

Ergebnisbericht

**Randomisierte, kontrollierte Feldstudie zur Wirksamkeit von
BEYOND MATTER bei subjektiver Beschwerdelast**

RECON – Research and Consulting Dr. Rainer Schneider
Hindenburgstr. 15
79331 Teningen

Datum der Berichterstellung: 09.12.2021



Inhalt

1. Forschungsauftrag	01
2. Methode	01
2.1 Stichprobe	01
2.2 Studiendesign	02
2.2.1 Zielparameter	02
2.2.2 Moderatorvariable	03
2.2.3 Intervention/Studienablauf	04
2.3 Analyse/Statistik	04
3. Ergebnisse	05
3.1 EMF-Exposition	05
3.2 Beschwerdesymptomatik (GBB-24)	06
3.2.1 Individueller Beschwerdeindex	06
3.2.2 Erschöpfung	06
3.2.3 Magen-/Darmbeschwerden	06
3.2.4 Gliederschmerzen	06
3.2.5 Herzbeschwerden	07
3.3 Beschwerdesymptomatik (BOSS II)	07
3.3.1 Körperliche Beschwerden	07
3.3.2 Kognitive Beschwerden	07
3.3.3 Emotionale Beschwerden	07
3.4 Weiterführende Analysen	09
4. Zusammenfassung und Diskussion	11
5. Literatur	12

1. Forschungsauftrag

Die in Auftrag gegebene Studie testete die Wirksamkeit dreier BEYOND MATTER-Produkten (Handy-Chip, Einlegesohlen und Anhänger). Im Vordergrund standen die Ermittlung der Wirksamkeit bei selbstberichteten gesundheitlichen Beschwerden sowie die Separierung spezifischer und unspezifischer Effekte.

2. Methode

2.1 Stichprobe

Insgesamt wurden N = 75 Personen gescreent, um die geplante Stichprobengröße von N = 60 Probanden zu erreichen. Diese waren im Schnitt 43,32 Jahre alt (22-64 Jahre). Als Hauptbeschwerden nannten 37 Personen (61,7%) Erschöpfungssymptome, 14 (23,3%) gaben Gliederschmerzen an und neun (15%) klagten über Magen-/Darmprobleme. Für die Teilnahme galten folgende Einschlusskriterien:

1. Volljährigkeit.
2. Keine Erkrankungen, die eine ärztliche Begleitung erfordern.
3. Keine Drogen- oder Medikamentenabhängigkeit.
4. Mindestens eine der folgenden Beeinträchtigungen:
 - a) Erschöpfung,
 - b) Magenbeschwerden,
 - c) Gliederschmerzen und
 - d) Herzbeschwerden seit mindestens vier Wochen.

Die Teilnahme war freiwillig und setzte die Aufklärung über Art und Zweck der Studie voraus, wie sie im Rahmen einer Doppelverblindung möglich ist. Voraussetzung für die Teilnahme war die schriftliche Einwilligung. Die Verblindung erfolgte nach dem Zufallsprinzip; die Hälfte der Probanden erhielt Placeboprodukte (Einlegesohlen mit Aktivkohle, Anhänger und Handy-Chip aus Perlmutter). Gemäß den einschlägigen Befunden der Placeboforschung wurde die Erwartung an die Intervention maximiert, d.h. die Probanden nahmen an, die Placebos seien wirksame Hilfen zur Linderung von Beschwerden. Die Studienarme wurden hinsichtlich des Geschlechts mittels Blockrandomisierung gematched, so dass eine gleiche Geschlechterbesetzung in den Studienarmen resultierte. Die Probanden in der Placebogruppe wurden nach Beendigung der Studie durch die Studienleitung vollständig aufgeklärt.

Die Studie orientiert sich an den CONSORT-Richtlinien und der Deklaration von Helsinki der World Medical Association (letzte Revision 2013). Für die Durchführung wurde keine

Ethikkommission angerufen, da die BEYOND MATTER-Produkte keine Medizinprodukte sind und deswegen keine externe Ethikkommission (z.B. die Landesärztekammer Baden-Württemberg) angerufen werden kann. Aus diesem Grund wurde die Studie auch nicht registriert (z.B. beim Deutschen Register Klinischer Studien [DRKS]).

2.2 Studiendesign

Bei der Untersuchung handelte es sich um eine prospektive, zweiarmige, randomisierte, kontrollierte, messwiederholte Feldstudie. Sie sollte die Frage beantworten, ob die tägliche Anwendung von BEYOND MATTER-Produkten über einen Zeitraum von sechs Wochen zu einer Reduzierung der subjektiven Beschwerdelast führt. Beiden Studienarmen (Verum, Placebo) ging eine Baseline- Erhebung voraus, um den Beschwerdegrad vor der Intervention zu erfassen. Die Datenerhebung fand in Eigenregie der Probanden in deren Lebensumfeld statt.

2.2.1 Zielparameter

Beschwerdesymptomatik

Der Gießener Beschwerdebogen (GGB-24) ist ein Verfahren zur Erfassung der psychosomatischen Bedingtheit körperlicher Beschwerden [1]. Der GGB kann verwendet werden, um zwischen organmedizinisch begründbarer, objektivierbarer Symptomatik und subjektiven Beschwerden zu unterscheiden, da Diskrepanzen zwischen diesen Bereichen wichtige diagnostische Hinweise geben können. Der GGB-24 erfasst körperliche Beschwerden in den Bereichen:

- a) Erschöpfung (z.B. Schwächegefühl, übermäßiges Schlafbedürfnis, rasche Erschöpfbarkeit),
- b) Magen-/Darmbeschwerden (z.B. Völlegefühl im Bauch, Übelkeit, Aufstoßen, Magenschmerzen),
- c) Gliederschmerzen (z.B. Rücken-, Schulter- oder Kopfschmerzen, Schwere in den Beinen),
- d) Herzbeschwerden (z.B. Herzjagen, Schwindel, Schmerzen in der Brust, Atemnot).

Das Antwortformat der 24 Beschwerde-Items ist fünfstufig (Likert-Skala) mit den Ankern „nicht“ (0), „kaum“ (1), „einigermaßen“ (2), „erheblich“ (3) und „stark“ (4).

Burnout-Symptomatik

Die Burnout-Screening-Skalen (BOSS) besteht aus drei Selbstbeurteilungsverfahren, die bei Personen mit Verdacht auf Vorliegen einer Burnout-Problematik eingesetzt werden [2]. Für die Studie wurde der Fragebogen BOSS II eingesetzt, der drei Dimensionen von Beschwerden erfasst, wie sie typischerweise im Rahmen eines Burnout-Syndroms auftreten:

- a) Körperliche Beschwerden (z.B. Muskelverspannung, Herzbeschwerden, Schlafstörungen),
- b) Kognitive (geistige) Beschwerden (z.B. Konzentration, Fehleranfälligkeit, Entschlusskraft)
- c) Seelische Beschwerden (z.B. Zukunftsängste, sozialer Rückzug, Antriebslosigkeit).

Der Fragebogen bezieht sich auf einen Beurteilungszeitraum von sieben Tagen. Das Format der drei Skalen à 10 Items ist sechsstufig mit den Ankern „trifft nicht zu“ (0), „trifft kaum zu“ (1), „trifft etwas zu“ (2), „trifft teilweise zu“ (3), „trifft überwiegend zu“ (4), „trifft zu“ (5) und „trifft stark zu“ (6).

Individualsymptomatik

Um dem Umstand Rechnung zu tragen, dass nur wenige Probanden auf allen Skalen hoch scorten, sondern vielmehr an einzelnen Symptomkomplexen litten, wurde ein individueller Beschwerdeindex aus dem GBB-24 gebildet (d.h. der jeweils höchste Wert, auf dem ein Proband scorte; siehe Punkt 2.1). Dieser galt als primärer Zielparameter; die anderen Skalen wurden als sekundäre Parameter definiert.

2.2.2 Moderatorvariable

EMF-Exposition

Um die Belastung der Probanden durch elektromagnetische Feldstärken einzuschätzen, kam eine selbstkonstruierte EMF-Checklist zum Einsatz. Die Wirksamkeit der BEYOND MATTER-Produkte konnte damit ggfls. daraufhin untersucht werden, ob die Beschwerdelast auch durch hohe EMF- Exposition moderiert war. Die Checklist wurde am Ende der Studie durch die Probanden ausgefüllt und enthielt folgende (offene) Items:

1. Die Zeit in Minuten, die in den vergangenen sechs Wochen im Schnitt täglich aktiv am Handy verbracht wurde (Telefonate, Internetsurfen etc.).
2. Die Zeit in Minuten, die das Handy im eingeschalteten Modus in den vergangenen sechs Wochen im Schnitt täglich am Körper getragen wurde.
3. Die Anzahl der Nächte, in denen sich das eingeschaltete Handy nachts in der Nähe des Bettes des Probanden befand (Antwortkategorie: nie, immer oder Anzahl).
4. Die Zeit in Minuten, die der Proband sich in den vergangenen sechs Wochen im Schnitt täglich in unmittelbarer Nähe zu einer EMF-Quelle aufhielt (z.B. WLAN-Router, DECT- Telefon, Laptop).
5. Die Zeit in Minuten, die der Proband in den vergangenen sechs Wochen im Schnitt in unmittelbarer Nähe zu einer EMF-Quelle schlief.

2.2.3 Intervention/Studienablauf

Vor Beginn der Erhebungen wurden die wahlfähigen Personen, die zunächst telefonisch gemäß den Einschlusskriterien selektiert wurden, zu einem Besuch bei RECON eingeladen, in welchem die Versuchsleiterin gemäß der Cover Story in die Rationale der Studie einführte. Produktname und Produktbeschaffenheit waren der Versuchsleiterin nicht bekannt. Im Erstbesuch wurde die Einverständniserklärung eingeholt und das Testmaterial sowie die BEYOND MATTER- bzw. Placeboprodukte ausgehändigt. Die Zuordnung zu den Studienarmen erfolgte durch verschlossene Briefumschläge, die die Studienleitung gemäß der Randomisierung vorbereitete und die die Versuchsleiterin im Beisein des Probanden öffnete. Im Zweitbesuch nach sechs Wochen nahm die Versuchsleiterin die ausgefüllten Fragebogen entgegen und entschädigte die Probanden für die Teilnahme. Die Intervention bestand in der täglichen Anwendung der Produkte nach folgendem Muster: der Anhänger wurde mindestens zwölf Stunden am Tag um den Hals gehängt und die Einlagen in Straßen- und Hausschuhen für die gleiche Dauer getragen. Der Handy-Chip wurde auf das Smartphone geklebt, welches – falls möglich – den ganzen Tag in Körpernähe getragen wurde.

2.3 Analyse, Statistik

Es wurden ausschließlich Effektmaße (Cohens d) [3] und 95%-Vertrauensintervalle [4] berechnet, um Fehlinterpretationen und Scheinevidenzen durch signifikanzbasierte Statistik zu vermeiden [5].¹ Der Fokus der Effektstärkenermittlung liegt auf Einzelvergleichen von je zwei Gruppen bzw. Mittelwerten. In der Studie wurden für den Vergleich der beiden Interventionen (Placebo versus BEYOND MATTER) die jeweiligen Baseline-Werte als Differenz verrechnet (Baseline 1 minus Intervention 1 versus Baseline 2 minus Intervention

¹ Hier seien einige Hauptargumente gegen Signifikanztestung (auch als „dirty little trick“ bezeichnet) genannt, die sich seit der vor kurzem für viel Aufhebens sorgenden Replikationskrise in der psychologischen Forschung als unstrittig etabliert haben:

1. Signifikante Ergebnisse sagen nichts über die Relevanz der zugrunde liegenden Effekte aus. Außerdem sind sie nicht über verschiedene Studien hinweg vergleichbar. Viele sogenannte „hochsignifikante“ Ergebnisse sind nur von geringer klinischer Bedeutung, wenn man sie in Bezug auf das tatsächliche Ausmaß der Effekte ausdrückt.
2. Jede Nullhypothese kann bei ausreichend großen oder homogenen Stichproben leicht verworfen werden. Außerdem können statistische Tests unabhängig von der Aussagekraft des untersuchten Zusammenhangs beliebig signifikant gemacht werden.
3. Viele statistische Annahmen im Zusammenhang mit NHST (Nullhypothesen-Signifikanztestung) werden trotz ritueller Verwendung falsch verstanden. Ein solcher fundamentaler Fehler ist die Kombination von Fishers Evidenzstatistik (p -Wert) und Neyman-Pearsons Fehlerschätzung (α). Trotz häufiger Missverständnisse sind sie in keiner sinnvollen Weise verknüpft, wenn angibt, dass $p < \alpha$.
4. Dies führt zu der Unlogik, die NHST begleitet: Zu glauben, dass ein unter der Annahme der Nullhypothese „unwahrscheinlicher“ Datensatz den Schluss zuließe, dass er wahrscheinlich falsch sei. Mit anderen Worten, wenn die Wahrscheinlichkeit von α kleiner als das angewendete Kriterium ist, wird angenommen, dass der Zufall das Ergebnis erzeugt habe, oder formal ausgedrückt: $P(\text{Daten} | H_0) \rightarrow P(H_0 | \text{Daten})$. Beide bedingten Wahrscheinlichkeiten sind jedoch unabhängig voneinander und können nicht vertauscht werden. Man ersetze D und H_0 mit „Regen“ und „Wolken“, um das zu veranschaulichen: Die Wahrscheinlichkeit von Regen bei Vorhandensein von Wolken ist eine andere als die Wahrscheinlichkeit von Wolken bei Vorhandensein von Regen.
5. Bis heute wird dieser Zusammenhang in statistischen Lehrbüchern falsch dargestellt und im statistischen Methodenunterricht unkritisch gelehrt. Darüber hinaus ist es ernüchternd, dass viele Statistiker und Wissenschaftler nicht in der Lage sind, ein sogenanntes „signifikantes“ Ergebnis richtig zu definieren.

2)². Für die Interpretation der Ergebnisse gelten folgende Richtlinien: Ein kleiner Effekt liegt vor, wenn $d = 0,2$, ein mittlerer Effekt entspricht $d = 0,5$ und ein großer Effekt entspricht $d \geq 0,8$. Effekte von $d \geq 1$ werden i.d.R. als praktisch bzw. klinisch relevant erachtet. Allerdings müssen bei der Betrachtung der statistischen Effekte immer auch die tatsächlichen absoluten Veränderungen der Mittelwerte in Betracht gezogen werden, um zu bestimmen, wie bedeutsam die differentiellen Effekte sind.

3. Ergebnisse

3.1 EMF-Exposition

Die durchschnittliche EMF-Exposition der Stichprobe war hoch, obgleich die Streubreite ebenfalls groß war. Die Teilnehmer verbrachten täglich ca. 119 Minuten aktiv am Handy bzw. Smartphone und trugen dieses täglich rund 298 Minuten am Körper. Dreißig Personen hatten es jede Nacht im eingeschalteten Modus in unmittelbarer Nähe zum Bett, 28 nie (eine Person tat das zehnmal, eine andere fünfzehnmal). Die Zeit, die tagsüber im Schnitt in unmittelbarer Nähe zu einer EMF-Quelle (z.B. WLAN-Router) verbracht wurde, betrug ca. 356 Minuten. Nachts war dies im Schnitt ca. 188 Minuten der Fall.

Für den Vergleich der beiden Studienarme ergaben sich hinsichtlich der EMF-Exposition keine statistischen Unterschiede. Die Verumgruppe nutzte das Handy aktiv durchschnittlich ca. 118 Minuten täglich, die Verumgruppe 120 Minuten ($d = 0,01$; KI: $-0,7 < d < 0,7$)³. Zwar trugen die Personen der Verumgruppe das Handy nominal länger am Körper als jene der Placebogruppe (341 versus 256 Minuten), dieser Unterschied war aber statistisch ein Nulleffekt ($d = 0,3$; KI: $-0,4 < d < 1,1$). Ein ähnliches Bild zeigte sich für die Zeit, die täglich durchschnittlich in unmittelbarer Nähe zu einer anderen EMF-Quelle als das Handy verbracht wurde. In der Verumgruppe waren dies ca. 375 Minuten, in der Placebogruppe 338 Minuten ($d = 0,2$; KI: $-0,5 < d < 0,8$). Der Unterschied für die Dauer, die nachts in unmittelbarer Nähe zu einer EMF-Quelle verbracht wurde, war mit 190 Minuten (Verum) und 186 (Placebo) ebenfalls statistisch unbedeutend ($d = 0,02$; KI: $-0,7 < d < 0,7$). Hinsichtlich der Angewohnheit, das Handy nachts im eingeschalteten Modus in unmittelbarer Nähe zum Bett zu platzieren, zeigte sich, dass dies fünfzehn Personen in der Verumgruppe taten und 18 in der Placebogruppe. Dieser Unterschied war mit $\chi^2 = 1,2$ ($df = 58$) statistisch jedoch klein ($d = 0,3$).

Insgesamt ergab sich damit für kein Kriterium ein Unterschied in der Exposition mit EMF-Feldern. Damit war auszuschließen, dass diese den Verlauf der Intervention beeinflussten oder modulierten. Aus diesem Grund wurde die Variable EMF-Exposition bei den weiteren Analysen nicht berücksichtigt.

² Diese Vorgehensweise ist identisch mit der formal vorliegenden 2x2 faktoriellen Varianzanalyse.

³ Ein statistisch bedeutsamer Effekt kann nicht zugleich positive und negative Werte annehmen. Liegt er innerhalb dieser Intervallgrenzen, handelt es sich um einen Nulleffekt.

3.2 Beschwerdesymptomatik (GBB-24)

3.2.1 Individueller Beschwerdeindex

Der primäre Zielparameter Individueller Beschwerdeindex, der vor der Intervention in beiden Gruppen nahezu identisch war (Mittelwert_{Verum}: 2,4 versus Mittelwert_{Placebo}: 2,2; siehe Tabelle 1 und Abbildung 1), sank in der Verumgruppe um 1,1 Skalenwerte, in der Placebogruppe hingegen um 0,25⁴. Dieser Unterschied war mit $d = 1$ groß (KI: $0,4 < d < 1,5$). Die Anwendung der BEYOND MATTER-Produkte führte somit zu einer Reduktion der Beschwerden von „einigermaßen bis erheblich“ auf „kaum“. Prozentual war die Symptomlinderung ca. 46%. Nach Anwendung der Placeboprodukte war sie ca. 11%. Der spezifische Effekt von BEYOND MATTER betrug damit 35%.

3.2.2 Erschöpfung

Für den Beschwerdekomples Erschöpfung zeigte der Vergleich der beiden Gruppen einen großen differentiellen Effekt von $d = 1,1$ (KI: $0,6 < d < 1,7$). Während die mittlere Erschöpfung der Verumgruppe von 2,3 Skalenpunkten zu Beginn der Studie auf 1,2 Skalenpunkte sank, blieb sie in der Placebogruppe mit 2,3 und 1,9 nahezu identisch.

3.2.3 Magen-/Darmbeschwerden

Die durchschnittlichen Magen-/Darmbeschwerden, die in beiden Gruppen vor der Intervention ebenfalls gleich waren (MW_{Verum} = 1, MW_{Placebo} = 0,9), wurden in der Verumgruppe am Ende der Studie auf die Hälfte reduziert, während sie in der Placebogruppe unverändert blieben (MW_{Verum}: 0,5, MW_{Placebo}: 0,9). Dieser Unterschied entsprach einem mittleren Effekt ($d = 0,7$; KI: $0,2 < d < 1,2$)⁵.

3.2.4 Gliederschmerzen

Ein ähnliches Bild zeigte sich für den Vergleich der Gliederschmerzen. Diese wurden in der Verumgruppe um 50% reduziert (Prä: MW_{Verum} = 2; Post MW_{Verum}: 1), in der Placebogruppe um ca. 13% (Prä: MW_{Placebo} = 1,5; Post: MW_{Placebo} = 1,3). Dieser differentielle Effekt war mit $d = 0,9$ groß (KI: $0,4 < d < 1,4$).

⁴ Die Skalensummenwerte wurden auf die Spannbreite der Antwortskala transformiert.

⁵ Die im Vergleich zum individuellen Beschwerdeindex geringeren Skalenwerte zeigen an, dass der erhobene Belastungsgrad in den Subskalen des GBB-24 vor allem in den Skalen Magen-/Darmbeschwerden und Gliederschmerzen deutlich geringer war, da sie in der Stichprobe prozentual weniger prävalent waren.

3.2.5 Herzbeschwerden

Die mittleren Herzbeschwerden sanken in der Verumgruppe von 0,9 auf 0,5, blieben in der Placebogruppe hingegen mit 0,8 und 0,7 nahezu gleich. Der Unterschied zwischen den beiden Studienarmen war groß ($d = 0,8$; KI: $0,2 < d < 1,3$)⁶.

3.3 Beschwerdesymptomatik (BOSS II)

3.3.1 Körperliche Beschwerden

Im Mittel war die Beschwerdelast der Dimension Körperliche Beschwerden in beiden Studienarmen nur mäßig ausgeprägt. Sie sank in der Verumgruppe jedoch deutlich stärker als in der Placebogruppe (Prä: $MW_{\text{Verum}} = 1,6$; Post: $MW_{\text{Verum}} = 1$ versus Prä: $MW_{\text{Placebo}} = 1,4$; Post: $MW_{\text{Placebo}} = 1,2$). Der Unterschied in den Reduktionen entsprach einem großen Effekt von $d = 1$ (KI: $0,4 < d < 1,5$).

3.3.2 Kognitive Beschwerden

Ein ähnliches Bild zeigte sich für die Dimension Kognitive Beschwerden. Während in der Verumgruppe die mittlere Ausprägung von 2 auf 1 am Ende sank, reduzierte sie sich in der Placebogruppe lediglich von 1,9 auf 1,6. Der Unterschied in der Symptomverbesserung war ebenfalls groß ($d = 1$; KI: $0,4 < d < 1,5$).

3.3.3 Emotionale Beschwerden

Bei der Dimension Emotionale Beschwerden fand sich ein vergleichbares Bild. Die Reduktion der Beschwerden entsprach nach Anwendung des BEYOND MATTER-Produkte einem Skalenwert (Prä: $MW_{\text{Verum}} = 1,8$; Post: $MW_{\text{Verum}} = 0,8$). Nach Anwendung der Perlmutter-Produkte sank sie hingegen von 2 auf 1,8. Der differentielle Effekt war ebenfalls groß ($d = 0,9$; KI: $0,4 < d < 1,5$).

⁶ Dieser „große“ Effekt verdeutlicht, dass tatsächliche Unterschiede vor dem Hintergrund ihrer Relevanz betrachtet werden müssen. Beide Gruppenmittelwerte waren so gering, dass von einer Belastung in dieser Dimension nicht gesprochen werden konnte. Der differentielle Effekt hatte damit keine praktisch relevante Bedeutung.

Tabelle 1: Mittelwerte und Standardabweichungen der Beschwerdeskalen

	Placebo		Blackstone Protection	
	Baseline	Intervention	Baseline	Intervention
Individueller Beschwerdeindex^a	2,2 (0,5)[†]	2 (0,9)	2,4 (0,9)	1,3 (0,8)
Erschöpfung^a	2 (0,8)	1,9 (0,9)	2,3 (0,9)	1,2 (0,8)
Magen-/Darmbeschwerden^a	1 (0,8)	0,9 (0,8)	1 (0,9)	0,5 (0,7)
Gliederschmerzen^a	1,5 (0,7)	1,3 (0,7)	2 (0,9)	1 (0,9)
Herzbeschwerden^a	0,8 (0,7)	0,7 (0,6)	0,9 (0,6)	0,5 (0,7)
Körperliche Beschwerden^b	1,4 (0,5)	1,2 (0,6)	1,6 (0,7)	1 (0,8)
Kognitive Beschwerden^b	1,9 (0,9)	1,6 (1)	2 (1,1)	1 (0,8)
Emotionale Beschwerden^b	2 (1)	1,8 (1)	1,8 (1,2)	0,8 (0,9)

[†] gerundete Werte; ^a Skala: 0-4 (GGB-24); ^b Skala: 0-5 (BOSS II)

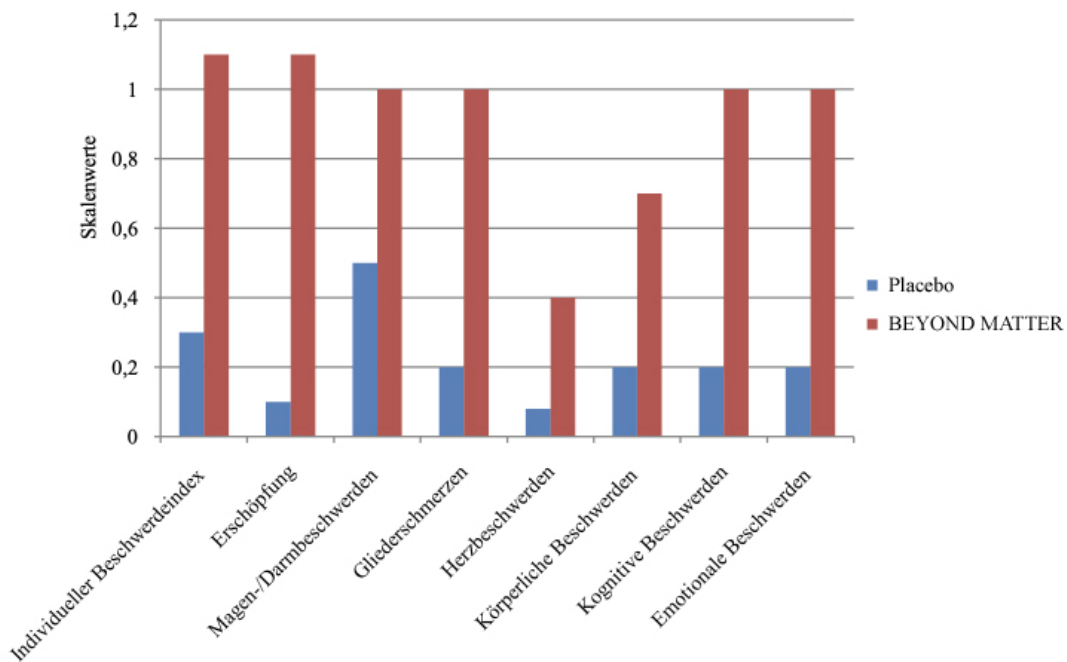


Abb. 1: Verbesserungsraten der Beschwerdedimensionen; Skalenwerte: 0-4 bzw. 0-5 (siehe Tabelle 1)

3.4 Weiterführende Analysen

Da Wirksamkeitsprüfungen i.d.R. anhand von Maßen der zentralen Tendenz (z.B. Mittelwerte) durchgeführt werden, bedeuten große Effekte selten, dass alle Probanden in gleichem Maße von einer bestimmten Maßnahme profitieren. Es ist z.B. bekannt, dass in klinischen Prüfungen in der Verumgruppe sog. Non-Responder-Raten von bis zu 40% auftreten. Auch bei wirkmächtigen Interventionen ist zu beobachten, dass Nulleffekte oder gegenläufige Effekte auftreten. Es war daher anzunehmen, dass manche Teilnehmer auf die BEYOND MATTER-Produkte nicht reagierten. Um diesen Sachverhalt zu testen, wurden bivariate Korrelationen getrennt für beide Studiengruppen gerechnet, bei welchen die Beschwerdelast zu Beginn der Studie mit der Symptomverbesserung (individueller Beschwerdeindex) in Beziehung gesetzt wurde.

Die Analysen zeigten, dass in der Verumgruppe ein starker Zusammenhang zwischen beiden Variablen bestand: Die Symptomlinderung war umso größer, je größer die individuellen Beschwerden zu Beginn der Studie waren ($r = 0,66$; $d = 1,8$). Demgegenüber bestand ein solcher Zusammenhang in der Placebogruppe nicht ($r = 0,15$; $d = 0,3$). Der Unterschied zwischen den beiden Korrelationen war hoch und entsprach einem großen Effekt (Cohens $q = 0,6$)⁷. Dieser Sachverhalt verdeutlicht Abbildung 2.

Aus Abbildung 2 ist ersichtlich, dass drei Personen in der Verumgruppe eine leichte Verschlechterung der Symptome erlebten. Die Besserungsrate der restlichen 27 Personen lag zwischen 8-87%. Die mittlere prozentuale Besserungsrate betrug 43%. Im Unterschied dazu berichteten zehn Personen in der Placebogruppe eine Verschlechterung der Symptome und zehn keine Veränderung. Die Spannbreite der Symptomveränderungen lag zwischen -36% und 100%. Die mittlere Besserungsrate war 11%.

⁷ Für Cohens q gilt die Einteilung: $< 0,1$ = kein Effekt; $0,1 - 0,3$ = kleiner Effekt; $0,3 - 0,5$ = mittlerer Effekt; $> 0,5$ = großer Effekt.

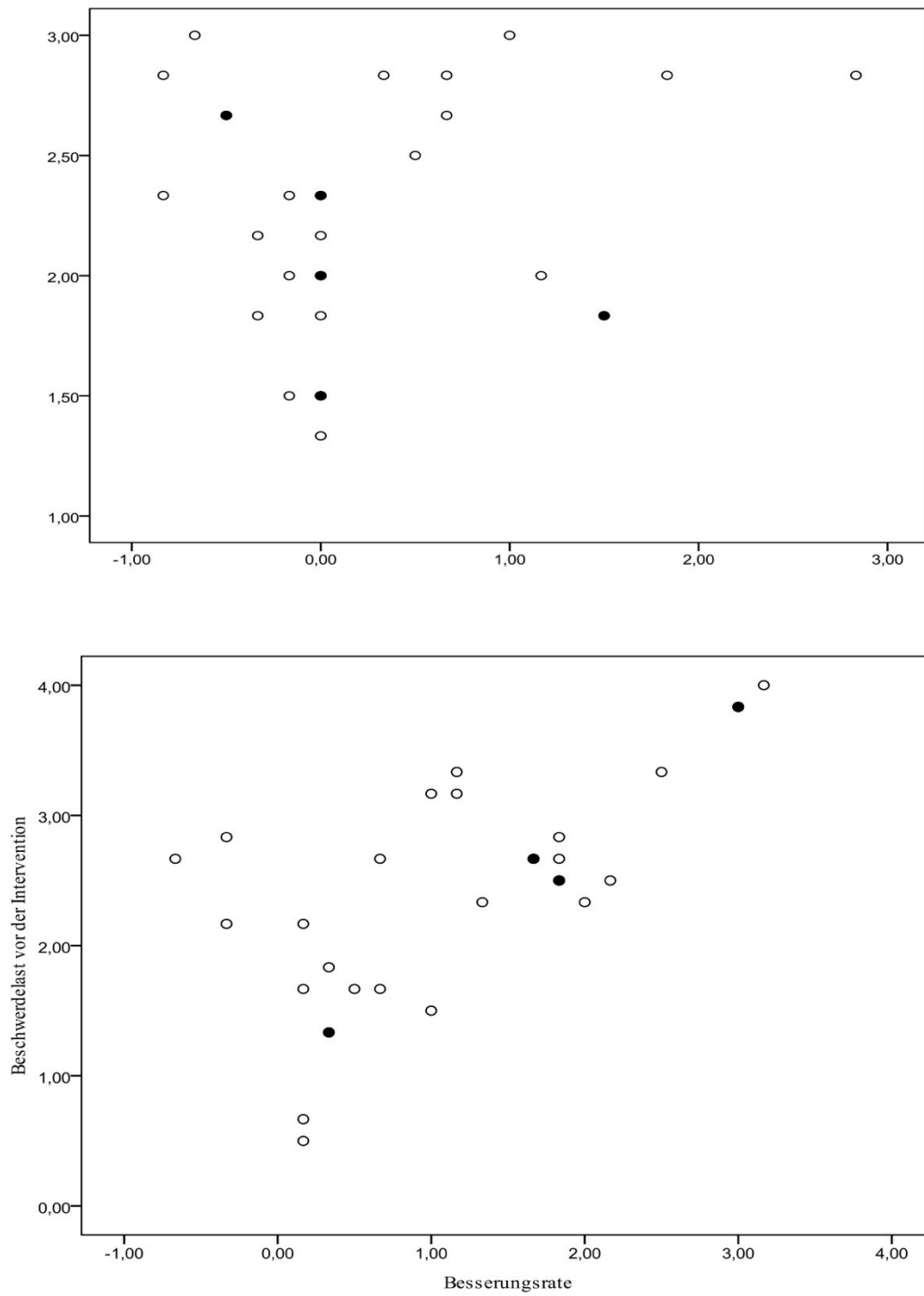


Abb. 2: Zusammenhang zwischen Beschwerdelast und Symptomverbesserung; oben: Placebogruppe; schwarze Kreise: Doppel- bzw. Dreifachüberlappungen

4. Zusammenfassung und Diskussion

Die vorliegende Feldstudie hatte zum Ziel, die Wirksamkeit der BEYOND MATTER-Produkte (Handy-Chip, Einlegesohlen, Anhänger) zu untersuchen. Die Anwendungszeit der Produkte war mit sechs Wochen vergleichsweise lange und erlaubte eine verlässliche phänomenologische Erfassung möglicher Beschwerdeveränderungen. Die Untersuchung war doppelblind, placebokontrolliert und randomisiert und entsprach damit formal einer klinischen Studie der Phase III. Im Unterschied zur klinischen Praxis wurde der Placeboeffekt allerdings maximiert, um spezifische Effekte von unspezifischen Effekten besser separieren zu können⁸.

In allen Vergleichen zeigte sich eine große Wirksamkeit von BEYOND MATTER. Die Effekte waren mit $d = 1$ hoch und entsprachen der Linderung von einer Standardabweichung. Im Mittel verbesserte sich die individuelle Beschwerdelast um 46% von „einigermaßen bis erheblich“ auf „kaum“, wobei die spezifische Wirksamkeit bei 35% lag. Im Vergleich zur Behandlung mit Placebo lag die Wirküberlegung bei über 300%. Statistisch betrachtet bedeutet ein Effekt von $d = 1$, dass ca. 84% der Teilnehmer der BEYOND MATTER-Gruppe eine Besserungsrate hatten, die über dem Mittelwert der Placebogruppe lag. Die Rate der sog. Non-Responder war mit 10% am unteren Ende des Bereichs, der für Interventionsstudien dieser Art zu erwarten ist.

Das Ausmaß der EMF-Exposition schien für die Verursachung der Beschwerdesymptomatik keine Rolle zu spielen, wobei einschränkend gesagt werden muss, dass diese Variable nicht experimentell-kausal, sondern nur deskriptiv bzw. quasiexperimentell erhoben wurde. Die insgesamt vergleichsweise hohe mittlere EMF-Exposition in beiden Studienarmen war aber kein Faktor, der die differentiellen Effekte unsystematisch verzerrte.

Die Analysen zeigten ferner, dass die Anwendung von BEYOND MATTER mit einem größeren Nutzen assoziiert war, je höher die Beschwerdelast der Anwender war. Dieser Nutzen zeigt sich vor allem beim Erschöpfungssyndrom, wobei aufgrund der Natur der Studie nicht gesagt werden kann, welche Rolle die Ätiologie der Symptome dabei spielte. Die Tatsache, dass Placeboeffekte zwar zum Gesamteffekt beitrugen, nominal aber von untergeordneter Bedeutung waren, war insofern wichtig, als die subjektive Befindlichkeit auch von Erwartungseffekten moduliert sein kann, und zwar umso mehr, je diffuser die Beschwerdesymptome sind [6].

⁸ Streng betrachtet wurde auf diese Weise die *Placeboreaktion* und nicht der *Placeboeffekt* ermittelt. Letzterer wäre z.B. nur durch Einschluss einer Nullkontrolle genau bestimmbar gewesen. Da im Vordergrund jedoch die Ermittlung der Wirküberlegenheit von BEYOND MATTER stand, spielte dieser Sachverhalt eine untergeordnete Rolle.

5. Literatur

[1] Brähler E, Hinz A, Scheer J W. GBB-24. Der Gießener Beschwerdebogen. Bern: Huber, 2006.

[2] Geuenich K, Hagemann W. Burnout-Screening-Skalen. Göttingen: Hogrefe, 2014.

[3] Cohen J. Statistical power analysis for the behavioral sciences. Hillsdale: Laurence Erlbaum Associates 2008.

[4] Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. Introduction to meta-analysis. Chichester: John Wiley & Sons 2009.

[5] Greenland S, Senn SJ, Rothmann KJ, Carlin JB, Poole C, Goodman SN, Altman D.G. Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations. *European Journal of Epidemiology* 2016, 31: 337-350.

[6] Schneider R, Kuhl J. Placebo forte: Ways to maximize unspecific treatment effects. *Medical Hypotheses* 2012, 78: 744-751.