausarztkontakte}_i = \beta_0 + \beta_1 * \text{Studiengruppe}_i + \beta_2 * \text{Center}_i + \beta_3 * \text{Anzahl_Kinder}_i + \beta_4 * \text{Bildung} + \beta_5 * \text{Stressniveau}_i + \beta_6 * \text{physische_Vorerkrankung}_i + \beta_7 * \text{Center}_i * \text{physische_Vorerkrankung}_i + \epsilon_i$$

Die Modellschätzung (Tabelle 42) gibt 1,261-mal so viele Hausarztkontakte für die IG im Vergleich zur KG an, wobei der Effekt nicht signifikant ist ($p = 0,079$). Eine Reduzierung der Hausarztkontakte durch die Intervention kann somit nicht bestätigt werden. Neben der Anzahl der Kinder und einer physischen Vorerkrankung stellt der Bildungsgrad einen relevanten Einflussfaktor dar. Hinsichtlich der Modellgüte sind die Abweichungen der Verteilung der geschätzten von den beobachteten Werten und die Überschätzung der Werte zu berücksichtigen.

Tabelle 43: Generalisiertes additives Modell mit negativer Binomialverteilung: Hausarztkontakte (N=235)

	Exp(β)	95 %-KI	p-Wert
Intercept	4,526	2,361; 8,679	<0,001*
Studiengruppe (IG)	1,261	0,975; 1,632	0,079
Studiencenter (Tübingen)	0,869	0,605; 1,248	0,448
Anz. der im Haushalt lebenden Kinder (1)	0,610	0,464; 0,803	<0,001*
Anz. der im Haushalt lebenden Kinder (2)	0,712	0,458; 1,108	0,134
Anz. der im Haushalt lebenden Kinder (3)	0,977	0,258; 3,704	0,973
Bildung - Mittlere Reife	1,389	0,776; 2,486	0,270
Bildung - Fachhochschulreife	2,017	1,096; 3,711	0,025*
Bildung - Abitur	1,176	0,678; 2,038	0,565
Stress (PHQ)	1,035	0,995; 1,078	0,088
Physische Vorerkrankung (Ja)	0,498	0,345; 0,718	<0,001*
Studiencenter (Tübingen) : Physische Vorerkrankung (Ja)	1,594	0,970; 2,620	0,067

*signifikant ($p < 0,05$)
 Globale Devianz: 1269,543
 AIC: 1295,543
 SBC: 1340,517
 N = 235
 Freiheitsgrade: 13
 Residuale Freiheitsgrade: 222

Die Betrachtung der Inanspruchnahme ambulanter Leistungen im zeitlichen Verlauf zeigt einen Schwerpunkt in der Zeit kurz vor der Entbindung. Der Verlauf wird entsprechend stark von den Facharztbesuchen der Frauenheilkunde und Geburtshilfe (D) bestimmt. Im Mittel zeigen sich in den zeitlich differenzierten Abbildungen (A-D) nur sehr geringfügige Unterschiede zwischen der IG und KG.

Erkrankungsspezifische Medikation

Neben der allgemeinen Arzneimittelinanspruchnahme wurde eine erkrankungsspezifische Medikation anhand der ATC-Gruppen N05 und N06 betrachtet. Die aufgrund der geringen Fallzahl zu einem binären Outcome aggregierte Auswertung (Tabelle 43) mittels Chi²-Test zeigt keinen signifikanten Unterschied zwischen IG und KG ($X^2 = 1,788$; $p = 0,181$). Bei der Interpretation gilt es zu berücksichtigen, dass im Rahmen der Analyse nicht für bereits vor Beobachtungsbeginn bestehende medikamentöse Behandlungen kontrolliert wurde.

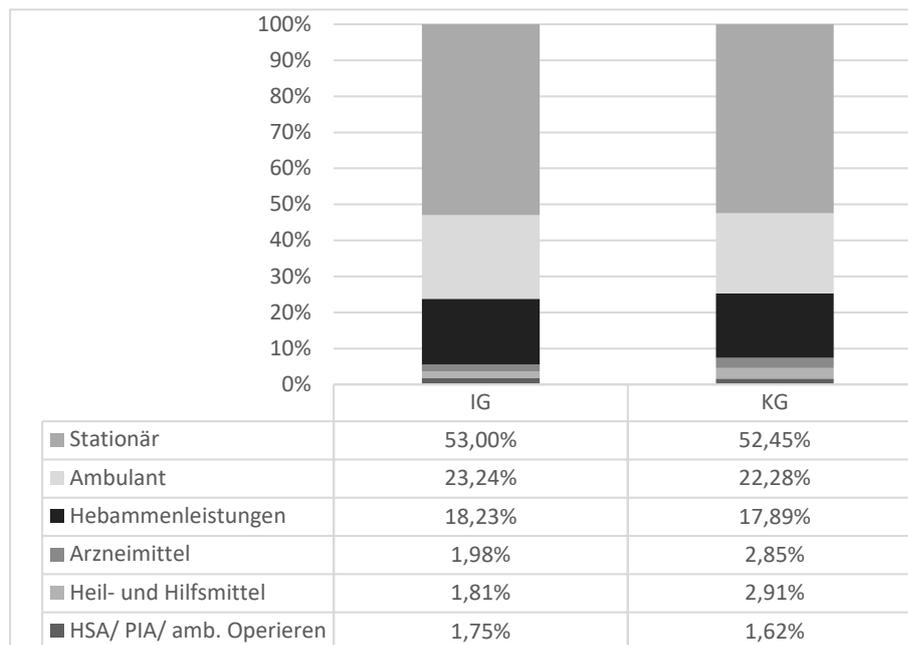
Tabelle 44: Erkrankungsspezifische Medikation (ATC N05/N06) (nein/ja) nach Studiengruppe (IG vs. KG)

Arzneimittel ATC-Gruppen N05/ N06	Intervention (n =117)	Kontrolle (n =141)	Teststatistik	p-Wert
Nein	102 (87,18 %)	131 (92,91 %)	$X^2 = 1,788$	0,181
Ja	15 (12,82 %)	10 (7,09 %)		
Mittelwert (SD = Standardabweichung), *signifikant ($p < 0,05$)				

Kostentreiberanalysen

Relativ betrachtet machen die stationären Leistungen den größten Anteil der durchschnittlichen Gesamtkosten in IG (53,00 %) und KG (52,45 %) aus (Abbildung 8). Es folgen in beiden Gruppen der ambulante Leistungsbereich sowie die Hebammenleistungen mit ca. 23 % und 18 %. Weniger relevante Kostenblöcke sind der Inanspruchnahme von Arzneimitteln, Leistungen aus dem Bereich HSA/ PIA/ ambulantes Operieren sowie Heil- und Hilfsmitteln zuzuordnen, wobei der Anteil der Heil- und Hilfsmittel in der KG - anders als in der IG - geringfügig höher ausfällt als der Anteil der mittleren Arzneimittelkosten.

Abbildung 8: Anteile der mittleren leistungsbereichsspezifischen Kosten an den durchschnittlichen Gesamtkosten pro Studienteilnehmerin der IG und KG (ohne Interventionskosten)



Neben dem Gruppenvergleich der relativen Anteile der Leistungsbereiche an den Gesamtkosten, wurden weitere Analysen zur Identifikation von Kostentreibern durchgeführt. Die Matrizen zeigen die Korrelationen und Assoziationen der Gesamtkosten der IG bzw. KG mit relevanten metrischen und dichotomen Variablen. Während in der IG die Anzahl der sonstigen Facharztkontakte (Arztkontakte_sonst; $\rho = 0,40$) die höchste Korrelation mit den Gesamtkosten aufweisen, sind in der KG die Anzahl der Besuche bei einem Facharzt der Neurologie, Psychiatrie oder Psychosomatik bzw. Psychotherapeuten (Arztkontakte_psy; $\rho = 0,50$) am stärksten mit den Kosten korreliert. Bei den dichotomen Merkmalen zeigt eine erkrankungsspezifische Verordnung (Ja/Nein) in beiden Gruppen die stärksten Assoziationen mit den Gesamtkosten (ATC_N05N06; IG: $R^2 = 0,36$; KG: $R^2 = 0,43$).

4. Schlussfolgerungen und Empfehlungen des Evaluators

Bei der Mind:Pregnancy Studie handelt es sich um eine der größten systematischen Studien zur Effektivität eines kombinierten Achtsamkeitsprogrammes in einer auf psychische Belastung gescreenten Schwangerenpopulation im ambulanten Bereich als Teil der Schwangerenvorsorge. In Deutschland existiert keine vergleichbare Studie. Sie bietet erstmals den Ansatz, Fragen nach dem Outcome eines Screenings für psychische Belastung als Teil der Schwangerenvorsorge und der Effektivität des Einsatzes einer schwangerschaftsspezifischen Achtsamkeitsintervention zur Geburtsvorbereitung und/oder zur Behandlung bei psychisch belasteten Schwangeren zu beantworten. Bei den adressierten Forschungsfragen handelt es sich um zentrale Aspekte bei der Versorgung Schwangerer, die bis jetzt nicht in den Mutterschaftsrichtlinien abgebildet sind und daher eine Versorgungslücke darstellen.

Zahlreiche Studien zeigen, dass mütterlicher Stress in der Schwangerschaft mit einer Vielzahl an ungünstigen physischen und psychischen Kurz- und Langzeitfolgen für Mutter und Kind einhergehen kann. Prospektive, klinische Studien, von denen die längste Jugendliche bis zum Alter von 15 Jahren untersucht hat, zeigen negative Effekte von Stress während der Schwangerschaft auf die Geburt, die Frühgeburtlichkeitsrate, Bindungsstörungen, auf die motorische und kognitive Entwicklung des Kindes sowie eine Assoziation mit kindlichen Verhaltensregulationsstörungen und ADHS (Guardino und Schetter 2014; Van den Berg et. al 2005). Postpartale Depressionen erschweren die Kontaktaufnahme und Entwicklung einer Beziehung zwischen Mutter und Kind (Papousek, 2001). So konnten bei Kindern von Müttern, die an einer Wochenbett-Depression erkrankt waren, emotionale Auffälligkeiten, Verhaltensstörungen sowie geistige Entwicklungsverzögerungen festgestellt werden. Eine postpartale Depression birgt außerdem die Gefahr der Chronifizierung. Diese Datenlage verdeutlicht die Notwendigkeit wirksamer Screening-, Präventions- und Interventionsprogramme, um Hinweise

frühzeitig zu erkennen und den betroffenen Frauen Unterstützungsmöglichkeiten und Therapieangebote zugänglich zu machen, sowohl im Sinne der Mutter als auch zur Vermeidung negativer Folgen für das Kind (Wimmer-Puchinger & Riechler-Rössler, 2006).

4.1 Wirksamkeitsevaluation (LMU)

Einordnung der Befunde

Das Ziel der vorliegenden Wirksamkeitsevaluation war es, die klinische Effektivität der online-basierten Achtsamkeitsintervention bei klinisch auffälligen Schwangeren zu untersuchen. Hierbei werden folgende inhaltliche Hypothesen geprüft: Die Intervention führt bei betroffenen Frauen zu

1. **Sinkender Symptombelastung bzw. einer niedrigeren Prävalenz peripartaler Depressionen**
2. **Sinkenden Niveaus neuroendokriner Stressmarker**
3. **Steigender Bewertung des Gesundheitszustandes**
4. **Subjektiv positiverem Erleben der Geburt und der Bindung zum Säugling**
5. **Einer Verringerung der Häufigkeit von primären und sekundären Kaiserschnitten sowie Kinderarztbesuchen**

Zusammenfassend konnte die Wirksamkeit der eMBI-Intervention teilweise und unter Berücksichtigung von Moderatoren bestätigt werden. Am deutlichsten zeigte sich die Wirksamkeit der Intervention im Hinblick auf den Anstieg des proximalen Maßes der Achtsamkeit (*Manipulation Check*) aber auch im Hinblick auf das Absinken der schwangerschaftsbezogenen Ängste.

Manipulation Check: Es kann als gesichert erachtet werden, dass die Achtsamkeitsintervention im statistischen Mittel tatsächlich die Achtsamkeit erhöht. Dies auch konsistent unter Berücksichtigung der betrachteten Moderatoren. Persönliches *Coaching* scheint aber auch unabhängig von der eMBI-Intervention die Achtsamkeit zu erhöhen.

Diskussion zu Hypothese 1: Die eMBI-Intervention konnte uneingeschränkt als wirksam im Hinblick auf eine Senkung der schwangerschaftsbezogenen Ängste beurteilt werden. Diese Wirksamkeit wurde konfirmatorisch nachgewiesen und kann somit als effektiv wirksam betrachtet werden.

Der Effekt auf depressive Symptome konnte unter Berücksichtigung klinischer Störungen und v.a. für die Subgruppe mit persönlichem Coaching, hier auch für die allgemeine Angst, nachgewiesen werden. Somit lässt sich die Intervention als eingeschränkt wirksam bzgl. der depressiven Symptomatik beurteilen. Ebenso zeigte sich ein signifikanter positiver Effekt auf die Rate der postpartalen Depressionen (Hypothese 17). Gerade der Moderator des „persönlichen Coachings“ scheint generell in der vorliegenden Studie eine tiefere Bedeutung zu haben. So konnte der Effekt auf die generellen Angstsymptome erst unter Berücksichtigung des Moderators „persönliches Coaching“ aufgedeckt werden. Hierbei zeigt sich auch, dass es vor allem die Kombination aus digitaler und persönlicher Intervention zu sein scheint, die wirksam Angstsymptome senkt, wobei sich dies eindeutiger bzgl. der situationsbezogenen im Vergleich zur merkmalsbezogenen Angst zeigt. Dies zeigt sich in Übereinstimmung zur aktuellen Studienlage. Während Auswirkungen auf die Geburtsangst bis zum jetzigen Zeitpunkt nicht untersucht worden sind, zeigen vergleichbare Studien vor allem in Face-to-Face-Situationen signifikante Reduktionen von Depressivität und allgemeiner Angst im Verlauf der Schwangerschaft (Pan et al, 48,49,50 im Hauptpaper). Dennoch kann festgehalten werden, dass die eMBI in Kombination mit einer Minimalanzahl an zusätzlichen persönlichen Kontakten (im Mittelwert 2,3) erreicht und somit deutlich ressourcenschonender ist als vergleichbare Studien, die 8 persönliche Interventionsmodule eingesetzt haben. Da die Moderatoranalysen als exploratorisch zu betrachten sind, bedürfen die in diesem Zusammenhang erworbenen Erkenntnisse die Bestätigung in Folgeuntersuchungen.

Diskussion zu Hypothese 2: Eine Wirksamkeit für die neuroendokrinen Stressmarker konnte in den Analysen nicht nachgewiesen werden. Bei der Abgabe der Proben gab es mehrere Limitationen. Zum einen wurden sensible Abnahmezeiten, vorgegeben durch die zirkadiane Cortisolrhythmik, nicht eingehalten, zum anderen kam eine geringere Anzahl an Proben als erwartet. Daher haben diese Analysen einen explorativen Charakter.

Diskussion zu Hypothese 3: Eine Wirksamkeit für die Gesamtbewertung des eigenen Gesundheitszustandes konnte nicht nachgewiesen werden. Für die 5 Skalen wurde je eine ANOVA mit dem Zwischensubjektfaktor „Interventionsgruppierung“ und den jeweils relevanten Kovariaten verwendet. Insgesamt zeigten sich auf keiner der Skalen Effekte der Intervention ($p \geq .051$). Somit lässt sich festhalten, dass die Intervention nicht zu einem Anstieg des Gesundheitszustandes führt; zumindest nicht gemäß den verwendeten Modellen und Messinstrumenten. Je höher allerdings die Stressbewertung und je niedriger das Bildungsniveau ist, desto niedriger wurde der Gesundheitszustand eingeschätzt – unabhängig von der Gruppenzugehörigkeit. Aktuelle Publikationen zur gesundheitsbezogenen Lebensqualität, die ebenfalls den EQ-5D eingesetzt haben, zeigen ähnliche Ergebnisse, u. a., dass die eigene Bewertung des Gesundheitszustandes relativ stabil ist während des Verlaufs der Schwangerschaft, mit einem Peak im 2. Trimenon (Wu et al. 2021).

Diskussion zu Hypothese 4: Es konnte eine Wechselwirkung zwischen der Interventionsgruppierung und den Visiten bzgl. der präpartalen Bindung an das Kind aufgedeckt werden; dies allerdings in eine kontraintuitive Richtung: So scheint vor allem die Kontrollgruppe hinsichtlich dieses Maßes zu profitieren, was sicherlich näherer Untersuchung bedarf. Zusammenfassend konnte eine Wirksamkeit für eine subjektiv positivere erlebte Bindung zum Säugling nicht nachgewiesen werden. Die existierende Literatur zeigt ebenso divergente Ergebnisse. Zum Teil kann in Studien die präpartale Bindung zum Kind durch Achtsamkeitsprogramme verbessert werden (Gheibi et al. 2020), zum Teil zeigt sich kein Einfluss oder teilweise auch negative Interaktionen ähnlich der unseren (Golmakani et al. 2021). Die Bindung zum ungeborenen Kind scheint ein multidimensionales Konstrukt zu sein, welches nur teilweise durch Achtsamkeit zu beeinflussen ist.

Diskussion zu Hypothese 5: Eine Wirksamkeit für die Häufigkeit von primären und sekundären Kaiserschnitten sowie Kinderarztvorstellungen nach Geburt konnte nicht nachgewiesen werden. Bereits eigene Vorstudien in einem ähnlichen Kollektiv konnten die international und deutschlandweit nachgewiesenen Auswirkungen von psychischen Erkrankungen auf die Sectiorate nicht nachvollziehen (Gawlik et al. 2014, Wallwiener et al. 2020, Review Grote et al.). Dies kann verschiedene Ursachen haben: Baden-Württemberg ist ein Bundesland, das sich von den anderen Bundesländern in vielen Aspekten unterscheidet. So ist die Arbeitslosenquote die 2. niedrigste in Deutschland mit 3,9%, die Kaufkraft liegt deutlich über dem bundesweiten Durchschnitt, was zu sehr günstigen sozio-ökonomischen Verhältnissen und einem positiven Zugang zum Gesundheitsbereich führt (www.baden-wuerttemberg.de). Dieses Thema bedarf weiterer Untersuchung in anderen Bundesländern, z. B. in Berlin oder Nordrhein-Westfalen. In der Moderatoranalyse mit zusätzlichen Coachings zeigte sich eine Abnahme der primären Kaiserschnitte in der Gruppe mit eMBI und Mehrfach-Gesprächen, dies war jedoch aufgrund der geringen Fallzahl nicht signifikant und bedarf ebenso weiterer Untersuchung.

Diskussion zu Hypothese 6: Eine Wirksamkeit für die Senkung postpartaler Depressionen konnte durch den Einsatz der digitalen Intervention, unabhängig von den Moderatoren, eine signifikant geringere Rate an postpartaler Depressionssymptomatik 4-6 Wochen nach Geburt nachweisen. So betrug die Rate an auffälligen Müttern zum Messzeitpunkt V6 in der Interventionsgruppe 27,96%, in der Kontrollgruppe 49,97%. Dieses Ergebnis ist statistisch signifikant und die Intervention kann somit als effektiv zur Senkung der postpartalen Depressionen bewertet werden.

Weitere Anmerkungen

Nach Anpassung des Evaluationskonzept konnte die geplante Fallzahl mit den Rekrutierungsbemühungen durch die koordinierenden Stellen erreicht werden. Diese Aussage bezieht sich allerdings auf die Zahl der in die Studienarme eingeschlossenen Frauen, nicht auf die Fallzahlen, die in den finalen konfirmatorischen und exploratorischen Analysen eingegangen sind. Durch fehlende Werte und dadurch bedingte listenweise Fallausschlüsse war die Fallzahl in diesen Analysen skalenspezifisch reduziert. Es konnte allerdings bzgl. der betrachteten Variablen nicht nachgewiesen werden, dass diese Fallausschlüsse zu einem *Bias* in der Stichprobe führen. Auch weisen die konfirmatorischen Analysen dennoch bzgl. großen und mittleren Effekte eine hohe Teststärke auf.

Lediglich kleine Effekte lassen sich mit den vorliegenden Daten im Falle non-signifikanter Ergebnisse meist nicht ausschließen.

Um in den vorliegenden Daten kleine Effekte möglichst nicht zu übersehen, wurden die konfirmatorischen Hypothesen zwei weiteren Verfahren unterzogen – zum einen den Mehrebenenanalysen und zum anderen den multiplen Imputationen. Beide diese Verfahren konnten zwar keine über die primären Analysen hinausgehende Wirksamkeit aufdecken, jedoch ist die Sicherheit bzgl. des Ausschlusses kleiner Effekte dadurch erhöht.

Nicht auszuschließen ist aber, dass sich auch bei hoher Teststärke einige Effekte in den vorliegenden Daten nicht entfalten konnten. Dies ist neben der reduzierten Fallzahl der primären Analysen auch auf die Imbalance der Interventionsgruppierung zurückzuführen, die sich aus unterschiedlichen Drop-Out-Raten ergab, was eine statistische Impräzision in der Interpretation von Wechselwirkungseffekten bedingen kann und sich in den Moderatoranalysen unter Berücksichtigung nicht adhoc geplanter Gruppierungsvariablen weiter verschärft. Darüber hinaus ist die Teststärke in der Untersuchung von dreifachen Wechselwirkungen bei gegebener Stichprobe ohnehin weiter reduziert. Dies konnte z.B. begründen, warum eben die dreifache Wechselwirkung in der multivariaten Analyse des STAI unter Berücksichtigung der klinischen Gruppierungsvariable statistisch bedeutsam war, in den univariaten post-hoc-Analysen allerdings nicht mehr. Auch die Tatsache, dass die zweifache Wechselwirkung zwischen der Interventionsgruppierung und des Visitenfaktors bei der Untersuchung des PRAQ-R unter Berücksichtigung der Gruppierung durch das persönliche *Coaching* nicht mehr bedeutsam war, könnte hiermit in Zusammenhang stehen.

Darüber hinaus ist von wenig Verzerrung durch mangelnde Reliabilität der Messinstrumente auszugehen. Die internen Konsistenzen der meisten Instrumente wurden direkt anhand der Studienstichprobe bestätigt und bewegen sich in einem guten bis sehr guten Bereich.

Zusammenfassung

Insgesamt kann die eMBI-Intervention uneingeschränkt als wirksam im Hinblick auf einen Anstieg der Achtsamkeit, eine Senkung der schwangerschaftsbezogenen Ängste sowie eine Senkung der Rate an postpartalen Depressionen beurteilt werden. Die Intervention lässt sich ferner als eingeschränkt wirksam bzgl. der depressiven Symptomatik beurteilen. Hier scheint der Effekt erst deutlich unter Berücksichtigung der klinischen Störungen und in Kombination mit zusätzlicher persönlicher Intervention, was allerdings einer Bestätigung in Folgeuntersuchungen bedarf. Ebenso lässt sich die Intervention als eingeschränkt wirksam im Hinblick auf die generelle Angstsymptomatik beurteilen. Diese scheint abhängig von zusätzlicher persönlicher Intervention, wenn dies auch weniger eindeutig der Fall war im Vergleich zur Wirksamkeit hinsichtlich der depressiven Symptomatik.

Zusammenfassend konnte die Wirksamkeit der eMBI-Intervention teilweise und unter Berücksichtigung von Moderatoren bestätigt werden, wenn auch die Effektstärken sich im kleinen bis mittleren Bereich bewegen. Am deutlichsten zeigte sich die Wirksamkeit der Intervention im Hinblick auf den Anstieg des proximalen Maßes der Achtsamkeit aber auch im Hinblick auf das Absinken der schwangerschaftsbezogenen Ängste.

Der Effekt auf depressive Symptome konnte unter Berücksichtigung klinischer Störungen und v.a. für die Subgruppe mit persönlichem Coaching nachgewiesen werden. Gerade letzterer Moderator scheint generell in der vorliegenden Studie eine tiefere Bedeutung zu haben. So scheint persönliches Coaching auch unabhängig von der eMBI-Intervention die Achtsamkeit zu erhöhen.

Der Effekt auf die generellen Angstsymptome konnten erst unter Berücksichtigung des Moderators „persönliches Coaching“ aufgedeckt werden. Hierbei zeigt sich auch, dass es vor allem die Kombination aus digitaler und persönlicher Intervention zu sein scheint, die wirksam Angstsymptome senkt, wobei sich dies eindeutiger bzgl. der situationsbezogenen im Vergleich zur merkmalsbezogenen Angst zeigt.

Beide Effekte des persönlichen Coachings bewirken eine Verbesserung der affektiven Symptomatik, welche mit der App allein nach einem einzigen Erstgespräch, nicht nachweisbar war. Bei der depressiven Symptomatik erscheint dieser Effekt zum teil unabhängig von der eMBI zu sein, bei der Angstsymptomatik nicht. Diese Auswirkung ist vereinbar mit der aktuellen Literatur, welche

mehrheitlich zeigt, dass im Rahmen von Face-To Face Situationen mit dem Einsatz achtsamkeitsbasierter Techniken die affektive Symptomatik positiv beeinflusst werden kann (Guo et al. 2021, Lönnberg et al. 2020) (.Unsere Auswertung zeigt, dass nicht alle wöchentlichen Module persönlich gehalten werden müssen, der Blended Care Ansatz aber in Kombination mit einer geringfügigen Anhalt an Coachings die Effektivität der eMBI steigern kann.

Des Weiteren scheint die Intervention im späteren Verlauf zu einer niedrigeren Rate depressiver Auffälligkeiten zu führen. Auch konnte eine Wechselwirkung zwischen der Interventionsgruppierung und der Visiten bzgl. der präpartalen Bindung an das Kind aufgedeckt werden; dies allerdings in eine kontraintuitive Richtung: So scheint vor allem die Kontrollgruppe hinsichtlich dieses Maßes zu profitieren, was sicherlich näherer Untersuchung bedarf.

Um in den vorliegenden Daten kleine Effekte möglichst nicht zu übersehen, wurden die konfirmatorischen Hypothesen zwei weiteren Verfahren unterzogen – zum einen den Mehrebenenanalysen und zum anderen den multiplen Imputationen. Beide diese Verfahren konnten zwar keine über die primären Analysen hinausgehende Wirksamkeit aufdecken, jedoch ist die Sicherheit bzgl. des Ausschlusses kleiner Effekte dadurch erhöht.

4.2 Gesundheitsökonomische Evaluation (Universität Bielefeld)

Auf Grundlage der dargestellten Ergebnisse lässt sich die zentrale Hypothese der gesundheitsökonomischen Evaluation nicht bestätigen, d.h. es kann keine Reduzierung der (direkten) Behandlungskosten durch die Intervention festgestellt werden. Sowohl das zentrale Gesamtkostenmodell aus GKV-Sicht als auch die Analyse aus gesellschaftlicher Perspektive und die Sensitivitätsanalysen zeigen keinen signifikanten Einfluss der Studiengruppe. Der Regressionsanalyse zufolge weist die IG im Vergleich zur KG nicht signifikant erhöhte direkte Gesamtkosten (0,9 %; $p = 0,889$) auf. Der Interaktionsterm aus Studiengruppe und -center deutet zudem an, dass die IG im Studiocenter Tübingen nicht signifikant höhere Kosten (16,7 %; $p = 0,075$) aufweist. Einen möglichen Erklärungsansatz stellt die Anzahl der Gespräche, die begleitend zur Studie in den Zentren stattgefunden haben, dar. Demnach wurden in Heidelberg (1,71) signifikant mehr Gespräche geführt als in Tübingen (1,08) ($W = 12354$, $p < 0,001^*$). Im Rahmen der Formulierung des Modellterms zeigte sich jedoch eine höhere Modellgüte bei Berücksichtigung der Center-Variable. Die Frage nach der Ursache des potentiellen „Center-Effektes“ kann somit nicht eindeutig beantwortet werden.

Ein statistisch signifikanter Effekt der Intervention auf die Gesamtkosten lässt sich nicht ableiten. Dies gilt auch für die Analyse der Gesamtkosten aus gesellschaftlicher Perspektive, deren Ergebnisse sich nur geringfügig von den direkten Gesamtkosten unterscheiden. Da die Analysen keinen Gruppenunterschied bzw. keine Mehrkosten der IG ergaben, wurde auf die Berechnung des ICER bzw. auf eine Kosten-Effektivitätsanalyse verzichtet. Signifikante Unterschiede der beiden Studiengruppen hinsichtlich der weiteren, aus gesundheitsökonomischer Perspektive relevanten, Outcomes sind ebenfalls nicht festzustellen. Demnach zeigen die unkontrollierten Analysen keine Unterschiede zwischen den Studiengruppen für die Anzahl der Arztkontakte, die ambulanten Kosten, die haus- und fachärztlichen Versorgung (Anzahl Kontakte und Kosten) sowie hinsichtlich der erkrankungsspezifischen Medikation im Beobachtungszeitraum. Lediglich die Anzahl der Arztkontakte wird, der Regressionsanalyse zufolge, signifikant vom Interaktionsterm aus Studiengruppe und psychischer Vorerkrankung beeinflusst ($\text{Exp}(\beta) = 1,407$; $p = 0,004$). Das Modell gibt für die Teilnehmerinnen der IG mit einer psychischen Vorerkrankung 27,3 % mehr Arztkontakte an. Ein isolierter Gruppeneffekt besteht jedoch nicht.

Die Relevanz digitaler Gesundheitsanwendungen hat in den vergangenen Jahren deutlich zugenommen. web-basierten Interventionen wird u. a. zugeschrieben, potentiell kosteneffektive Alternativen zur Behandlung von Depressionen und Angststörungen darstellen zu können. Die Kontrollbedingungen und Kostenkomponenten der Evaluationsstudien variieren jedoch stark (Donker et al., 2015; Manoharan et al., 2019; Mitchell et al., 2021). Die gesundheitsökonomische Perspektive wurde in bisherigen Evaluationen digitaler Interventionen für perinatale Depressionen oder Angststörungen nur selten berücksichtigt (Hussain-Shamsy et al., 2020). Die Kosteneffektivität analoger Interventionsansätze erscheint dagegen besser erforscht, wobei die Heterogenität der Designs und Settings auch hier zu Unsicherheiten führt (Camacho & Shields, 2018). Die Einordnung der dargestellten Ergebnisse in den Kontext der Literatur ist somit eingeschränkt. Monteiro et al. (2022) untersuchten eine web-basierte selbstgesteuerte kognitive Verhaltenstherapie zur Förderung der mentalen Gesundheit von Müttern mit geringem Risiko für eine postpartale Depression mittels

Wartelistenvergleichs. Demnach führt der Ansatz zu nicht signifikanten Einsparungen und einer nicht signifikanten Erhöhung der QALYs (quality adjusted life years). Aus ihren Analysen leiten die Autoren ab, dass die Implementation der Intervention aus gesellschaftlicher Perspektive einen kosteneffektiven Ansatz zur Verbesserung der perinatalen mentalen Gesundheit darstellen könnte. Die vorliegenden Ergebnisse stimmen mit dieser Studie insofern überein, dass keine signifikanten Unterschiede zwischen den Kosten der IG und KG gefunden wurden. Zheng et al. (2022) vergleichen in ihrer dreiarmligen RCT zwei psychoedukative Interventionen (web-basiert/ von zu Hause) mit einer Kontrollgruppe. Die Studie fokussiert Erstgebärende in der frühen postpartalen Phase und kommt zu dem Ergebnis, dass die web-basierte Intervention beide Alternativen dominiert und die höchste Wahrscheinlichkeit der Kosteneffektivität aufweist. Sie weisen jedoch auf Unsicherheiten hin und berichten keine inferenzstatistischen Analysen der durchschnittlichen Gesamtkosten der drei Studiengruppen. Analog zu den Ergebnissen der vorliegenden Evaluation, lassen sich auf Grundlage der Studie von Zheng et al. (2022) keine uneingeschränkten Schlussfolgerungen ableiten.

Für die Einordnung der Ergebnisse sind zudem die folgenden Stärken und Limitationen der gesundheitsökonomischen Evaluation zu berücksichtigen: Die zentrale Datenbasis der Evaluation bilden die GKV-Routinedaten, welche zahlreiche Informationen aus den Diagnose- und Leistungsdaten der verschiedenen Versorgungsbereiche zur Verfügung stellen und sich im Vergleich zu Primärdaten durch ihre Verfügbarkeit und Objektivität auszeichnen. Zu bedenken ist jedoch, dass die Daten nicht primär zu Forschungszwecken erhoben werden und somit einige Einschränkungen, wie z. B. die Abhängigkeit vom Codierungsverhalten der Leistungserbringer, mit sich bringen (Neubauer et al., 2017; Zeidler & Braun, 2012).

Die Plausibilität der Routinedaten wurde mittels umfangreicher Prüfungen weitgehend sichergestellt, wobei ein Einfluss kassenspezifischer Dokumentationsformate nicht gänzlich ausgeschlossen werden kann. Dies gilt insbesondere für Bereiche wie Leistungen der Haushaltshilfe, welche aus der Analyse der Gesamtkosten ausgeschlossen wurden. Darüber hinaus wurde eine Einordnung der Ergebnisse anhand verfügbarer Statistiken und Analysen vorgenommen. Hierzu zählen u. a. die Honorarberichte der Kassenärztlichen Bundesvereinigung (2019), die durchschnittlichen Erlöse relevanter (geburtsbezogener) DRGs (Reimbursement Institute, 2022) und die Auswertungen des WIdO zum Arzneimittelverbrauch (Wissenschaftliches Institut der AOK, 2021). Unter Berücksichtigung der Länge des Beobachtungszeitraumes und der besonderen Versorgungssituation im Kontext der Schwangerschaft, wurden keine auffälligen Abweichungen identifiziert. Hinsichtlich der Berechnung der Produktivitätsverluste ist jedoch zu berücksichtigen, dass sich das auf Basis der Erwerbstätigenrechnung (Destatis, 2022a) und der volkswirtschaftlichen Gesamtrechnung (Destatis, 2022b) ermittelte durchschnittliche Entgelt pro Arbeitnehmer pro Tag auf die gesamte Arbeitnehmerschaft bezieht. Da die weibliche im Vergleich zur männlichen Arbeitnehmerschaft einen insgesamt geringeren Beschäftigungsumfang und ein niedrigeres Gehaltsniveau aufweist (Destatis, 2021), ist davon auszugehen, dass die Produktivitätsverluste der betrachteten weiblichen Studienpopulation tendenziell überschätzt wurden.

Die angesetzten Interventionskosten basieren auf den zu Projektbeginn kalkulierten Kosten der online-basierten Achtsamkeitsintervention. Der durchschnittliche Preis sowie der potenzielle Höchstbetrag vergleichbarer Anwendungen (Indikationsgruppe Depressionen und Angststörungen) sind im DiGA-Verzeichnis (Bundesinstitut für Arzneimittel und Medizinprodukte [BfArM]) verfügbar. Angepasst an die Dauer der Anwendung der vorliegenden Intervention, läge der durchschnittliche tatsächliche Preis bei 154,93 €, der potenzielle Höchstbetrag bei 285,60 € sowie ein möglicher Abschlag bei 6 %. Die Kosten der evaluierten online-basierten Achtsamkeitsintervention in Höhe von 216,86 € liegen damit 69,74 € unter dem maximalen und 71,22 € über dem minimalen Betrag und sind als angemessen zu betrachten. In zwei Sensitivitätsanalysen wurden mögliche Abweichungen der Interventionskosten (+/- 30 %) berücksichtigt, wodurch sich jedoch keine deutlichen Änderungen der Evaluationsergebnisse ergaben.

Weitere methodische Limitationen beziehen sich u.a. auf die reduzierte Stichprobe. Die Ausschlüsse waren erforderlich, da Teilnehmerinnen, deren Beobachtungszeitraum nicht vollständig in den GKV-Routinedaten abgebildet war, nicht in die gesundheitsökonomische Evaluation einbezogen werden konnten. Im Rahmen einer Sensitivitätsanalyse wurden daher Teilnehmerinnen aufgenommen, für die lediglich ein Monat der Beobachtungszeit fehlte. Das Ergebnis der Analyse änderte sich hinsichtlich der Studiengruppe nicht wesentlich. Als weitere Einschränkung sind die fehlenden Werte zu benennen, welche zu Fallausschlüssen in den Regressionsanalysen führten. Da jedoch im Vorfeld

eine zufällige Verteilung der Missings festgestellt wurde, ist davon auszugehen, dass hiermit keine Beeinflussung der Ergebnisse verbunden ist. Mit dem gewählten Studiendesign (RCT), welches in der wissenschaftlichen Forschung als Goldstandard gilt (Guyatt et al., 1995), wurde möglichen Verzerrungen Rechnung getragen. Ungeachtet dessen ist zu berücksichtigen, dass in den dargestellten Analysen keine Vorjahresdaten der Inanspruchnahme und der Kosten von Gesundheitsleistungen der Teilnehmerinnen berücksichtigt wurden. Darüber hinaus können weitere unbeobachtete Faktoren die Ergebnisse beeinflussen. Mit Blick auf die Anwendung der Modellselektion anhand des AIC ist anzumerken, dass diese einen möglichen Bias durch multiples Testen nicht vollständig auffängt. Demnach könnte das Verfahren zu einer Überschätzung der Signifikanz der Ergebnisse führen. Da die Ergebnisse für das interessierende Merkmal der Studiengruppe in den meisten Analysen p-Werte deutlich über 5 % aufweisen, ist dies jedoch zu vernachlässigen.

Insgesamt deuten die Ergebnisse darauf hin, dass mit dem erzielten Nutzen der Intervention keine Reduktion der (direkten) Gesamtkosten, jedoch auch keine signifikant höheren Kosten aus Sicht der GKV verbunden sind. Dies erscheint, angesichts der verhältnismäßig niedrigen Kosten der online-basierten Achtsamkeitsintervention, plausibel. Dennoch sind die beschriebenen Einschränkungen zu berücksichtigen und einzubeziehen, dass die Power-Kalkulation der RCT nicht auf die gesundheitsökonomische Evaluation ausgerichtet war. Wenngleich entsprechende Schlussfolgerungen damit einer gewissen Unsicherheit unterliegen, liefern die Ergebnisse einige relevante Erkenntnisse zur Inanspruchnahme von Gesundheitsleistungen und zu den damit assoziierten Kosten der betrachteten Population. Zum einen ist die höhere Anzahl ambulanter Arztkontakte von Teilnehmerinnen der IG mit einer psychischen Vorerkrankung zu nennen. Zum anderen liefern die Kostentreiberanalysen Informationen über für die betrachtete Zielgruppe besonders relevante Leistungsbereiche. Erwartungsgemäß machen stationäre und ambulante Leistungen den größten Anteil der Gesamtkosten aus, gefolgt von den Hebammenleistungen. Im ambulanten Leistungsbereich entfallen die meisten Kontakte auf die Gynäkologie bzw. Geburtshilfe und sonstige Facharztgruppen. Besuche bei Neurologen, Psychiatern usw. kommen im Mittel seltener vor, weisen allerdings die höchsten durchschnittlichen Fallkosten auf. Aus gesundheitsökonomischer Perspektive wird somit die Relevanz innovativer Versorgungsansätze für peripartale Depressionen und Angststörungen betont. In zukünftigen Studien sollte möglichst auch die für die gesundheitsökonomische Evaluation erforderliche Stichprobengröße berücksichtigt werden. Darüber hinaus sind Optimierungen der Datenbasis der Evaluation (z. B. Vorjahresdaten, Selbstzahlerleistungen) und längere Nachbeobachtungszeiträume anzustreben, um mögliche Effekte einer Intervention auf die Kosten aus GKV- und gesellschaftlicher Perspektive bewerten zu können.

5. Literaturverzeichnis

Areskog, B, Uddenberg, N & Kjessler, B (1981). Fear of childbirth in late pregnancy. *Gynecol Obstet Invest*, 12(5), 262-266.

Bergant, A., Nguyen, T. & Heim, K. (1998). Deutschsprachige Fassung und Validierung der „Edinburgh postnatal depression scale“. *Dtsch Med Wochenschr*, 123, 35–40.

Bundesinstitut für Arzneimittel und Medizinprodukte. (2022). DiGA-Verzeichnis. <https://diga.bfarm.de/de>

Camacho, E. M. & Shields, G. E. (2018). Cost-effectiveness of interventions for perinatal anxiety and/or depression: a systematic review. *BMJ open*, 8(8), e022022. <https://doi.org/10.1136/bmjopen-2018-022022>

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (rev. ed.). Lawrence Erlbaum Associates, Inc. <http://search.ebscohost.com/login.aspx?direct=true&db=psych&AN=1987-98267-000&site=ehost-live>

Cumming, G. (2009). Inference by eye: reading the overlap of independent confidence intervals. *Statistics in medicine*, 28(2), 205–220. <https://doi.org/10.1002/sim.3471>

- Deligiannidis KM, Meltzer-Brody S, Gunduz-Bruce H, et al. (2021) Effect of Zuranolone vs Placebo in Postpartum Depression: A Randomized Clinical Trial. *JAMA Psychiatry*.78(9):951–959. doi:10.1001/jamapsychiatry.2021.1559
- Diego, M. A., Jones, N. A., Field, T., Hernandez-Reif, M., Schanberg, S., Kuhn, C. & Gonzalez-Garcia, A. (2006). Maternal Psychological Distress, Prenatal Cortisol, and Fetal Weight. *Psychosomatic Medicine*, 68(5), 747–753. <https://doi.org/10.1097/01.psy.0000238212.21598.7b>
- Dimidjian, S., Goodman, S. H., Felder, J. N., Gallop, R., Brown, A. P. & Beck, A. (2016). Staying well during pregnancy and the postpartum: A pilot randomized trial of mindfulness-based cognitive therapy for the prevention of depressive relapse/recurrence. *Journal of Consulting and Clinical Psychology*, 84(2), 134–145. <https://doi.org/10.1037/ccp0000068>
- Donker, T., Blankers, M., Hedman, E., Ljótsson, B., Petrie, K. & Christensen, H. (2015). Economic evaluations of Internet interventions for mental health: a systematic review. *Psychological medicine*, 45(16), 3357–3376. <https://doi.org/10.1017/S0033291715001427>
- Dubber, S., Reck, C [C.], Müller, M. & Gawlik, S. (2015). Postpartum bonding: The role of perinatal depression, anxiety and maternal–fetal bonding during pregnancy. *Archives of Women's Mental Health*, 18(2), 187–195. <https://doi.org/10.1007/s00737-014-0445-4>
- Dunn (1961). Multiple Comparisons among means. *Journal of the American Statistical Association*, 56, 52–64.
- Enders, C. K. (2010). *Applied missing data analysis*. Guilford Press. <http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2010-13190-000&site=ehost-live>
- Eriksson, M. C. M., Kivi, M., Hange, D., Petersson, E.-L., Ariai, N., Häggblad, P., Ågren, H., Spak, F., Lindblad, U., Johansson, B. & Björkelund, C. (2017). Long-term effects of Internet-delivered cognitive behavioral therapy for depression in primary care—The PRIM-NET controlled trial. *Scandinavian Journal of Primary Health Care*, 35(2), 126–136. <https://doi.org/10.1080/02813432.2017.1333299>
- Fahrmeir, L., Kneib, T., Lang, S. & Marx, B. (2013). *Regression: Models, Methods and Applications*. Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-642-34333-9>
- Falkei, P. & Wittchen, H.-U. (2015). *Diagnostisches und statistisches Manual psychischer Störungen DSM-5 (1. Aufl.)*. Hogrefe.
- Faul, F., Erdfelder, E., Buchner, A. & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, 41(4), 1149–1160.
- Faul, F., Erdfelder, E., Lang, A.-G. & Buchner, A. (2007). GPower 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191. <http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2007-11814-002&site=ehost-live>
- Field, T, Diego, M & Hernandez-Reif, M (2006). Prenatal depression effects on the fetus and newborn: a review. *Infant Behav Dev*, 29(3), 445-455. doi: 10.1016/j.infbeh.2006.03.003
- Field, T., Hernandez-Reif, M., Diego, M., Figueiredo, B., Schanberg, S. & Kuhn, C. (2006). Prenatal cortisol, prematurity and low birthweight. *Infant Behavior & Development*, 29(2), 268–275. <https://doi.org/10.1016/j.infbeh.2005.12.010>
- Gawlik, S, Waldeier, L, Muller, M, Szabo, A, Sohn, C & Reck, C (2013). Subclinical depressive symptoms during pregnancy and birth outcome--a pilot study in a healthy German sample. *Arch Womens Ment Health*, 16(2), 93-100. doi: 10.1007/s00737-012-0320-0
- Gensorowsky, D., Witte, J., Batram, M. & Greiner, W [Wolfgang] (2022). Market access and value-based pricing of digital health applications in Germany. *Cost effectiveness and resource allocation : C/E*, 20(1), 25. <https://doi.org/10.1186/s12962-022-00359-y>

Gerste, B. & Roick, C. (2014). Prävalenz und Inzidenz sowie Versorgung depressiver Erkrankungen in Deutschland. In J. Klauber, C. Günster, B. Gerste, B.-P. Robra, N. Schmacke & (Keine Angabe) (Hrsg.), Versorgungs-Report 2013/2014: Schwerpunkt: Depression. Schattauer. https://www.wido.de/fileadmin/Dateien/Dokumente/Publicationen_Produkte/Buchreihen/Versorgungsreport/2013-2014/Kapitel%20mit%20Deckblatt/wido_vsr2013-2014_gesamt.pdf

Gesetz für eine bessere Versorgung durch Digitalisierung und Innovation - Digitale-Versorgung-Gesetz (2019). https://www.bgbl.de/xaver/bgbl/start.xav?startbk=Bundesanzeiger_BGBL&start=%2F%2F%2A%5B%40attr_id%27bgbl119s2562.pdf%27%5D#__bgbl__%2F%2F%5B%40attr_id%3D%27bgbl119s2562.pdf%27%5D__1657198559500

Gräfe, K., Zipfel, S [Stephan], Herzog, W [Wolfgang] & Löwe, B. (2004). Screening psychischer Störungen mit dem 'Gesundheitsfragebogen für Patienten (PHQ-D)': Ergebnisse der deutschen Validierungsstudie = Screening for psychiatric disorders with the Patient Health Questionnaire (PHQ) Results from the German validation study. *Diagnostica*, 50(4), 171–181. <https://doi.org/10.1026/0012-1924.50.4.171>

Greiner, W [Wolfgang]. (2012). 4 Der EQ-5D der EuroQol-Gruppe. In O. Schöffski & J.-M. Graf von der Schulenburg (Hrsg.), *Gesundheitsökonomische Evaluationen* (S. 411–422). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-21700-5_15

Greiner, W [Wolfgang] & Damm, O. (2012). 3 Die Berechnung von Kosten und Nutzen. In O. Schöffski & J.-M. Graf von der Schulenburg (Hrsg.), *Gesundheitsökonomische Evaluationen* (S. 23–42). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-21700-5_3

Guo P, Zhang X, Liu N, Wang J, Chen D, Sun W, Li P, Zhang W: Mind-body interventions on stress management in pregnant women: A systematic review and meta-analysis of randomized controlled trials. *J Adv Nurs* 2021, 77(1):125-146.

Guyatt, G. H., Sackett, D. L., Sinclair, J. C., Hayward, R., Cook, D. J. & Cook, R. J. (1995). Users' guides to the medical literature. IX. A method for grading health care recommendations. Evidence-Based Medicine Working Group. *JAMA*, 274(22), 1800–1804. <https://doi.org/10.1001/jama.274.22.1800>

Hassdenteufel K, Müller M, Abele H, Brucker SY, Graf J, Zipfel S, Bauer A, Jakubowski P, Pauluschke-Fröhlich J, Wallwiener S. Using an Electronic Mindfulness-based Intervention (eMBI) to Improve Maternal Mental Health during Pregnancy: Results from a Randomized Controlled Trial. *Psychotherapy and Psychosomatics*. 2022 Under Review.

Heidenreich, T., Ströhle, G. & Michalak, J. (2006). Achtsamkeit: Konzeptuelle Aspekte und Ergebnisse zum Freiburger Achtsamkeitsfragebogen = Mindfulness: Conceptual aspects and results of the Freiburg Mindfulness Inventory. *Verhaltenstherapie*, 16(1), 33–40. <https://doi.org/10.1159/000091521>

Heslin, M., Jin, H., Trevillion, K., Ling, X., Nath, S., Barrett, B., Demilew, J., Ryan, E. G., O'Connor, S., Sands, P., Milgrom, J., Bick, D., Stanley, N., Hunter, M. S., Howard, L. M. & Byford, S. (2022). Cost-effectiveness of screening tools for identifying depression in early pregnancy: a decision tree model. *BMC health services research*, 22(1), 774. <https://doi.org/10.1186/s12913-022-08115-x>

Hirani, S. P., Beynon, M., Cartwright, M., Rixon, L., Doll, H., Henderson, C., Bardsley, M., Steventon, A., Knapp, M., Rogers, A., Bower, P., Sanders, C., Fitzpatrick, R., Hendy, J. & Newman, S. P. (2014). The effect of telecare on the quality of life and psychological well-being of elderly recipients of social care over a 12-month period: The Whole Systems Demonstrator cluster randomised trial. *Age and ageing*, 43(3), 334–341. <https://doi.org/10.1093/ageing/aft185>

Hussain-Shamsy, N., Shah, A., Vigod, S. N., Zaheer, J. & Seto, E. (2020). Mobile Health for Perinatal Depression and Anxiety: Scoping Review. *Journal of medical Internet research*, 22(4), e17011. <https://doi.org/10.2196/17011>

Akronym: Mind:Pregnancy
Förderkennzeichen: 01NVF17034

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (2020). Allgemeine Methoden - Version 6.0, 2020. https://www.iqwig.de/methoden/allgemeine-methoden_version-6-0.pdf?rev=144030

Kassenärztliche Bundesvereinigung. (2021). Richtlinie der Kassenärztlichen Bundesvereinigung nach § 75 Absatz 7 SGB V zur Vergabe der Arzt-, Betriebsstätten- sowie der Praxisnetznummern. https://www.kbv.de/media/sp/Arztnummern_Richtlinie.pdf

Kassenärztlichen Bundesvereinigung. (2019). Honorarbericht: Quartal 4/2019. <https://www.kbv.de/html/honorarbericht.php>

Kirschbaum, C. (1991). Cortisolmessung im Speichel, eine Methode der biologischen Psychologie. Huber.

Kirschbaum, C. & Hellhammer, D. H. (2007). Salivary cortisol. In G. Fink (Hrsg.), *Encyclopedia of Stress* (Bd. 3, S. 405–409). Academic Press.

Kirschbaum, C (1991). Cortisolmessung im Speichel, eine Methode der biologischen Psychologie. Cortisolmessung im Speichel, eine Methode der biologischen Psychologie.

Kirschbaum, C & Hellhammer, DH (1994). Salivary cortisol in psychoneuroendocrine research: recent developments and applications. *Psychoneuroendocrinology*, 19(4), 313-333.

Khoury, B, Lecomte, T, Fortin, G, Masse, M, Therien, P, Bouchard, V, Chapleau, MA, Paquin, K & Hofmann, SG (2013). Mindfulness-based therapy: a comprehensive meta-analysis. *Clin Psychol Rev*, 33(6), 763-771. doi: 10.1016/j.cpr.2013.05.005

Klainin-Yobas, P, Cho, MA & Creedy, D (2012). Efficacy of mindfulness-based interventions on depressive symptoms among people with mental disorders: a meta-analysis. *Int J Nurs Stud*, 49(1), 109-121. doi: 10.1016/j.ijnurstu.2011.08.014

Lan, K. K. & DeMets, D. L. (1989). Changing frequency of interim analysis in sequential monitoring. *Biometrics*, 45(3), 1017–1020.

Laux, L., Glanzmann, P., Schaffner, P. & Spielberger, C. D. (1981). *Das State-Trait-Angstinventar* (Testmappe mit Handanweisung, Fragebogen STAI-G Form X 1 und Fragebogen STAI-G Form X 2). Beltz.

Little, R. J. (1988). A Test of Missing Completely at Random for Multivariate Data with Missing Values. *Journal of the American Statistical Association*, 83(404), 1198–1202. <https://doi.org/10.1080/01621459.1988.10478722>

Lönnerberg G, Jonas W, Unternaehrer E, Bränström R, Nissen E, Niemi M: Effects of a mindfulness based childbirth and parenting program on pregnant women's perceived stress and risk of perinatal depression-Results from a randomized controlled trial. *J Affect Disord* 2020, 262:133-142.

Löwe, B., Spitzer, R. L., Zipfel, S [S.] & Herzog, W [W.]. (2002). Gesundheitsfragebogen für Patienten [Prime MD Patient Health Questionnaire (PHQ; Spitzer, R.L., Williams, J.B., Kroenke, K., Linzer, M., deGruy, F.V. III, Hahn, S.R., Brody, D. & Johnson, J.G., 1994) - German version].

Lucas, R. E. (2022). It's Time To Abandon The Cross-Lagged Panel Model. <https://doi.org/10.31234/osf.io/pkec7>

Maas, C. J. M. & Hox, J. J. (2005). Sufficient Sample Sizes for Multilevel Modeling. *Methodology*, 1(3), 85–91. <https://doi.org/10.1027/1614-2241.1.3.85>

Manoharan, A., McMartin, K., Young, C., Gajic-Veljanoski, O., Ali, A. & Walter, M. (2019). Internet-Delivered Cognitive Behavioural Therapy for Major Depression and Anxiety Disorders: A Health Technology Assessment. *Ontario health technology assessment series*, 19(6), 1–199.

Mitchell, L. M., Joshi, U., Patel, V., Lu, C. & Naslund, J. A. (2021). Economic Evaluations of Internet-Based Psychological Interventions for Anxiety Disorders and Depression: A Systematic Review. *Journal of affective disorders*, 284, 157–182. <https://doi.org/10.1016/j.jad.2021.01.092>

Akronym: Mind:Pregnancy
Förderkennzeichen: 01NVF17034

Monteiro, F., Antunes, P., Pereira, M., Canavarro, M. C. & Fonseca, A. (2022). Cost-utility of a web-based intervention to promote maternal mental health among postpartum women presenting low risk for postpartum depression. *International journal of technology assessment in health care*, 38(1), e62. <https://doi.org/10.1017/S0266462322000447>

Neubauer, S., Zeidler, J [Jan], Lange, A. & Graf von der Schulenburg, J.-M [Johann-Matthias]. (2017). *Prozessorientierter Leitfaden für die Analyse und Nutzung von Routinedaten der Gesetzlichen Krankenversicherung* (1. Auflage). Nomos Verlagsgesellschaft mbH & Co. KG. <http://nbn-resolving.org/urn:nbn:de:bsz:31-epflicht-1046499>

Pruessner, J. C., Kirschbaum, C., Meinlschmid, G. & Hellhammer, D. H. (2003). Two formulas for computation of the area under the curve represent measures of total hormone concentration versus time-dependent change. *Psychoneuroendocrinology*, 28(7), 916–931. <http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2003-07222-007&site=ehost-live>

R Core Team. (2020). *R: A language and environment for statistical computing* (Version 4.0.3) [Computer software]. R Foundation for Statistical Computing. Wien, Österreich. www.R-project.org

Rahmenvereinbarung nach § 134 Abs. 4 und 5 SGB V zwischen dem GKV-Spitzenverband und den Verbänden der Hersteller von digitalen Gesundheitsanwendungen 16.12.2021). https://www.gkv-spitzenverband.de/media/dokumente/krankenversicherung_1/telematik/digitales/DiGA_Rahmenvereinbarung_nach_134_Abs_4_und_5_SGB_V_Stand_16.12.2021.pdf

Reck, C [Corinna], Klier, C. M., Pabst, K., Stehle, E., Steffenelli, U., Struben, K. & Backenstrass, M. (2006). The German version of the Postpartum Bonding Instrument: Psychometric properties and association with postpartum depression. *Archives of Women's Mental Health*, 9(5), 265–271. <https://doi.org/10.1037/t01756-000>;

Reck, C, Struben, K, Backenstrass, M, Stefenelli, U, Reinig, K, Fuchs, T, Sohn, C & Mundt, C (2008). Prevalence, onset and comorbidity of postpartum anxiety and depressive disorders. *Acta Psychiatr Scand*, 118(6), 459-468. doi: 10.1111/j.1600-0447.2008.01264.x

Reimbursement Institute. (2022). DRG Kataloge: Fallpauschalen. <https://app.reimbursement.info/drgs>

Rigby, R. A [R. A.] & Stasinopoulos, D. M. (2005). Generalized additive models for location, scale and shape (with discussion). *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 54(3), 507–554. <https://doi.org/10.1111/j.1467-9876.2005.00510.x>

Saß, H., Wittchen, H.-U., Zaudig, M. & Houben, I. (2003). *Diagnostisches und Statistisches Manual Psychischer Störungen - Textrevison (DSM-IV-TR)*. Hogrefe. <http://search.ebscohost.com/login.aspx?direct=true&db=pdx&AN=0159500&site=ehost-live>

Seidl, H., Bowles, D., Bock, J.-O., Brettschneider, C., Greiner, W [W.], König, H.-H. & Holle, R. (2015). FIMA--Fragebogen zur Erhebung von Gesundheitsleistungen im Alter: Entwicklung und Pilotstudie [FIMA--questionnaire for health-related resource use in an elderly population: development and pilot study]. *Gesundheitswesen (Bundesverband der Ärzte des Öffentlichen Gesundheitsdienstes (Germany))*, 77(1), 46–52. <https://doi.org/10.1055/s-0034-1372618>

Shea, AK, Streiner, DL, Fleming, A, Kamath, MV, Broad, K & Steiner, M (2007). The effect of depression, anxiety and early life trauma on the cortisol awakening response during pregnancy: preliminary results. *Psychoneuroendocrinology*, 32(8-10), 1013-1020. doi: 10.1016/j.psyneuen.2007.07.006

Sozialgesetzbuch Fünftes Buch - Gesetzliche Krankenversicherung (Zuletzt geändert durch Art. 8 G v. 2022 I 969).

Stalder, T., Kirschbaum, C., Kudielka, B. M., Adam, E. K., Pruessner, J. C., Wüst, S., Dockray, S., Smyth, N., Evans, P., Hellhammer, D. H., Miller, R., Wetherell, M. A., Lupien, S. J. & Clow, A. (2016). Assessment of the cortisol awakening response: Expert consensus guidelines. *Psychoneuroendocrinology*, 63, 414–432. <https://doi.org/10.1016/j.psyneuen.2015.10.010>

Akronym: Mind: Pregnancy
Förderkennzeichen: 01NVF17034

Stasinopoulos, M. D., Rigby, R. A [Robert A.], Heller, G. Z., Voudouris, V. & Bastiani, F. de. (2017). Flexible regression and smoothing: Using GAMLSS in R. A Chapman & Hall book. CRC Press Taylor & Francis Group a Chapman & Hall book.

Statistisches Bundesamt. (2021). Arbeitsmarkt und Verdienste: Auszug aus dem Datenreport 2021. https://www.destatis.de/DE/Service/Statistik-Campus/Datenreport/Downloads/datenreport-2021-kap-5.pdf?__blob=publicationFile

Statistisches Bundesamt. (2022a). Erwerbstätigkeit: Erwerbstätige und Arbeitnehmer nach Wirtschaftsbereichen (Inlandskonzept) 1000 Personen. <https://www.destatis.de/DE/Themen/Arbeit/Arbeitsmarkt/Erwerbstaetigkeit/Tabellen/arbeitsnehmer-wirtschaftsbereiche.html>

Statistisches Bundesamt. (2022b). Volkswirtschaftliche Gesamtrechnungen: Bruttonationaleinkommen, verfügbares Einkommen und Volkseinkommen. <https://www.destatis.de/DE/Themen/Wirtschaft/Volkswirtschaftliche-Gesamtrechnungen-Inlandsprodukt/Tabellen/Irvgr04.html#242556>

Takuya Yanagida. (2022). Miscellaneous Functions 'T. Yanagida': Package 'misty'. <https://cran.r-project.org/web/packages/misty/misty.pdf>

Techniker Krankenkasse (2017). Geburtenreport. Eine Routinedatenanalyse zu Kaiserschnitt und Frühgeburt. from <https://www.tk.de/centaurus/servlet/contentblob/951938/Datei/90325/Geburtenreport-2017.pdf%5B18.07.2018>

Turkstra, E., Creedy, D. K., Fenwick, J., Buist, A., Scuffham, P. A. & Gamble, J. (2015). Health services utilization of women following a traumatic birth. Archives of Women's Mental Health, 18(6), 829–832. <https://doi.org/10.1007/s00737-014-0495-7>

Verbeke, E., Bogaerts, A., Nuyts, T., Crombag, N. & Luyten, J. (2022). Cost-effectiveness of mental health interventions during and after pregnancy: A systematic review. Birth (Berkeley, Calif.), 49(3), 364–402. <https://doi.org/10.1111/birt.12623>

Walach, H., Buchheld, N., Buttenmüller, V., Kleinknecht, N. & Schmidt, S. (2006). Measuring mindfulness--The Freiburg Mindfulness Inventory (FMI). Personality and Individual Differences, 40(8), 1543–1555. <https://doi.org/10.1016/j.paid.2005.11.025>

Wilkinson, A., Anderson, S. & Wheeler, S. B. (2017). Screening for and Treating Postpartum Depression and Psychosis: A Cost-Effectiveness Analysis. Maternal and child health journal, 21(4), 903–914. <https://doi.org/10.1007/s10995-016-2192-9>

Wissenschaftliches Institut der AOK. (2021). Der GKV-Arzneimittelmarkt: Klassifikation, Methodik und Ergebnisse 2021. <https://doi.org/10.4126/FRL01-006429816>

Wissenschaftliches Institut der AOK. (2022). Arzneimittel-Stammdatei des GKV Arzneimittelindex.

WISTA: Wirtschaft und Statistik (3. Aufl.). (2018). Statistisches Bundesamt.

Wu H, Sun W, Chen H, Wu Y, Ding W, Liang S, Huang X, Chen H, Zeng Q, Li Z, Xiong P, Huang J, Akinwunmi B, Zhang CJP, Ming WK (2021) Health-related quality of life in different trimesters during pregnancy. Health Qual Life Outcomes. 2021(07)21;19(1):182. doi: 10.1186/s12955-021-01811-y.

Zeidler, J [J.] & Braun, S. (2012). Sekundärdatenanalysen. In O. Schöffski & J.-M. Graf von der Schulenburg (Hrsg.), Gesundheitsökonomische Evaluationen. Springer Berlin Heidelberg.

Zheng, Q., Shi, L., Zhu, L., Jiao, N., Chong, Y. S., Chan, S. W.-C., Chan, Y. H., Luo, N., Wang, W. & He, H. (2022). Cost-effectiveness of Web-Based and Home-Based Postnatal Psychoeducational Interventions for First-time Mothers: Economic Evaluation Alongside Randomized Controlled Trial. Journal of medical Internet research, 24(3), e25821. <https://doi.org/10.2196/25821>

6. Anhang

-

7. Anlagen

Anlage 1: Kostenwochenbuch

Anlage 2: Kostentreiberanalysen

Anlage 3: Auswertungsmethoden

Anlage 4: Analyse Gesamtkosten

Anlage 5: Sensitivitätsanalyse1

Anlage 6: Sensitivitätsanalyse2

Anlage 7: Analyse Gesamtkosten

Anlage 8: Analyse amb Arztkontakte

Anlage 9: Analyse Hausarztkontakte



Kostenwochenbuch zur Erfassung selbst finanzierter Gesundheitsleistungen

Studien-ID:

Das **Kostenwochenbuch** dient der Erfassung Ihrer selbst finanzierten Gesundheitsleistungen. Wir bitten Sie, Ihre Ausgaben ab Beginn der Nutzung der *mindmom*-App bis fünf Monate nach der Geburt Ihres Kindes regelmäßig und wahrheitsgemäß einzutragen. Es gibt dabei keine richtigen oder falschen Angaben.

Wichtig! Bitte senden Sie das Kostenwochenbuch mit den letzten Cortisol-Teströhrchen (fünf Monate nach der Geburt Ihres Kindes) zurück.

Vielen Dank für Ihre Unterstützung!

Bei Fragen zum Kostenwochenbuch wenden Sie sich bitte an:
mind-pregnancy@uni-bielefeld.de

1 Alternative diagnostische Verfahren und Heilmethoden

Unter alternativen diagnostischen Verfahren und Heilmethoden sind hier Gesundheitsleistungen zu verstehen, die von Ihrer gesetzlichen Krankenkasse nicht finanziert und daher von Ihnen selbst bezahlt werden müssen, wie z. B.:

Alternative diagnostische Verfahren:

- Zusätzlicher 2D-, 3D- oder 4D-Ultraschall
- Streptokokken-Test
- Test auf Zytomegalievirus (CMV)
- Test auf Toxoplasmose
- Ultraschall der Eierstöcke oder der Brust zur Krebsfrüherkennung
- usw.

Alternative Heilmethoden:

- Akupunktur
- Osteopathie
- Ernährungstherapie
- Bewegungstherapie
- Hydrotherapie
- Mutter-Kind-Yoga
- usw.

Ausfüllhinweise (siehe Tabelle Seite 3):

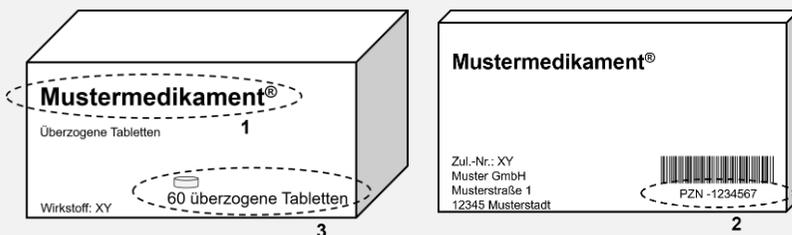
1. Bitte geben Sie zunächst an, um welche Art der Leistung es sich handelt (z. B. Akupunktur).
2. Tragen Sie nun ein, wann Sie die Leistung in Anspruch genommen haben. Bei mehreren Leistungen geben Sie das erste Datum an.
3. Tragen Sie anschließend die Anzahl der Leistungen (Sitzungen, Anwendungen o. Ä.) ein.
4. Geben Sie, falls dieser bekannt ist, den Einzelpreis der Leistung an.
5. Bitte berechnen Sie, wenn möglich, den Gesamtpreis, indem Sie die Anzahl der Leistungen mit deren Einzelpreis multiplizieren.

Achtung! Die genannten Leistungen sind lediglich als Beispiele und keineswegs als Empfehlungen zu verstehen.

2 Nicht-verschreibungspflichtige, homöopathische oder naturheilkundliche Arzneimittel

Unter nicht-verschreibungspflichtigen, homöopathischen oder naturheilkundlichen Arzneimitteln sind hier Medikamente zu verstehen, die ohne Rezept des Arztes in der Apotheke gekauft werden können, wie z. B.: Vitamin-Präparate, Folsäure-Präparate, magensäurebindende Arzneimittel, Globuli usw.

Ausfüllhinweise (siehe Tabelle Seite 4):



1. Bitte tragen Sie den vollständigen Medikamentennamen ein.
2. Tragen Sie die PZN-Nummer zur eindeutigen Identifikation des Medikamentes ein. Diese finden Sie unter dem Strichcode auf einer der Seiten der Packung. Die Buchstabenfolge „PZN“ steht immer vor der 7- oder 8-stelligen Nummer.
3. Geben Sie anschließend die Packungsgröße mit Einheitsbezeichnung an (z. B.: 60 Tabletten, 50 ml oder 5 Spritzen).
4. Tragen Sie die Dosierung ein (z. B.: 1 Tablette täglich, 3 mal 20 Tropfen täglich, 1 Spritze pro Woche oder bei Bedarf).
5. Geben Sie an, seit wann die Einnahme erfolgt (Monat/Jahr). Wenn Sie das Medikament schon länger als ein Jahr einnehmen, ist das Jahr ausreichend.
6. Bitte geben Sie an, welche Kosten Ihnen für den Kauf des Medikamentes entstanden sind. Falls Sie ein Medikament über einen längeren Zeitraum einnehmen, tragen Sie es bitte jeweils in der Woche ein, in der Sie mit der aktuellen Packung beginnen.

Achtung! Die genannten Arzneimittel sind lediglich als Beispiele und keineswegs als Empfehlungen zu verstehen.

WOCHE 1

1 Alternative diagnostische Verfahren und Heilmethoden

Leistung	Datum	Anzahl	Einzelpreis, in Euro	Gesamtpreis, in Euro
<i>Beispiel: Akupunktur</i>	01.01.19	4	30,00	(4 x 30,00€) = 120,00

WOCHE 2

1 Alternative diagnostische Verfahren und Heilmethoden

Leistung	Datum	Anzahl	Einzelpreis, in Euro	Gesamtpreis, in Euro
<i>Beispiel: Akupunktur</i>	01.01.19	4	30,00	(4 x 30,00€) = 120,00

WOCHE 3

1 Alternative diagnostische Verfahren und Heilmethoden

Leistung	Datum	Anzahl	Einzelpreis, in Euro	Gesamtpreis, in Euro
<i>Beispiel: Akupunktur</i>	01.01.19	4	30,00	(4 x 30,00€) = 120,00

WOCHE 4

1 Alternative diagnostische Verfahren und Heilmethoden

Leistung	Datum	Anzahl	Einzelpreis, in Euro	Gesamtpreis, in Euro
<i>Beispiel: Akupunktur</i>	01.01.19	4	30,00	(4 x 30,00€) = 120,00

WOCHE 5

1 Alternative diagnostische Verfahren und Heilmethoden

Leistung	Datum	Anzahl	Einzelpreis, in Euro	Gesamtpreis, in Euro
<i>Beispiel: Akupunktur</i>	01.01.19	4	30,00	(4 x 30,00€) = 120,00

WOCHE 6

1 Alternative diagnostische Verfahren und Heilmethoden

Leistung	Datum	Anzahl	Einzelpreis, in Euro	Gesamtpreis, in Euro
<i>Beispiel: Akupunktur</i>	01.01.19	4	30,00	(4 x 30,00€) = 120,00

WOCHE 7

1 Alternative diagnostische Verfahren und Heilmethoden

Leistung	Datum	Anzahl	Einzelpreis, in Euro	Gesamtpreis, in Euro
<i>Beispiel: Akupunktur</i>	01.01.19	4	30,00	(4 x 30,00€) = 120,00

WOCHE 10

1 Alternative diagnostische Verfahren und Heilmethoden

Leistung	Datum	Anzahl	Einzelpreis, in Euro	Gesamtpreis, in Euro
<i>Beispiel: Akupunktur</i>	01.01.19	4	30,00	(4 x 30,00€) = 120,00

WOCHE 11

1 Alternative diagnostische Verfahren und Heilmethoden

Leistung	Datum	Anzahl	Einzelpreis, in Euro	Gesamtpreis, in Euro
<i>Beispiel: Akupunktur</i>	01.01.19	4	30,00	(4 x 30,00€) = 120,00

WOCHE 12

1 Alternative diagnostische Verfahren und Heilmethoden

Leistung	Datum	Anzahl	Einzelpreis, in Euro	Gesamtpreis, in Euro
<i>Beispiel: Akupunktur</i>	01.01.19	4	30,00	(4 x 30,00€) = 120,00

WOCHE 13

1 Alternative diagnostische Verfahren und Heilmethoden

Leistung	Datum	Anzahl	Einzelpreis, in Euro	Gesamtpreis, in Euro
<i>Beispiel: Akupunktur</i>	01.01.19	4	30,00	(4 x 30,00€) = 120,00

WOCHE 14

1 Alternative diagnostische Verfahren und Heilmethoden

Leistung	Datum	Anzahl	Einzelpreis, in Euro	Gesamtpreis, in Euro
<i>Beispiel: Akupunktur</i>	01.01.19	4	30,00	(4 x 30,00€) = 120,00

WOCHE 15

1 Alternative diagnostische Verfahren und Heilmethoden

Leistung	Datum	Anzahl	Einzelpreis, in Euro	Gesamtpreis, in Euro
<i>Beispiel: Akupunktur</i>	01.01.19	4	30,00	$(4 \times 30,00\text{€}) = 120,00$

WOCHE 16

1 Alternative diagnostische Verfahren und Heilmethoden

Leistung	Datum	Anzahl	Einzelpreis, in Euro	Gesamtpreis, in Euro
<i>Beispiel: Akupunktur</i>	01.01.19	4	30,00	(4 x 30,00€) = 120,00

WOCHE 17

1 Alternative diagnostische Verfahren und Heilmethoden

Leistung	Datum	Anzahl	Einzelpreis, in Euro	Gesamtpreis, in Euro
<i>Beispiel: Akupunktur</i>	01.01.19	4	30,00	(4 x 30,00€) = 120,00

WOCHE 18

1 Alternative diagnostische Verfahren und Heilmethoden

Leistung	Datum	Anzahl	Einzelpreis, in Euro	Gesamtpreis, in Euro
<i>Beispiel: Akupunktur</i>	01.01.19	4	30,00	(4 x 30,00€) = 120,00

WOCHE 19

1 Alternative diagnostische Verfahren und Heilmethoden

Leistung	Datum	Anzahl	Einzelpreis, in Euro	Gesamtpreis, in Euro
<i>Beispiel: Akupunktur</i>	01.01.19	4	30,00	(4 x 30,00€) = 120,00

WOCHE 20

1 Alternative diagnostische Verfahren und Heilmethoden

Leistung	Datum	Anzahl	Einzelpreis, in Euro	Gesamtpreis, in Euro
<i>Beispiel: Akupunktur</i>	01.01.19	4	30,00	(4 x 30,00€) = 120,00

WOCHE 21

1 Alternative diagnostische Verfahren und Heilmethoden

Leistung	Datum	Anzahl	Einzelpreis, in Euro	Gesamtpreis, in Euro
<i>Beispiel: Akupunktur</i>	01.01.19	4	30,00	(4 x 30,00€) = 120,00

WOCHE 22

1 Alternative diagnostische Verfahren und Heilmethoden

Leistung	Datum	Anzahl	Einzelpreis, in Euro	Gesamtpreis, in Euro
<i>Beispiel: Akupunktur</i>	01.01.19	4	30,00	(4 x 30,00€) = 120,00

WOCHE 23

1 Alternative diagnostische Verfahren und Heilmethoden

Leistung	Datum	Anzahl	Einzelpreis, in Euro	Gesamtpreis, in Euro
<i>Beispiel: Akupunktur</i>	01.01.19	4	30,00	(4 x 30,00€) = 120,00

WOCHE 24

1 Alternative diagnostische Verfahren und Heilmethoden

Leistung	Datum	Anzahl	Einzelpreis, in Euro	Gesamtpreis, in Euro
<i>Beispiel: Akupunktur</i>	01.01.19	4	30,00	(4 x 30,00€) = 120,00

WOCHE 25

1 Alternative diagnostische Verfahren und Heilmethoden

Leistung	Datum	Anzahl	Einzelpreis, in Euro	Gesamtpreis, in Euro
<i>Beispiel: Akupunktur</i>	01.01.19	4	30,00	(4 x 30,00€) = 120,00

WOCHE 26

1 Alternative diagnostische Verfahren und Heilmethoden

Leistung	Datum	Anzahl	Einzelpreis, in Euro	Gesamtpreis, in Euro
<i>Beispiel: Akupunktur</i>	01.01.19	4	30,00	(4 x 30,00€) = 120,00

WOCHE 27

1 Alternative diagnostische Verfahren und Heilmethoden

Leistung	Datum	Anzahl	Einzelpreis, in Euro	Gesamtpreis, in Euro
<i>Beispiel: Akupunktur</i>	01.01.19	4	30,00	(4 x 30,00€) = 120,00

WOCHE 28

1 Alternative diagnostische Verfahren und Heilmethoden

Leistung	Datum	Anzahl	Einzelpreis, in Euro	Gesamtpreis, in Euro
<i>Beispiel: Akupunktur</i>	01.01.19	4	30,00	(4 x 30,00€) = 120,00

WOCHE 29

1 Alternative diagnostische Verfahren und Heilmethoden

Leistung	Datum	Anzahl	Einzelpreis, in Euro	Gesamtpreis, in Euro
<i>Beispiel: Akupunktur</i>	01.01.19	4	30,00	(4 x 30,00€) = 120,00

WOCHE 30

1 Alternative diagnostische Verfahren und Heilmethoden

Leistung	Datum	Anzahl	Einzelpreis, in Euro	Gesamtpreis, in Euro
<i>Beispiel: Akupunktur</i>	01.01.19	4	30,00	(4 x 30,00€) = 120,00

WOCHE 31

1 Alternative diagnostische Verfahren und Heilmethoden

Leistung	Datum	Anzahl	Einzelpreis, in Euro	Gesamtpreis, in Euro
<i>Beispiel: Akupunktur</i>	01.01.19	4	30,00	(4 x 30,00€) = 120,00

WOCHE 32

1 Alternative diagnostische Verfahren und Heilmethoden

Leistung	Datum	Anzahl	Einzelpreis, in Euro	Gesamtpreis, in Euro
<i>Beispiel: Akupunktur</i>	01.01.19	4	30,00	(4 x 30,00€) = 120,00

WOCHE 33

1 Alternative diagnostische Verfahren und Heilmethoden

Leistung	Datum	Anzahl	Einzelpreis, in Euro	Gesamtpreis, in Euro
<i>Beispiel: Akupunktur</i>	01.01.19	4	30,00	(4 x 30,00€) = 120,00

WOCHE 34

1 Alternative diagnostische Verfahren und Heilmethoden

Leistung	Datum	Anzahl	Einzelpreis, in Euro	Gesamtpreis, in Euro
<i>Beispiel: Akupunktur</i>	01.01.19	4	30,00	(4 x 30,00€) = 120,00

WOCHE 35

1 Alternative diagnostische Verfahren und Heilmethoden

Leistung	Datum	Anzahl	Einzelpreis, in Euro	Gesamtpreis, in Euro
<i>Beispiel: Akupunktur</i>	01.01.19	4	30,00	(4 x 30,00€) = 120,00

Bitte prüfen Sie, ob alle Angaben vollständig sind.

Vielen Dank und alles Gute für Sie!

Kontakt: mind-pregnancy@uni-bielefeld.de

Anlage 2: Kostentreiberanalysen

Kostentreiberanalysen

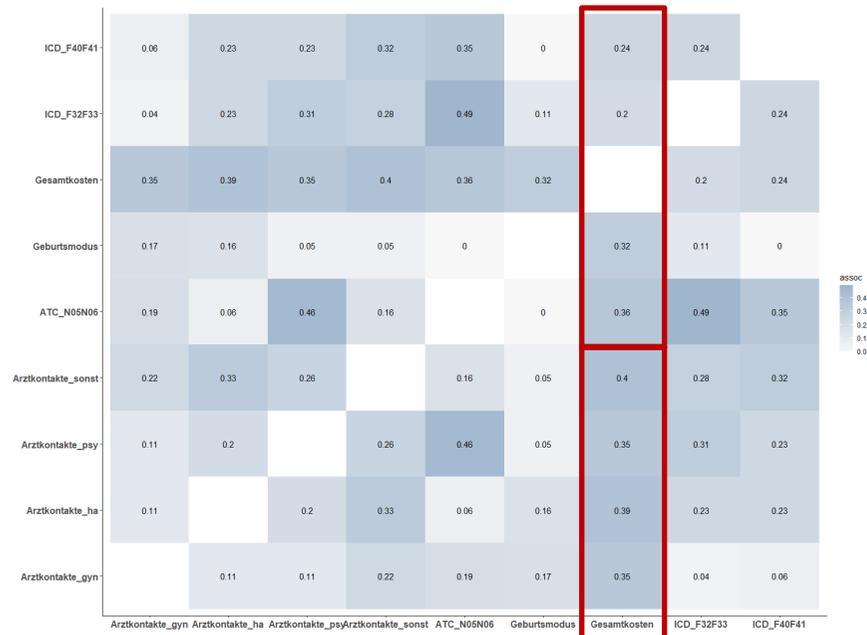


Abbildung 1: Korrelationen und Assoziationen Gesamtkosten (IG)

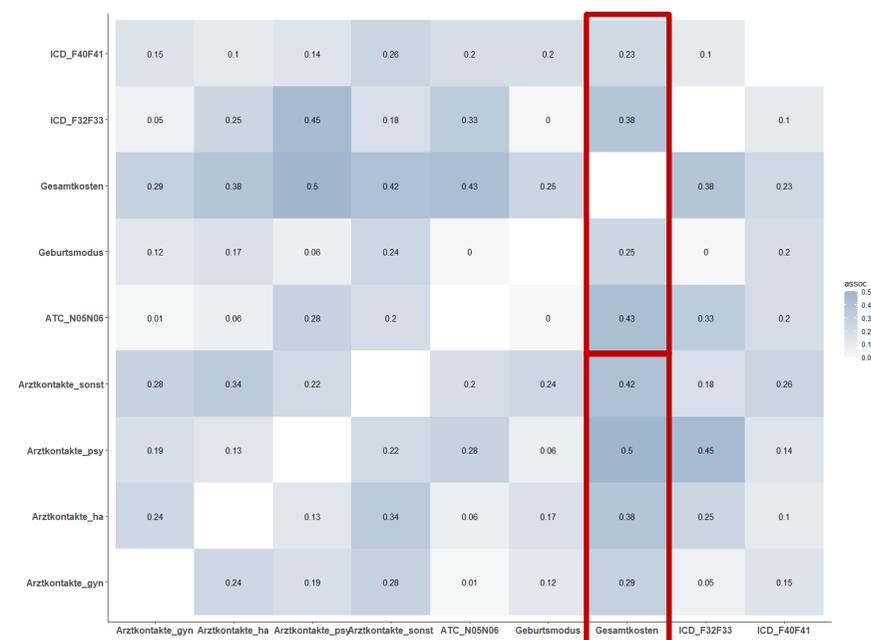


Abbildung 2: Korrelationen und Assoziationen Gesamtkosten (KG)

Anlage 3: Auswertungsmethoden

Auswertungsmethoden

Tabelle 1: Übersicht Leistungsbereiche/ Outcomes und Analysemethoden

Outcome/ Leistungsbereich	Operationalisierung	Test	Modell
Baseline			
Alter	Jahre bei Einschluss	t-Test	
Bildung	kein Abschluss; Volks-/Hauptschule; Mittlere Reife; Fachhochschulreife; Abitur	Wilcoxon-Mann-Whitney-U-Test	
Einkommen	< 1.500 €; 1.500 - 2.999 €; 3.000 - 4.999 €; 5.000 - 8.000 €	Wilcoxon-Mann-Whitney-U-Test	
Anzahl der im Haushalt lebenden Kinder	0; 1; 2; 3	Wilcoxon-Mann-Whitney-U-Test	
Stress	PHQ-Stress-Skala (0-20)	t-Test	
Depressivität	EPDS Score (0-30)	t-Test	
Psychische Vorerkrankung	Ja/Nein	X ² -Test	
Physische Vorerkrankung	Ja/Nein	X ² -Test	
Studiencenter	Heidelberg/ Tübingen	X ² -Test	
Anzahl der Gespräche	Anzahl der Gespräche im Studiencenter	Wilcoxon-Mann-Whitney-U-Test	
Hauptanalysen			
Gesamtkosten (GKV-Perspektive)	Summe aus: <ul style="list-style-type: none"> – Ambulante Kosten – Stationäre Kosten – Kosten HSA/ PIA/ ambulantes Operieren – Arzneimittelkosten – Heil- und Hilfsmittelkosten – Hebammenkosten 	Wilcoxon-Mann-Whitney-U-Test	Generalisiertes additives Modell mit inverser Gammaverteilung
Gesamtkosten (gesellschaftliche Perspektive)	Summe aus: <ul style="list-style-type: none"> – Ambulante Kosten – Stationäre Kosten – Kosten HSA/ PIA/ ambulantes Operieren – Arzneimittelkosten – Heil- und Hilfsmittelkosten – Hebammenkosten – Produktivitätsverluste 	Wilcoxon-Mann-Whitney-U-Test	Generalisiertes additives Modell mit inverser Gammaverteilung
Ambulante Leistungen	Arztkontakte	Wilcoxon-Mann-Whitney-U-Test	

	Kosten		
	Fallkosten		
Stationäre Leistungen	Fälle	Wilcoxon-Mann-Whitney-U-Test	
	Aufenthaltsdauer (LOS)		
	Kosten		
	Fallkosten		
HSA/ PIA/ ambulantes Operieren	Fälle	Wilcoxon-Mann-Whitney-U-Test	
	Kosten		
Arzneimittel	Verordnungen	Wilcoxon-Mann-Whitney-U-Test	
	DDDs		
	Kosten		
	Kosten pro Verordnung		
Heilmittel	Positionsnummern	Wilcoxon-Mann-Whitney-U-Test	
	Kosten		
Hilfsmittel	Positionsnummern		
	Kosten		
Hebammenleistungen	Tage		
	Kosten		
Produktivitätsverluste	AU-Tage		
	Kosten		
Haushaltshilfe	Positionsnummern		
	Kosten		
Mutterschaftsgeld	Tage		
	Kosten		
Ergänzende Analysen			
Ambulante Leistungen (alle)	Arztkontakte	Wilcoxon-Mann-Whitney-U-Test	Generalisiertes additives Modell mit negativer Binomialverteilung
Hausärztliche Leistungen	Arztkontakte		Generalisiertes additives Modell mit negativer Binomialverteilung
	Kosten		
Fachärztliche Leistungen - Neurologie, Psychiatrie, Psychosomatik, Psychotherapie	Arztkontakte		
	Kosten		

- Gynäkologie und Geburtshilfe - Sonstige			
Erkrankungsspezifische Medikation	Verordnung mindestens eines Medikamentes der ATC- Gruppen N05/ N06 (Nein/ Ja)	X ² -Test	

Kosten der neuen Versorgungsform

Um die Festlegung der Kosten der online-basierten Achtsamkeitsintervention im Hinblick auf die Regelversorgung zu überprüfen, wurden Informationen aus dem aktuellen Versorgungsgeschehen vergleichbarer Digitaler Gesundheitsanwendungen (DiGA) genutzt. Der Einsatz von DiGA ist seit dem Inkrafttreten des Digitale-Versorgung-Gesetzes (DVG) (DVG, 2019) auf Grundlage einer Bewertung von Studiendaten erstattungsfähig. Den tatsächlichen Preis der können Hersteller laut § 134 SGB V, Zuletzt geändert durch Art. 8 G v. 2022 I 969 im ersten Jahr frei festlegen. Laut Rahmenvereinbarung (Rahmenvereinbarung nach § 134 Abs. 4 und 5 SGB V zwischen dem GKV-Spitzenverband und den Verbänden der Hersteller von digitalen Gesundheitsanwendungen/16.12.2021) gelten jedoch Höchstbeträge, welche mittels der Bildung von Indikationsgruppen bestimmt werden. Nach dem ersten Jahr wird ein Vergütungsbetrag mit dem Spitzenverband Bund der Krankenkassen (GKV-Spitzenverband) verhandelt, wobei neben dem tatsächlichen Preis insbesondere die positiven Versorgungseffekte der Anwendung zugrunde liegen. Im DiGa-Verzeichnis des Bundesinstituts für Arzneimittel und Medizinprodukte (BfArM) sind alle DiGA, die das Bewertungsverfahren erfolgreich durchlaufen haben, gelistet. Darin finden sich neben Informationen zur medizinischen Zweckbestimmung auch die Kosten für die GKV, mögliche Abschläge sowie ggf. Zuzahlungsbeträge für Patienten. In Anlehnung an Gensorowsky et al. (2022) wurde ein durchschnittlicher tatsächlicher Preis, ein durchschnittlicher Höchstbetrag sowie ein mittlerer Abschlag für Apps des Indikationsbereichs Depressionen und Angststörungen ermittelt (Kapitel **Fehler! Verweisquelle konnte nicht gefunden werden.**, **Fehler! Verweisquelle konnte nicht gefunden werden.**). Diese dienen als Grundlage für die Berechnung eines potenziellen Höchst- und Mindestbetrags für den Einsatz der online-basierten Achtsamkeitsintervention in der Regelversorgung, zur Einordnung der angenommenen Interventionskosten im Rahmen der Analyse der Gesamtkosten.

Tabelle 2: Ermittlung eines potenziellen Vergütungsbetrages anhand vergleichbarer Anwendungen im DiGA-Verzeichnis (Eigene Darstellung nach Gensorowsky et al. (2022))

DiGA-Charakteristika				Vergütung					
DiGA	Indikation	Interventionsansatz	Anwendungsdauer	1. Jahr nach Listung			Ab 13. Monat nach Listung		Anmerkung Preismodell
				Preis ¹	Potenzieller Höchstbetrag	Relativer Abschlag	Vergütungsbetrag ²	Abschlag ggü. Preis im 1. Jahr	
Selfapy (Gen. Angststörung)	Angststörungen	Kognitive Verhaltenstherapie	90 Tage	540,00 €	459,00 €	15,00 %	/	/	Vergütung ab 2 Lektionen
velibra	Angststörungen	Kognitive Verhaltenstherapie	90 Tage	476,00 €	540,00 €	0,00 %	/	/	/
deprexis	Depressionen	Kognitive Verhaltenstherapie	90 Tage	297,50 €	540,00 €	0,00 %	/	/	/
Novego	Depressionen	Kognitive Verhaltenstherapie	90 Tage	249,00 €	459,00 €	0,00 %	/	/	/
Selfapy (Depression)	Depressionen	Kognitive Verhaltenstherapie	90 Tage	540,00 €	459,00 €	15,00 %	/	/	Vergütung ab 2 Lektionen
Durchschnitt:				420,50 €	491,40 €	6,00 %	/	/	/
¹ Wurde der Herstellerpreis im ersten Jahr der Listung im DiGA-Verzeichnis geändert, wird an dieser Stelle der letzte Preisstand berichtet. ² Berichtet wird der letzte verfügbare Vergütungsbetrag ohne Berücksichtigung etwaiger vorheriger Preisänderungen.									
Anwendung der Ergebnisse auf die vorliegende Anwendung									
Online-basierte Achtsamkeitsintervention	Depressionen/ Angststörungen in der Schwangerschaft	Achtsamkeitsintervention	ca. 56 Tage	154,93 €	285,60 €	6,00 %	/	/	/

Anlage 4: Analyse Gesamtkosten

Analyse Gesamtkosten (GKV-Perspektive)

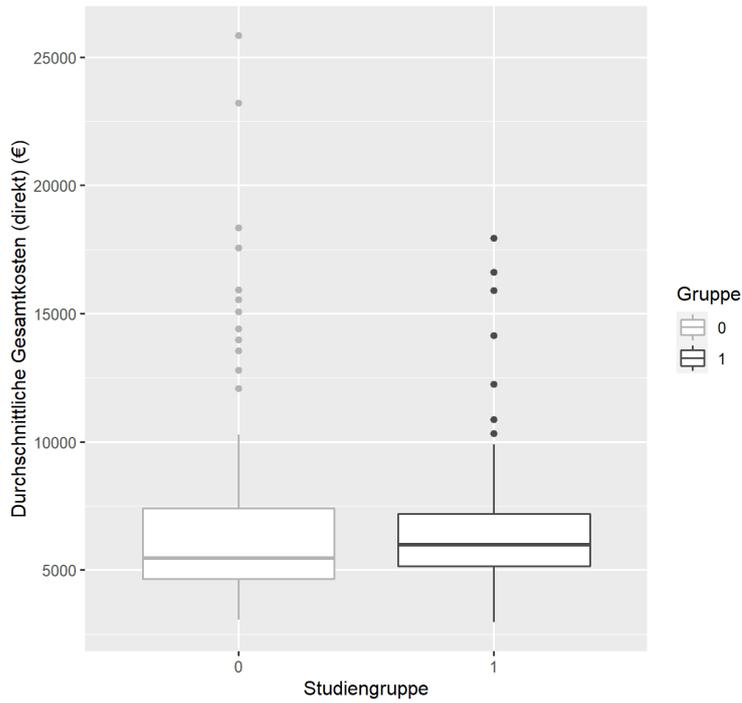


Abbildung 3: Boxplot der Gesamtkosten (GKV-Perspektive) in IG und KG

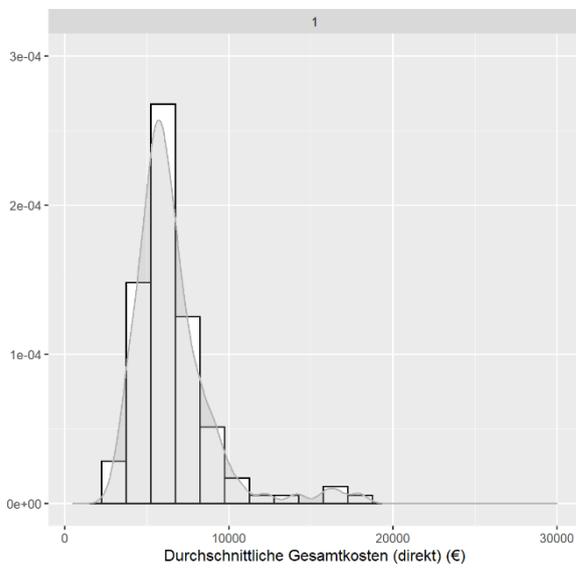


Abbildung 4: Histogramm und Dichtekurve der Gesamtkosten (GKV-Perspektive) der IG

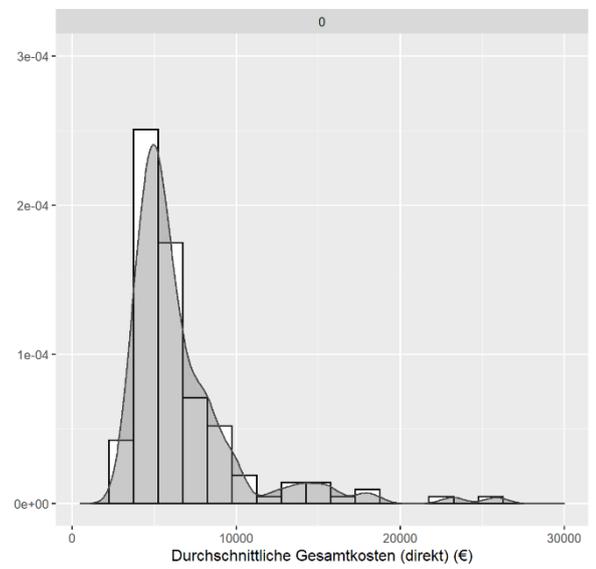


Abbildung 5: Histogramm und Dichtekurve der Gesamtkosten (GKV-Perspektive) der KG

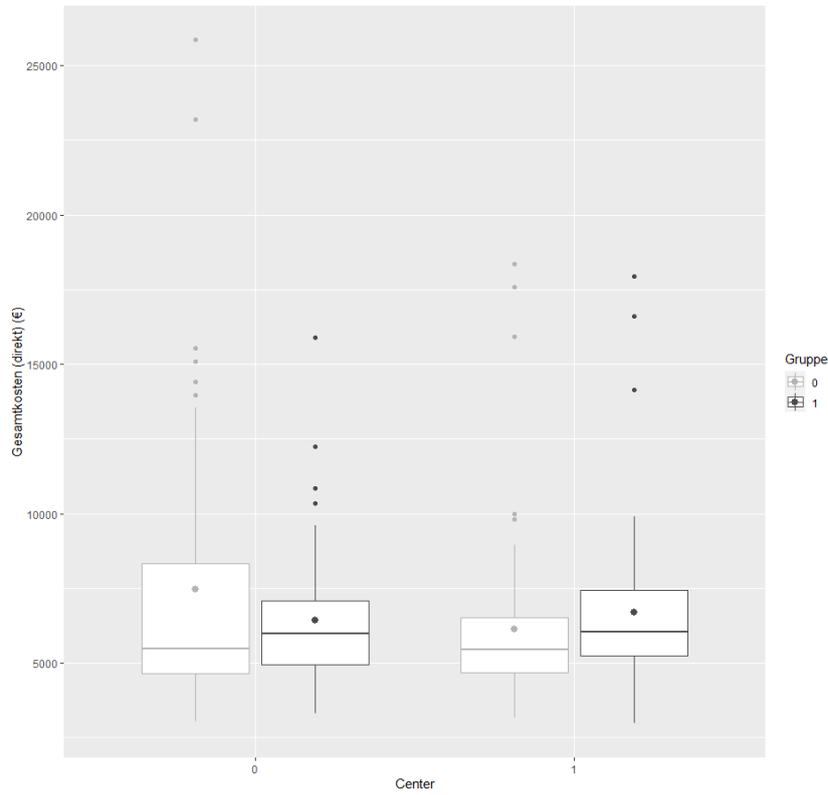


Abbildung 6: Boxplot der Gesamtkosten (GKV-Perspektive) in IG und KG differenziert nach Center (HD/TU)

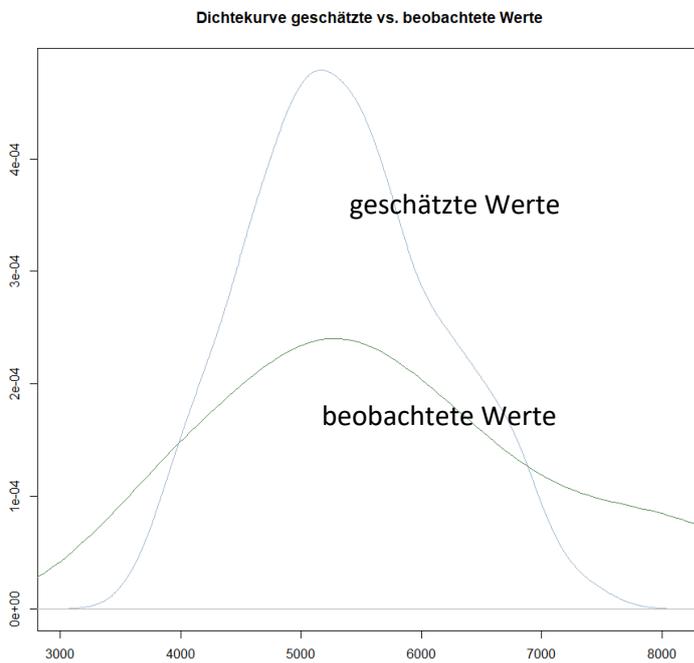


Abbildung 7: Modellgüte Analyse der Gesamtkosten (GKV-Perspektive): geschätzte vs. beobachtete Werte

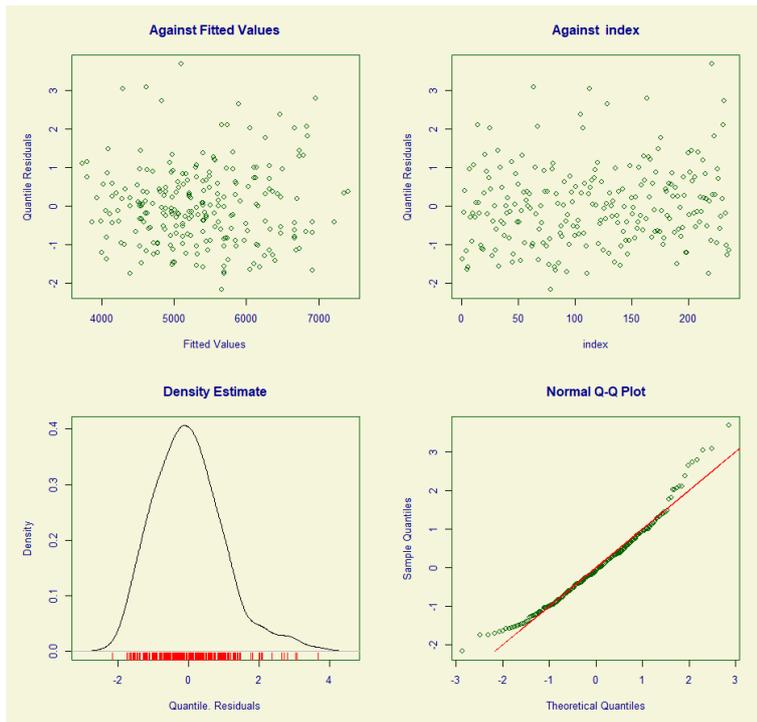


Abbildung 8: Modellgüte Analyse der Gesamtkosten (GKV-Perspektive): Residuenplots, QQ-Plot

Anlage 5: Sensitivitätsanalyse1

Sensitivitätsanalyse I: Anpassung der Interventionskosten: +/-30%

Tabelle 3: Sensitivitätsanalyse I: Wilcoxon-Mann-Whitney-U-Test Gesamtkosten (GKV-Perspektive)

Kosten	Intervention (n = 117)	Kontrolle (n = 141)	Teststatistik	p-Wert
SUMME Interventionskosten + 30 % (€)	6.628,55 (SD=2.474,05)	6.692,21 (SD = 3.608,22)	W = 7120	0,059
SUMME Interventionskosten - 30 % (€)	6.498,43 (SD=2.474,05)	6.692,21 (SD = 3.608,22)	W = 7493	0,206
Mittelwert (SD = Standardabweichung), *signifikant (p<0,05)				

Tabelle 4: Sensitivitätsanalyse I: Ausschnitt generalisiertes additives Modell mit inverser Gammaverteilung Gesamtkosten (GKV-Perspektive)

	Exp(β)	95 %-KI	p-Wert
Interventionskosten + 30%			
Studiengruppe (IG)	1,021	0,902; 1,155	0,748
Studiengruppe (IG) : Center (TU)	1,167	0,985; 1,382	0,075
*signifikant (p<0,05)			
Globale Devianz: 4260,113			
AIC: 4284,113			
SBC: 4325,679			
N = 236			
Freiheitsgrade: 12			
Residuale Freiheitsgrade: 224			
Interventionskosten - 30%			
Studiengruppe (IG)	0,997	0,880; 1,130	0,965
Studiengruppe (IG) : Center (TU)	1,168	0,985; 1,385	0,076
*signifikant (p<0,05)			
Globale Devianz: 4258,726			
AIC: 4282,726			
SBC: 4324,292			
N = 236			
Freiheitsgrade: 12			
Residuale Freiheitsgrade: 224			

Anlage 6: Sensitivitätsanalyse2

Sensitivitätsanalyse II: Erweiterung der Studienpopulation

Tabelle 5: Sensitivitätsanalyse II: Generalisiertes additives Modell mit inverser Gammaverteilung Gesamtkosten (GKV-Perspektive) (N=250)

	Exp(β)	95 %-KI	p-Wert
Intercept	5.614,513	4879,125; 6460,74	<0,001*
Studiengruppe (IG)	0,996	0,882; 1,123	0,942
Center (TU)	0,869	0,753; 1,001	0,053
Anz. der im Haushalt lebenden Kinder (1)	0,788	0,721; 0,861	<0,001*
Anz. der im Haushalt lebenden Kinder (2)	0,848	0,734; 0,98	0,026*
Anz. der im Haushalt lebenden Kinder (3)	0,723	0,461; 1,133	0,158
Stress (PHQ)	1,015	1,003; 1,028	0,018*
Physische Vorerkrankung (Ja)	0,844	0,748; 0,954	0,007*
Psychische Vorerkrankung (Ja)	1,089	1,000; 1,185	0,050
Studiengruppe (IG) : Center (TU)	1,170	0,993; 1,378	0,061
Center (TU) : Physische Vorerkrankung (Ja)	1,134	0,963; 1,335	0,133
*signifikant ($p < 0,05$) Globale Devianz: 4506,775 AIC: 4530,775 SBC: 4573,032 N = 250 Freiheitsgrade: 12 Residuale Freiheitsgrade: 238			

Anlage 7: Analyse Gesamtkosten

Analyse Gesamtkosten (gesellschaftliche Perspektive)

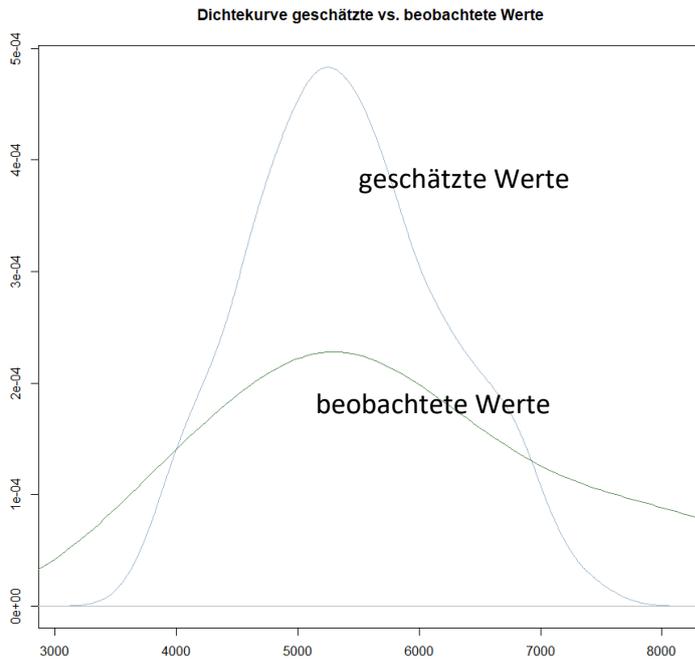


Abbildung 9: Modellgüte Analyse der Gesamtkosten (gesellschaftliche Perspektive): geschätzte vs. beobachtete Werte

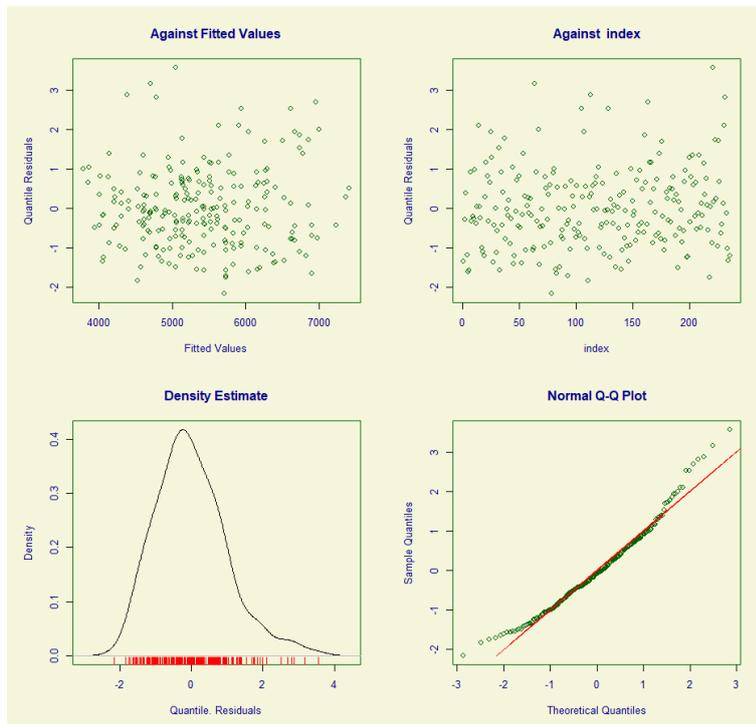


Abbildung 10: Modellgüte Analyse der Gesamtkosten (gesellschaftliche Perspektive): Residuenplots, QQ-Plot

Anlage 8: Analyse amb Arztkontakte

Analyse ambulante Arztkontakte

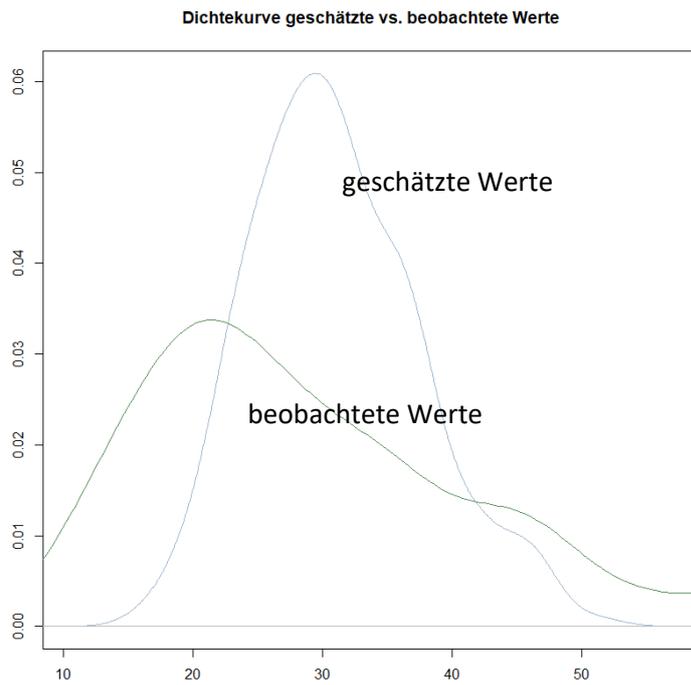


Abbildung 11: Modellgüte Analyse der ambulanten Arztkontakte: geschätzte vs. beobachtete Werte

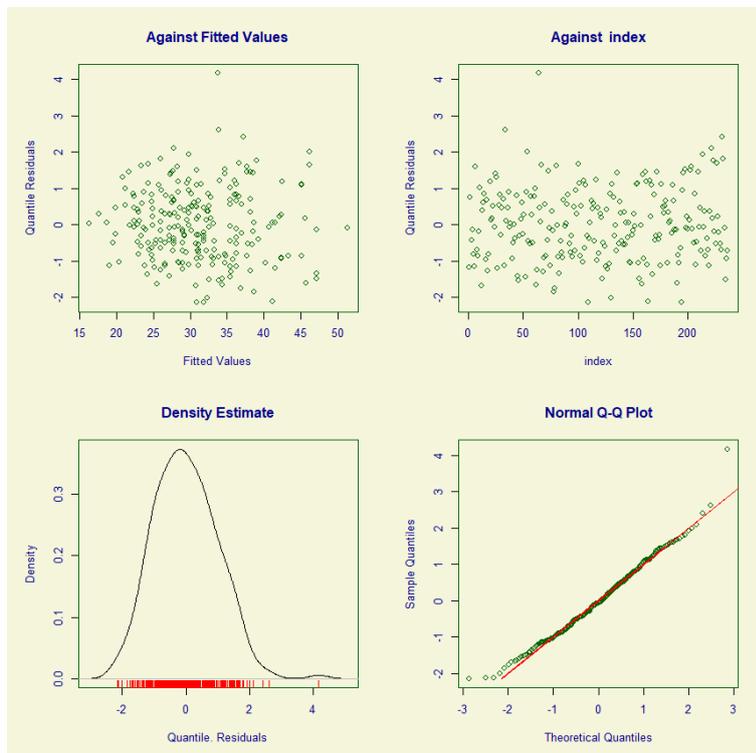


Abbildung 12: Modellgüte Analyse der ambulanten Arztkontakte: Residuenplots, QQ-Plot

Anlage 9: Analyse Hausarztkontakte

Analyse Hausarztkontakte

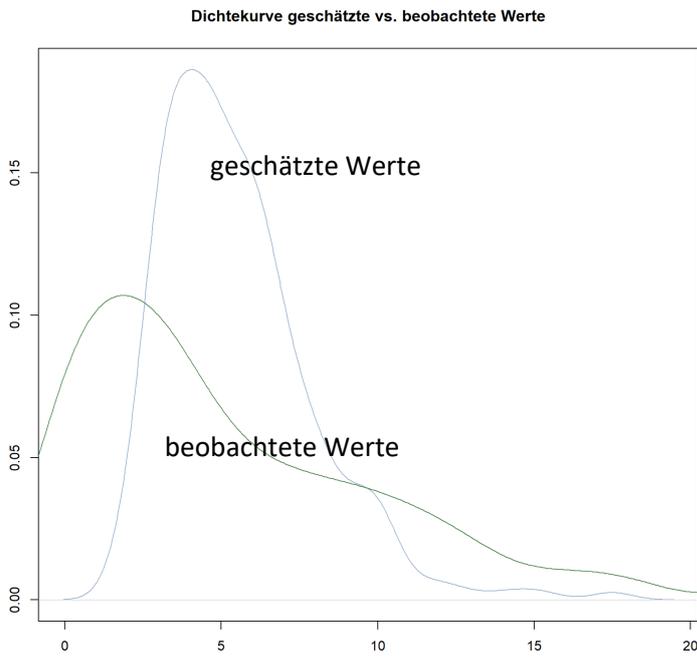


Abbildung 13: Modellgüte Analyse der Hausarztkontakte: geschätzte vs. beobachtete Werte

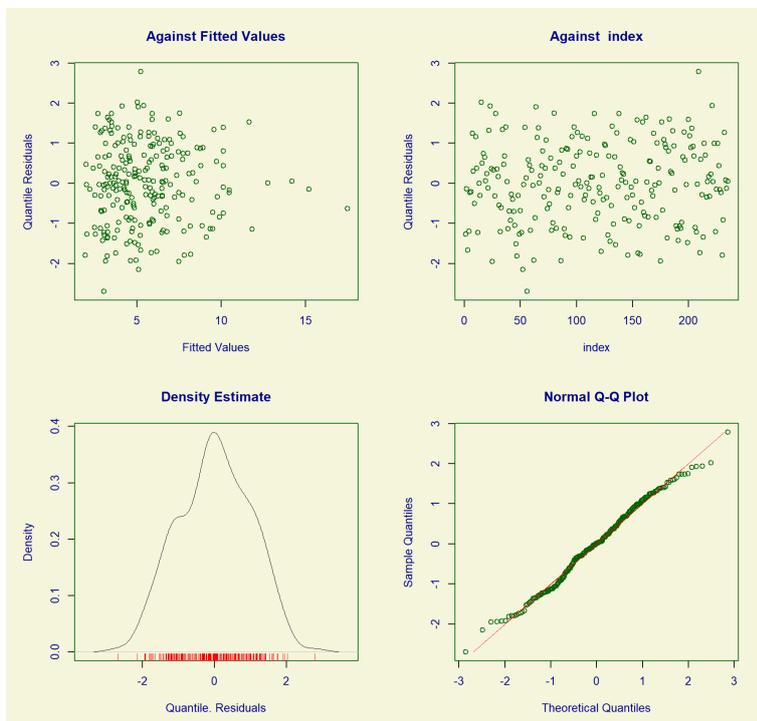


Abbildung 14: Modellgüte Analyse der Hausarztkontakte: Residuenplots, QQ-Plot