Informationsmedizin auf dem Prüfstand

Ein Kristall, der Energie schenkt. Ein Diode, die gesund macht. Ein Chip, der Schmerzen reduziert. Im Bereich der Alternativmedizin scheint nichts unmöglich, wenn es darum geht, schulmedizinische Pfade zu verlassen und unsere Vorstellung von Gesundheit und Gesundungsprozessen neu zu definieren. Doch was ist dran an informationsmedizinischen Technologien? **Halten sie einer wissenschaftlichen Prüfung stand** oder sind sie nicht doch bloß das, was allopathische Vertreter schon lange behaupten: Humbug oder Placeboeffekt. **Dr. phil. Dipl. Psych. Rainer Schneider** geht dieser Frage auf den Grund.

Innerhalb der Komplementärmedizin erfreuen sich sogenannte Biofeld-, Energie- oder Informationstherapien wachsender Beliebtheit. Es ist nicht ganz einfach, diese Therapieformen einheitlich zu klassifizieren. Der gemeinsame Nenner ist die Annahme von in Wechselwirkung stehenden (elektromagnetischen) Feldern, von denen angenommen wird, sie bestimmten alles biologische Leben. Informationstherapien postulieren nun, dass biologische Felder von außen verändert, wiederhergestellt oder vitalisiert werden können, wenn sie gestört sind.

In Resonanz mit dem Gewebe

Die Idee ist eigentlich schon sehr alt. Dennoch beginnt die wissenschaftliche Erforschung der Informationsmedizin erst Anfang des vergangenen Jahrhunderts. In der Literatur wird sie bevorzugt mit der Pionierarbeit des amerikanischen Arztes Albert Abrams in Verbindung gebracht. Abrams machte u.a. die Beobachtung, dass seine Patienten mit entnommenem Krebsgewebe „in Resonanz gingen". Das äußerte sich dadurch, dass vibrierende Krebsgewebemoleküle mit Muskelkontraktionen einher gingen. Abrams baute auf Grundlage seiner Beobachtungen diagnostisch-therapeutische Apparaturen, die ihm – trotz entsprechender Heilerfolge – wenig Anerkennung seiner Kollegen bescherten. In Kritikerkreisen ging er als „Dekan der Quacksalber des 20. Jahrhunderts" in die Geschichte ein.

Interessanterweise gibt es eine Reihe von Wissenschaftlern, die just zu dieser Zeit ähnlich Phänomene untersuchten. Einer davon war der russische Biologe Alexander G. Gurwitsch. Er beobachtete, dass lebende Systeme ultraschwache Photonen

Wirksamkeitsprüfung am Beispiel des Vita Chip E

emittieren; ein Phänomen, das er „mitogenetische Strahlung" nannte. Erst viele Jahre später griff der deutsche Biophysiker Fritz Albert Popp dieses Phänomen wieder auf und belegte es mit dem Begriff Biophotonik. Popps Postulat, dass ein „zelluläres Licht" Kommunikation und Regulierung in und zwischen den Zellformationen regle, löste zunächst wiederum Spott der etablierten Wissenschaft aus. Heute wissen wir, dass viele, wenn auch nicht alle, der Behauptungen Popps wissenschaftlich bestätigt sind. Nicht nur das: Erst durch Popps Erweiterung der Biologie durch Biophotonik auf Grundlage der Quantenphysik ist unser Verständnis der Zellkommunikation ein wesentlich umfassenderes.

Ein weiterer, posthum prominent gewordener Kollege Abrams, ist der kroatische Ingenieur und Physiker Nikola Tesla, der im Jahre 1900 ein Patent auf Grundlage der „Transmission elektrischer Ener-

gie über natürliche Medien" erhielt. Tesla gelang es zu Lebzeiten nicht, eine schlüssige Theorie als Erklärung seiner Beobachtungen zu finden. Trotzdem war er überzeugt, dass hochfrequente elektrische Ströme eine nicht unwesentliche medizinische Bedeutung haben. Erst etwa einhundert Jahre später wurden „Tesla-Wellen" durch die erweiterte Maxwell'sche Feldtheorie in eine stimmige physikalische Theorie eingebettet. Inzwischen ist man experimentell in der Lage, sogenannte Skalarwellen, eine spezifische Klasse von Longitudinalwellen, als Signalträger zwischen Zellen und Geweben zu nutzen: Skalarwellen, die man z.B. mit der biologischen Information eines Fungizids (Clotrimazol) informiert, hemmen das Wachstum einer Hefekultur. [1]

Resistentes Grundsatzproblem

Trotz solcher Befunde ist sowohl die Anwendung, als auch der Nutzen informationsmedizinischer Technologien innerhalb der Schulmedizin im Allgemeinen kontrovers. Die Gegner akzeptieren weder die behauptete Wirkungsweise (d.h. Übertragung durch Informationssignale), noch die Wirkung selbst. Das gängige Argument ist der Mangel an schlüssigen Beweisen (z.B. durch Metaanalysen oder wissenschaftlich saubere Wirksamkeitsstudien). Diese Kritik ist nicht unberechtigt, denn viele der auf dem Markt (ehemals) erhältlichen Geräte und Apparaturen grenzen eher an Wunschdenken, manche sogar an Betrug. Es gibt durchaus positive Erfahrungsberichte hinsichtlich der Wirkung und Wirksamkeit; sie sind jedoch meist anekdotisch und/oder entstammen informellen oder persönlichen Testaten. Gerade die Kritik, dass der empirische Nachweis der Wirkung von informationsmedizinischen Interventionen in nicht-isolierten Zellsystemen (d.h. in komplexen Organismen wie Säugetieren oder Menschen), sowie anhand methodisch sauberen Humanstudien fehlten, ist weitestgehend zutreffend. Dies dürfte einer der Gründe sein, warum Kritiker die bestehenden empirischen Befunde überwiegend als Placeboeffekte abtun. Ein anderer Grund ist natürlich, dass die Akzeptanz gegenüber neuen Paradigmen in der Wissenschaft eher gering ist.

"Es verwundert aber kaum (oder doch?), dass selbst statistische Experten nicht in der Lage sind, eine korrekte Definition eines „signifikanten" Ergebnisses zu geben."

Ich möchte das Argument des Placeboeffekts aufgreifen, denn es ist ein wichtiges. Eines vorweg: Placeboeffekte sollten nicht nur bei der Evaluierung bislang noch nicht verstandener Therapieformen in Betrachtung gezogen werden. Sie spielen ganz allgemein eine Rolle, wann immer so genannte unspezifische Effekte am Werk sind. Die Liste der Krankheiten und Störungen, wo dies der Fall ist, ist lang. Prominente Beispiele sind Schmerz, Atemwegserkrankungen, Depression, sexuelle Dysfunktionen, Parkinson'sche Krankheit, kardiovaskuläre Erkrankungen, affektive Störungen oder gastrointestinale Beschwerden, um nur einige zu nennen. Erwartungen, Konditionierungen, Bedürftigkeit und Glauben sind Faktoren, die zusätzlich zu so genannten spezifischen Effekten (d.h. solche, die z.B. auf ein pharmakologisches Agens zurückgehen) wirken und bisweilen sogar in größerem Maße zu Verbesserungsraten beitragen können.[2] Dies zeigt, dass das Gewahrsein und die Bedeutung einer Intervention wichtige psychologische Faktoren seitens der Behandelten sind, denn sie lösen einen Effekt aus oder modulieren ihn.[3] Nicht spezifische Effekte umfassen aber auch nicht psychologische Faktoren (Placeboeffekte). Hierzu zählen z.B. der natürliche Krankheitsverlauf, zeitgleiche andere Interventionen, biologische Rhythmen, Messeffekte oder Regressionseffekte; Diese Effekte müssen bei jeder Behandlung berücksichtigt werden, was verdeutlicht, dass bei der Evaluierung therapeutischer Effekte sogenannte natürliche Verkaufskontrollen sehr wichtig sind (allerdings wird in klinischen Prüfungen so gut wie nie gegen eine natürliche Kontrolle getestet). Wenn nun der Ruf nach Wirksamkeitsprüfungen gerade bei unorthodoxen Therapieformen wie der Informationsmedizin laut wird, müssen ganz unterschiedliche Wirkfaktoren berücksichtigt werden, um zu einem aussagekräftigen Verdikt zu kommen.

Ein fundamentales Grundsatzproblem: Das schmutzige kleine Geheimnis der Beweisführung

Wie bereits angedeutet, genügt es bei der Behauptung, Informationsmedizin wirke, nicht, ausschließlich Erfahrungsberichte als Beweis zu zitieren. Sogenannte anekdotische Evidenz ist zwar wichtig, aber sie unterliegt zahlreichen Verzerrungs- und Störfaktoren. Sie haben als wissenschaftlicher Beweis die geringste Validität und erlauben keine Generalisierung. Innerhalb des wissenschaftlichen Kontexts sucht man nach allgemeingültigen Gesetzmäßigkeiten, zumindest dann, wenn es um Interventionen und Therapien geht, von denen möglichst viele Menschen profitieren sollen. Der zu erbringende Wirksamkeitsnachweis, anhand dessen man einen klinischen Nutzen einschätzt, muss anders aussehen.

In der experimentellen und klinischen Forschung erbringt man den Nachweis für gewöhnlich auf Basis von Signifikanztests. Diese beruhen auf der sogenannten statistischen Nullhypothese. Sie besagt, dass z.B. zwischen einer Placebogruppe und einer Verumgruppe kein Unterschied besteht. Wird nun ein Ergebnis „signifikant" oder gar „hoch signifikant", verwirft man die Nullhypothese und behauptet einen Effekt. Denn: Die Wahrscheinlichkeit des Ergebnisses ist kleiner als ein Kriteri-

um, auf das sich die Wissenschaftliche Gemeinschaft per Konvention geeinigt hat (i.d.R. 5%). Ist also die Wahrscheinlichkeit eines Studienergebnisses kleiner als 5%, schließt man, dass zwei Testgruppen nicht mehr derselben Grundgesamtheit entstammen, also verschieden sind. Wenn jetzt in der Verumgruppe eine insgesamt größere Besserungsrate zu beobachten ist als in der Placebogruppe, sagt man, das Verum sei signifikant besser als Placebo. In sich klingt das schlüssig und logisch. Und diese Art zu denken, war und ist der praktische Alltag von Generationen von Forschern. Es gilt das Motto: Wenn es alle tun, muss es auch richtig sein. Und tatsächlich: Abertausende von Forschungsarbeiten beruhen auf der Signifikanztestung. Die Formel „Signifikanz = Beweis" gehört zum (fast religiösen) Glaubensbekenntnis der Wissenschaft.

Es gibt da nur ein kleines Problem, das bei genauer Betrachtung umso größer wird, je genauer man hinschaut. Die sogenannte Nullhypothesentestung ist statistischer Nonsense, wenn auch einer, dem sehr viele Wissenschaftler auf den Leim gehen.[4] Wohlgemerkt: Man darf dieser Ignoranz nicht unbedingt Absicht unterstellen, denn dieses Glaubensbekenntnis gehört in der universitären Lehre sowie im Prozess des wissenschaftlichen Publikationsbetriebs zur Indoktrination.[5] Warum sollte man also auf die Idee kommen, dieses Bekenntnis in Frage zu stellen? Dieser Irrglaube wurde allerdings schon früh und zu einer Zeit erkannt, in der Statistik noch Handarbeit war…

Die Praxis der Signifikanztestung wurde bereits in den 1930er Jahren harsch kritisiert. Leider fand dies wenig Gehör. Man kann sich leicht vorstellen, warum am als „kleines schmutziges Geheimnis" bezeichneten Irrglauben festgehalten wurde.[5] Der Psychologie (von dort nahmen Signifikanztestungen ihren Ursprung) gaben signifikanzstatistische Prüfungen den Nimbus der Objektivität und damit der Wissenschaftlichkeit. Folglich gewannen sie zunehmend an Bedeutung und folglich wurde es immer schwerer, von ihnen abzulassen. Umso problematischer wurde die Situation, umso mehr statistische „Nachweise" erbracht wurden. Nicht nur die empirische Psychologie ist inzwischen in einer ziemlich prekären Situation: Es herrschen große Zweifel darüber, über welches gesicherte Wissen man wirklich verfügt. Man weiß jedoch eines: In vielen Fällen liegen eher Scheinbeweise vor.[6,7] Viele Insider fürchten deswegen, dass so manches Paradigma und so manche Theorie wie ein Kartenhaus zusammen fällt, wenn dieses Geheimnis gelüftet wird.

"Die ausschließliche Nutzung von Effektmaßen und Vertrauensintervallen als Grundlage der Bestimmung der Evidenz macht die Einschätzung der Studienergebnisse direkt vergleichbar."

An dieser Stelle sollen einige Gründe genannt werden, warum Signifikanztests wenig mit wissenschaftlichem Nachweis zu tun haben.

(1) Die meisten der selbst „hoch signifikanten" Ergebnisse sind von eher geringer praktischer oder klinischer Relevanz, weil die zugrunde liegenden Effekte i.d.R. sehr klein sind. Mathematischen Schätzungen zufolge sind z.B. erschütternde 80% der Befunde in der Medizin mit hoher Wahrscheinlichkeit Scheinevidenzen.

(2) Ein weiterer Grund, der Signifikanztestungen als Instrument einer Scheinwissenschaft enttarnt, ist die Tatsache, dass nahezu jede Nullhypothese mit Leichtigkeit verworfen werden kann. Man muss dazu z.B. nur die Stichprobe hinreichend groß machen (was man a priori mit sog. Power-Analysen machen kann). Statistische Simulationen haben gezeigt, dass selbst völlig sinnfremde Zusammenhänge signifikant werden, so lange man munter irgendwelche Variablen zusammenwürfelt. Egal, was man rechnet, irgendetwas wird immer signifikant.

(3) Der zugrunde liegende fundamentale Fehler bei der Signifikanztestung – weil rituell übernommen und nicht weiter hinterfragt – ist die Verknüpfung des Wahrscheinlichkeitswerts (p) mit der Fehlerschätzung (α). Beide sind mathematisch nicht miteinander assoziiert; daran ändert auch die Tatsache nichts, dass dies (falsch) in Lehrbüchern steht oder in universitären Curricula als Fakt vermittelt wird.

(4) Viele der Anwender von Signifikanztests haben das Prinzip der bedingten Wahrscheinlichkeit nicht verstanden. Formal lässt sich ein statistisches Ergebnis darstellen als p = (D|H0), d.h. die Wahrscheinlichkeit der Daten bei Gültigkeit der Nullhypothese. Dies ist nicht identisch mit p = (H0|D), also der Wahrscheinlichkeit der Nullhypothese bei gegebener Datenlage. Genau das macht die Logik der Signifikanztestung: Sie verkehrt bei „signifikantem" Ergebnis die Verhältnisse und stellt das in Frage (die Hypothese), auf dessen Grundlage sie die Daten berechnet. Man kann das besser verstehen, wenn man D und H0 einmal mit „Regen" und „Wolken" ersetzt. Die Wahrscheinlichkeit von Regen bei Vorhandsein von Wolken ist eine andere als die von Wolken bei Vorhandensein von Regen.

(5) Generell muss man anzweifeln, wie realistisch die Annahme einer Nullhypothese überhaupt ist. Zwischen zwei Testgruppen gibt es immer eine nahezu unendliche Zahl von Faktoren, die mitwirken, und zur

Unterschiedlichkeit von Gruppen beitragen.

Es verwundert aber kaum (oder doch?), dass selbst statistische Experten nicht in der Lage sind, eine korrekte Definition eines „signifikanten" Ergebnisses zu geben (Definition: Die Wahrscheinlichkeit des vorliegenden Ergebnis unter der als gültig vorausgesetzten Nullhypothese). Man bringt ein über Jahrzehnte gefestigtes Missverständnis eben nur schwer aus den Köpfen. Es ist vielfach auch ein wissenschaftssoziologisches, wenn nicht sogar ein existenzielles Problem: Gerade etablierten oder renommierten Wissenschaftlern dürfte es schwer fallen, einen als bewiesen geglaubten Sachverhalt, zu dem sie vielleicht selbst aktiv beigetragen haben, als Scheinevidenz enttarnt zu sehen. Der allgemeine Wissenschaftsbetrieb macht daher in der Regel weiter „Business as usual". Allerdings erkennen mehr und mehr Kollegen, dass sich diese Praxis ändern muss und erste Fachzeitschriften gehen andere Wege[8] oder zweifeln öffentlich die Glaubwürdigkeit klinischer Studien an [9, 10].

Einer dieser Wege besteht darin, ausschließlich Effektmaße und Vertrauensintervalle als Grundlage der Bestimmung der Evidenz zu nehmen. Effektmaße sind dimensionslos und daher direkt miteinander vergleichbar. Bei der Bestimmung ihrer Höhe orientiert man sich zwar auch an Konventionen, sie vermitteln jedoch wesentlich direkter die klinische oder praktische Bedeutsamkeit eines Befundes. Effektmaße sichert man durch die Wahrscheinlichkeitsbestimmung von Vertrauensintervallen ab, bei denen die untere und obere Grenze eines Effektmaßes festgelegt wird. Effektmaße sind auch das, was in Metaanalysen gerechnet wird, in denen möglichst viele zu einem Sachverhalt erhältliche Studienergebnisse in einer Analyse bewertet werden. Das macht die Einschätzung eines einzelnen Studienergebnisses direkt vergleich- und einschätzbar.

Informationstechnologie und Wirksamkeitsnachweis am Beispiel des Vita Chip E

Ich möchte an dieser Stelle zwei Studien vorstellen, die die Wirksamkeit eines neuen, auf den Grundlagen informationstechnologischer Prinzipien entwickelten „Chips" auf den Prüfstand stellten. Es handelt sich um prospektive, randomisierte, messwiederholte Dokumentationsstudien mit Personen, die unter chronischem Schmerz litten. Im Vordergrund dieser Studien stand nicht die Frage, wie das Verfahren wirkt, sondern, ob es wirkt (das ist die typische Frage von Wirksamkeitsprüfungen). Diese Unterscheidung ist wichtig für alle Verfahren, deren Wirkprinzipien umstritten oder noch nicht vollständig verstanden sind.

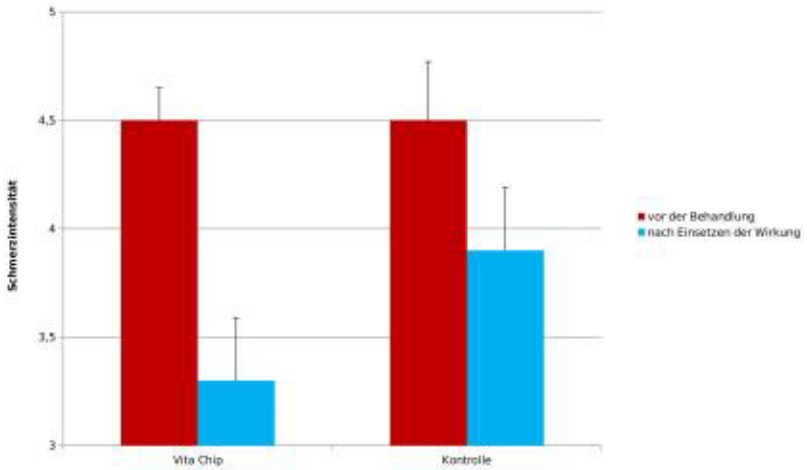


Abbildung : Schmerzreduktionen in Studie 1 (Skala: 1 [kein Schmerz] bis 7 [unerträglich])

In diesen Studien dokumentierten die Teilnehmer unter Alltagsbedingungen auftretende Schmerzen anhand eines standardisierten Schmerz- und Befindlichkeitsprotokolls für die Dauer von jeweils einer Woche. Als Zielvariable dienten die Schmerzintensität, die Latenz bis zur spürbaren Schmerzlinderung, die Schmerzdauer und das allgemeine Wohlbefinden. An den Untersuchungen nahmen volljährige Erwachsene teil, die außer chronischen Schmerzen (jedweder Art!) keine weiteren gesundheitlichen Beeinträchtigungen hatten. Das Verfahren, der Frequenzchip Vita Chip E, besteht aus mikroskopisch kleinen Kristallen und Mineralien. Laut Hersteller werden seine Bestandteile genutzt, um darauf „Heilfrequenzen" zu speichern und zu übertragen. Er ist er sehr klein (1 cm^2) und leicht an elektronischen Geräten anzubringen. Letztere werden als Verstärker der Informationssignale genutzt, die auf den Vita Chip E aufmoduliert sind (analog eines CD-Rohlings auf den Musik aufgespielt wird). Bislang lag zur Wirksamkeit lediglich anekdotische Evidenz vor; diese ist eindrücklich und schließt sehr heterogene Klassen von Störungen ein (z.B. Schlafstörungen, Allergien, Schmerzen).

Studie 1

Diese Studie war eine Crossover-Studie, d.h. die 30 Teilnehmer dokumentierten ihre Schmerzen für jeweils eine Woche mit und eine Woche ohne Vita Chip E. Die Wirksamkeit wurde damit anhand des natürlichen Schmerzverlaufs ermittelt. Die Hälfte der Teilnehmer setzte die Vita Chips E in der ersten Woche ein, die andere in der zweiten. Die Intervention bestand im Auflegen des Handys (mit Vita Chip E) auf die Schmerzstelle (z.B. Schulter) für eine Stunde, sowie dem Aufkleben eines zweiten Vita Chip E in der Körpermitte (z.B. Halswirbelsäule). Für jede Schmerzart wurden analog je bis zwei Referenzstellen definiert. Weder Teilnehmer, noch

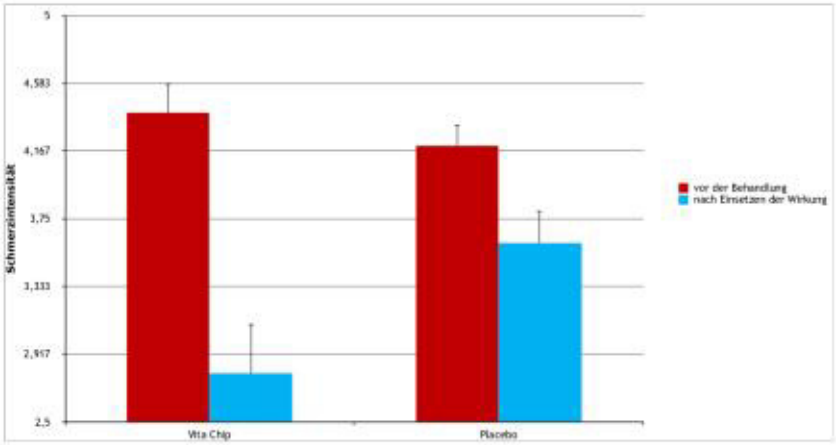


Abbildung: Schmerzreduktionen in Studie 2

Versuchsleiterin hatten Wissen darüber, woraus der Vita Chips im Detail bestand und wie er genau funktionierte.

Ergebnisse

Abbildung 1 zeigt die Schmerzintensitäten zu den gemessenen Zeitpunkten. Wie ersichtlich war der Schmerzrückgang bei Einsatz der Vita Chips E sehr deutlich ausgeprägt. Der statistische Effekt der Senkung (Erstmessung versus Zweitmessung) war sehr groß ($d = 1,3$; Vertrauensintervall: $0,7 < d < 1,9$) und entsprach einer Reduktion von „stark" auf „gering" (großer Effekt: $d \geq 0,8$). Demgegenüber war die Reduktion in der Kontrollwoche deutlich geringer ($d = 0,5$) und statistisch nicht aussagekräftig, weil das Vertrauensintervall in den negativen Bereich ragte ($-0,1 < d < 1,1$). Nach acht Stunden war der Unterschied im Schmerzrückgang zwischen Vita Chip E und Kontrolle mit $d = 1,8$ (VI: $0,6 < d < 3,0$) noch größer. Auch in den zeitlichen Parametern zeigten sich Unterschiede: Unter Einsatz der Vita Chips E trat die schmerzlindernde Wirkung nach ca. 66 Minuten ein, ohne Vita Chips E erst nach ca. 173 Minuten. Dieser Effekt war ebenfalls groß ($d = 0,8$; VI: $0 < d < 1,6$). Auch der Schmerzdauer war mit dem Vita Chip E deutlich verkürzt (von 173 auf 130 Minuten). Allerdings war dieser Effekt klein und fiel in negative Vertrauensintervallgrenzen). Ein klarer Effekt zeigt sich bezüglich des Allgemeinbefindens, das nach Anwendung der Vita Chips E um 15% anstieg ($d = 1,1$; $0,4 < d < 1,8$), wohingegen sich in der Kontrollwoche keine Veränderung ergab ($d = 0,1$). (siehe Abbildung 1)

Studie 2

In dieser Studie wurde im Gegensatz zur ersten statt der natürlichen Kontrolle gegen ein Pseudoplacebo getestet, um zu ermitteln, inwieweit die Effekte des Vita Chip E auf einen Placeboeffekt zurückzuführen sind. Zwanzig Teilnehmer erhielten das Verum (zwei informierte Vita Chips E) und zwanzig Teilnehmer erhielten zwei sogenannte Pseudoplacebos. Diese waren uninformierte Vita Chips E, d.h. Rohlinge mit denselben Mineralanteilen.

Ergebnisse

In beiden Gruppen sank der Schmerz nach Auflegen der Vita Chips. In der Verumgruppe war der Effekt jedoch deutlich größer und entsprach einer Senkung von „mittelstark" auf „schwach" ($d = 1,4$, $0,7 < d < 2,1$). In der Placebogruppe sank er nur von "mittel" auf „mittel bis gering" ($d = 0,8$, $0,2 < d < 1,4$). Der Unterschiede der Wirkung zwischen den beiden Bedingungen war groß ($d = 1,0$; $0,4 < d < 1,6$). Nach einer Stunde (Zeit des Einsatzes der Vita Chips) war der Unterschied ebenfalls groß ($d = 0.8$; $0 < d < 1.6$), sank dann ab und bestand nicht mehr ($d \leq$

Fußnoten

[1] Ebbers, J. A. & Meyl, K. (2014). Drug effects in yeast mediated by scalar waves. Medical Science, 8, 58-62.

[2] Schneider, R. (2007). The psychology of the placebo effect: Exploring meaning from a functional account. Journal of Mind and Behavior, 18, 11-17.

[3] Schneider, R. & Kuhl, J. (2012). Placebo forte: Ways to maximize unspecific therapy effects. Medical Hypotheses, 78, 744-751.

[4] Branch, M. (2014). Malignant side effects of null-hypothesis significance testing. Theory and Psychology, 24, 256-277.

[5] Gigerenzer, G. (2004). Mindless statistics. The Journal of Socio-Economics, 33, 587-606

[6] Ioannidis, J. P. A. (2005). Why most published research findings are false. PLoS Medicine, 2, 0696-0701.

[7] Lambdin, C. (2012). Significance tests as sorcery: Science is empirical - significance tests are not. Theory and Psychology, 22, 67-90.

[8] Schneider, R. (in press). There is something in the air: Testing the efficacy of a new olfactory stress relief method (AromaStick®). Stress & Health.

[9] Marcovitsch, H. (2010). Editors, publishers, impact factors, and reprint income. PLoSMed,Oct; 7(10): e1000355.

[10] Horton, R. (2015). Offline: What is medicine's 5 sigma? The Lancet, 385.650397

0,5; -0,8 < d < 1,2). Auch die zeitlichen Schmerzparameter unterschieden sich, ließen sich aber statistisch nicht absichern. In der Verumgruppe setzte die schmerzlindernde Wirkung nach ca. 26 Minuten ein, in der Placebogruppe nach ca. 57 Minuten (d = 0,7; -0,1 < d < 1,5). Ebenso dauerte der Schmerz in der Placebogruppe ca. 82 Minuten länger an (402 Minuten) als in der Verumgruppe (320 Minuten); dieser Effekt war aber statistisch gesehen nicht abzusichern (d = 0,3; - 0,6 < d < 1.4). Unter Einsatz der Verum-Vita-Chips war auch die Befindlichkeit besser, aber dieser mittlere Effekt von d = 0,5 fiel ebenfalls in negative Vertrauensintervallgrenzen. (siehe Abbildung 2)

Effektstärken

Eines der bekanntesten Effektmaße ist Cohen's d. Per Konvention gilt ein kleiner Effekt als d = 0,2, ein mittlerer Effekt als d = 0,5 und ein großer Effekt als d ≥ 0,8. Die Maßeinheit von d ist die gemittelte Standardabweichung. Ein d von 1 bedeutet z.B., dass zwei Gruppen genau eine Standardabweichung auseinander liegen (innerhalb einer Standardabweichung liegen etwas mehr als 68% der beobachteten Werte). Zur Bestimmung von d sind die Grenzen des Vertrauensintervalls wichtig. Diese dürfen nicht gleichzeitig positiv und negativ sein. Bei kleinen Stichproben sind die Grenzen oft so weit auseinander, dass der Effekt nicht abgesichert werden kann. Erst wenn die Effekte sehr groß werden, liegt auch das Vertrauensintervall in einem verlässlichen Bereich. Effektmaße sollten neben dem Vertrauensintervall immer anhand der Qualität der Studie und der praktischen Relevanz des Effekts beurteilt werden. In der Regel kann davon ausgegangen werden, dass Effekte größer d = 1 praktisch relevant sind. Je größer d, desto weiter liegen die getesteten Gruppen auseinander bzw. desto weniger Überlappungsbereich besteht zwischen ihnen. Bei einem Effekt von d = 1, 4 (Studie 2) liegen die Werte der Erstmessung (vor der Behandlung) und Zweitmessung (nach der Behandlung) soweit auseinander, dass 92% der Probanden in der Zweitmessung oberhalb des Mittelwertes der Erstmessung liegen.

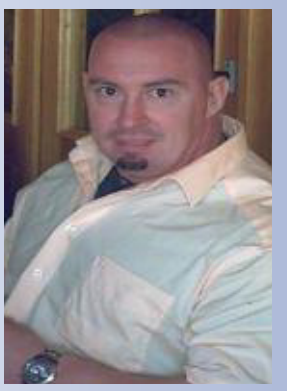
Fazit

In beiden Studien ging nach einstündigem Einsatz zweier Vita Chips E der chronische Schmerz deutlich zurück. Diese Reduktion war nicht erklärbar durch den natürlichen Schmerzrückgang oder eine rein psychologische Wirkung. Die Tatsache, dass diese Wirkung nicht vollständig auf einen Placeboeffekt zurückgeführt werden konnte, ist wichtig. Das Verhältnis von spezifischem zu unspezifischem Effekt war 61 Prozent zu 39 Prozent. Damit war der spezifische Anteil um mehr als 150 Prozent größer als der unspezifische. Auch die Latenzzeit der schmerzlindernden Wirkung sowie die Gesamtdauer der Schmerzen wurden durch die Vita Chips reduziert, wobei die absoluten und prozentualen Reduktionen sich nicht durchgängig in robusten Effekten äußerten. Dies hing im Wesentlichen damit zusammen, in welchem Maß die Anwender auf die Vita Chips reagierten.

Die Wirkung des Vita Chips E setzte aber erwartungsgemäß im Wesentlichen während der tatsächlichen Anwendung ein; bei den meisten Probanden kam es in dieser einen Stunde zu den deutlichsten Schmerzrückgängen. Neben der schmerzlindernden Wirkung schien der Vita Chip auch eine stimmungsaufhellende Wirkung zu entfalten. Sie hängen aber davon ab, wie das Allgemeinbefinden ist (in Studie 2 war sie generell hoch und konnte kaum weiter gesteigert werden). Die Rate sogenannter Non-Responder (Personen, die nicht auf den Vita Chip reagieren), lag bei etwa 20 Prozent (in beiden Studien), die Rate der Non-Responder bei Placebo bei 55 Prozent. Auch hier bestehen beträchtliche interindividuelle Unterschiede: Manche Personen berichten eine unmittelbare Wirkung nach wenigen Minuten bis zur völligen Schmerzfreiheit. Manche spüren eher eine allmähliche Linderung, die deutlich länger dauert.

Insgesamt zeigen die beiden Studien, dass sich die Wirksamkeit informationsmedizinischer Verfahren wie dem Vita Chip E empirisch nachweisen lassen. Selbst sehr unterschiedliche Schmerzen, die im Schnitt mindestens seit sechs Monaten bestanden und die die Probanden nur leidlich in den Griff bekamen, können damit deutlich reduziert werden. Damit ist zumindest dieses spezielle Verfahren weder Hokuspokus, noch reiner Placeboeffekt.

Der Autor



Rainer Schneider ist promovierter Psychologie. Nach langjähriger Arbeit in der experimentellen Psychologie sowie in der Evaluation Alternativer Heilmethoden am Universitätsklinikum Freiburg gründete er 2007 sein eigenes Unternehmen RECON - Research and Consulting. Im Jahr 2010 erhielt er seine Lizenz als funktionalanalytischer Coach (IMPART/Universität Osnabrück). Schneider koordiniert nationale und internationale Forschungsprojekte und forscht im Bereich spezifischer und unspezifischer Effekte. Er ist Autor zahlreicher Ratgeber und Publikationen, Gutachter wissenschaftlicher Zeitschriften und wissenschaftlicher Berater von Stiftungen und Organisationen (u.a. Bundesärztekammer).