In meinem persönlichen Blog habe ich über eine Tagung berichtet, die sich mit dem sog. Decline-Effekt oder Absinkungseffekt beschäftigte. Dahinter verbirgt sich folgendes Phänomen: Oft zeigt sich am Anfang einer Forschungsreihe ein besonders deutlicher oder spannender Effekt, zumindest legen Publikationen das nahe. Wenn andere diese Ergebnisse wiederholen wollen, finden sie oft weniger deutliche Effekte. Oft nehmen die Effekte ab oder lassen sich nicht bestätigen. Solche Absinkungs-Effekte sind in der gesamten biomedizinischen, aber auch in der psychologischen, biologischen und vor allem parapsychologischen Forschung zu beobachten und lange bekannt.

Ursachen für den Decline-Effect
Dies kommt – vor allem, aber nicht ausschließlich – von folgendem Phänomen: Oft führen Grundlagenforscher – vor allem diese, denn sie arbeiten mit relativ rasch und einfach durchzuführenden Assays und Experimenten – in neuen Gebieten ein paar Testversuche durch. Die fruchten nicht und werden in die Tonne gesteckt. Dann bastelt wer ein bisschen herum und plötzlich kommt ein signifikantes Ergebnis heraus. Das wird publiziert. Die negativen werden natürlich nicht erwähnt. Wen interessiert das schon? Denn die Zeitschriften, die Editoren, die Wissenschaftler, überhaupt die Öffentlichkeit ist an positiven Befunden interessiert, nicht an negativen. Nun ist also ein neues Phänomen geboren. Weil es neu und vielleicht sogar spektakulär ist, wird es in den bekannten Zeitschriften mit hohem Impact veröffentlicht. Denn deren Geschäft ist es, solche neuen, aufwühlenden Befunde unter die Leute zu bringen. Nun kommen ein paar kritische Geister, lesen das und machen es nach. Manche ohne Erfolg. Sie denken sich dann: „wir haben vielleicht was falsch gemacht“ und stecken das negative Ergebnis in die Tonne. Oder, wenn sie hartnäckiger sind, glauben sie nicht daran, dass das ursprünglich publizierte Ergebnis richtig war, führen mehrere Replikationen und Abwandlungen durch und versuchen dann ihren negativen Befund zu publizieren. Das wird mit Sicherheit schwieriger werden, als den ursprünglich positiven Befund zu publizieren. Vielleicht müssen sie mehrere Journals anschreiben, ihren Text aufgrund von kritischen Gutachten überarbeiten, noch ein paar Experimente nachliefern. Denn es ist wahrscheinlich, dass Freunde, Bekannte oder Kollegen der ursprünglichen Forschungsgruppe, wenn nicht gar Mitglieder dieser Gruppe selbst zu den Gutachtern der negativen Replikationsstudie gehören. Also dauert es, bis die negativen Befunde publiziert sind, wenn sie überhaupt das Licht der Öffentlichkeit erblicken.

Und so werden Mythen gebildet
Die ursprünglich positiven Befunde schaffen es ins Bewusstsein der Öffentlichkeit: Selektive Serotonin Reuptake Inhibitoren sind wirksam zur Behandlung der Depression! Sagt man dann. Bis nach Jahrzehnten dann auch die negativen Studien bekannt werden und man sieht: so wirksam sind sie nun auch wieder nicht. Bis dann ist der Mythos bereits geboren. Oder: Aufmerksamkeits-Defizit (ADHS) ist eine Gehirnkrankheit mit einem deutlichen Problem in den Dopamintransportern der Basalganglien, wird verkündet. Naheliegend, dass

das nur pharmakologisch zu beheben ist. Bis dann die entsprechenden Nachfolgestudien kommen, die den ursprünglichen Befund nicht bestätigen.

Aus genau diesem Grund ist es methodisch wichtig sich genau zu überlegen, an welcher Stelle des Forschungsprozesses sich eine Studie befindet. Ist sie die allererste, die einen neuen Befund behauptet? Dann ist Vorsicht geboten. Ist es eine Studie, die bereits vorliegende Daten repliziert? Dann ist es wichtig, sich die Effektgröße anzusehen. Ist sie etwa so groß wie die ursprüngliche? Dann ist das Ergebnis robust. Ist sie wesentlich niedriger? Dann überschätzt vermutlich der ursprüngliche Befund den Effekt. Aus genau dem Grund sind auch große Studien sicherer und Meta-Analysen noch sicherer in der Schätzung von Effekten. Aber auch diese können nicht darüber hinwegtäuschen dass es ein Problem dann gibt, wenn anfänglich negative Befunde unterschlagen wurden, oder wenn spätere negative Befunde mit großer Verzögerung oder gar nicht publiziert werden. Dann wird ein Effekt suggeriert, wo gar keiner vorhanden ist.

Genau das hat Ioannidis (1) schon vor einiger Zeit gezeigt und damit eine heftige Debatte losgetreten. Er hat nämlich behauptet, die meisten publizierten Forschungsergebnisse seien falsch, genau aus dem hier oben beschriebenen Grund.

Repräsentanz wissenschaftlicher Ergebnisse in den Medien
Nun ist in zwei aktuellen Artikeln dieses Thema wieder aufgegriffen worden, aber mit zwei sehr beunruhigenden Zungenschlägen (2, 3). Gonon und Kollegen (2) zeigen in ihrer Studie, dass genau dieser Vorgang auch den öffentlichen Diskurs dominiert. Sie verwenden das Beispiel von ADHS, suchen die sog. „Top 10“ der Studien aus, über die am meisten berichtet wurde in der öffentlichen Presse und verfolgen ihr Schicksal. Alle diese „Top 10“ Studien hatten spektakuläre neue Berichte über „Fortschritte“ zum Gegenstand, die die Wissenschaft bei ADHS angeblich gemacht hatte. Verfolgt man die Berichte weiter, so blieb von diesen 10 Fortschrittsmeldungen gerade mal eine stabil. Die anderen wurden entweder später widerlegt, oder substanziell abgeschwächt. Was nun das Besorgniserregende ist: die Presse berichtete nur über die anfängliche Euphorie ausführlich. Die Nachfolgestudien wurden kaum mit Aufmerksamkeit bedacht. Sie erschienen ja auch in weniger hochkarätigen Zeitschriften. Und in vielen Fällen geistert die anfängliche positive Meinung noch immer in den Köpfen der Öffentlichkeit herum, obwohl sie schon längst widerlegt wurde. Nur keiner hat’s gemerkt, weil die Presse dies nicht mehr berichtet. Ist ja auch peinlich, wenn man seine eigene Euphorie revidieren muss. Ich empfehle allen, die online einsehbare Studie selber zu lesen bzw. sich die darin enthaltenen Grafiken anzusehen: Decline-Effekte in Hülle und Fülle und vom Feinsten. Das stimmt nicht sehr zuversichtlich, was die Mainstream-Haltung zur Behandlung von ADHS mit Ritalin angeht.

Die zweite Studie (3), die ein verwandtes Thema behandelt zeigt, dass unsere Presse nicht

sonderlich gut darin ist, Fehlinterpretationen zu entdecken, die Autoren ihren Studien mitgeben, wenn nicht das gewünschte Ergebnis herausgekommen ist. Die Autoren haben fast 500 Pressemeldungen von 70 randomisierten Studien analysiert. In knapp der Hälfte der Studien fanden sich ins Rosarot verzerrende Darstellungen der Studienergebnisse im Abstract oder im Text der Studie. Die Autoren ließen die Daten besser und robuster erscheinen, als sie in Wirklichkeit waren. Dadurch entsteht der gleiche Effekt, wenn die Presse diese Meinung so transportiert. Man meint, man hätte ein positives Ergebnis gefunden, wo in Wirklichkeit gar keines da ist. Und siehe da: die angeblich so kritischen Journalisten der Zeitungen waren offenbar außerstande, den sogenannten „Spin", der sich aus einer allzu wohlwollenden Interpretation der Studienergebnisse ergibt zu entdecken und transportierten ihn weiter in ihren Meldungen. In einer Regressionsanalyse war der einzige Prädiktor, der vorhersagen konnte, ob in einer Pressemeldung „Spin" auftauchte, also ein positiver Anstrich eines ansonsten gar nicht so spektakulären Ergebnisses, die Frage, ob ein solcher „Spin" in den Schlußfolgerungen des Abstracts der entsprechenden Studie vorhanden war.

Was folgert der kritische Leser daraus? Richtig: Journalisten sind viel zu beschäftigt, um eine Studie genau zu lesen; vielleicht sind sie auch nicht wirklich kompetent, das könnte auch noch sein. Sie folgen lieber den Schlußfolgerungen, die der Autor seiner Studie selber mitgibt, und zwar im Abstract. Vielleicht lesen sie sogar nur das Abstract. Auf jeden Fall sind die meisten offenbar außerstande, Studien wirklich kritisch zu analysieren und zu lesen. Und so entsteht über die Medien transportiert ein Hype über Daten und Ergebnisse, die sich mit großer Wahrscheinlichkeit später als nicht haltbar herausstellen werden.

Was lernen wir draus? Drei Lektionen:
1. Eine Schwalbe macht keinen Sommer. Immer abwarten, ob Nachfolgestudien anfängliche Ergebnisse bestätigen.
2. Der Sommer kommt sowieso selten und wenn, dann sehr spät. Wir leben, wissenschaftlich gesehen, weit nördlich des Polarkreises und haben weniger wirkliche Befunde und Durchbrüche, als wir glauben, zumindest was die Medizin und die Gesundheitswissenschaften angeht.
3. Was auch immer in der Wissenschaftspresse steht: es ist eine gute Heuristik, erst mal auch das Gegenteil dessen, was berichtet wird, für wahr zu halten.

Literatur

1. Ioannidis, J. P. (2005). Why most published research findings are false. PLoS Medicine, 2(8), e124. http://www.plosmedicine.org/article/info:doi/10.1371/journal.pmed.0020124
2. Gonon, F., Konsman, J.-P., Cohen, D., & Boraud, T. (2012). Why most biomedical

findings echoed by newspapers turn out to be false: The case of Attention Deficit Hyperactivity Disorder. PLoS ONE, 7(9), e44275. http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0044275

3. Yavchitz, A., Boutron, I., Bafeta, A., Marroun, I., Charles, P., Mantz, J., et al. (2012). Misrepresentation of randomized controlled trials in press releases and news coverage: A cohort study PLoS Medicine, 9(9), e1001308. http://www.plosmedicine.org/article/info%3Adoi%2F10.1371%2Fjournal.pmed.1001308