

Replik auf Prof. Dr. phil. Harald Walachs Antwort auf den Offenen Brief „Doppelblindstudien, Rationalität und Homöopathie“

von Dr. med. Hanspeter Seiler

Juli 2006

Lieber Herr Professor Walach,

für Ihre detaillierte Antwort danke ich Ihnen und entschuldige mich für die späte Replik. Da mir Kritik aufgrund des etwas negativen Beigeschmacks, welcher ihr auch bei konstruktiver Ausrichtung immer anhaftet, als Hauptfach nicht so liegt, wollte ich zuerst noch ein paar Projekte rein aufbauenden Inhaltes etwas weiterverfolgen. - Bevor ich nun aber Punkt für Punkt mit der notwendigen Ausführlichkeit auf Ihr Schreiben eingehe, möchte ich für eilige LeserInnen meine generelle Stellungnahme hierzu kurz zusammenfassen:

Zusammenfassung

Das einzige zu einer Korrektur Anlass gebende Resultat Ihrer Antwort ist, dass bereits während des Vorlaufes Ihrer Studie ein teilweiser Kaffee-Entzug erfolgte und nicht erst in der Therapiephase. Dies ändert aber nichts am Gesamtergebnis meiner Analyse. Im übrigen scheint mir Ihre Argumentation trotz einiger erfreulicher Gemeinsamkeiten deutlich zu oberflächlich, zu ausweichend und teilweise auch sachlich falsch, wie im Folgenden detailliert gezeigt werden soll. In jedem Fall bleiben die drei Grundpfeiler meiner Kritik davon unberührt:

- 1. Überalterung und die heikle Vermischung von Anfalls- und DauerpatientInnen machen Ihre Studie wesentlich reaktionsträger als die Vergleichstudien.*
- 2. Aufgrund von bisher nicht erkannten Randomisierungsartefakten reagierten Placebo- und Verumgruppe klar unterschiedlich, wodurch sich eine erhebliche Verzerrung Ihrer Resultate ergibt.*
- 3. Im Gegensatz zu Ihrer Interpretation der Therapiereaktionen, welche u.a. auch durch eine Datenverwechslung erheblichen Ausmasses beeinträchtigt wurde, machte die Verum-Gruppe spezifische und teilweise wahrscheinlich auch lange anhaltende homöopathische Verschlimmerungen durch.*

Diese Faktoren genügten, um die Feststellung des homöopathischen Heileffektes Ihrer Studie erheblich abzuschwächen und schliesslich im statistischen Endresultat völlig zum Verschwinden zu bringen. - Auf meine grundsätzliche Widerlegung Ihrer magischen Homöopathie-Interpretation gehen Sie leider kaum ein.

Diskussion Ihrer einzelnen Kritikpunkte

Ich zitiere Ihren Originaltext in Normalschrift, *meine Antwort darauf folgt kursiv. Die Antwort auf einzelne zusammengehörige Punkte Ihrer Aufzählung habe ich zusammengefasst; bei anderen habe ich mir erlaubt, getrennte Aussagen zur besseren Übersicht alphabetisch zu unterteilen (z.B. 11 a., 11 b. etc.). Seitenzahlen in Klammern beziehen sich, wenn nichts anderes angegeben ist, immer auf meine Originalarbeit, welche ja ebenfalls auf dieser Netz-Seite publiziert ist. Gekürzte Quellenangaben (z.B. Walach 1999) beziehen sich ebenfalls auf das Literaturverzeichnis dieser Arbeit.*

0. (Einleitung) Walach: Wenn Sie 60 Seiten darauf verwenden, eine Studie zu kritisieren, ohne die entsprechende Diskussion darüber gelesen zu haben, und dies dann auch noch als „prägnant“ bezeichnen, ist etwas faul.

Seiler: Ich gebe zu, dass ich nicht die ganze Sekundärliteratur zu Ihrer Studie systematisch aufgearbeitet habe, und dass dies ein gewisser Mangel ist. Als viel beschäftigter Praktiker habe ich lernen müssen, mich bei meinen wissenschaftlichen Arbeiten auf das Allerwesentlichste zu beschränken. Meist genügt es nämlich auch, die Originalquellen genau anzuschauen - dafür dann aber wirklich genau! - Im Übrigen habe ich meine Arbeit nicht selbstherrlich als „prägnant“ bezeichnet, sondern lediglich festgehalten, dass ich mich um diese Eigenschaft bemühen werde.

1. Die „Magie-Interpretation“ der Homöopathie beruht nicht allein auf unserer Kopfschmerzstudie, sondern auf einer jahrelangen, sehr intensiven Sichtung der Literatur, auf den Daten von Arzneimittelprüfungen von mir und von vielen anderen Autoren.

Diese Berichtigung nehme ich gerne zur Kenntnis, zumal sie mir auch schon vorher durchaus bewusst war (ich habe ja auch Ihren Magie-Artikel gründlich gelesen!). Sicher aber scheint mir, dass die Münchner Kopfschmerzstudie als Ihre bekannteste Arbeit eine der wichtigeren Grundlagen Ihrer Magie-Interpretation der Homöopathie ist.

2. Die Kopfschmerzstudie wurde von mir überhaupt nicht mit Medienrummel vorgetragen. Dass die Medien einen Rummel daraus gemacht haben, liegt leider in der Natur der Sache.

O.K., nehme ich auch zur Kenntnis. Jetzt aber läge es an Ihnen, mit Ihrem professoralen Hintergrund auch die erwiesenen Fehler und wissenschaftlich fundierten Kritikpunkte Ihrer Studie entsprechend publik zu machen.

3. Es ist leider ein weit verbreiteter Fehler, von einer gelungenen Studie dazu überzugehen, die „Homöopathie als bewiesen“ zu vermelden. Ich weise konstant auf das Doppelproblem hin: Es gibt zu viele Daten, um Homöopathie als Unsinn abzutun, aber viel zu wenig *konsistente* Daten, um sie als kausal und lokal wirksam zu belegen. Siehe dazu unseren neuesten Überblick, in dem die Argumentation ausgeführt ist.

Hier hingegen bin ich anderer Ansicht. Eine grosse Anzahl homöopathischer Studien geht von ungeeigneten oder fehlerhaften Ansätzen aus und/oder wird fehlinterpretiert. Ihre Studie ist für mich ein historisches Musterbeispiel dafür. Eine fachgerechte homöopathische Studienkultur steckt noch ganz in den Kinderschuhen, da die homöopathische Studien-Szene heute noch immer mehrheitlich von Nicht-Medizinern sowie einem zu einseitigem Kentismus einerseits und von zu sehr der Rezept-Homöopathie verpflichteten Firmen andererseits dominiert wird. Negative Erfahrungen wie Ihre Münchner Studie sind aber wichtige Wegweiser in die Zukunft.

Vorläufig aber gilt nach wie vor: „Auch wenn unter hundert Studien nur eine einzige mit einem eigenständigen Design ein seriös dokumentiertes positives Resultat zeigt, hat sie zumindest für diesen spezifischen Bereich für sich allein schon Beweiskraft, und zwar

*solange, bis sie bei **genauer** Reproduktion lege artis widerlegt wird“ (S. 6). Nur diese Position kann m.E. als wissenschaftlich bezeichnet werden, nicht aber ein nach rein schulmedizinisch-statistischen Kriterien zusammengewürfelter Studienmix, wie er für viele Metaanalysen typisch ist (vgl. die Egger-Studie und teilweise auch Ihre Zusammenstellungen). Natürlich muss man dann auch bei den als negativ klassifizierten Studien möglichst klar aufzeigen können, wie und warum sie zu dieser Aussage kommen konnten. Dies habe ich für die von mir analysierten Kopfschmerz-Untersuchungen m.E. zur Genüge getan.*

4. Die in der Homöopathie-Szene euphorische Rezeption der Reilly-Studien ist verständlich, hält aber bei genauerem Hinsehen nicht stand: Keine einzige der Studien ist eine wirklich genaue Replikation. Wenn man genau hinsieht, ist bei jeder Studie was verändert, und diejenigen Studien, die sich als genaue Replikation verstanden, haben das Phänomen eben genau nicht repliziert.

Da bin ich mit Ihnen zumindest teilweise völlig einverstanden, s. auch oben. Ihr Aussage gilt dann aber gerade auch für die Reilly ebenfalls nur ungenau reproduzierende Lewith-Studie, welche seine beiden positiven Studien angeblich widerlegen soll (S. 5 und 6). Und natürlich ist Reillys zweite Studie mit Asthma-PatientInnen keine einfache Reproduktion seiner vorangegangenen Arbeit über Heuschnupfen, so war sie auch nie intendiert. Hingegen bestätigt sie unabhängig vom unterschiedlichen Patientengut die für beide Studien geltenden Resultate:

- *Zumindest kurzfristig und relativ schwach ausgeprägt ist Reillys isopathischer Nosodenansatz bei Heuschnupfen und allergischem Asthma signifikant wirksam.*
- *Daraus folgt als wichtigste Schlussfolgerung, dass Hochpotenzen grundsätzlich klinisch wirksam sind.*

5. Es ist ein komplettes Missverständnis, mich als „Gegner einer rationalen Begründbarkeit“ zu titulieren. Dagegen wehre ich mich. Die Tatsache, dass Sie das tun, zeigt, dass Sie weder meine Arbeiten ausreichend zur Kenntnis genommen haben, um sich ein solches Urteil zu erlauben und zweitens ein Verständnis von „rationaler Begründbarkeit“ haben, das in einem Wissenschaftsverständnis des 19. Jahrhunderts begründet ist. Dies ist an sich für mich nicht satisfaktionsfähig. Daher sollte ich jetzt eigentlich aufhören zu schreiben. Da ich nicht als dialogunfähig gelten will, tue ich dies bewusst nicht.

Ich will gerne zugeben, dass an dieser Stelle eine saubere Begriffsdefinition von kausal und akausal, determiniert und undeterminiert sowie von Magie, Rationalität und Irrationalität angesagt wäre. Aber einmal mehr muss ich sagen, dass ich naturwissenschaftlicher Praktiker und nicht geisteswissenschaftlicher Professor bin. Selbst vom akademisch gebildeten Normalverbraucher werden ja Magie, Nicht-Kausalität und Irrationalität oft etwas undifferenziert in den gleichen esoterischen Kübel geworfen, dies aber keineswegs völlig zu Unrecht. Und wenn Sie sich schon ausdrücklich als Vertreter einer magischen Interpretation der Homöopathie positionieren (und da möchte ich wirklich nicht in Ihrer Haut stecken!), müssen Sie sich halt von der Basis derartige Bezeichnungen gefallen lassen, ohne gleich betupft zu reagieren.

Wenn Sie wie zu vermuten mit Ihrem Vorwurf des überholten Wissenschaftsverständnisses aus dem 19. Jahrhundert das letztlich in der Aufklärung begründete Glaubensbekenntnis zu einer kausalen Rationalität meinen, ist dies für mich eher ein Kompliment als ein Vorwurf. Mit diesem Argumentarium ist nämlich Hahnemann gegen die damalige Schulmedizin losgezogen, und es hatte auch noch für Einstein seine volle Gültigkeit, als er sich zeitlebens gegen die Kopenhagener Interpretation der Quantenmechanik gewehrt hat, auf welcher heute Esoteriker und religiöse Schwärmer jeglicher Provenienz ihr schales Süppchen kochen. Hüten Sie sich davor, hier als Zugesel vor den falschen Karren gespannt zu werden, indem Sie den nicht-kausalen vermutlichen Irrweg der Kopenhagener Interpretation auch in der wissenschaftlichen Homöopathie salonfähig machen wollen - es wäre schade um Ihre unbezweifelten akademischen Qualitäten!

Das Ende des lokalen Kausalitätsbegriffes ist für mich das Ende des wissenschaftlichen

Denkens schlechthin, zumindest in dem Sinne, wie es die Menschheit in den letzten dreitausend Jahren gehandhabt hat. Natürlich müssen wir das monokausale Denken eines verknöcherten Vulgärmaterialismus durch einen multifaktoriellen und auch die feinstofflich-lebensenergetischen Realitätsebenen erfassenden Denkansatz ersetzen. Aber dies ist, wie Mesmer, Reich und Hahnemann gezeigt haben, noch lange kein Abschied vom Kausalitätsprinzip! Dies werden schlussendlich auch die QuantenphysikerInnen der Kopenhagener Schule lernen müssen - nicht zuletzt auch von der bioenergetischen Ganzheitsmedizin!

6. Ich bin ziemlich Ihrer Meinung, was die Egger-Studie betrifft und habe meine entsprechenden Argumente öffentlich kundgemacht. Daher spare ich mir eine Wiederholung. Allerdings ist es ein Missverständnis zu glauben, die wissenschaftliche Methodik würde der Homöopathie im Prinzip nicht gerecht. Dies stimmt aus meiner Sicht nicht.

Ganz meine Meinung. Ich glaube bei Ihnen ganz generell in den letzten Monaten einen tendenziellen Positionswechsel zurück auf die rational-kausale Ebene feststellen zu dürfen, welcher mich – wie schon in der Schlussfassung meiner Arbeit ausdrücklich festgehalten (S. 7) - sehr freut. Nun aber kehren Sie gar den Spiess um und bezeichnen mich als Gegner einer wissenschaftlichen Begründbarkeit der Homöopathie! Ich bin ja weisgott nicht grundsätzlich gegen die Doppelblindstudie, wie aus meiner Arbeit wohl zur Genüge ersichtlich ist. Ich habe lediglich wissenschaftlich begründet, warum diese einer ganzheitsmedizinischen Methode wie der klassischen Homöopathie nur partiell gerecht werden kann.

7. Die Frei-Studie ist ein denkbar schlechtes Beispiel. Wenn man sie nämlich genau analysiert, sieht man, dass eigentlich gar kein therapeutischer Effekt in dem Sinne abgebildet ist, sondern eine differentielle Reaktion auf Placebo. Was das bedeutet, ist eine komplexe Frage. Jedenfalls gilt auch hier: Jede *neue* Studie zeigt irgendetwas Interessantes. Die Frage ist: Lässt sich ein Effekt stabil reproduzieren?

Ich habe diesen Passus überhaupt nicht verstanden, und Heiner Frei, dem ich ihn ebenfalls vorgelegt habe, ebenfalls nicht. Was meinen Sie damit genau? Es ist ja sicher, dass die ADS-Patienten der Frei-Studie nicht nur mit Plazebo behandelt wurden und dass das Resultat signifikant positiv ist. Dass dies nur genau für das Setting dieser Studie gilt, müssen wir nach dem Obigen wohl nicht mehr wiederholen.

8. Das Argument gegen „Massenstudien“ ist schwach. Es ist zwar pragmatisch gesehen richtig. Sachlich wäre aber denkbar, dass eine riesige Studie mit lauter individualisierten Behandlungen durchgeführt wird. Unsere Kopfschmerzstudie war ein erster Anfang, und andere Studien, die diesen Ansatz wählten, sind alle nicht sehr weit gekommen (etwa Elly de Lange, Straumsheim, um nur einige Beispiele zu nehmen). Es gibt auch gute Gegenbeispiele, aber keines mit einer stabilen Replikation.

Theoretisch wären individualisierte Massenstudien natürlich denkbar, aber wirklich nur theoretisch. An die Teilnehmerzahlen der von mir kritisierten indischen Euphrasia-Massenstudie von über 1300 wird man mit einem guten klassisch-homöopathischen Konzept auch mittelfristig nicht herankommen. Hierfür müssten die beteiligten ÄrztInnen ja möglichst allesamt Spitzenleute sein. Wie sollte man diese für ein derart grosses Patienten-Kollektiv zusammenbekommen? - Im Übrigen zielt meine Kritik an den Massenstudien vor allem gegen die Egger-Studie, über welche wir uns - wie Sie ja auch betonen - erfreulicherweise grundsätzlich einig sind. Diese Frage ist deshalb für unsere Auseinandersetzung nicht so relevant wie z.B. Ihre Ablehnung der Frei-Studie.

9. Das Beispiel der Jacobs-Studie übersieht, dass das Ergebnis gerade nicht replizierbar war. Die letzte Studie ist nicht signifikant und nur eine Meta-Analyse behält das Ergebnis bei.

Auch hier verstehe ich Ihre Argumentation nicht. Nach meiner Beurteilung wurde die Jacobs-Studie mit absolut korrekten Methoden aus drei Studien zusammengesetzt, womit sich ein zwar nur diskretes, aber doch eindeutig signifikantes positives Resultat nachweisen lässt. Dieses wurde denn auch in einer hochoffiziellen pädiatrischen Fachzeitschrift publiziert. Wenn die Fallzahl einer einzelnen Studie für eine positive Signifikanz nicht ganz ausreicht, wieso soll man dann nicht – korrektes Vorgehen vorausgesetzt - wie bei einer multizentrischen Arbeit mehrere Studien zusammenfassen dürfen?

10. Bei den Kopfschmerzen haben Sie erwiesenermaßen Unrecht: Die neue Dokumentation von Witt et al. zeigt, dass Kopfschmerzen häufig in der homöopathischen Praxis behandelt werden und auch mit Erfolg. (Witt, C. M., Lüdtke, R., Baur, R., & Willich, S. N. (2005). Homeopathic medical practice: Long term results of a cohort study with 3981 patients. *BMC Public Health*, 5, 115.) Ich habe mich im Vorfeld mit sehr vielen, damals weitherum bekannten Homöopathen genau auf Kopfschmerzen als potenziell einfaches Syndrom verständigt.

Ich habe natürlich nie behauptet, dass Kopfschmerzen – und auch solche chronischer Natur - in einer homöopathischen Praxis nicht mit einer anständigen Erfolgschance behandelt werden können. In schwer chronifizierten Fällen muss man hierfür allerdings einen realistischen, nicht selten auch mehrjährigen Zeitrahmen einräumen. Deshalb auch mein Hinweis auf Hahnemanns Anglesea-Fall (S. 13), wo sich der Altmeister selbst - wenn auch keineswegs völlig erfolglos - über Jahre hinweg und mit aus Patientensicht letztlich ungenügendem Gesamtergebnis abgemüht hat. Die grundsätzliche homöopathische Therapierbarkeit auch chronischer Kopfschmerzen beweisen übrigens ja zumindest partiell auch alle drei anderen von mir analysierten Migräne-Studien. Ihre Münchner Studie stellt hier, wie ich zu zeigen versucht habe, lediglich eine gut erklärbare Ausnahme dar. - Nicht chronifizierte Kopfschmerzen etwa im Rahmen einer akuten psychovegetativen Belastungssituation oder infolge eines akuten Infektes sind zudem auch tatsächlich relativ einfach und in kurzer Zeit homöopathisch therapierbar. Wenn aber homöopathische ÄrztInnen behaupten, dass mehrjährig chronifizierte Kopfschmerzen zu den leicht behandelbaren Krankheiten gehören, kann ich besagte KollegInnen nur als naiv-unerfahren oder grosssprecherisch bezeichnen. Wenn sich die KonsiliarärztInnen Ihrer Studie tatsächlich in diesem Sinne geäußert haben, haben Sie sich wirklich die falschen Leute ausgesucht. Tatsache aber ist, dass die real an der Studie beteiligten FachärztInnen zumindest post festum ein deutlich anderes Urteil abgegeben haben und ausdrücklich festhalten, dass "die im Rahmen der Studie behandelten Patienten zu den schwierigsten gehörten, die sie in ihrer Praxis behandelt hatten und dass für diese schwierigen Patienten die Studiendauer viel zu kurz war" (S. 30).

- 11 a. Warum die Brigo-Studie so erfolgreich war, ist schwer zu sagen. Post-hoc lässt sich alles sagen. Tatsache ist, dass ich viele Versuche unternommen habe, die originalen Autoren zu kontaktieren, auch um zu erfahren, warum diese Studie nie ordentlich publiziert wurde. Weder die Autoren noch die Sponsoren (Boiron) haben mir irgendeine Auskunft gegeben. Meine private Schlussfolgerung lautet: Diese Daten sind wissenschaftlich nicht brauchbar. Aber da kann man anderer Meinung sein.

Die Brigo-Studie wurde 1991 in „The Berlin Journal on Research in Homeopathy“ publiziert, wo damals Ihr Fachkollege Prof. J. Hornung federführend war, den ich persönlich gut kenne und schätze. Ihre obige Behauptung ist ein Affront gegen ihn. Was soll denn an dieser Publikation nicht in Ordnung sein? Sie ist nach meiner Praktiker-Beurteilung sogar formal die beste von allen vier Kopfschmerzstudien, die ich untersucht habe, und weist z.B. weniger sachliche Fehler und Unklarheiten auf als Ihre Publikationen zur Münchner Studie. Dies sagt allerdings noch nichts aus über die Qualität der Daten, deren Erhebung Hornung m.W. nicht persönlich kontrollieren konnte. Da mag es vielleicht noch offene Fragen geben, da das Studienresultat - wie von mir des längeren diskutiert - tatsächlich „ein fast allzu schönes“ ist (S. 18). Persönlich hat mir allerdings Brigo einen sehr guten

Eindruck gemacht. Ich konnte ihn problemlos telefonisch kontaktieren und er hat mir bereitwilligst alle Fragen beantwortet.

- 11 b. Die Glasgower Migräne-Studie war als Replikation von Brigo geplant. Dass einzelne Parameter anders waren, liegt in der Natur der Sache, weil im klinischen Kontext eine komplette Replikation ohne Variation Unsinn und Geldverschwendung ist. Die entscheidende Tatsache ist, dass ein der Sache nach sehr ähnliches Therapiekonzept sich nicht replizieren ließ. Dies lag sicher nicht am mangelnden Enthusiasmus oder an der mangelnden Kompetenz der Autoren. Man kann über die Art, Ähnlichkeit zu definieren, sehr geteilter Meinung sein. Aber es ist nicht einzusehen, warum etwas, das bei Brigo funktioniert hat, plötzlich in Glasgow nicht mehr funktionieren soll. Ihre Argumente sind aus meiner Sicht hilflose Versuche, diese simple Tatsache zu ignorieren.

Sie versuchen, Ihre eindeutige Falschmeldung, dass die Whitmarsh-Studie eine genaue Reproduktion der Brigo-Studie sein soll („direct replication“, S. 19), mit einigen recht fadenscheinigen Unverbindlichkeiten zu zerreden. Ich habe detailliert aufgezeigt, dass die Whitmarsh-Studie vor allem im homöopathischen Konzept gegenüber Brigo wesentlich verändert wurde. Damit und aufgrund eines ebenfalls leicht nachweisbaren Randomisierungsartefaktes habe ich die zwar von Brigo erheblich abweichenden, aber dennoch zumindest partiell positiven Resultate der Glasgower-Untersuchung recht gut erklären können. Hierzu möchte ich klare Gegenargumente hören. (Vgl. hierzu übrigens auch Ihr überraschendes Zusatz-Statement zu den positiven Teilresultaten der Whitmarsh-Studie unter Punkt 12!) - Zudem dürfte nach dem Obigen klar sein, dass bei einer wichtigen, aber umstrittenen Studie eine möglichst exakte Wiederholung wissenschaftlich wünschenswert und damit sehr wohl ihr Geld wert ist.

- 11 c. Es ist übrigens falsch zu glauben, Placebo-Effekte würden rasch wieder zurückgehen und alles was stabile Veränderung ist, kann daher nur ein Verum-Effekt sein. Diese Aussage ist sachlich und empirisch nicht haltbar. Die Verlaufskurven und Besserungen von Placebo und Verum sind schon in konventionellen Studien hochkorreliert (mit $r = .78$, wie eine von uns durchgeführte Meta-Analyse ergab; siehe Walach, H., Sadaghiani, C., Dehm, C., & Bierman, D. J. (2005). The therapeutic effect of clinical trials: understanding placebo response rates in clinical trials - A secondary analysis. BMC Medical ResearchMethodology, 5, 26.).

Ihre obige statistische Arbeit zum Plazebo-Effekt habe ich durchgelesen und - wie ich hoffe - zumindest in den Grundaussagen verstanden. Sie beweist aber m.E. nirgends, dass ernsthafte chronische Krankheiten jemals durch die unspezifisch-ungezielten Plazebo-Wirkungen schulmedizinischer Studien wirklich dauerhaft geheilt wurden, wie man dies aus Ihrer obigen Äusserung fast schliessen könnte. Sie beweist lediglich, dass Plazebo- und Verum-Ansprechraten in Doppelblindstudien um einiges enger korreliert sind als von der Schulmedizin bisher angenommen. Dies gilt aber bezeichnenderweise vor allem für prophylaktische Studien. Untersucht man hingegen die Therapieresultate einzelner schon bestehender organischer Krankheitsgruppen, wird die Plazebo-Ansprechraten deutlich geringer. Dies gilt insbesondere auch für neurologische Erkrankungen, wozu Migräne und chronifizierte Kopfschmerzen ja gehören (s. Ihre oben zitierte Arbeit S. 1 und 4). Damit ergibt sich kein Widerspruch zu den Studien von Brigo und Whitmarsh, wo die Plazebo-Wirkungen ja tatsächlich - wie bei Ihnen übrigens auch - gegen Schluss der Behandlungsphase deutlich nachlassen.

Doch finde ich es natürlich sehr gut, dass Sie im Schlussabschnitt Ihrer Arbeit den „Plazebo-Effekt“ ausdrücklich aufwerten. Wie sehr zu Unrecht wird er mit dem verächtlichen Namen „Plazebo“ (= Einbildungsfaktor) belegt! Denn Ein-Bildung, die Arbeit mit inneren Bildern, kann bei gezieltem und bewusstem Einsatz ein mächtiger Heilfaktor sein. Genau so auch Mesmers Heilmagnetismus, d.h. die Übertragung aetherisch-feinstofflicher Heilkräfte von einem Menschen auf den andern, oder auch nur die von einem bestimmten „energetisierten“ materiellen Umfeld (Praxisraum, Tempel, Mesmerisches Bacquet, Reichscher Orgon-Akumulatur, Amulett) ausgehenden

feinstofflichen Heilwirkungen. Man sollte also anstatt von Plazebo-Effekt besser von der feinstofflichen (d.h. nicht-(grob)materiellen) bzw. der lebensenergetischen Heilwirkung eines bestimmten Studienkontextes sprechen. In diesem Sinne muss der nach Ihrer obigen Arbeit zentrale „Plazebo“-Faktor in einer Studie entweder bewusst maximal eingesetzt oder minimalisiert werden. In jedem Fall ist er in Planung und Auswertung mit hoher Priorität zu berücksichtigen.

Nicht einverstanden bin ich natürlich mit Ihrer magisch-nichtkausalen Interpretation dieser feinstofflichen Effekte im Sinne der Kopenhagener Quantenphysik (S. 10 und 11 Ihrer obigen Arbeit). Mesmer hat uns hierfür eine viel menschengerechtere und damit sehr wahrscheinlich auch der Natur näher kommende Erklärung gegeben: Es handelt sich um sehr konkrete, lokal-kausal einwirkende „Ein-Flüsse“ im Sinne von Ätherströmungen (für Details s. das Schlusskapitel meines offenen Briefes).

12. Ich habe übrigens nie behauptet, dass die Whitmarsh-Studie komplett negativ war. Ich habe in der von mir durchgeführten und in meiner Habil-Arbeit enthaltenen Meta-Analyse die Effektstärke berechnet und weiß sehr wohl, dass sie einen kleinen positiven Effekt enthält, der auch eine gewisse signifikante Tendenz enthält. Der springende Punkt ist, dass die Effektstärke dramatisch gesunken ist.

Das habe ich von Ihnen bisher aber völlig anders verstanden. In ihrem Magie-Artikel schreiben Sie nämlich ganz explizit und ohne jede Einschränkung, dass die Whitmarsh-Studie „zu den neueren Arbeiten mit klar negativem Resultat“ gehöre (Walach 1999 S. 292). Damit stehen Sie offensichtlich in Widerspruch zu Ihrer eigenen Publikation.

13. In der norwegischen Studie wurde individualisierte Homöopathie eingesetzt. Die unterschiedlichen Placebo-Reaktionsraten in Studien als Beweis für irgendetwas einzusetzen, geschweige denn dafür, dass die Daten falsch sein müssen, ist Zeichen von wissenschaftlicher Ignoranz (aus meiner Sicht). Dies ist völlig unhaltbar, widerspricht der Literatur und ist aus meiner Sicht ein klarer Beleg für die Unfähigkeit oder Unwilligkeit, die Wirklichkeit so zu sehen, wie sie sich präsentiert: variabel.

Tja, auch da kann ich Ihnen nicht ganz beipflichten, auch wenn ich sicher einiges weniger von Statistiker-Latein verstehe als Sie. Aber wiederum möchte ich versuchen, den einfachen Menschenverstand des Praktikers sprechen zu lassen: Ich habe die Plazebo-Besserungsraten der Kopfschmerzfrequenz in der Brigo- und der Whitmarsh-Studie sowie zusätzlich auch noch in einer schulmedizinischen Migräne-Studie miteinander verglichen und festgestellt, dass diese Werte mit etwa 20%, 15% bzw. 14% sehr nahe beieinander liegen (S. 26). Bei Ihrer auch andere schwer chronifizierte Kopfschmerz-Typen umfassenden Studie lag die Plazebo-„Frequenzsenkung“ aus gut verständlichen Gründen (Näheres hierzu s. unten) mit nur etwa 6% wesentlich tiefer. Natürlich muss man sehr vorsichtig sein, wenn man mit den unterschiedlichen Plazebo-Ansprechraten verschiedener Studien argumentiert, die Vergleichbarkeit ist heikel und muss zumindest generell gegeben sein. Wenn nun aber innerhalb einer diagnostisch eng gefassten Gruppe von Studien (alles chronische Migränefälle) das Resultat von Straumsheim mit 30% Plazebo-Effekt sehr deutlich vom obigen Durchschnitt abweicht, ist dies m.E. erklärungsbedürftig.

Die Autoren der Straumsheim-Studie geben ja offen zu, dass ein unerklärter massiver Anstieg der Tagebuch-Anfallsfrequenz in der Vorlauf-Phase ein wesentlicher Mangel Ihrer Studie war (Straumsheim S. 7). Da ist also in der Vorlauf-Phase, zumindest so wie sie im Tagebuch erfasst wurde, sicher etwas schief gelaufen. Die durch einen erheblichen Störfaktor beeinflussten Patientenaufzeichnungen haben dann zusammen mit anderen Faktoren das zwar nur diskrete, aber doch ärztlich objektivierbare homöopathische Plus-Resultat der norwegischen Studie offensichtlich bis zur Unkenntlichkeit zu verschleiern vermocht (S. 25 - 26; vgl. hierzu auch den nachfolgenden Abschnitt).

Ich glaube also nicht, dass meine Argumentation in dieser Sache so völlig abwegig ist wie von Ihnen dargestellt.

14. Der springende Punkt der Straumsheim-Studie ist: Der eigentlich anerkannt sensitivste Parameter, das Tagebuch, ergibt einen paradox negativen Effekt, während ein unsensibler Parameter einen positiven Effekt zeigt. Auf jeden Fall wurde die a priori Hypothese nicht bestätigt. Schade drum, in der Tat.
15. Ihre Betonung des ärztlichen Urteils gegenüber dem Tagebuch ist zwar aus Ihrer Sicht verständlich, verkennt aber die Ergebnisse von etwa 2 Dekaden Forschung, mehreren hundert Studien zur Validität und Reliabilität solcher Daten und die Tatsache, dass sich die gesamte Beurteilung von Schmerzsyndromen und ihrer Besserung in allen Bereichen ausschließlich auf Patientenurteile stützt, weil allen mittlerweile klar ist, dass es nur einen Experten für Schmerz gibt: den Patienten selber. Und es gibt nur eine reliable Art, dies zu messen: möglichst häufig nachfragen. Dabei sieht man dann, dass Tagebücher, oder idealerweise mehrfach am Tage getätigte Abfragen andere Ergebnisse liefern als retrospektive Befragungen von Patienten über Fragebögen oder Arzturteile.

Ich stimme mit Ihnen völlig überein, dass der beste Experte für seine Symptome einzig und allein der Patient ist - Hahnemann lässt grüssen! Damit wird die Tagebuch-Erfassung zu einem zentralen Instrument für kontrollierte Studien, das streite ich nirgends ab. Aber ganz offensichtlich ist gerade bei Kopfschmerzpatienten das Tagebuch zumindest zu Beginn der Erfassungsphase ein heikles und störungsanfälliges Instrument, welches noch der Verbesserung bedarf (z.B. ähnlich wie von Ihnen vorgeschlagen durch noch besser supervidierte Vorlaufzeiten, vgl. hierzu auch die nachfolgende Kaffee-Kontroverse!). Wenn aber das Tagebuch bei Straumsheim (und noch mehr bei Ihnen!) derart hohe, bereits dem Endresultat der Studie entsprechende (oder dieses sogar noch deutlich übertreffende!) Anfangsschwankungen zeigt, ist eine ärztliche Parallelkontrolle oder zumindest eine ausführliche Diskussion dieses Tatbestandes in der Studienauswertung nicht nur sinnvoll, sondern unabdingbar. Beides fehlt in Ihrer Studie; Ihr Nachuntersucher Straumsheim hat hier - nicht zuletzt vielleicht auch aufgrund Ihrer paradoxen Resultate - seriöser gearbeitet.

Ich kritisiere also nicht das Instrument des Patienten-Tagebuchs an sich, sondern nur dessen mangelhafte Anwendung und Interpretation in konkreten Fällen.

16. Wir haben die Brigo-Studie als generelle Argumentationsgrundlage verwendet, nie als Vorbild. Daher auch die Ausweitung auf chronische Kopfschmerzen. Die Epidemiologie gibt uns hier Recht, und eigentlich die Pragmatik (siehe Witt et al. 2005) auch. Man kann darüber natürlich geteilter Meinung sein.

Grundsätzlich einverstanden. Ich möchte nur daran erinnern, dass Sie im Einleitungsabschnitt Ihrer Hauptpublikation als Begründung Ihrer Studie einzig und allein sowie auch detailliert die Arbeit von Brigo anführen. Deshalb ist es für mich nach wie vor schlecht nachvollziehbar, weshalb Sie deren Konzept noch wesentlich weitergehend verändert haben als Whitmarsh.

17. Ich habe noch nie abgestritten, dass unsere Patienten möglicherweise ein sehr schwieriges Kollektiv waren und auch nicht, dass dies eine mögliche Ursache für unser negatives Ergebnis war. Sicherlich haben die Ärzte dies so gesehen und wenn ich mich nicht täusche, ist auch ein entsprechender Hinweis in unserer Studienbeschreibung enthalten.

Meine Worte, s. oben unter Punkt 10, wo auch das entsprechende Zitat zu finden ist! Dort bezeichnen Sie Kopfschmerzen allerdings noch ohne ausdrückliche Relativierung als „potentiell einfaches Syndrom“ (Punkt 10).

18. Es ist allerdings falsch zu sagen, dass dadurch unsere Messung in irgendeiner Form gelitten hat. Denn wir haben nicht die Frequenz von Migräneanfällen als Parameter genommen, sondern die Häufigkeit, die Dauer und die Intensität von Kopfschmerzen, als drei gleichwertige Parameter. Auch das ist klar beschrieben.

19. Es ist richtig, dass die Schmerzintensität möglicherweise ein sensiblerer Parameter wäre. Deswegen haben wir ja die beiden anderen Parameter hinzugezogen. Der springende Punkt ist: In keinem Parameter gab es einen Unterschied. Es ist übrigens falsch zu sagen, wir hätten die Anfallsfrequenz als wichtigsten Parameter definiert. Wir hatten drei.

Wieso sprechen Sie dann schon im Summary Ihrer Hauptpublikation erstrangig und weitaus am ausführlichsten von der „headache-frequency“ (S. 29)? Zudem spielt es keine Rolle, ob Sie von Kopfschmerz-Frequenz oder -Häufigkeit sprechen: Beides ist für Ihre Studie nicht korrekt, da bei Ihrem relativ hohen Prozentsatz von Dauerschmerz-PatientInnen der Begriff des für Migräne typischen Anfallsschmerzes klinisch sehr heikel und in der statistischen Aufarbeitung irreführend wird. Ich habe ausführlich dargestellt, auf welcher fragwürdigen Weise Ihre DauerpatientInnen in Anfallskranke verwandelt wurden, um Ihre konzeptuell wesentlich andersartige Studie mit dem wichtigsten Parameter von Migräne-Untersuchungen vergleichbar zu machen (S. 27 - 29 und 37 - 38). - Natürlich haben Sie dann an anderer Stelle korrekter vom Patientenprozentsatz mit Kopfschmerzen pro Tag gesprochen, jedoch diesen schwierig durchschaubaren Parameter immer wieder mit den Frequenzangaben der Migränestudien in Beziehung gebracht.

Zudem wird bei Ihnen nirgends diskutiert, dass sich bei dieser statistischen Akrobatik aufgrund der erheblichen Anzahl von Dauerschmerz-PatientInnen geradezu schwindelerregend hohe durchschnittliche „Anfallsfrequenzen“ ergeben. Diese erreichen bei Ihrer sensitiveren Placebo-Gruppe (Näheres hierzu s. unten) schliesslich nicht weniger als etwa 18 „Anfälle“ pro Monat (S. 38)! - Auch die noch wichtigere, von Ihnen oben zumindest teilweise zugegebene Tatsache, dass der „Frequenz“-Parameter Ihrer Studie trotz seiner artifiziellen Anfangs-Überhöhung therapeutisch träger und schwächer reagiert als derjenige von Migränestudien, kommt in Ihrer Hauptpublikation nirgends zur Sprache. Die bereits erwähnte, im Vergleich zu den mehr als doppelt so hohen Placebo-Frequenzsenkungen der reinen Migränestudien wesentlich geringere Besserung von nur 6% ist eine logische Folge davon. Damit schwächt sich die Aussagekraft Ihrer Studie zumindest für diesen wichtigsten Migräne-Parameter bereits schon ganz erheblich ab.

20. Meine Aussagen beruhen auf einer Responder-Analyse. Wir haben die Responder mit den Nicht-Respondern verglichen und untersucht, in welchen Parametern sie sich unterscheiden. Interessanterweise waren die Responder eigentlich die komplizierteren Patienten: Sie hatten längere Krankheitsdauer und häufiger Doppeldiagnosen. Sie können es glauben oder bleiben lassen. Das ist ihr Problem. Die Daten und die Wirklichkeit waren in diesem Falle so gelagert. Ich entscheide mich im Zweifelsfall für die empirische Wirklichkeit und frage mich, was ich daraus für Schlussfolgerungen zu ziehen habe. Ein ideologisch motivierter Denker tut das Gegenteil. Mir scheint, Ihre Argumentationsstruktur ist ideologisch, deshalb sollte ich eigentlich aufhören zu schreiben und meine Zeit sinnvoller verwenden. Aber da ich auch immer an die Belehrbarkeit von Menschen glaube, unternehme ich halt noch weitere Dialogversuche.

Auf Ihre Responder-Analyse bin ich recht ausführlich eingegangen (S. 29). Ich wiederhole nur kurz: Nach Ihrer Statistik waren nur gerade 21 der insgesamt 98 PatientInnen der Münchner Studie ResponderInnen. Dies ist für eine statistische Aussage innerhalb dieser Gruppe bereits eine sehr kleine Zahl. Von diesem kleinen Kollektiv gehörte zudem sicher ein guter Teil zu den Placebo-PatientInnen, da dieses gesamthaft ja sogar leicht besser abschnitt als die Homöopathie-Gruppe. Wenn nun Ihr Patientenkollektiv zufällig z.B. eine auch nur sehr kleine Anzahl positiver Placebo-ResponderInnen mit überdurchschnittlich langer Leidensdauer aufwies, hätte man aus dieser kleinen Fallzahl bereits den statistischen Fehlschluss ziehen können, dass PatientInnen mit längerer Krankheitsdauer positiver auf die Therapie ansprechen würden! Eine derartige Aussage, möge Sie statistisch auch noch so korrekt abgeleitet sein, stünde natürlich in völligem Widerspruch zur alltäglichen Praxiserfahrung (vgl. hierzu auch den nachfolgenden Abschnitt).

21. Ihre Argumentation übersieht einen Punkt: Gegen die von Ihnen angeführten möglichen Zufallsschwankungen ist nichts zu sagen. Dies wird aber genau durch die

Irrtumswahrscheinlichkeit quantifiziert. Wir wissen, dass wir uns möglicherweise irren. Mit einer Wahrscheinlichkeit von 1 in tausend Fällen. Die Tatsache, dass der Effekt trotz der kleinen Fallzahl signifikant ist, zeigt, wie stark er ist. Basale Statistik. Sie sollten sich erst kundig machen, bevor Sie zu so starken Worten greifen.

Es ist mir wie schon oben erwähnt schleierhaft, wie Sie bei dieser kleinen Fallzahl auf eine so hohe statistische Sicherheit kommen können. In keiner Ihrer Publikationen sind nähere Angaben zu den diesbezüglichen Ausgangswerten zu finden. Dies habe ich bereits in meiner Arbeit ausdrücklich bedauert (S. 29), und Sie bleiben nun auch in Ihrer Antwort die statistischen Details schuldig. - In jedem Fall aber gilt wie gesagt: Wenn in einer wie auch immer gearteten Studie der Schluss gezogen wird, dass schwerer chronifizierte Kopfschmerzfälle besser auf die Therapie ansprechen würden als weniger langfristig erkrankte, oder wenn auch nur behauptet wird, dass der Chronifizierungsgrad von Kopfschmerzen für deren Behandelbarkeit irrelevant sei, kann dies nur auf einem statistischen Ausrutscher beruhen - und sei dessen theoretische Wahrscheinlichkeit auch noch so klein! Eine allein dieser Fragestellung gewidmete grössere Studie würde diese klinische Binsenwahrheit mit Sicherheit bestätigen.

22. Ihre Ausführungen zu unserem Randomisierungsprozess zeigt, dass sie weder die Grundlagen noch das Ziel dieses Vorgangs wirklich verstanden haben: Wir haben eine unbeschränkte Randomisation gewählt. Das Ziel dieses Verfahrens war es, erstens einen wirklich reinen Zufallsprozess zu erzeugen, was übrigens die wenigsten Studien tun, und zweitens die konfundierenden Variablen möglichst optimal zu balancieren. Dass uns das fast ganz gelungen ist, zeigt ein Blick auf die Tabelle der Ausgangswerte. Dass dies unserer Statistik überhaupt nichts anhat, ergibt sich schon daraus, dass wir gar keine gerechnet haben. Erst wenn wir einen Effekt in die vorhergesagte Richtung gesehen hätten, dann hätte uns die ungleiche Patientenzahl vielleicht in Schwierigkeiten gebracht, weil dann die statistische Mächtigkeit reduziert worden wäre, aber sicher nicht irgendetwas anderes. Der beobachtete Baseline-Unterschied in der Ausgangsmedikation ist zwar sichtbar, aber nicht signifikant. Ich habe an anderer Stelle ausführlich auf die Medikationsunterschiedshypothese geantwortet und wiederhole mich hier nicht (Walach, H. (2002). Response to Vithoulkas: Homeopathic fantasies about science, a metacritique. *Homeopathy*, 91, 35-39.)

Ihre Ausführungen zur Studienmedikation in Ihrer Antwort an Vithoulkas', dessen Kritik trotz einiger guter Ansätze leider zu oberflächlich geblieben ist, habe ich durchgelesen. Sie sprechen hier sehr ausführlich über die Relativierung des klassisch-homöopathischen Antidot-Denkens, womit ich wie auch mit anderem grundsätzlich einverstanden bin. Hahnemann wäre sicher der letzte gewesen, der sich einer offenen, ideologiefreien Diskussion seiner Thesen widersetzt hätte. Zu dem von Ihnen oben angesprochenen Baseline-Unterschied in der Medikation hingegen habe ich in dem Artikel allerdings kaum etwas gefunden, ausser der hiermit nur indirekt zusammenhängenden Feststellung, dass der Medikamentenkonsum während der Studie in den beiden Vergleichsgruppen kaum unterscheidbar gewesen sei (Ihre oben zitierte Publikation S. 37). Dies ist für mich aber kein so wichtiger Punkt.

Umso wichtiger ist mir dafür aber der Baseline-Unterschied im vorangegangenen Medikamentenkonsum: Wie Sie in Ihrer Hauptpublikation offener als oben zugeben (Walach 1997 S. 121), ist es zumindest auffällig, dass vor Studienbeginn 75% der Verum-PatientInnen allopathische Medikamente einnahmen, während es in der Vergleichsgruppe nur 58% waren. Diese wichtige Tatsache wird noch ergänzt durch den komplementären (von Ihnen aber offenbar nicht der Erwähnung wert gehaltenen) Umstand, dass vorgängig zur Studie in der Placebo-Gruppe fast doppelt so viele PatientInnen (27% vs. 14%) primär auf der psychischen Ebene ansetzenden Therapien angewandt hatten.

Die Erklärung hierfür liefert der wohl wichtigste, von Ihnen aber ebenfalls übersehene oder einfach unter den Tisch gewischten Randomisierungsartefakt Ihrer Studie: Aufgrund eines fast unglaublichen Zufalls mit einer Wahrscheinlichkeit von nur etwa 1:350 landeten sämtliche im kreativ-künstlerischen Bereich tätigen StudienteilnehmerInnen ausnahmslos

in Ihrer kleinen Plazebo-Gruppe, wo sie dann auch gleich die zweitgrösste Berufsgruppe stellen (S. 31, vgl. auch Graphik hier im Anhang, S. 21)! Da KünstlerInnen naturgemäss sensitiver sind - und auch sein müssen - als NormalbürgerInnen, ist es keineswegs verwunderlich, dass Ihre Plazebo-Gruppe auch einen psychotherapeutischen Ansatz einer nur symptomdämpfenden chemischen Medikation bevorzugte. Vor allem aber ist die ausgeprägtere psychosomatische Empfindlichkeit Ihrer Vergleichsgruppe sehr wahrscheinlich für die auffälligen Schwankungen Ihrer Plazebo-Resutate verantwortlich, welche Sie zugegebenermassen nicht erklären können (Walach 1997 S. 125). Ihre nachweisbar nüchternere Verum-Gruppe (S. 31) wies im Vergleich hierzu viel flachere Verlaufskurven auf.

Zudem wird, wie ich in meiner Publikation leider erst in einer Anmerkung angedeutet habe (S. 31), dieser Trend zu erhöhter Empfindlichkeit Ihrer Plazebo-Gruppe wahrscheinlich auch noch durch ein deutliches Überwiegen der medizinischen Hilfs- und Pflegeberufe (14% vs. nur 2% in der Verum-Gruppe!) wesentlich verstärkt. Von den 6 Angehörigen dieser Gruppe wurden immerhin 5 in die 37-köpfigen Vergleichsgruppe eingeteilt. Diese ist zwar deutlich weniger unwahrscheinlich als der erstgenannte Randomisierungsartefakt, aber trotzdem noch sehr auffällig (vgl. Graphik im Anhang, S. 21). Obwohl es ebenfalls etwas weniger eindeutig ist, dass KrankenpflegerInnen und PraxisassistentInnen deutlich sensitiver sein sollen/sind als der Durchschnitt, ist dies gerade im medizinischen Bereich nach meiner Erfahrung durchaus wahrscheinlich. Damit erhöht sich der Anteil psychosomatisch potentiell empfindlicherer StudienteilnehmerInnen in Ihrer Plazebo-Gruppe von 6 auf 11 bzw. auf ziemlich genau 30% (s. auch Graphik im Anhang, S. 21)!

23. Die Sistierung von Prophylaxe-Medikation fand übrigens schon vor Beginn der Behandlung statt, da dies ein Einschlusskriterium war. Deshalb ist das kein wirklich stichhaltiges Argument.

Ich glaube, dass Sie hier Ihr eigenes Protokoll nicht mehr ganz richtig im Kopf haben: Zumindest die definitive Umstellung der bisherigen allopathischen Medikation, auf deren Gesamtheit sich meine Argumentation natürlich bezieht, erfolgte ganz explizit erst mit Beginn der Therapiephase (Walach 1997 S. 124). Zu diesem Zeitpunkt erfolgte ja auch erst das ärztliche Erstgespräch, von welchem dann überhaupt erst die definitive Aufnahme in die Studie abhing. Zudem erwähnen Sie ausdrücklich, dass zumindest einige wenige PatientInnen sogar die prophylaktische Medikation bis zum Beginn der Therapiephase einnahmen und erst dann sistierten (Walach 1997 S. 124).

Damit wäre es einzig noch denkbar, dass einige PatientInnen auch ihre sonstige allopathische Medikation schon während des Vorlaufes an die Bedingungen der Therapiephase anzupassen begannen. Hierüber ist allerdings m.W. nichts dokumentiert. Deshalb habe ich diesen Umstand, welcher je nach Ausmass der vorweggenommenen Medikamentenreduktion ebenfalls zu dem deutlichen „Frequenz“-Anstieg Ihrer Vorlauf-Phase beigetragen haben könnte, in meiner Analyse bisher noch nicht berücksichtigt. Doch würde durch eine allenfalls notwendige Korrektur das Gesamtergebn meiner Analyse nicht beeinflusst: Mit Sicherheit erfolgte nämlich wie erwähnt die definitive und eingreifendste Umstellung der allopathischen Gesamtmedikation erst nach der ärztlichen Eintrittsuntersuchung und Beratung. Damit ist v.a. bei der vermehrt medikamentenorientierten Verum-Gruppe zumindest zu Beginn der Therapiephase nach wie vor mit mehr oder weniger starken Entzugs- und Umstellungssymptomen zu rechnen.

24. Das Argument bezüglich des Koffein-Entzuges ist natürlich bedenkenwert. Aber nicht schlagkräftig: Entzugssymptome dauern normalerweise maximal 6 Tage. Außerdem durften die Patienten auch auf andere Formen von Koffeinersatz (Tee, Cola) ausweichen. Und überhaupt haben die Patienten dies bereits zu Baselinebeginn getan.

Auch hierzu muss ich zumindest zwei grosse Fragezeichen setzen:

1. Seit Jahrzehnten versuche ich, bei meinen PatientInnen einen Stop zumindest des täglichen Konsums von Kaffee und Schwarztee zu erreichen, obwohl ich mir sicher bin, dass diese Drogen die homöopathischen Mittel keineswegs völlig zu antidotieren ver-

mögen. Hingegen bin ich von einer gewissen Bremswirkung und vor allem von einem generellen langfristigen Gesundheitsrisiko dieses Suchtmittels überzeugt. Bei eingefleischten KaffeetrinkerInnen und vor allem natürlich in Kombination mit Kopfschmerzen ist aber selbst ein moderater Entzug nur sehr schwer zu bewerkstelligen. Bei einigen habe ich dieses Ziel auch über Jahre hinweg nicht erreichen können, wobei allerdings nur mit kontinuierlicher Überzeugungsarbeit und nicht mit massivem Druck vorgegangen wurde.

Ihre dieser klinischen Erfahrung scheinbar widersprechende obige Behauptung, dass Kaffee-Entzugssymptome maximal nur etwa eine Woche dauern sollen (und damit implizit als harmlos zu betrachten sind), findet sich auch gleich zweimal in Ihrer ausführlichen Studiendokumentation (Walach 2000 S. 32 und S. 160). Ich bin nur deshalb bisher nicht näher darauf eingegangen, weil ich nicht noch ein weiteres, zumindest auf den ersten Blick eher nebensächlich erscheinendes Kritik-Thema anschneiden wollte. Nun aber habe ich Ihre hierzu zuerst angeführte generelle Quellenangabe (Walach 2000 S. 32, Silverman et al.) durchgelesen - und etwas überrascht nicht einmal einen Hinweis auf Ihre Aussage gefunden! Die betreffende Studie untersucht lediglich kurze zweitägige Entzugsphasen bei chronischen Kaffeetrinkern. Hierbei lassen sich bei den doppelblind untersuchten psychologischen Parametern erwartungsgemäss deutliche Entzugssymptome objektivieren; von einer längerfristigen Untersuchung fehlt aber, wenn ich mich nicht ganz schwer täusche, jede Spur. Haben Sie hier etwa die Quellenangabe verwechselt?

Eine zweite Quellenangabe zu diesem Thema, welche Sie beim zweiten Zitat zusätzlich noch anführen (Walach 2000 S. 160, Murray), brachte mich auch nicht viel weiter. Immerhin aber steht dort in einem Nebensatz zu lesen: „However, discontinuation of caffeine after seven days of 250 mg of caffeine three times a day did not lead to noticeable withdrawal symptoms“ (Murray S. 579). Ob Sie wohl diese Passage gemeint haben? - Diese kann sich bei genauer Lektüre aber nur auf eine lediglich 7-tägige Einnahmepériode beziehen, nach welcher keine Entzugssymptome beobachtet wurden. Dies sagt natürlich überhaupt nichts aus über die physische und psychische Entzugsdauer bei chronischem Konsum!

Um ganz sicher zu sein, habe ich hierzu auch noch die von Murray zu dieser Nebenbemerkung angegebene Quelle konsultiert (Robertson et al.: „Tolerance to the Humoral and Hemodynamic Effects of Caffeine in Man“, Journal of Clinical Investigation 1981, Bd. 67/4, S. 1111 - 17). Erwartungsgemäss bestätigt diese, dass die Studie nicht mit gewohnheitsmässigen Koffein-Konsumenten (S. 1111 und 1115) durchgeführt wurde. Hierbei stellten sich unter doppelblind-plazebokontrollierter Koffeinzufuhr über 7 Tage hinweg sehr schnell und deutlich die bekannten Effekte auf Kreislauf und Hormonsystem (Katecholamine, Renin) ein. Diese wurden aber schon nach kurzer Zeit gegenreguliert, und beim Absetzen zeigten sich auch keine nachweisbaren Nachschwankungen oder Entzugssymptome. Die Studie erfasste lediglich die obgenannten physiologischen Parameter, psychosomatische Aspekte wurden nicht berücksichtigt. Sie zeigt lediglich, dass Kaffee-Effekte sehr schnell korrigiert werden können und zumindest bei kurzfristiger Gabe keine physiologischen Nachwirkungen oder gar Entzugssymptome hinterlassen. Wie schon anfangs erwähnt, hat dies aber mit der Problematik eines Langzeit-Entzugs bei chronisch Abhängigen gar nichts zu tun, geschweige denn mit den hierbei bei chronischen Kopfschmerz-PatientInnen zu erwartenden Schwierigkeiten!

Doch hat mich die Sache weiterhin nicht in Ruhe gelassen, und schliesslich bin ich in einer sehr umfassenden neueren Übersichtsarbeit (Juliano LM, Griffiths RR, A critical review of caffeine withdrawal: Empirical validation of symptoms and signs, incidence, severity, and associated features., Psychopharmacology, Oct. 2004;176(1):1-29) endlich fündig geworden: Dort findet sich nämlich tatsächlich die Aussage, dass das akute Kaffee-Entzugssyndrom nach 12-24h Abstinenz seinen Anfang nimmt, nach 1 - 2 Tagen das Maximum erreicht und schliesslich nach 2 - 9 Tagen wieder abklingt (S. 1 der obigen Quelle). Bemerkenswerterweise sind Kopfschmerzen hierbei eindeutig das Hauptsymptom. Wahrscheinlich also beruht Ihre obige Aussage, dass Entzugssymptome nur wenige Tage dauern, auf einer der Quellen dieser grossen Meta-Analyse. In diesem Sinne ist sie also trotz mangelhafter Literaturangaben durchaus korrekt. Jedoch beziehen sich, wie mir der Leiter der Studie, Prof. R. Griffiths von der Universität Baltimore, schriftlich bestätigt hat,

diese Angaben erwartungsgemäss nur auf die Dauer der „severe objective signs“ des Kaffee-Entzugs, und: „Some individuals claim that feeling of fatigue and impaired concentration may continue for weeks or months; relapse to caffeine use is very common“ (Mail von Griffiths). Er berichtet auch, dass es bisher leider keine gute Studien zum Langzeit-Entzug von Kaffee gäbe. In seiner obgenannten Übersichtsarbeit ist aber immerhin eine Studie angeführt (Strain et al., Caffeine Dependence Syndrome, JAMA 272, Okt. 1994, S. 1043 - 48), laut welcher unter chronischen Koffein-Konsumenten durchaus auch Leute zu finden sind, auf welche die sonst nur für eine schwere Drogenabhängigkeit geltenden diagnostischen Voraussetzungen durchaus erfüllt sind. Für unsere Fragestellung sind vor allem die folgenden beiden Kriterien sehr eindrücklich, welche für die überwiegende Mehrzahl dieser Fälle von schwerwiegender Koffein-Abhängigkeit zutreffen (S. 1044 - 45 der obigen Quelle):

- Anhaltender Wunsch oder wiederholtes Fehlschlagen von Versuchen, abstinent zu werden oder den Konsum zu limitieren.

- Fortgesetzter Abusus trotz Kenntnis der Tatsache, dass hierdurch ein bestehendes oder immer wiederkehrendes physisches oder psychisches Gesundheitsproblem ausgelöst oder verschlechtert wird.

Diese Beobachtung entspricht durchaus auch meiner oben erwähnten jahrzehntelangen klinischen Erfahrung. - Fazit: Für chronische KonsumentInnen kann der Koffein-Entzug auch über längere Zeit hinweg ein erhebliches, ohne Hilfe von aussen manchmal sogar unlösbares Problem sein. Und dies gilt notabene bereits schon für Leute, die nicht chronisch an Migräne oder Kopfschmerzen leiden!

Und wie sieht es nun bei dieser letztgenannten Patientengruppe aus? - Griffiths teilte mir schon in seinem obgenannten Mail mit, dass bisher seines Wissens noch keinerlei Studien zum Koffein-Entzug bei chronischen Kopfschmerzen gemacht wurden. Dies bedauert auch der von ihm zu diesem Thema freundlicherweise extra noch einmal beigezogene Spezialist Barry Spencer, welcher mir ebenfalls persönlich geschrieben hat. Nach seiner Vermutung sind Migräne und Koffein-Abusus noch enger miteinander verbunden, als dies bisher angenommen wurde. Dementsprechend sind bei Entzug - wie auch schon in meinem Originalpapier kurz erwähnt (S. 17, Anm. 62) - nach einer anfänglich in den allermeisten Fällen zu erwartenden Verschlechterung mittel- und langfristig zumindest manchmal auch Besserungen der Kopfschmerz-Symptomatik möglich. Dadurch wird der potentielle Störeinfluss dieses Faktors auf eine längerfristige Kopfschmerzstudie natürlich noch unberechenbarer!

Die von Ihnen oben erwähnte Umstiegsmöglichkeit auf andere Koffeinprodukte ist vom grundsätzlichen klinisch-homöopathischen Standpunkt her, womit die ganze Kaffee-Umstellung ursprünglich ja begründet wurde, bereits einmal sehr heikel. Und mehr als eine gewisse Milderung der Entzugsproblematik ist neben potentiell neuen Umstellungssymptomen davon kaum zu erwarten. Dies beweist die im Schlussabschnitt dieses Kapitels ausgeführte Tatsache, dass Probleme mit der Kaffee-Abstinenz das wichtigste Ausstiegsmotiv aus der Therapiephase Ihrer Studie waren!

Für homöopathische Kopfschmerzstudien ergibt sich hieraus nur eine logische Konsequenz: Entweder man belässt - wie dies Brigo in weiser Voraussicht getan hat - den Faktor Kaffee während der Studie möglichst unverändert; oder man akzeptiert nur PatientInnen, welche seit mindestens etwa einem Jahr stabil abstinent sind! Die letztere Variante dürfte selbst bei Akzeptanz von Gelegenheitskonsum recht schwierig zu realisieren sein. Im ersteren Fall muss man aber den Kaffee als Konsumgut und allopathisch-phytotherapeutisches Heilmittel akzeptieren, genauso wie anderweitige diätetische Störfaktoren und sonstige Analgetika. Dies ist gemäss Brigos Resultaten und auch in der praktischen Alltagserfahrung ein durchaus gangbarer, aber oft steiniger Weg.

2. Auch Ihre zweite Aussage zu diesem Punkt, dass nämlich der Kaffee-Entzug in Ihrer Studie bereits während der Baseline-Periode vollzogen worden sei, bedarf gerade auch im Hinblick auf die erheblichen potentiellen Akut- und Langzeit-Wirkungen dieser Massnahme zumindest einer deutlichen Relativierung. Die PatientInnen hatten ja wie erwähnt zu diesem Zeitpunkt noch gar keinen beratenden Arztkontakt gehabt! Zudem ist es unklar, wie zwingend der Kaffee-Entzug bereits in dieser provisorischen Studienphase war: An einer Stelle Ihrer Dokumentation erklären Sie zwar, dass die PatientInnen schon in der

Vorlaufphase auf Kaffee verzichten mussten, relativieren dies aber in Gegensatz zu Ihrem obigen Statement umgehend: „Manche Patienten werden dies wohl erst nach endgültiger Aufnahme in die Studie getan haben“ (Walach 2000 S. 160). An anderer Stelle sprechen Sie dann sogar nur noch von einer Uebergangsphase: „Die Patienten wurden bereits zu Beginn der Baseline auf diese Massnahme hingewiesen und hatten Zeit, sich auf diese Veränderung einzustellen“ (Walach 2000 S. 32).

Es besteht also für den Kaffee-Entzug während Ihrer Vorlauf-Phase ähnlich wie für die allopathische Begleitmedikation offensichtlich eine gewisse Grauzone, welche aber für meine Analyse nicht relevant ist. Lediglich muss beim Kaffeekonsum noch deutlicher als bei der Begleitmedikation festgehalten werden, dass zusätzlich zum Tagebuch-Effekt vermutlich bereits auch Kaffee-Entzugssymptome für den massiven Baseline-Anstieg des Patientenprozentsatzes mit Kopfschmerzen mit verantwortlich waren. Dies werde ich in der Schlussfassung meiner Arbeit in diesem Sinne korrigieren.

Aufgrund des obigen ist aber mit Sicherheit davon auszugehen, dass selbst die ab Beginn des Vorlaufes abstinenter chronischen KaffeetrinkerInnen Ihres Studien-Kollektivs in der Therapiephase zumindest teilweise noch immer mit Problemen zu kämpfen hatten - von den erst später damit beginnenden ganz zu schweigen! Beweisend hierfür ist die bereits erwähnte Tatsache, dass bei mindestens der Hälfte der allerdings nur sehr wenigen Aussteiger aus der Therapiephase Ihrer Studie die Kaffeeabstinenz der ausschlaggebende Faktor war (S. 39) und wir die Klage eines durch Kaffee-Entzug verschlechterten Zustandes zumindest als Einzelfall auch unter den im Tagebuch spezifizierten Therapie-Nebenwirkungen ausdrücklich verzeichnet finden (S. 44).

25. Selbstverständlich sieht man trotz optimaler Randomisation immer einige Unterschiede, da eine vollständige Homogenisierung nur bei sehr großen Kollektiven von etwa 500 und mehr zu erwarten ist. Diese Zahlen waren für uns nicht realisierbar. Dennoch würde ich behaupten, dass wir gemessen an der kleinen Zahl von Patienten eine optimale Homogenisierung erreichen konnten. Die geringen Unterschiede, die verblieben sind, lassen sich sicher nicht ausschlichten, um das Ergebnis vollständig zu negieren. Wenn man schon so argumentiert, dann müsste man das Argument des multiplen Testens einführen und die Anzahl der möglichen Tests ins Feld führen. Dann würde Ihre Baseline-Auffälligkeit, die sie mit einer Wahrscheinlichkeit von $p = 0.0029$ benennen, etwa zu einer Wahrscheinlichkeit von $p = 0.14$ werden, da wir von etwa 50 Ausgangsvariablen und Tests ausgehen müssen.

Auch ich bin überzeugt, dass Ihr Spezialist Gaus und Sie sich um eine gute Randomisierung bemüht haben. Wie Sie aber ebenfalls erwähnen, kann hierbei trotzdem einmal ein relevanter Fehler passieren - zu dem man dann halt auch stehen muss! Ein solcher Fehler ist m.E. die völlig einseitige Einteilung der künstlerisch-kreativen (und partiell auch der medizinischen) Berufsgattung in die Plazebo-Gruppe (s. auch Punkt 22 und die Graphik im Anhang, S. 21). Ich glaube nicht, dass Sie hier ein statistischer Spagat retten kann: In Ihrer nur 37-köpfigen Plazebo-Gruppe stellen die 6 VertreterInnen aus dem künstlerisch-kreativen Tätigkeitsbereich mit einem Anteil von 16% nämlich gleich auch die zweitgrösste Berufsgruppe dar! Wenn also jedes sechste Mitglied der Plazebo-Gruppe zu der sensitivsten Berufsgruppe Ihrer Studie gehört, ist es gut denkbar, dass hierdurch das Verhalten der ganzen Gruppe wesentlich geprägt wird. Auf die für Ihre Studie zusätzlich sehr ungünstige Potenzierung dieser Dysbalance durch die Fehlverteilung der medizinischen Berufe habe ich bereits hingewiesen (s. Punkt 22 und Graphik im Anhang, S. 21).

Zur Illustration wollen wir aber nicht allein in trockener Statistik verharren, sondern ein bildhaft-lebendiges Gleichnis versuchen (s. Graphik im Anhang, S. 21): Stellen Sie sich vor, Sie hätten als Gärtner 98 Blumenzwiebeln (= die Gesamtzahl Ihrer PatientInnen) möglichst gleichmässig auf zwei Gartenbeete zu verteilen, von denen das eine fast doppelt so gross ist wie das andere (= Ihre beiden Vergleichsgruppen mit 61 bzw. 37 TeilnehmerInnen). Die Blumenzwiebeln sehen alle etwa gleich aus, werden aber ein gutes Dutzend verschiedene, sich in Form und Farbe unterscheidende Blumenarten in unterschiedlicher Anzahl pro Sorte hervorbringen (= Ihre 14 Berufsgruppen nach Gaus 1994 S. 16).

Wenn nun durch einen fast unglaublichen Zufall alle 6 blauen Lilien (= Ihre künstlerisch-kreative Berufsgruppe) und von den ebenfalls 6 violetten Trichterwinden (= die medizinischen Hilfsberufe) auch noch gleich 5 im kleineren Plazebo-Gartenbeet landen, wird man bei der Blüte die auffällige blau-violette Dominanz dieser beiden Blumensorten im kleineren Beet schon auf den ersten Blick erkennen (s. Graphik im Anhang, S. 21): Die verwandten Farben Blau und Violett entsprechen der für beide Gruppen anzunehmenden überdurchschnittlichen psychosomatischen Sensitivität, die eindeutig nüchterneren Berufssparten sind in Rottönen gehalten. Auf den zweiten Blick kann man dann leicht nachkontrollieren, dass im grossen Beet tatsächlich keine einzige der blauen Lilie zu finden ist, und von den violetten Trichterwinden nur eine einzige! - Ist da dem Gärtner nicht ein ganz eigenartiges Missgeschick passiert?

26. Wir haben sogar in den meisten Fällen längere Tagebuchvorlaufdaten und auch längere Nachbeobachtungen. Diese Auswertungen sind in meinem Buch enthalten. Sie unterscheiden sich nicht von den Aussagen, die die protokollgemäße Auswertung der offiziellen Publikation ergab. Dies ist wichtig, denn alle Argumente, dass unsere Beobachtung zu kurz war, die Vorbeobachtung zu kurz etc., sind aufgrund dieser zusätzlichen Informationen nicht haltbar.

Ich denke auch nicht, dass die zu kurze Studiendauer das Hauptargument gegen Ihre Studie ist, dagegen sprechen ja auch die zumindest partiell signifikant positiven Aussagen aller übrigen Vergleichsstudien von ähnlicher Dauer.

27. Man kann die Individualisierung so durchführen, wie wir sie gemacht haben, oder so, wie sie bei Brigo/Straumsheim gemacht wurde. Immer ist entscheidend: Wurde wirklich individualisiert und wurde das richtige Mittel getroffen? Ich gehe davon aus, dass unsere Ärzte ihr Bestes gegeben haben, um die richtige Wahl zu treffen. Ich gehe auch davon aus, dass sie fähig waren. Immerhin gehörte und gehört ihre Praxis zu den frequentiertesten in Deutschland. Was Ihre Argumentation übersieht: Die Arzneimittelwahl war immer die eines individuellen Arztes. Das Kollektiv entschied nur darüber, ob aufgrund der Falldarstellung die Arzneimittelwahl potenziell richtig war, und vor allem, ob der Fall klar genug war. Denn:

28. Das wichtigste Element wird gern übersehen: Wir haben nur die Patienten aufgenommen, bei denen die Ärzte sich sicher waren, dass eine von zwei ausgewählten Arzneien wirklich passt. Alle anderen wurden gar nicht eingeschlossen. Ich finde es auch typisch und bezeichnend für Ihre selektive Lesart, dass Ihnen anscheinend dieses entscheidende Designdetail entgangen ist.

Dieses Detail habe ich nicht übersehen. Unter anderem gerade auch deshalb glaube ich ja auch nicht, dass homöopathische Fehldiagnosen der wesentliche Grund für das negative Resultat Ihrer Studie sind. Die ungünstige Patientenauswahl, die von mir aufgeklärten Randomisierungsartefakte und das Übersehen von spezifischen Reaktionen der Verum-Gruppe sind viel wichtigere Negativfaktoren. Fehler von Seiten Ihrer ÄrztInnen bei der Anwendung der Q-Potenzen haben aber nach meiner Beurteilung die z.T. sehr ausgeprägten Therapiereaktionen der Verum-Gruppe wahrscheinlich wesentlich mit verursacht

- 29 a. Ihre Ausführungen zu Q-Potenzen in Ehren, aber diese Argumente finde ich albern. Unser Anliegen war ja nicht, die Wirksamkeit und Wirklichkeit immaterieller Potenzen zu beweisen, sondern die klinische Brauchbarkeit von klassischer Homöopathie zu untersuchen, so wie sie praktiziert wird.

So einfach können Sie den Kopf nicht aus der Schlinge ziehen. In dem von Ihrem Kollegen Gaus gemeinsam mit Ihnen publizierten Studienkonzept ist nämlich im Gegensatz zu Ihrer obigen Aussage die kritische Überprüfung von immateriellen Arzneiwirkungen ein zentrales

Anliegen: „Hochpotenzen enthalten nach naturwissenschaftlichem Verständnis von der Urtinktur und der „Wirksubstanz“ kein Molekül mehr. Die Homöopathie beansprucht jedoch eine Arzneimittelwirkung gerade der Hochpotenzen. Im Folgenden wird ein Studienprotokoll zum Nachweis der klassischen Homöopathie bei chronischen Kopfschmerzen vorgelegt...“ (Gaus 1992 S. 134). Auch für Brigos Migränestudie, welche Sie ja zugegebenermassen zumindest als Argumentationsgrundlage verwendeten (Punkt 16), war der Nachweis der immateriellen Arzneiwirkung ein Hauptanliegen (Brigo S. 99). Und im Gegensatz zu Ihnen wurde die italienische Untersuchung dann auch korrekt nur mit rein nicht-materiellen Hochpotenzen durchgeführt!

- 29 b. ... Und da gehört die Verwendung von Q-Potenzen eben mit dazu. Wenn man übrigens Q-Potenzen richtig herstellt, dann ist das Ausgangsmaterial zur C3 verrieben, hat also einen rechnerischen Gehalt von 10-6, wird dann bei der Q1 etwa auf einen rechnerischen Endgehalt von 10-10 bis -12 weiterverdünnt, etc. Wenn eine solche Verdünnung wirksam wäre, würde sich jeder Pharmazeut verwundert fragen. Abgesehen davon sind in den konventionell hergestellten C200 Potenzen, die in unserem Falle Korsakov-Potenzen waren, in den meisten Fällen noch molekulare Spuren nachweisbar. Nur bei Einglaspotenzen funktioniert die Algebra. Ich halte das aber eigentlich für einen komplett irrelevanten Nebenschauplatz.

Auch hier versuchen Sie, einen offensichtlichen Fehler zu zerreden, anstatt ihn offen zuzugeben. Ihre Falschmeldung, dass die Grenze zur immateriellen Arzneigabe bereits ab der Q 2 überschritten wird (S. 33), wird dadurch nicht richtiger. Und Hahnemann verwendete natürlich immer die Mehrglas-Methode (§ 270 des Organon). Ebenso unwiderlegt bleibt damit auch die penible Tatsache, dass sie in einer Studie, welche laut obigem Zitat (s. Punkt 29 a) klar als Hochpotenz-Nachweis konzipiert war, homöopathische Arzneidosen verwenden, welche teilweise noch im materiellen Bereich liegen. Aber trotzdem halte auch ich dies alles, wie schon im Originaltext betont, eher für einen Nebenschauplatz. Dieser ist aber kennzeichnend für die über weite Strecken mangelhafte klinisch-homöopathische Kompetenz ihrer Studie.

30. Die Q-Potenzen wurden nach Herstellerangaben eingenommen und verdünnt. Dazu gehört, dass die Tropfflaschen nochmals in einem Glas Wasser verdünnt werden, etwa 3- 5 Tropfen in einem Glas Wasser. Allein dadurch schon erübrigt sich die Haarspalterei ziemlich. Die Anwendung, die Indikation und die Häufigkeit der Gabe lag den Ärzten frei und entsprach deren normaler Praxis.

Es wäre besser, Sie würden das Organon zitieren anstatt Ihre Herstellerfirma, welche Hahnemanns klare diesbezügliche Vorschriften teilweise missachtet. Zudem steht Ihre obige Angabe in Widerspruch zu Ihrer Originalpublikation, wo Sie in Ergänzung zum oben angeführten Verdünnungsprozedere ausdrücklich festhalten: „Manchmal wird die Arznei auch direkt verabreicht“ (Walach 2000 S. 33). Da Sie diese Angabe nicht als irrtümlich bezeichnen, bleibt meine Feststellung unwiderlegt, dass die Q-Potenzen Ihrer Studie teilweise auch unverdünnt verabreicht wurden (S. 34 und 35). - Doch nicht nur in unverdünnter, sondern auch noch in um Größenordnungen stärkerer Verdünnung als von Ihnen beschrieben können Q-Potenzen massive Reaktionen auslösen, was eine engmaschige Kontrolle erforderlich macht (s. hierzu auch den nachfolgenden Abschnitt 31). Dies hat schon Hahnemann beschrieben, und diese erstaunliche Tatsache entspricht auch meiner Erfahrung. Deshalb ist und bleibt es durchaus wahrscheinlich, dass bei der in Ihrer Studie praktizierten Verordnungsweise (für weitere Details hierzu s. unter 32) und der ebenfalls nicht der Hahnemannschen Praxis entsprechenden, zu weitmaschigen Kontrollfrequenz zumindest die Q-Potenz-PatientInnen Ihrer Verum-Gruppe erhebliche und vor allem auch lange anhaltende Therapiereaktionen durchmachten.

31. Es ist falsch zu sagen, die Patienten und Ärzte hätten die Modulationsfähigkeit der Q-Potenzen nicht genutzt. Die Ärzte gaben den Patienten präzise Anweisungen mit und die Patienten konnten sich melden, falls sie Fragen oder Unklarheiten hatten.

Wie ich aber gezeigt habe, wurde diese Möglichkeit deutlich weniger genutzt, als es bei Anwendung der Q-Potenzen gemäss Hahnemannscher Praxis bei einem vergleichbaren Patientengut angezeigt gewesen wäre (S. 35). Dies lag wohl kaum an den PatientInnen, sondern an deren Instruktion. Bei täglicher Symptomatik verlange ich auch bei chronischen PatientInnen eine Rückmeldung nach 4 - 5 Tagen, wenn keine Besserungstendenz eintritt, und bei Therapiereaktionen immer eine sofortige Kontaktaufnahme.

32. Die beklagte mangelnde Verordnungsdynamik kann und will ich nicht kommentieren. Ich gehe davon aus, dass dies damit zusammenhängt, dass durch den Einschluss von Patienten, bei denen sich die Ärzte sicher waren, dass die Arznei stimmt, die Tendenz gegeben war, einmal möglichst lange abzuwarten, was ja durchaus mit dem klassischen Vorgehen kompatibel ist. Was sicher richtig ist, ist die Tatsache, dass die Ärzte sich auf eine eher konservativ-vorsichtige Arzneimittelwahl gestützt haben.

Der Begriff des Abwartens ist hier etwas unklar: Bei der Kentschen Technik bedeutet dies ja, dass nach der Einzelgabe ohne weitere Mitteleinnahme abgewartet wird, was durchaus dem klassischen Vorgehen entspricht. Sie beziehen den Begriff aber offensichtlich auf die von mir kritisierte mangelnde Verordnungsdynamik bei den Q-Potenzen. Diese wurden aber bei Ihren PatientInnen, wie mir einer Ihrer Studienärzte ausdrücklich bestätigt hat, bei in der Regel täglicher Gabe zumindest in einzelnen Fällen über das ganze Kontroll-Intervall von 6 Wochen durchgezogen (S. 35)! Diese Art des „Abwartens“ mit demselben Mittel entspricht nun aber keinesfalls dem klassischen Vorgehen nach Hahnemann, zumal wenn sich keine Besserung zeigt. Wenn man bei fehlender Heilwirkung in dieser Weise über sechs Wochen beim gleichen Mittel in gleicher Dosierung verharrt, verpasst man bestenfalls ohne weiteren Schaden die bei der Q-Potenztechnik sonst durchaus gegebene Chance, durch eine rechtzeitige Korrektur von Mittelwahl und/oder Dosierung den Heilungsprozess mit minimalem Zeitverlust doch noch in Gang zu setzen. Schlimmstenfalls aber drückt man die PatientInnen durch die unkontrolliert wiederholten Mittelgaben in eine schleichende Arzneimittelprüfsymptomatik, welche gerade bei nur subakuter Verschlechterung oder bei lediglich geringgradigen neuen Symptomen nur allzu oft nicht erkannt wird. Letzteres kann sehr ernsthafte therapeutische Probleme nach sich ziehen. Meine detaillierte Analyse der Therapiereaktionen weist darauf hin, dass dieser Fehler in Ihrer Studie zumindest gelegentlich vorgekommen sein dürfte.

- 33 a. Es ist völlig verfehlt, nur auf die Häufigkeit von Kopfschmerzen abzuheben. Das Interessante ist: Alle Parameter zeigen die gleiche Tendenz. Sie können sich auch die gesamte Gedankenakrobatik sparen: Der verblüffendste Befund ist, dass sich eben überhaupt kein Unterschied zeigt, im Gegenteil, dass Placebo sogar leicht besser ist als Verum.

Ich musste in meiner Analyse zuerst recht ausführlich aufzeigen, dass der wichtigste Migräneparameter „Anfallsfrequenz“ in Ihrer Studie fragwürdig wird und auch mit Sicherheit wesentlich schwächer reagiert als in den Vergleichsstudien. Diese in Ihrer Statistik geschickt getarnte und in Ihrer Studienanalyse völlig ausgeklammerte Tatsache war bisher nicht bekannt. Entsprechendes gilt aber auch für Ihren Schmerzintensitäts-Parameter, welchen ich entgegen Ihrem obigen Vorwurf ebenfalls genau untersucht habe. Wie in Ihrer Analyse leider ebenfalls unerwähnt bleibt, zeigt dieser Parameter verglichen mit den Migränestudien nun im Gegensatz zur „Anfallshäufigkeit“ ganz abnorm tiefe Werte, welche aber trotzdem deutlicher (allerdings auch deutlicher paradox) reagiert haben (S. 47ff).

- 33 b. Man kann natürlich, wie Vithoulkas, das Ganze als ein Verschlimmerungsphänomen abtun. Ich halte dies für unlauter, zumal wir auch aus den Nachbeobachtungen, der Langzeitstudie und den Einzelfallverläufen keinerlei Hinweise darauf haben, dass die Studie etwas anderes abbildet.

Ich verstehe nicht ganz genau, was Sie hier meinen. Aber in jedem Fall habe ich ja bereits festgehalten, dass Vithoukas' Kritik auch nach meiner Beurteilung trotz guter Ansätze zu sehr an der Oberfläche bleibt. Die von mir wesentlich detaillierter analysierten Verum-Verschlechterungen (s. den nächsten Punkt) sind zudem nur einer der drei Grundpfeiler meiner Kritik.

34. Umgekehrt müssten dann aber die teilweise sehr spezifischen Nebenwirkungssymptome der anderen Gruppe als Placebosymptome interpretiert werden. Sie können es drehen und wenden wie Sie wollen: Das Entscheidende ist, dass Sie spezifische und unspezifische Symptome in beiden Gruppen gleichermaßen finden.

Dies ist eine recht fadenscheinige Antwort auf meine differenzierte Analyse der Therapiereaktionen. Nach Aufdeckung Ihrer fast vollständigen diesbezüglichen Datenverwechslung (!), zu welcher Sie bezeichnenderweise keinen Kommentar abgeben, habe ich die immerhin 47 detaillierten Therapiereaktionen beider Vergleichsgruppen nach klassisch-homöopathischen Kriterien in 7 hierarchisch gegliederte Qualitätsgruppen eingeordnet und hierbei bei den hochwertigen Symptomen ein deutliches Überwiegen der Verum-Gruppe gefunden (S. 41 - 46). Natürlich könnte man diese Gruppeneinteilung der Therapiereaktionen auch in etwas anderer Weise vornehmen, und die einzelnen Gruppenzahlen sind wirklich nur klein. Jedoch wird wohl immer etwa dieses Resultat herauskommen. Wenn Sie nun aber trotzdem auf Ihrer obigen gegenteiligen Position beharren wollen, müssten Sie dies zumindest ansatzweise durch eine konkrete Alternative belegen können. Dies geschieht aber jetzt ebenso wenig wie zuvor.

Gar keinen Kommentar geben Sie dann wiederum zu der fast noch auffälligeren und m.E. nun sogar auch statistisch signifikanten Tatsache, dass die Therapiereaktionen der Verum-Gruppe fast ausschliesslich in der ersten Therapiephase Ihrer Studie auftreten. Dieses homöopathietypische Muster, welches in Ihrer Darstellung wiederum nur versteckt erscheint, weist einen wesentlichen Anteil der Verum-Reaktionen nochmals als echte homöopathische Mittelreaktionen aus. Wo bleibt da Ihre Antwort?

35. Es ist durchaus denkbar, dass wir in den 3 Monaten Therapie nur Verschlimmerung abgebildet haben. Dagegen sprechen aus meiner Sicht folgende Beobachtungen und Daten: Es gibt keine einzige offene Beobachtungsstudie, die annähernd plausibel machen würde, dass man in einem Migränekollektiv mit korrekter Behandlung *ausschließlich* 12 Wochen lang Verschlimmerungen sieht. Die wenigen Daten, die es zu homöopathischen Erstverschlimmerungen gibt, legen nahe, dass diese nach 10, 14, allerlängstens 28 Tagen vorbei sind. Was wir erwartet hätten, ist, dass sich wenigstens gegen Ende der Studie ein Trend gezeigt hätte. Ein wirklich reagierender Patient in der Verum-Gruppe oder deutlich gebesserte hätten genügt, um einen statistisch signifikanten Effekt zu erreichen. Der wirklich spannende Befund ist, dass genau dies nicht, aber auch nicht ansatzweise, zu beobachten war. Es gab nämlich wirklich dramatische Verbesserungen. Dummerweise waren die alle in der Placebogruppe. Die einzige Patientin, die komplett beschwerdefrei wurde, war in der Placebogruppe. Es ist genau diese Konfundierung der Verhältnisse, die das Ergebnis so erstaunlich macht.

Ich habe nie behauptet, dass man bei Ihnen ausschliesslich nur Verschlimmerungen sehen würde. Ihre obigen Daten zur Dauer von Erstverschlimmerungen mögen, wenn auch eine homöopathische Reaktion in Ausnahmefällen nach meiner Erfahrung noch wesentlich dauern kann, als Faustregel durchaus richtig sein. Sie übersehen aber, dass bei der angewandten Dosierungsweise der Q-Potenzen (kontinuierliche Einnahme über das ganze Beobachtungsintervall hinweg!) homöopathische Mittelreaktionen natürlich ohne weiteres auch bis zum Studienende und darüber hinaus andauern konnten. - Nach meiner Analyse haben aber - wie schon mehrfach erwähnt - neben den homöopathischen Therapiereaktionen auch Randomisierungsartefakte sowie medikamentöse Entzugs- und Umstellungsreaktionen mindestens ebenso wesentlich zur völligen Maskierung bzw. paradoxen Umkehr des wahrscheinlich auch bei Ihnen vorhandenen diskreten

homöopathischen Heileffektes beigetragen (s. hierzu auch die zusammenfassende Analyse des Studienverlaufes am Schluss dieser Replik).

36. Man kann nun eben, wie Sie es tun, viel Aufwand ins Wegerklären investieren. Mit viel Mühe wird das immer gelingen. Man übersieht dabei, dass man der Phänomenologie Gewalt antut. Vor allem, wenn man die Daten auf einen breiteren Hintergrund tut, und wenn man die Daten aus anderen Forschungsbereichen und Studien mitberücksichtigt.

Dieser Passus scheint mir inhaltlich nicht ganz klar, aber ich glaube zu verstehen, was Sie meinen: Das mit dem breiteren Hintergrund ist ja sehr wahrscheinlich positiv gemeint. Da bin ich durchaus einverstanden, möchte den breiteren Hintergrund aber neben den von Ihnen zu Recht genannten Bereichen in erster Linie als erweiterte medizinische und klinisch-homöopathische Fachkompetenz verstanden wissen.

37. Sie als Praktiker können gar nicht anders, als sich vor dem Rutsch Ihres Weltbildes zu sichern. Soll mir recht sein. Dann bleiben Sie aber bitte bei Ihrem Leisten und wildern nicht in Gegenden, von denen Sie vergleichsweise wenig Ahnung haben. Ich komme ja auch nicht und pfusche Ihnen in Ihr Handwerk und sage Ihnen, wie Sie Ihre Patienten zu behandeln haben. In diesem Sinne kommentiere ich Ihre Aussagen zu den anderen Aspekten nicht mehr. Im Gegensatz zu Ihnen will ich niemand von meiner Sicht der Dinge überzeugen. Ich stelle meine Erkenntnisse zur Diskussion und bin gerne bereit, mit jedem und jederzeit über divergierende Ansichten, vor allem aber über divergierende Daten zu reden. Aus Ihren Ausführungen habe ich keinen einzigen Punkt entnommen, der mich wirklich dazu verführen könnte, Ihre Ansicht zu übernehmen. Ich habe vielmehr vielfältige Bemühungen gesehen, die Augen vor der Wirklichkeit zu verschließen. Dies ist, wie gesagt, Ihr gutes Recht. Aber stellen Sie dies bitte um Himmels willen nicht als Wissenschaft, und schon gar nicht als Wahrheit dar. Ich habe vielmehr wahrgenommen, dass Sie Ihre Wahrnehmung willentlich, unwissentlich, absichtlich oder unbewusst auf einen vergleichsweise kleinen Ausschnitt der Wirklichkeit begrenzen. Sie haben z.B. den ganzen Diskussionsprozess um unsere Studie nicht zur Kenntnis genommen, auch nicht die anderen Daten, die ich und andere produziert haben. Auch nicht meine weiterführenden Überlegungen. Müssen Sie auch nicht. Aber dann treten Sie doch bitte mit ein klein wenig mehr Bescheidenheit auf. Fände ich jedenfalls angemessener. Ich bin gerne bereit zu weiterem Dialog. Das wird aber das letzte Mal sein, dass ich mir ein 60-seitiges Elaborat aus Ihrer Feder durchsehe und kommentiere. Alles Weitere ließe sich nur in einem wirklichen und konstruktiven Dialog klären. Nach allem, was ich gelesen habe, sind Sie daran anscheinend nicht interessiert, sondern nur am Beweis der Richtigkeit Ihres Standpunktes. An solchen Diskussionen bin nun ich wiederum nicht interessiert, weil sie nicht weiterführen, weil sie nichts zu neuer Erkenntnis beitragen und weil sie allen fundamentalen Prinzipien zuwider sind, die auch für Hahnemann wichtig waren: sich nämlich aufgrund der Erfahrung eines Besseren belehren zu lassen und die Theorie entsprechend anzupassen. So funktioniert nämlich Wissenschaft. Nicht anders.

Das scheint mir alles recht ideologisch gefärbt. Entscheidend aber ist die Frage: Wer hat die besseren Argumente? Wenn meine in dieser Diskussion nun nochmals neu dargelegten und punktuell auch vertieften Kritikansätze nicht besser widerlegt werden können, bleibt es dabei: Die Analyse Ihrer Studie muss neu geschrieben werden!

Zum Abschluss fasse ich meine Interpretation des Studienverlaufes nochmals kurz zusammen:

a) Verlauf bei der Placebo-Gruppe

Gegen Ende der Vorlaufphase erreicht die für den Tagebuch-Effekt und möglicherweise auch für den bereits zu diesem Zeitpunkt beginnenden Kaffee-Entzug sensitivere Placebo-Gruppe in allen drei relevanten Parametern einen gegenüber Verum deutlich erhöhten Ausgangswert. Dieser geht erst kurz vor Beginn der Therapiephase wahrscheinlich infolge

des relativ schnellen Abklingens des Tagebucheffektes, wie er uns von der Glasgower-Studie her bekannt ist (S. 21), in eine leichte Besserungstendenz über. Diese Tendenz verstärkt sich noch deutlich mit Beginn der Therapiephase, in erster Linie nun aufgrund des Plazebo-Effektes, der in dieser Gruppe ja wahrscheinlich ebenfalls etwas deutlicher wahrgenommen wird. Medikamentös bedingte Entzugs- und Umstellungsreaktionen, welche zu Beginn der Therapiephase auch bei dem weniger arzneiorientierten Plazebo-Kollektiv als negative Einflüsse zu erwarten wären, treten offensichtlich höchstens als Bremsfaktoren dieser deutlichen Besserungstendenz in Erscheinung. Erst mit dem zu erwartendem Abklingen des Plazebo-Effektes gegen Ende der Therapiephase zeigt sich wieder eine deutliche Verschlechterung. Da diese aber die überhöhten Ausgangswerte nicht erreicht, ergibt sich im Endresultat für Plazebo trotzdem eine leichte scheinbare Verbesserung. Diese ist aber wesentlich geringer als bei reinen Migräne-Studien, was sich durch das sehr ungünstige Studiensetting mit hohem Durchschnittsalter und dem Einbezug schwer chronifizierter Dauerschmerz-PatientInnen erklären lässt.

b) Verlauf bei der Verum-Gruppe

Die tendenziell nüchternere Verum-Gruppe zeigt hingegen in der Vorlaufphase einen fast stationären Verlauf, der sich dann aber überraschenderweise auch in der Therapiephase kaum merklich ändert. Offensichtlich wird hier der Plazebo-Effekt, der bei dieser Gruppe - wenn auch in etwas schwächerem Ausmass - ebenfalls zu erwarten ist, durch die deutlich nachweisbaren homöopathischen Therapiereaktionen und wahrscheinlich auch durch medikamentöse Entzugs- und Umstellungsreaktionen, welche in diesem stärker arzneiorientierten Kollektiv vermehrt zu erwarten sind, gerade etwa kompensiert. Dies gilt auch für den vor allem in der zweiten Therapieperiode zu erwartenden, allerdings sicher nur sehr schwachen homöopathischen Heileffekt. Es resultiert deshalb für Verum nur eine minimale statistische Besserung, welche von der Plazebo-Gruppe sogar recht deutlich übertroffen wird.

Mit freundlichen Grüßen

Hanspeter Seiler

Dr. med. Hanspeter Seiler
FMH Allgemeine Medizin
ehem. Chefarzt der Bircher-Benner-Klinik Zürich
Im Dörfli
CH-8124 Maur

Telefon: 044 - 980 47 80
Telefax: 044 - 980 42 69
E-Mail: praxisseiler@bluewin.ch

Verteilung der Berufsgruppen auf Verum und Plazebo

