



Einsichten eines Wissenschaftsnarren (47)

Mehr Handys, mehr Dicke?

Eine Studie über vermeintliche Gesundheitsfolgen von Mobiltelefonen versammelt beispielhaft all die Probleme, die die biomedizinische Forschung seit Jahren plagen.

„Voll verstrahlt: Daddeln macht dick!“ – so titelte kürzlich die *BILD!* Der originellen Alliteration wegen ignorierte der Reporter sogar, dass es im Artikel um das Telefonieren und nicht um das Spielen mit Handys ging.

Im Bericht bezog die *BILD* sich auf eine Pressemitteilung der Uni Lübeck. Diese hatte die Studie eines Wissenschaftler-Teams aus der dortigen Psychoneurobiologie angepriesen, in der angeblich der „Einfluss von Handystrahlung auf die Nahrungsaufnahme nachgewiesen“ wurde. Demnach ging das so: In einem „durchdachten Versuchsdesign“ hatte man 15 junge Männer in einem Abstand von zwei Wochen insgesamt dreimal einbestellt. Im Experiment wurden die Probanden dann mit zwei verschiedenen Handys bestrahlt beziehungsweise einer Scheinbestrahlung als Kontrolle ausgesetzt. Und im Anschluss durften sie sich für eine definierte Zeit an einem Buffet bedienen. Gemessen wurde die spontane Nah-

»Mit 15 Studienteilnehmern kann man nur bestätigen, was man im Grunde eh schon weiß.«

rungsaufnahme, der Energiestoffwechsel des Gehirns via Phosphor-Magnetresonanz-Spektroskopie (MRS) sowie verschiedene Blutwerte vor und nach der Bestrahlung. Und siehe da: Nach 25 Minuten Handy am Ohr verzehrten die Probanden sage und schreibe rund ein Viertel mehr Kalorien zum Frühstück als die „Scheinbestrahlten“. Das entspricht kalorisch etwa einem halben Liter Bier oder einem Stück Apfeltorte!

Die Pressemitteilung fand dieses Ergebnis „erstaunlich“, und den Effekt „überraschend deutlich“. Dies kann man getrost als massi-

ves Understatement bezeichnen. Denn sollte dies stimmen, müssten wir uns wohl alle wundern, warum wir den etwa dreißig Minuten nach solchen Mobiltelefonaten einsetzenden Heißhunger bisher kaum an uns selbst bemerkt haben. Wichtiger aber noch: Der in *Nutrients* (14(2): 339) veröffentlichte Befund hätte nicht abzuschätzende Implikationen für die Menschheit. Dies blieb auch dem Autorenteam nicht verborgen, ja war sogar die Motivation für ihre Studie. Bereits im ersten Satz des Abstracts wird ein Zusammenhang zwischen gestiegener Mobilfunknutzung und der weltweiten Adipositas-Epidemie suggeriert.

Tatsächlich waren laut WHO im Jahr 2016 mehr als 1,9 Milliarden Erwachsene ab 18 Jahren übergewichtig, 650 Millionen davon sogar fettleibig. Wie wir wissen, erhöht dies das Risiko für Diabetes mellitus, Herz-Kreislauf-Erkrankungen, Bluthochdruck und Schlaganfall sowie bestimmte Krebsarten deutlich.

Sollte also die Nutzung von Mobiltelefonen ein wichtiger Grund für eine der wichtigsten vermeidbaren Ursachen von weltweiter Morbidität und Mortalität sein, wäre das eine Entdeckung vom Format Penicillin, Polio-Impfung oder *Helicobacter* und Magengeschwür. Der Nobelpreis wäre für die Entdecker in greifbarer Nähe. Deshalb ist es umso erstaunlicher, dass, wie die Pressemitteilung stolz vermeldet, die Studie zwar von *RTL*, der Fachzeitschrift *Elektrosmog* und dem Portal *diagnose:funk* – Letzteres ein Organ für Elektrosmog-Obskuranten –, aber bisher von keinem der seriösen deutschen oder internationalen Medien beachtet wurde.

Hunderte von kleinen und methodisch problematischen Studien haben in den letzten Dekaden einen Zusammenhang von elektromagnetischer Handy-Strahlung und Krebs, Frühgeburten, Depressionen sowie jeder Menge anderer Gesundheitsschäden nahegelegt – und nach deren Interpretation natürlich sogar bewiesen. Große, gut gemachte Studien haben aber all dies widerlegt. Dass eine sehr kleine Studie nun einen akuten Effekt von Handy-Strahlung auf das Gehirn – konkret eine massive, einfach nachzuweisende Verände-

itung im Essverhalten – nachweisen kann, ist demnach bemerkenswert. Also hat der Wissenschaftsnarr sich den *Nutrients*-Artikel mal genauer angeschaut. Auch mit dem Hintergrund der Frage, was die DFG denn so alles fördert. Denn letztlich wurde diese Studie von uns Steuerzahlern finanziert, und zwar über den Transregio SFB 134.

Um es vorwegzunehmen: Diese Studie ist ein Lehrstück für all die Probleme, die die Biomedizin seit Jahren plagen – und dem Narren leider unerschöpflichen Stoff für diese Kolumne bieten. Nichts wurde hier ausgelassen. Die Probleme fangen beim Studiendesign an und hören erst beim Reporting, also der eigentlichen Veröffentlichung, auf. Aber immer der Reihe nach.



Foto: BIH/Thomas Rafalzyk

Ulrich Dirnagl

ist experimenteller Neurologe an der Berliner Charité sowie Gründungsdirektor des QUEST Center for Transforming Biomedical Research am Berlin Institute of Health. Für seine Kolumne schlüpft er in die Rolle eines „Wissenschaftsnarren“ – um mit Lust und Laune dem Forschungsbetrieb so manche Nase zu drehen.

Sämtliche Folgen der „Einsichten eines Wissenschaftsnarren“ gibt es unter www.laborjournal.de/rubric/narr

Zunächst einmal gilt in der Wissenschaft der Grundsatz: Außergewöhnliche Aussagen benötigen außergewöhnliche Evidenz. Und die ist mit 15 jungen Männern, auch in einem Cross-over-Design, einfach nicht zu haben. Das kann man ganz einfach statistisch begründen, das sagt einem aber auch der gesunde Menschenverstand. Ob nun ein paar Hundert oder ein paar Tausend Probanden nötig wären und wie oft das dann von anderer Stelle repliziert werden müsste, darüber lässt sich trefflich biometrisch fachsimpeln. Klar ist nur: 15 Probanden und ohne weitere Bestätigung – das geht gar nicht! Dies umso mehr, da hier eine heterodoxe, sehr unwahrscheinliche Hypothese untersucht wird.

Mit 15 Studienteilnehmern kann man nur bestätigen, was man im Grunde eh schon weiß. Zum Beispiel, dass ein neues Medikament aus einer umfassend untersuchten Substanzklasse, die nachgewiesenermaßen und zweifelsfrei den Blutdruck senkt (beispielsweise Sartane), dies eben auch macht. Statistisch ausgedrückt besteht das Problem unter anderem darin, dass mit abnehmender „Prior Probability“ der Hypothese („Vortestwahrscheinlichkeit“, „Base Rate“) die Zahl der falsch positiven Resultate zunimmt. Bei unwahrscheinlicher Hypothese, niedriger statistischer Power (15 Probanden!) und nicht sehr stringentem Typ-I-Fehler-Signifikanzniveau (hier 5 Prozent) werden falsch positive Ergebnisse immer wahrscheinlicher.

Nur wenn man den p-Wert mit dem positiven (beziehungsweise negativen) Vorhersagewert verwechselt – und viele Kollegen tun dies leider! –, kann man sich unter diesen Bedingungen mit einem „statistisch signifikanten Ergebnis“ wohl fühlen. Alle anderen werden unbeeindruckt stärkere Evidenz fordern – vor allem diejenigen, die sich noch die diesbezügliche Wissenschaftsnarretei, „Brüder, zur Sonne, dem p-Wert ein Ende, Brüder,

zum Lichte empor!“ an dieser Stelle erinnern (LJ 10/2019: 24-5).

Ganz nebenbei sei erwähnt, dass in der Physik für eine „Entdeckung“ – und zu einer solchen sollte man das von den Lübeckern beschriebene Phänomen getrost zählen – eine statistische Signifikanz auf 5-Sigma-Niveau gefordert wird. Das bedeutet, einen Typ-I-Fehler mit einer Wahrscheinlichkeit von über 1 zu 3,3 Millionen zu akzeptieren, ein p-Wert mit

samen Website tylervigen.com/spurious-correlations ist eine Vielzahl solcher „Spurious Correlations“ gelistet. Mein Liebling darunter ist die nahezu perfekte Korrelation der US-Ausgaben für Wissenschaft, Raumfahrt und Technologie mit den Selbstmorden durch Erhängen, Strangulieren und Ersticken. Anyway ...

Die Lübecker Studie war übrigens *nicht* präregistriert. Wir wissen also nicht, was alles in deren Verlauf angepasst wurde, welche Analysen vorgesehen waren und welche dann tatsächlich gemacht wurden, welche Daten Verwendung finden sollten und welche es tatsächlich in die Auswertung geschafft haben und so weiter („Undisclosed Flexibility“). Ich will hier gar nichts unterstellen, aber es macht die Interpretation einer Studie so viel eindeutiger, wenn alles auf dem Tisch liegt, bevor es losgeht. Selbstverständlich hätten die Lübecker dies mit einer Sperrfrist bis zur Veröffentlichung der Resultate tun können. Das ist bei solchen Studien sogar die Regel. Damit niemand die tolle Idee klauen und vorher publizieren kann.

Laut den Autoren war das Studiendesign „single-blinded“. Man kann nur annehmen, dass die Verblindung sich hier auf die Studienteilnehmer bezog. Demnach wussten diese nicht, was die eigentliche Fragestellung der Studie war, noch welches Handy sendete oder nur „sham“ war. Das ist vorbildlich. Nur waren dann die Untersuchenden selbst, die die Daten auswerteten, offenbar nicht verblindet. Ich denke, es braucht keine weitere Erläuterung, welche Auswirkungen unbewusster Bias in solch einem Fall auf eine Studie haben kann.

Bleiben noch die Originaldaten, beziehungsweise deren (Nicht-)Verfügbarkeit. Im Artikel steht, dass diese auf Anfrage zur Verfügung gestellt werden könnten. Das ist an sich schon eine Enttäuschung. Denn wer je versucht hat, „data on reasonable request“ von

»Nicht nur wurde die Weltbevölkerung dicker, auch der Thunfisch im Mittelmeer hat abgenommen.«

mehr als 7 Nullen. Wer sich bereits mit dem 5-Prozent-Niveau begnügt, akzeptiert einen Fehler in zwanzig Fällen. Ronald Fisher, der „Erfinder“ des p-Werts, charakterisierte Befunde auf 5-Prozent-Niveau daher lapidar als höchstens „Worth a Look“.

Die statistischen Probleme der Handy-Studie gehen aber noch weiter. Zwar machen die Autoren eine Vielzahl von Vergleichen, definieren aber keinen davon als „primären Endpunkt“ – und kennzeichnen die anderen Vergleiche dann als explorativ. Deshalb hätten die Typ-I-Fehler-Niveaus auch entsprechend korrigiert werden müssen (für Fortgeschrittene: zum Beispiel mit einer Bonferroni-Korrektur).

Und natürlich muss auch die sich im Artikel andeutende – beziehungsweise sogar unmittelbar suggerierte – Verwechslung von Korrelation und Kausalität hinsichtlich „Mehr Handys, mehr Dicke!“ erwähnt werden. Schließlich sind zeitgleich mit der Zunahme der Handy-Nutzung viele Dinge passiert. Nicht nur wurde die Weltbevölkerung dicker, auch haben etwa Elektroautos zu- und der Thunfisch im Mittelmeer abgenommen. Auf der sehr unterhalt-



MICRONIC DT510 SINGLE-TUBE-READER

Der neue DT510 liest schnell und zuverlässig den 1D-/2D-Barcode sowohl von Tubes, als auch von Racks. Es ist ein **Plug & Play Lesegerät**, das keine Softwareinstallation erfordert und überaus einfach zu bedienen ist. Das spezielle **High-End-Dekodierungssystem** sorgt für einen äußerst zuverlässigen Dekodierungsprozess. Testen Sie selbst!

Wir werden den DT510 LIVE auf der diesjährigen **ANALYTICA** in München vorführen. Besuchen Sie uns dort vom 21. bis 24. Juni 2022 an unserem Messestand 318A.



LABORJOURNAL



**Inhalte
verantworten**

**Fakten
erkennen**

**Propaganda
entlarven**

**Sprache
beherrschen**

Freie Presse
Wissen, wen man liest.

Studienautoren zu bekommen, weiß, dass dies in den seltensten Fällen gelingt. Falls man überhaupt eine Antwort auf die Anfrage bekommt, spricht meist irgendetwas gegen eine Herausgabe. Manchmal ist sogar die Festplatte kaputtgegangen oder der Zuständige ist schlachtweg unbekannt verzogen.

Auch hier will ich nichts unterstellen. Fest steht aber, dass die Daten von öffentlich geförderten Studien auch der Öffentlichkeit zur Verfügung gestellt werden müssen. Sagt das nicht sogar der Fördergeber dieser Studie, die DFG? Warum können wir die Daten dieser Studie nicht einfach von ZENODO oder einem anderen Repositorium herunterladen?

Dabei hätte zumindest der Narr sich die Originaldaten sehr gerne angesehen, denn deren Repräsentation in der deskriptiven Statistik des Artikels lässt leider zu wünschen übrig. Dort finden wir statt Dot- und Box-Plots sowie vernünftigen Varianzmassen leider nur die allgegenwärtigen, aber die wahren Varianzen beschönigenden und echten Verteilungen verschleiernden Bar-Graphen mit Standardabweichungen (SEM). Bei der Darstellung der neuroenergetischen Daten fehlen die SEM für die „Sham“-Ergebnisse gleich ganz. Der hier offenbar gezeigte Gruppen-Mittelwert führt denn auch zu einer starken Überbetonung der Effektstärke. MR-Spektroskopiker hätten sich außerdem auch sehr über ein paar Originalspektren zur Beurteilung der Messqualität gefreut. All dies sind Hinweise darauf, dass die Reviewer von *Nutrients* hier wohl nicht so genau hingeschaut haben.

Nun ist MR-Spektroskopie aufwendig und teuer. Allerdings hätte man das Experiment durchaus mit einer größeren Zahl von Probanden ganz ohne MR-Spektroskopie replizieren können – nur unter Messung der Kalorienaufnahme. Da wären dann nur noch die Kosten für das Frühstücksbuffet angefallen. Auch hätte man damit gleich ausschließen können, dass es zu einer wie auch immer gearteten Wechselwirkung des Leistungseintrags in das Gehirn beim Senden über die Mobilfunkantenne und dem Leistungseintrag durch die HF-Pulse während der Magnetresonanz-Spektroskopie gekommen ist. Vor allem wenn man von 15 Probanden auf die Weltbevölkerung schließen will, würde das viel Sinn machen, denn die wenigsten von uns liegen vor und nach einem Handy-Telefonat im Tomographen.

Manche Leserin oder Leser wird sich nun wohl fragen, warum ein derart sensationeller und potenziell weltbewegender Befund nicht in *Nature* oder *Science* veröffentlicht wurde, sondern stattdessen in *Nutrients* – einer Zeitschrift, die schon mehrfach durch Skandale aufgefallen war (für Interessierte: Googeln Sie zum Beispiel mal „Australian Paradox“). Ich halte diesen Aspekt allerdings für wenig

relevant. Eine Studie muss auf der Basis ihres Designs, ihrer Ergebnisse und deren Interpretation bewertet werden, nicht nach dem Journal, in dem sie publiziert wurde. Klar ist aber – und viele von uns kennen das aus eigener Erfahrung –, dass sich dieser Artikel in einer Kaskade von Journals mit abnehmendem Journal-Impact-Faktor abwärts bewegt haben muss. Für den Narren ist das folglich wieder mal ein Hinweis darauf, dass letztlich alles irgendwo publiziert werden kann und wird – und dass der Review-Prozess nicht die derzeit viel beschworene Filterfunktion hat, unter anderem auch wegen der Zunahme von Preprints. Dass ein Artikel in einem Journal mit einem Impact-Faktor von über 5 und damit aus dem oberen Viertel der Kategorie „Nutrition & Dietetics“ veröffentlicht wird, ist für sich erst mal gar kein Ausweis für Qualität. Man muss sich immer noch die Mühe machen, den Artikel zu lesen und ihn inhaltlich und methodisch bewerten.

**»Ist der Ansatz der Autoren total esoterisch?
Natürlich nicht!«**

Was also bleibt? Sollte man hochgradig unwahrscheinliche Hypothesen erst gar nicht untersuchen? Ist die Untersuchung einer möglichen Wirkung von elektromagnetischer Strahlung auf das Gehirn und unser Verhalten grundsätzlich abzulehnen? Ist der Ansatz der Autoren total esoterisch? Natürlich nicht! „To boldly go where no man has gone before“ ist eine der vornehmsten und spannendsten Aspekte der Wissenschaft. Nur müssen wir methodisch solide bleiben und unsere Ergebnisse nicht überinterpretieren. Eine Lübecker Handy-Studie, prä registriert, mit ein paar hundert Probanden (nur Frühstück!), doppelt verblindet, inklusive dem versuchten Nachweis einer Dosis-Wirkungsbeziehung (15 Minuten Handy-Telefonat vs. 30 Minuten), repliziert in einem anderen Labor, die Daten frei verfügbar bei ZENODO – das wäre einer Berichterstattung in *Nature*, *New York Times* und *Neue Zürcher Zeitung* würdig. Auch wenn herauskäme, dass man nach Handy-Bestrahlung gerade mal ein Löffelchen Müsli mehr essen würde.

Der Wissenschaftsnarr dankt Prof. Dr. Harald Möller vom Max-Planck-Institut für Kognitions- und Neurowissenschaften in Leipzig für seine sachdienlichen Hinweise zur MR-Spektroskopie in der Studie. Weiterführende Literatur und Links finden sich wie immer unter: <http://dirnagl.com/lj>.