



Die Fitness-Empfehlungen sind wissenschaftlich gesichert [?]

Aus: Udo Pollmer, Susanne Warmuth, Gunter Frank: „Lexikon der Fitness-Irrtümer: Missverständnisse, Fehlinterpretationen und Halbwahrheiten von Aerobic bis Zerrung“ Seite 148 - 158 PIPER 2003

"Die Wahrscheinlichkeit für gute Ideen ist und bleibt leider eine große Unbekannte. Und auch die inhärenten methodischen Probleme randomisierter Studien bleiben weiterhin bestehen. Wir sind der Überzeugung, dass die Statistik, zumindest so, wie sie heute angewandt wird, eine Sackgasse für die Forschung ist."

Prof. Dr. Hans Peter Beck Bornholdt u. Priv.-Doz. Hans Hermann Dubben
„Der Schein der Weisen – Irrtümer und Fehlurteile im täglichen Denken“
Seite 255, ROWOHLT 2003

Sollte man eigentlich meinen. Zumal, wenn dergleichen seit Jahren gebetsmühlenhaft von Autoritäten aus Politik, Gesundheitswesen und Wissenschaft behauptet wird.

Der wissenschaftliche Nachweis gilt in der Öffentlichkeit immer noch als stärkstes Argument für eine Position oder Empfehlung.

Dabei gehen viele Steuerzahler ganz selbstverständlich davon aus, dass Forscher vor allem für den Wissenszuwachs arbeiten, neutral und unvoreingenommen handeln und damit die besten Ratgeber für die Weiterentwicklung unserer Gesellschaft darstellen.

Doch auch der Wissenschaftler lebt von seinem Einkommen und braucht für seine Karriere Forschungsaufträge. Und schon gilt die alte Stammtischregel: „**Wer zahlt, hat recht**“. Sei es der (politische) Dienstherr, der nur „richtige“ Ergebnisse akzeptiert, als auch andere Förderer, die auf den fraglichen Gebieten ihre wirtschaftlichen Interessen wahrnehmen.

Wer Professor Marlboro [der Tabakindustrie] glaubt, dass Rauchen nicht schadet, ist doch selbst schuld, wenn er Lungenkrebs bekommt, könnte man achselzuckend kommentieren.

Doch **erstens** ist vielfach nicht erkennbar, in wessen Diensten der Spezialist steht, der mit ernster Miene in die Kamera blickt. Aus Kostengründen wenden sich die Medien zunehmend an Experten, die ihr Honorar von einem anderen Auftraggeber erhalten.

Und **zweitens** werden Universitätsprofessoren und öffentliche Institutionen nach wie vor aus Steuergeldern finanziert, also mit unseren sauber verdienten Euros, Weihnachts- und Urlaubsgeld sowie Pensionen inklusive.

Deshalb hat die Gesellschaft einen berechtigten Anspruch darauf, von wissenschaftlichen Würdenträgern seriös über den aktuellen Kenntnisstand informiert zu werden.

Wenn man jedoch – wie wir – einen Blick hinter die Kulissen öffentlicher Empfehlungen und Begründungen für gesundheitspolitische Maßnahmen wirft und die Daten prüft, kommen rasch Zweifel auf.

W a r u m ?

Wenn Erkenntnisse als wissenschaftlich seriös gelten wollen, müssen sie 3 Voraussetzungen erfüllen.

Erstens: Das Versuchsergebnis muss objektiv messbar sein [Analysier-/Quantifizierbarkeit].

Zweitens: Bei wiederholten Messungen muss immer dasselbe herauskommen [Reproduzierbarkeit].

Drittens: Das Resultat muss auch im wirklichen Leben eine Bedeutung besitzen.

Man spricht von **Objektivität** (Sachlichkeit), **Reliabilität** (Zuverlässigkeit) und **Validität** (Gültigkeit).

Erst wenn diese Voraussetzungen erfüllt sind, darf man davon ausgehen, dass die neue Erkenntnis einen praktischen Nutzen besitzt.

Ein Beispiel: Angenommen, im Tierreich wird festgestellt, dass Hamster, die viel im Laufrad rennen, selten Lungenkrebs bekommen. Gilt dann für Menschen dasselbe? Nein, nicht zwangsläufig, denn der Hamster ist ein ganz anderes Lebewesen, das sein Leben in einem langweiligen Käfig fristet, statt zur Arbeit zu gehen, zu feiern und Steuern zu zahlen.

Doch selbst wenn Beobachtungen vorlägen, dass Jogger seltener an Lungenkrebs erkranken, dürfte man daraus nicht schließen, dass man die Leute nur in den Wald scheuchen müsste, um ihre Lungen vor bösartigen Wucherungen zu schützen.

Die Jogger könnten sich ja auch noch in anderer Hinsicht von anderen Zeitgenossen unterscheiden – beispielsweise seltener rauchen als Nichtjogger.

Es geht also darum, eine **Hypothese** [I], die aus Tierversuchen oder aus Beobachtungen am Menschen abgeleitet wurden, in Studien zu überprüfen, die so angelegt sein müssen (man spricht hier vom „Studiendesign“ [II]), dass sie die genannten 3 Kriterien möglichst gut erfüllen.

Erst wenn die Ergebnisse einer sinnvoll angelegten und sauber ausgewerteten **Studie** die formulierte Hypothese (hier: Joggen schützt vor Lungenkrebs) bestätigen, kann man von einem wissenschaftlichen Beleg sprechen. Und nur dann besteht Hoffnung, dass der untersuchte Effekt auch im wirklichen Leben eintritt.

Fast alle Studien, die wir für dieses Buch durchgesehen haben – darunter nicht wenige von anerkannten Autoritäten -, lassen in Hinblick auf die wissenschaftliche Arbeitsweise sehr zu wünschen übrig.

Besonders beliebt in der Medizin: die Spezialdisziplin **Datenmassage**, in der sie regelmäßig zu olympischen Höchstleistungen aufläuft. Da wird so lange geknetet, geschoben und gedrückt, bis die Statistik ein „passendes“ Ergebnis liefert. Und zur Not wird auch mal wesentliches weggelassen oder dazugedichtet.

Beispiele gefällig? Siehe Kasten.

T r a u , s c h a u , w e m - e i n B l i c k i n d i e T r i c k k i s t e

Der Trick mit den Abstracts: In wissenschaftlichen Arbeiten ist es üblich, die Ergebnisse in einer Zusammenfassung, dem so genannten Abstract [III], kurz und präzise darzustellen. Roy Pitkin, Herausgeber einer medizinischen Fachzeitschrift [„Obstetrics & Gynecology“], ärgerte sich darüber, dass ein Drittel der wissenschaftlichen Artikel, die ihm zur Veröffentlichung zugesandt wurden, fehlerhafte Abstracts enthielten, beispielsweise Aussagen, die in der Studie überhaupt nicht enthalten waren. In anderen Fällen wurden entscheidende Ergebnisse zugunsten belangloser Feststellungen weggelassen. Zusammen mit Kollegen [7] analysierte Pitkin je 44 Abstracts aus den 6 weltweit führenden medizinischen Fachblättern.

Das Ergebnis fiel erschütternd aus: Von 264 Abstracts waren 104 mangelhaft. In einer der 6 Fachzeitschriften stimmten sogar 68 % nicht.

Fatalerweise lesen viele Ärzte und andere Fachleute nur noch diese Abstracts, um sich rasch einen Überblick über den aktuellen Wissensstand zu verschaffen. Und das sind noch die überdurchschnittlich gut Informierten – die meisten halten sich angesichts der Informationsflut und der Arbeitsüberlastung allenfalls mit standespolitischen Vereinspostillen auf dem laufenden, die ihnen den Inhalt der Fachpresse mund-, umsatz- und ideologierecht aufbereiten.

Wenn schon im Kleinen so offensichtlich manipuliert wird, was passiert wohl in den Bereichen, in denen weder die Herausgeber einer Fachzeitschrift noch die Leser die Möglichkeit haben, die Echtheit der Daten nachzuprüfen?

Der Äpfel und Birnen-Trick: Man teilt die Versuchspersonen für Fitnessstests zwar fein säuberlich in Gesunde und Kranke und, je nach Leistung auf dem Laufband, in „Fitte“ und „Unfitte“. Bei der Auswertung kommen aber alle Teilnehmer in einen Topf, und die Autoren freuen sich über phantastische Unterschiede in Lebenserwartung und Herz-Kreislauf-Sterblichkeit.

Was aber nur daran liegt, dass die Kranken die Kategorie „unfit“ bevölkern, während sich bei den „Fitten“ die Gesunden tummeln. Damit ist aber leider nur bewiesen, dass Kranke nicht so gesund sind wie Gesunde.

Der Rohdaten-Trick: Wie eine Meta-Analyse [2] von Studien zum Thema körperliche Aktivität zeigt, gibt es nicht wenige ältere Untersuchungen, bei denen nicht einmal eine Alterskorrektur vorgenommen wurde. Das heißt, hier muss sich auch ein 75-jähriger Rheumatiker mit einem durchtrainierten Jüngling messen. Klar, dass der junge Sportsfreund eine höhere Lebenserwartung hat als der gehbehinderte Alte. Auch die Bereinigung um den Faktor Rauchen, der sich unbestritten auf die Sterblichkeit auswirkt, wurde keinesfalls immer durchgeführt. Dabei ist es nichts Neues, dass sich unter den Unsportlichen mehr Raucher befinden. Wenn die Rohdaten besser passen, stellt man diese ohne weitere Korrekturen als Beweis vor. Passen sie nicht, werden sie so lange massiert, bis sie passen, und die verräterischen Rohdaten fallen unter den Tisch.

Der Trick mit den Kategoriegrenzen: In einer schwedischen Sportstudie [3] wird der Lebensabschnitt zwischen 20 und 60 in zwei Einheiten unterteilt. Warum ausgerechnet in 20 bis 38 und 39 bis 60, verraten die Autoren nicht. Logisch wäre beispielsweise 20 bis 40 und 41 bis 60. In solchen Fällen keimt zwangsläufig der Verdacht auf, dass schon eine Verschiebung der Kategoriegrenzen um ein Jahr andere, offenbar unpassende Ergebnisse gebracht hätte. Mit diesen Kategoriegrenzen lässt sich statistisch viel herausholen, vor allem dann, wenn die Daten sehr uneinheitlich sind. Man macht eine Einteilung in drei, vier, fünf oder sechs Gruppen und spielt das Ganze nach Ergebnis, nach Ausgangsdaten, nach Zahl der Teilnehmer usw. am Computer durch. Irgendwann ergeben auch die wertlosesten Daten „passende“ Ergebnisse.

Der Trick mit dem Relativen Risiko: Wenn man etwa eine Gruppe von 3.000 jungen Frauen täglich eine Stunde joggen lässt, und es erkrankt innerhalb eines Jahres 2 von ihnen an Brustkrebs, während in der gleich großen Vergleichsgruppe ohne Sport 3 Fälle auftreten, dann hat der Sport das relative Brustkrebsrisiko rein rechnerisch um 33 % gesenkt. Das ist für den einen Menschen, der vom Krebs verschont wird, durchaus relevant, der Effekt für die gesamte Gruppe hält sich allerdings in Grenzen, denn er liegt in absoluten Zahlen nur bei 0,33 Promille.

Stirbt aber eine Frau durch die sportlichen Aktivitäten am plötzlichen Herztod, holen sich 3 andere einen Kreuzbandriss und ein weiteres Dutzend erleidet die typischen Ermüdungsbrüche, von denen [- 12 -] sich anschließend 1 Patientin durch den Gips eine Unterschenkelthrombose mit Lungenembolie und Todesfolge einhandelt, dann sieht die Gesamtbilanz sogar negativ aus. Trotzdem kann man immer noch zu Recht behaupten, dass Sport gesund sei, weil er die Brustkrebsrate (Angst!) um ein Drittel [rel. % Angabel] gesenkt habe.

Der Trick mit der Gesamtmortalität: Gerne wird der vorbeugende Effekt einer Maßnahme vor einer bestimmten Krankheit oder Todesursache hervorgehoben.

Häufig fehlt dabei die Angabe der Gesamtsterblichkeit [IV]. Aber nur die erlaubt es, die Aussage zu beurteilen.

Denn oft hängt der vermeintliche Nutzen schlicht davon ab, dass in der einen Gruppe ein paar mehr Menschen am Herzinfarkt, dafür aber weniger am Krebs gestorben sind, und in der anderen Gruppe war es genau umgekehrt.

Wenn die Gesamtsterblichkeit in beiden Gruppen aber gleich war, dann hat die Maßnahme aufs Ganze gesehen nichts gebracht. Dieser Verdacht besteht immer dann, wenn diese Daten in der Publikation fehlen, obwohl sie in den meisten Studien erhoben wurden.

Manchmal gehen die Autoren sogar soweit, mit dieser Methodik eine Erhöhung der Sterblichkeit allen Ernstes auch noch als Erfolg zu verkaufen.

Angenommen, eine allseits empfohlene Maßnahme erhöht die Krebsrate, und die Lebenserwartung sinkt. Dann ist eine Folge davon, dass andere Todesursachen, beispielsweise Herzinfarkt, seltener auftreten – denn wer bereits an Krebs gestorben ist, kann keinen Herzschlag mehr erleiden.

Unter der Schlagzeile „xy schützt vor Herzinfarkt, der Todesursache Nr. 1“ geht das Ergebnis durch die Medien. Sie könnten genau so gut ihre Mitarbeiter erschießen und anschließend erklären, sie hätten eine ebenso billige wie wirksame Vorbeugung von Krebs, AIDS und Alzheimer durchgeführt. Zynisch? Solange es einem „guten Zweck dient“, egal ob gesunde Ernährung, Fitness oder Prävention, ist diese Art der Argumentation in der Gesundheitspresse – das gilt für die Fachwelt ebenso wie für bunter Blätter – gar nicht so selten.

Die Interpretation von Studienergebnissen und damit die Umsetzung in konkrete Empfehlungen hängt in hohem Maße von der **Qualität einer Studie** [V] ab - Qualität beim Konzipieren, Durchführen und Analysieren.

Da der Nutzen einer Studie in der Vergangenheit von jedem Experten anders beurteilt wurde – und weil auch vielen Medizinern der Unsinn bei den therapeutischen Ratschlägen über die Hutschnur ging -, wurde seit den achtziger Jahren ein neues Bewertungssystem eingeführt, mit dem die Beweiskraft einer Studie oder einer Analyse mit vier Noten (A bis D) [V] beurteilt wird. Man spricht von *evidence based medicine*, eingedeutscht „evidenzbasierte Medizin (EBM)“ und übersetzt „auf wissenschaftlichen Belegen beruhende Medizin“ (siehe s. 156) [V]

A ist dabei die Stufe, mit der der vorausgesagte Nutzen sehr wahrscheinlich zutrifft (zum Beispiel placebokontrollierte doppelblinde Interventionsstudien), das Schlusslicht, die Note **D**, gilt für reine Spekulationen (zum Beispiel Konsensuskonferenzen) [V].

Für das Thema Bewegung und Gesundheit liegen keine beweiskräftigen Studien im Sinne der Evidenzbasierten Medizin vor (Note A bzw. 1). Es mag wohl schwierig sein, einer großen Zahl von Menschen entweder ein langjähriges Fitness-Programm zu verordnen oder sie auf Dauer auf der Couch zu halten. Trotzdem fragt man sich, warum dies in 50 Jahren Sportwissenschaft – bei gleichzeitigen Lobpreisungen des Nutzens von Sport und Bewegung – nicht möglich war?

Die heutigen Empfehlungen gründen sich im wesentlichen auf Beobachtungsstudien, noch dazu ohne Zufallsgruppierung. Dabei wächst auf internationaler Ebene das Bewusstsein durchaus, dass Privatmeinungen von Lehrstuhlinhabern auch im Sport als wissenschaftliche Belege nicht mehr ausreichen.

Wie dringend in der Sportwissenschaft brauchbare Daten benötigt werden, macht das folgende Zitat aus dem *British Journal of Sports Medicine* deutlich. Unter der Überschrift „Wissenschaft oder Show Business?“ [5] hält der Allgemein- und Sportmediziner Domhnall MacAuley von der Universität Belfast seinen Kollegen eine akademische Standpauke [Moralpredigt]: „Nicht einmal für die Basisbehandlung von Weichteilverletzungen mit Eis und Druckverbänden haben wir klare Wirksamkeitsnachweise. Wenn wir uns als ernst zu nehmende Disziplin etablieren wollen, müssen wir es zulassen, dass unsere klinische Praxis auf den Prüfstand gestellt wird. Der Vortragsreisende in Sachen Sport muss qualitativ hochwertige Beweise vorlegen und nicht bloß hübsche Bilder (...) So lange wir unser ärztliches Handeln nicht mit wissenschaftlichen Belegen untermauern können, zaubern wir lediglich Kaninchen aus dem Zylinder“.

Doch die mahnenden Worte aus den eigenen Reihen halten die Mehrheit der Sportwissenschaftler nicht davon ab, unverzagt weiter die üblichen Gesundheitsempfehlungen zu predigen, die – rein zufällig – ihre Branche neuen Umsatz verschaffen.

Sie tun es sogar, wenn sie selbst feststellen, dass sich die vorliegenden Erkenntnisse, wenn man sie nach den Regeln der Evidenzbasierten Medizin bewertet, am Rande der Spekulation bewegen.

Deutlich wird dies beispielsweise im Sonderheft einer renommierten wissenschaftlichen Sportzeitschrift, in dem über die Ergebnisse eines internationalen Symposiums zur „Evidenz-Lage“ in Sachen Bewegung und Gesundheit berichtet wird.

Die Autoren des Kapitels „*Körperliche Aktivität und Gesamtsterblichkeit*“ [6] geben nach Durchsicht der wichtigsten internationalen Studien für die These, dass Bewegung und Fitness die Lebenserwartung erhöhen, korrekt den **Empfehlungsgrad C** [V] an. Das heißt es liegen nur Studien vor, deren Ergebnisse man als „gehobene Spekulation“ bezeichnen könnte.

Dennoch können die Autoren der Versuchung nicht widerstehen, die gravierenden Mängel in den vorliegenden Studien im Text ein wenig zu relativieren: „Heute gibt es eindeutige Beweise für diese Hypothese“.

Für die deutsche Sportmedizin bot sich damit eine gute Gelegenheit, den bescheidenen **Evidenzgrad C** [V] unauffällig zu „revidieren“: In einem viel beachteten Fachbuch [8] zum Thema Sport und Evidenzbasierte Medizin wird ebenfalls die These diskutiert, dass Sport das Leben verlängert. Der Autor bezieht sich auf die oben erwähnte Veröffentlichung und gibt statt des dürftigen Evidenzgrades C (entspricht bei der Feineinteilung der Note 4) die deutlich bessere Note 2b an! [siehe V]. Dies ist schon keine Irreführung mehr sondern gemahnt an Täuschung, vorausgesetzt, der Autor hat die Fachliteratur gelesen und auch verstanden.

Nachdem auf diesem Wege aus der Vermutung, Sport verlängert das Leben, mir nichts dir nichts quasi eine Tatsache wurde, ist der Weg frei für professorale Gesundheitslyrik: In schönen Tabellen werden nun sämtliche Klischees für die gesundheitsfördernde Wirkung des Sport aufgelistet, vom Autor selbst mit Evidenzgraden geschmückt.

Das wäre nun alles nicht weiter bemerkenswert, wäre dieses Werk [8] nicht von den Krankenkassen als wissenschaftliche Grundlage für ihre Investitionen in Fitnessprogramme benannt worden.

Das erklärt so manches.

Der Pharma-Industrie sind diese nationalen Besonderheiten natürlich auch nicht verborgen geblieben – und sie stellt sich entsprechend darauf ein. Friedrich Schwartz [9], ehemaliger Vorsitzender des Sachverständigenrats für die Konzertierte Aktion [VII] im Gesundheitswesen, berichtet von einem Gespräch mit einem Manager eines Weltkonzerns. In dessen Haus würden Skandinavien, Großbritannien und die Niederlande als der europäische *evidence belt* [„Beweis-Gürtel“] bezeichnet, weil neue Produkte nur einzuführen seien, wenn man qualitativ hochwertige Studien vorweisen könne. In Deutschland und Österreich dagegen genüge es, die Meinungsführer der Ärzteschaft zu überzeugen, weshalb die deutschsprachigen Länder unter *eminence belt* [„Meinungsführer (Eminenz)-Gürtel“] liefen ...

Was bleibt?

Auf die Frage, ob medizinische Forschungsergebnisse in großem Stil manipuliert werden, rückt der Statistiker Hans-Peter Beck-Bornholdt [10, VI] die Maßstäbe zurecht: „Manipulation würde ich diese Vorgehensweise nicht nennen. das ist eher anerkannter Standard in der Forschung“.

Mit Blick auf die Sportwissenschaft könnt man auch sagen: Wertlose Daten multipliziert mit hohem Unsicherheitsfaktor ergeben Kraft eines akademischen Titels hochsignifikante Fitness-Empfehlungen.



Text:

Udo Pollmer, Susanne Warmuth, Gunter Frank: „Lexikon der Fitness-Irrtümer: Missverständnisse, Fehlinterpretationen und Halbwahrheiten von Aerobic bis Zerrung“ Seite 148 - 158 PIPER 2003

Quellen:

- [1] S. N. Blair et al.: „Changes in Physical Fitness and All-Cause Mortality. A prospective study of healthy and unhealthy men“ JAMA. 1995 Apr 12;273(14):1093-8.
- [2] Jesse A. Berlin, Graham A. Colditz: „A Meta-Analysis of Physical Activity in the Prevention of Coronary Heart Disease“ American Journal of Epidemiology 1990/132/S. 612 ff.
- [3] Lauren Lissner et al.: „Physical Activity Levels and Changes in Relation to Longevity. A Prospective Study of Swedish Women“ American Journal of Epidemiology 1996/143/S. 54-62
- [4] R. Simon: „Patient Subjects and Variation in Therapeutic Efficacy“ British Journal of Clinical Pharmacology 1982/14/S. 473 ff.
- [5] Domhnall MacAuley: „Science or show business? Br J Sports Med. 1999 June; 33(3): 147–148
- [6] I. M. Lee, P. J. Skerrett: „Physical activity and all-cause mortality: what is the dose-response relation?“ Medicine & Science in Sports & Exercise 2001 Jun;33(6 Suppl):S. 459-71; discussion S493-4.
- [7] Roy M. Pitkin, Mary Ann Branagan, Leon F. Burmeister: „Accuracy of Data in Abstracts of Published Research Articles“ JAMA – Journal of the American Medical Association 1999;281(12):1110-1111
- [8] G. Samitz, G. B. M. Mensink (Hrsg.): „Körperliche Aktivität in Prävention und Therapie. Evidenzbasierter Leitfaden für Klinik und Praxis“ Hans Marseille Verlag München, 2002
- [9] Harro Albrecht, Friedrich Wilhelm Schwartz: „Das ist der Tod der wissenschaftlichen Beratung. Ein prominenter Sachverständiger tritt als Vorsitzender zurück, weil er die Unabhängigkeit seines Gremiums gefährdet sieht“ Interview von Harro Albrecht. Die Zeit 5.9.2002, S. 35
- [10] Cornelia Stolze, Hans-Peter Beck-Bornholdt: „Die Sicherheit ist nur vorgegaukelt“ Interview. Die Woche 22.2.2003, s. 27

Ergänzungen:

[I] Eine **Hypothese** (vom Altgriechischen über das Spätlateinische, wörtlich für „Unterstellung“) ist eine Aussage, deren Gültigkeit man für möglich hält, die aber nicht bewiesen oder verifiziert ist. Für Hypothesen ist es üblich, dass die Bedingungen angegeben werden, unter denen sie gültig sein sollen.

In positivistischen wissenschaftstheoretischen Strömungen ist die Hypothese eine Vorstufe einer Theorie, zu der sie durch verifizierende Beobachtungen werden kann. Anders als im allgemeinen Sprachgebrauch bezeichnet „Theorie“ in der Wissenschaft also im Unterschied zur Hypothese eine Aussage oder eine voneinander abhängige Gruppe von Aussagen, die durch empirische Erfahrungen bestätigt wurden (z. B. Relativitätstheorie, Gravitationstheorie, Evolutionstheorie). Eine Behauptung oder ein Leitsatz, der durch wissenschaftliche Beweisführung bewiesen oder widerlegt werden soll, heißt These (z.B. "Wirtschaften an sich ist frei von moralischem Gehalt").

Kritisch-rationale Ansätze vertreten allerdings die Auffassung, **Theorie**, **Spekulation** und **Hypothese** seien gleichwertig, da theoretische Annahmen grundsätzlich nicht verifiziert werden könnten (Prinzip der Falsifizierbarkeit). Quelle: Internet

[II] Siehe www.draloisdengg.at ->Infos -> 9.) Statistik Glossar & Allerlei >>>> Studiendesign usw.
[Meine Ergänzungen] Dr.med. Alois Dengg, Hollenzen 100, A-6290 Mayrhofen, www.draloisdengg.at

[III] Ein **Abstract** [„Zusammenfassung am Anfang“] ist eine *prägnante Inhaltsangabe*, ein *Abriss ohne Interpretation und Wertung* einer wissenschaftlichen Arbeit. In DIN 1426 wird das (oder auch der) Abstract als Kurzreferat (auch als Synonym verwandt) zur Inhaltsangabe beschrieben. Die Definition des American National Standards Institute (ANSI) lautet: „An abstract is defined as an abbreviated accurate representation of the contents of a document.“ –

„Ein Abstract ist definiert als eine gekürzte präzise Darstellung des Inhalts eines Dokuments.“

Das Abstract sollte immer am Anfang des Originaldokuments stehen. Die bibliografischen Angaben zum Dokument sollten unmittelbar nach dem Abstract folgen. Die Länge sollte vom Inhalt und nicht von der Dokumentenlänge abhängen.

Abstract - Allgemeine Merkmale:

- Objektivität: Es soll sich jeder persönlichen Wertung enthalten.
- Kürze: Es soll so kurz wie möglich sein.
- Verständlichkeit: Es weist eine klare, nachvollziehbare Sprache und Struktur auf.

- Vollständigkeit: Alle wesentlichen Sachverhalte sollen enthalten sein.
- Genauigkeit: Es soll genau die Inhalte und die Meinung der Originalarbeit wiedergeben. Quelle: Internet

[IV] **Gesamt mortalität/-sterblichkeit:** Die Sterblichkeit an allen Todesursachen zusammengenommen, bezogen auf eine bestimmte Population - z.B. 100.000 Einwohner - in einem definierten Zeitraum.

Die Mortalität, Sterblichkeit oder Sterberate ist ein Begriff aus der Demografie (Bevölkerungswissenschaft). Sie bezeichnet die Anzahl der Todesfälle, bezogen auf die Gesamtanzahl der Individuen oder, bei der spezifischen Sterberate, bezogen auf die Anzahl der betreffenden Population, meist in einem bestimmten Zeitraum. Quelle: Internet

[V] **EbM - Klassen:** Nach Centre for Evidence-based Medicine 2009, www.cebm.net/ , www.ebm-netzwerk.de/

Klasse **Ia:** Evidenz durch Meta-Analysen (Systematische Übersichtsarbeit, systematic Review) von mehreren randomisierten, kontrollierten Studien.

Klasse **Ib:** Evidenz aufgrund von mindestens einer randomisierten, kontrollierten Studie. Randomisierte, kontrollierte Einzelstudie mit engem Konfidenzintervall (Vertrauensintervall, Mutungsintervall)

Klasse **Ila:** Evidenz aufgrund von mindestens einer gut angelegten, jedoch nicht randomisierten und kontrollierten Studie. Systematischer Review von Kohortenstudien.

Klasse **Ilb:** Evidenz aufgrund von mindestens einer gut angelegten quasi-experimentellen Studie. Einzelne Kohortenstudie/randomisierte, kontrollierte Studie von geringer Qualität.

Klasse **III:** Evidenz aufgrund gut angelegter, nicht-experimenteller deskriptiver Studien wie etwa Vergleichsstudien, Korrelationsstudien oder Fall-Kontroll-Studien.

Klasse **IIla:** systematischer Review von Fall-Kontroll-Studien.

Klasse **IIlb:** einzelne Fall-Kontroll-Studie.

Klasse **IV:** Nichtexperimentelle Beobachtungsstudien / Kohorten- und Fall-Kontroll-Studien von geringer Qualität.

Klasse **V:** Evidenz aufgrund von Berichten der Experten-Ausschüsse oder Expertenmeinungen / Konsensusverfahren bzw. klinischer Erfahrung anerkannter Autoritäten.

"Wir Menschen bestehen nun einmal nicht aus statistischen Mittelwerten, sondern die Natur möchte eine Streuung von Merkmalen, das heißt, wir sind einfach unterschiedlich. Deshalb empfiehlt die EbM auch nicht, stur nach Statistik zu therapieren, sondern spricht von einem Behandlungskorridor. Liegt Empfehlungsgrad A vor, dann ist die Wahrscheinlichkeit, Patienten mit dieser Therapie richtig zu behandeln, hoch. Aber nicht 100%. Ist der Korridor sehr weit, weil es nur B oder C gibt, dann entscheidet oft die persönliche Erfahrung des Therapeuten wirkungsvoller, welche Therapie die richtige ist. Ein guter Arzt wird deshalb bei jedem einzelnen Patienten den Erfolg einer Therapie nicht nur am Erreichen von Normwerten messen, sondern an der Beantwortung der Frage, ob es dem Patienten mit der Therapie besser geht. Ist die Antwort nein, kann es in solchen begründeten Ausnahmefällen sehr sinnvoll sein, sich trotz bester "Beweislage" gegen die Therapie zu entscheiden." Aus: Dr. med. Gunter Frank: „Schlechte Medizin: Ein Wutbuch“ Teil I: Schlechte Medizin in der täglichen Behandlung. Kapitel: Gute Medizin: Klare Regeln sind nicht verhandelbar. Weitere Fallstricke. Seite 67f. KNAUS 5. Auflage 2012

Vier Evidenz-Empfehlungsgrade: ("Studien TÜV")

A (hoch = Nutzen der Empfehlung gut belegt):

Es ist unwahrscheinlich, dass weitere Forschungsarbeiten unser Vertrauen in die Richtigkeit unserer Aussage erschüttern werden.

- ▶ Es liegen mehrere wissenschaftlich einwandfreie Studien (1a) mit konsistenten Ergebnissen vor.
- ▶ In besonderen Fällen genügt eine große und qualitativ hochwertige multizentrische Studie.

B (mäßig hoch = Nutzen der Empfehlung statistisch gut möglich):

Weitere Forschungsarbeiten haben wahrscheinlich einen signifikanten Einfluss auf unser Vertrauen in die Richtigkeit unserer Aussagen und könnten uns zu einer Korrektur derselben veranlassen.

- ▶ Eine qualitativ hochwertige Studie.
- ▶ Mehrere Studien (1a, 1b, 2a), für die schwere Einschränkungen (Blutdruckstudie für Diabetiker) gelten.

C (niedrig = Empfehlung eher spekulativ):

Weitere Forschungsarbeiten haben höchst wahrscheinlich einen signifikanten Einfluss auf unser Vertrauen in die Richtigkeit unserer Aussagen und werden uns wahrscheinlich zu einer Korrektur derselben veranlassen.

- ▶ Eine oder mehrere Studien (2a, 2b) , für die schwer wiegende Einschränkungen angezeigt sind.

D (sehr niedrig = Keine statistische Belege für die Richtigkeit vorhanden):

Alle Aussagen sind mit einer großen Ungewissheit behaftet.

- ▶ Reine Expertenmeinung; Begründung anhand "guter ärztlicher Praxis"
- ▶ Keine aus Studien gewonnene direkte Evidenz.
- ▶ Eine oder mehrere Studien ("minderer statistischer Qualität"), für die schwere Einschränkungen gelten.

Aus: <http://www.ebm-guidelines.at/>, Gunter Frank: "Schlechte Medizin. Ein Wutbuch" Seite 60. KNAUS 2012

Anfang der achtziger Jahre ging einigen Wissenschaftlern und Ärzten der offensichtliche Unsinn, der in der Medizin als Wissenschaft etikettiert wurde, endgültig über die Hutschnur. Eine kanadische Arbeitsgruppe entwickelte deshalb ein System, das Licht in das Dunkel undurchsichtiger Studien bringen sollte. Die so genannte Evidenzbasierte Medizin bewertet die *Faktenlage* (englisch *evidence*) zu einer Fragestellung nach wissenschaftlichen Kriterien. Im Mittelpunkt des Interesses steht dabei die Qualität der Beweise für einen vermuteten Sachverhalt. (Nebenbei bemerkt: wenn man in der Medizin wirklich wissenschaftlich arbeiten würde, bräuchte man keine eigene Fachrichtung zur Qualitätskontrolle. Aber hier offenbart sich der Unterschied zwischen Medizin und Wissenschaft.)

Die Aussagekraft von Studienergebnissen hängt in hohem Maß vom *Studiendesign* ab, also davon, wie eine Untersuchung angelegt ist. Es ist keineswegs egal, ob ich Patienten rückwirkend (*retrospektiv*) über ihre Jugend befrage oder ob ich sie mit Blick in die Zukunft (*prospektiv*) über viele Jahre begleite und dabei das Auftreten von Krankheiten beobachte. Und es macht einen Unterschied, ob bei einer Studie Arzt und Patient informiert sind, wer das Placebo bekommt, oder ob die Studie „doppelblind“ angelegt wurde, so dass es keiner von beiden weiß. Ergebnisse besitzen ein anderes Gewicht, wenn mit Statistikprogrammen Daten gegeneinander gerechnet werden, bis etwas Passendes dabei ist (man würfelt so lange, bis ein Pasch fällt), oder wenn vor Beginn der Studie definiert wird, welchen Effekt man erwartet. Tritt er ein, war die These richtig, tritt er nicht ein, war sie falsch.

Für die **Bewertung der Evidenz** [„Faktenlage“] gibt es **4 Stufen**:

A ist dabei die Stufe, bei der der vorausgesagte Effekt einer Maßnahme sehr wahrscheinlich eintritt, **D** die Stufe, deren Wahrscheinlichkeit die der reinen Spekulation nicht übersteigt (siehe Tabelle).

Die Evidenzbasierte Medizin beantwortet übrigens nicht die Frage, ob eine Aussage richtig oder falsch ist. Angenommen, jemand behauptet, dass Ananas die Potenz steigert. Niemand weiß es, es wäre aber immerhin denkbar. Mit Hilfe von wissenschaftlichen Studien kann man nun die **Wahrscheinlichkeit** ermitteln, mit der Ananasesser mit einer höheren Manneskraft gesegnet sind.

Grad der Empfehlung	Evidenz-/Faktenlage-Stufe	Studientyp	Wertung für die Praxis
A	1a	Mehrere oder einzelne qualitativ hochwertige Studien. Diese Studien haben Kontrollgruppen, und die Verteilung auf die Gruppen erfolgt nach Zufall. Es gibt keine wesentlichen statistischen Fehler. Die Studien sind so angelegt, dass die zu untersuchende Hypothese vorher formuliert wurde und durch das Ergebnis eindeutig bestätigt wird.	Aussagekräftig. Man sollte ernsthaft darüber nachdenken, sein Verhalten daran auszurichten.
	1b		
B	2a	Mehrere oder einzelne Studien, ähnlich wie A, jedoch mit methodischen Mängel. Studien mit einwandfreier statistischer Methodik aber ohne Kontrollgruppe Fallberichte mit Kontrollen.	Lohnt sich, sich damit näher zu befassen.
	2b		
	3a		
	3b		
C	4	Studien ohne Zufallsverteilung, reine Beobachtungsstudien, Studien mit großem statistischen Mängeln, Fallbericht.	Noch nah der Spekulation
D	5	Konsensuskonferenzen, Expertenmeinungen	Spekulation

Die Anlage der Studie entscheidet über die Genauigkeit, mit der diese Wahrscheinlichkeit angegeben werden kann, sie ist ausschlaggebend für die Qualität der wissenschaftlichen Begründung – und genau dies drücken die **Empfehlungsgrade** aus. Eine Meinung kann also jeder haben, auch ein Wissenschaftler. Ob diese Meinung aber wissenschaftlich ist, hängt nicht vom akademischen Grad ab, sondern von einer hohen Evidenzstufe. Erst wenn eine solche vorliegt, ist es auch wissenschaftlich gerechtfertigt, den Menschen über Kampagnen zu irgendwelchen Änderungen ihrer Lebensgewohnheiten anzuraten

Die höchste Wahrscheinlichkeitsstufe, also **A**, erfordert folgendes Vorgehen. Eine Erhebung hat gezeigt: Die Frankfurter essen mehr Wiener Würstchen als die Berliner und bekommen weniger Krebs. Daraus wird die Hypothese abgeleitet, dass Würstchenverzehr vor Krebs schützt. Doch bis jetzt liegt noch kein Beleg vor, sondern eben nur eine statistische Korrelation, also ein paralleles Auftreten von zwei Faktoren. Um die Hypothese zu testen, führt man eine „*prospektive kontrollierte und randomisierte Interventionsstudie*“ durch. Dahinter verbirgt sich folgendes: Die Berliner werden in 2 Gruppen gewürfelt. Die eine darf weiter essen wie bisher, die

andere muss mehr Wiener Würstchen auf den Speiseplan nehmen. Wenn dann nach einer bestimmten Zeit in der Würstchengruppe tatsächlich weniger Krebsfälle festgestellt werden, liegt für die formulierte Hypothese ein Beleg mit dem hohen evidenzgrad a vor.

Selbstverständlich wäre es schön, wenn wir für jede Frage klare Beweise hätten. Aber das ist häufig nicht der Fall. Und oft muss gehandelt werden, noch ehe die allerletzten Fragen geklärt sind. Dann hilft es, wenigstens zu wissen, auf welchem **(Un-)Sicherheitsniveau** man sich bewegt. Natürlich gibt es auch **Erfahrungswerte**, die sinnvoll erscheinen, obwohl sie vielleicht nicht durch Studien abgesichert sind. **Und nicht zuletzt bedeutet eine hohe statistische Wahrscheinlichkeit keineswegs, dass etwas immer und in 100 % der Fälle eintritt. Eine letzte Ungewissheit bleibt. Daher kann selbst eine Maßnahme mit hohem Evidenzgrad im Einzelfall für einen Patienten unangebracht, ja sogar falsch sein. Wir sind schließlich immer noch Individuen und kein repräsentativer Mittelwert einer Gesamtpopulation.**

Die Evidenzbasierte Medizin fordert deshalb dazu auf, **Erfahrung** und **statistisch abgesichertes Wissen** zu kombinieren und dadurch nachvollziehbare und sinnvolle therapeutische Entscheidungen zu treffen. anhand der Bewertungskriterien kann der Arzt abschätzen, ob eine medizinische Maßnahme wissenschaftlich gesichert einen Nutzen verspricht. **Weil sie individuelle Besonderheiten und statistische Grenzen anerkennt, kann die Evidenzbasierte Medizin jedoch nie verbindliche Vorschrift sein.** Vielmehr liefert sie die Grundlage, anhand der Patient und Arzt zusammen entscheiden können, ob eine bestimmte Maßnahme im konkreten Einzelfall die richtige ist. Schließlich geht der Patient ja zum Arzt und nicht zum Statistiker. Aus: Udo Pollmer, Susanne Warmuth, Gunter Frank: „Lexikon der Fitness-Irrtümer: Missverständnisse, Fehlinterpretationen und Halbwahrheiten von Aerobic bis Zerrung“ Seite 156 - 158 PIPER 2003

Quellen: David Lawrence Sackett: „Levels of Evidence and grades of Recommendations“ Oxford Centre for Evidence-based Medicine 2001. In: <http://cebmljr2.ox.ac.uk/docs/levels.html>, www.ebm-netzwerk.de

[VI] Hans Peter Beck Bornholdt, Hans Hermann Dubben: „Der Hund, der Eier legt – Erkennen von Fehlinformation durch Querdenken“ rororo 2001, „Der Schein der Weisen – Irrtümer und Fehltritte im täglichen Denken“ Hoffmann & Campe 2001

[VII] Die **Konzierte Aktion** bezeichnet einen Abstimmungsprozess der Interessen zwischen unterschiedlichen wirtschaftspolitischen Akteuren, um unter Hintanstellung divergierender kurzfristiger oder nachrangiger Zielsetzungen ein mittel- oder langfristig besseres Gesamtergebnis zu erreichen. Quelle: Internet

Manfred Deix , österreichischer Cartoonist: „Arnold Schwarzenegger - Die nackte Wahrheit“ (2007)
Aus: <http://www.tagesspiegel.de/arnold-schwarzenegger-wie-der-karikaturist-ihn-sieht-/1116214.html>

[Meine Ergänzungen] Dr.med. Alois Dengg, Hollenzen 100, A-6290 Mayrhofen, www.dr.aloisdengg.at