Ich hatte ja vorgeschlagen, dass Sie sich als nächsten Schritt ein paar konkrete Beispiele suchen, wie Daten und wissenschaftliche Information in der Medizin verwendet werden. Dazu kommen wir dann, um alles etwas zu konkretisieren, in einem der nächsten Blog-Beiträge. Falls Sie Ihre Hausaufgabe schon gemacht haben, ist das sehr gut, denn dann verstehen Sie das, was ich jetzt sagen werde, noch besser.

Dann werden Sie nämlich gemerkt haben: „Evidence Based Medicine“ (EBM) unterstellt eine Hierarchie von Erkenntnis. Sie geht davon aus, dass an der untersten Stufe, sozusagen als Basis, klinische Anekdoten und unkontrollierte, klinische Erfahrung, also Expertenmeinungen stehen. Dann kommen Fallserien und Fallkontrollstudien, die schon etwas systematischer sind, aber eben immer noch Einzelfälle. Schliesslich kommen Studien an natürlichen Gruppen von Patienten, sogenannte Kohortenstudien, bei denen Patienten, die mit einer Methode behandelt wurden mit anderen verglichen werden, die anders behandelt wurden. Das Entscheidende daran: die Patienten, oder ihre Ärzte, haben die entsprechende Behandlung ausgesucht. Das sind also keine reinen Experimente, denn hier wird die Zuteilung auf die Bedingungen rein zufällig vorgenommen. Das hat, zumindest theoretisch, den Vorteil, dass dann die Patientenmerkmale zufällig verteilt werden und am Ende die beiden Gruppen besser vergleichbar sind. Damit wird es wahrscheinlicher, dass ein gefundener Unterschied zwischen den Gruppen auf die zu testende Intervention zurückzuführen ist und nicht auf zufällige Schwankungen in der Population.

Weil das so ist geht man auch davon aus, dass solche experimentellen Studien bei denen der Computer die Zuteilung der Patienten auf Gruppen per Zufall übernimmt – sog. „randomisierte Studie“ von engl. „random-zufällig“ – besser sind, und deswegen stehen sie in der Gunst der Wissenschaftler und auch der Herausgeber wissenschaftlicher Zeitschriften höher. Das hat seine Berechtigung; dazu kommen wir später. Schliesslich kann man alle Daten die man so in einzelnen Studien gefunden hat zu einer Meta-Analyse, also einer zusammenfassenden Studie, zusammenführen und erhält eine Schätzung des therapeutischen Effekts. Damit wird die Erkenntnis, so meint man, immer besser. Viele der Leitlinien, das werden Sie gesehen haben, stützen sich auf solche Daten.

Was steckt dahinter?

Das Leitbild von der Hierarchie der Erkenntnis („hierarchy of evidence“)
Hinter dieser Vorstellung steht ein explizit hierarchisches Verständnis von Wissen oder Erkenntnis („evidence“). Je stärker kontrolliert die Studie, je experimenteller das Vorgehen, umso sicherer sind die Schlussfolgerungen, die man aus einer Studie ziehen kann. Das ist zweifelsfrei der Fall. Daher geht man davon aus, dass das einzige, was wirklich wichtig ist, diese Sicherheit der Schlussfolgerungen ist. Das nennt man im Fachjargon die interne Validität, also die Gültigkeit einer Studie aufgrund ihrer methodischen Merkmale der

Kontrolle, der Zufallszuteilung und der Verblindung. Wenn man die Prämisse gelten lässt, dass diese Sicherheit der Schlussfolgerungen das einzig wirklich interessante und wichtige Merkmal von wissenschaftlichen Studien ist, dann kann man, in einem zweiten Schritt, sagen: alle anderen Studien haben weniger Gültigkeit, sind also weniger „gut“, können daher auch vernachlässigt werden, wenn solche Studien hoher „Güte“ vorliegen.

Dann kann man alle anderen Studien nur als Vorstufen und Vorbereitungen solcher endgültiger Studien verstehen und dann, wenn solche endgültigen, randomisierten Studien vorliegen, alle anderen vernachlässigen. In der Folge werden nur noch diese randomisierten Studien zusammengefasst: denn oftmals schwanken auch die Ergebnisse solcher „guter“ Studien beträchtlich und es ist schwierig, den Überblick zu behalten. Dies leistet dann die formale Meta-Analyse. Sie gibt ein statistisches Verfahren vor, wie man die Ergebnisse verschiedener Studien mathematisch zusammenführt und daraus eine Schätzung der vermutlich „wahren“ Effekgrösse gewinnt, die sich hinter einer Intervention in der Population verbirgt. Im Idealfall sind die Effekte der einzelnen Studien nämlich Schätzungen dieses „wahren“ Effektes, um den sie streuen.

Aus dieser Theorie leitet sich dann das Vorgehen ab, wie es in der Regel zur Generierung von „Evidenz“ im modernen, wissenschaftlichen Sinne verwendet wird. Man überprüft die vorhandenen formalen Merkmale der Studien und ordnet sie nach vermeintlicher „Güte“, also danach, ob sie experimentell sind oder nicht, und verwendet zur Einordnung noch andere Kriterien, zu denen wir später kommen werden. Liegen viele Befunde vor, wird die Lage leicht unübersichtlich, denn oft widersprechen sich Studien. Daher wenden Ersteller von Übersichtsarbeiten oder Reviews, bzw. Meta-Analytiker, gerne folgendes Verfahren an: sie formulieren Einschlusskriterien für ihre Arbeit, z.B.: „nur randomisierte Studien, die verblindet sind, in englischer Sprache verfasst und mit mindestens 50 Patienten pro Gruppe“. Dadurch verringert sich der Aufwand, wissenschaftliche Literatur zu beschaffen und zu sichten beträchtlich. Die Komplexität wird geringer und es wird leichter, eine klare Aussage zu treffen. Wir werden zu den einzelnen Schritten und ihren Implikationen später kommen. Hier interessiert zunächst nur das generelle Vorgehen, die Vorraussetzungen, die dahinter stecken und die Plausibilität des gesamten Ansatzes.

Wir sehen sehr leicht: Hinter diesem Vorgehen verbirgt sich die Leitmetapher der „Hierarchie“ oder der „Pyramide“, an deren Spitze das wirklich Wertvolle, das tatsächlich Interessante steht, nämlich die „wahre Effektgrösse“ einer Intervention, herausdestilliert durch die vielen einzelnen Schritte, die „unten“ getan wurden, die schliesslich zu den „wirklich guten“ Studien, den randomisierten, geführt haben und dadurch geholfen haben, den „echten Effekt“ zu finden. Der Rest ist unwichtig, wissenschaftliche Maische sozusagen, die nur dazu gedient hat, die Distille der Wahrheit zu bedienen; so ähnlich wie man ein 1000 Liter Fass Kirschen braucht um ein paar Liter Kirschschnaps zu brennen. Die Maische

kriegen die Säue, der Schnaps wird verkauft oder getrunken.

Wie plausibel und stichhaltig ist diese Theorie?
Wir analysieren diese Voraussetzung in drei Schritten. Erstens: Wir sehen uns die paradigmatischen Voraussetzungen ein bisschen genauer an. Zweitens: Wir analysieren die methodischen Hintergründe, die Entscheidungen und die Konsequenzen. Drittens: Wir kehren zurück zu unserem Beispiel der leitliniengestützten Therapie und fragen, ob dieses Vorgehen in der Praxis wirklich zu brauchbaren Ergebnissen führt.

1. Ist die Wirklichkeit hierarchisch?
Die Voraussetzung, die offensichtlich bei diesem Vorgehen gemacht wird, ist die der Hierarchie. Die Wirklichkeit, auch die Erkenntnis und die Einsicht, so unterstellt man, sei hierarchisch. Denken wir für einen Moment mal nach: Woher kennen wir Hierarchien, Pyramiden, und dergleichen? Offensichtlich sind alle feudalen Systeme der Vergangenheit hierarchisch geordnet. An der Spitze steht ein König oder Kaiser, ein Pharao oder Gott, von dem alle anderen ihre Rechte und Pflichten haben und an den alle anderen wieder in ihrer Verantwortlichkeit zurück Bericht erstatten und dem gegenüber sie verantwortlich sind. Feudale Systeme, das hat die Geschichte gezeigt, sind nicht die flexibelsten. Sie sind anfällig für Schwächen der Einzelnen. Solange ein sehr weiser, sehr starker und sehr gesunder Führer an der Spitze steht, kann das System gedeihen. Dann hat der Eine immense Lasten zu tragen und kann es vielleicht auch, und alle profitieren.

Historisch gesehen sind lange Friedenszeiten unter grossen Königen und Führern Beispiele dafür, wenn wir etwa an das Preussen Friedrichs des Grossen denken, an das Österreich unter Maria-Theresia, an Frankreich unter Ludwig dem Heiligen oder Ludwig dem XIV., an Rom unter Augustus oder Hadrian. Sobald ein Führer schwach ist, kränklich, nicht lange lebt, geistesgestört oder etwas dümmlich, wird es brenzlig. Nazideutschland ist ein Paradebeispiel dafür, wohin hierarchische Systeme tendieren, wenn es an der Spitze krankt. Daher haben politische Systeme in der Neuzeit Kontrollmechanismen eingebaut, die absolute Macht beschränken und die Hierarchie bzw. die damit verbundene Machtposition beschränken. Moderne politische Systeme, die funktionstüchtig sind, haben zwar eine hierarchische Struktur, vor allem um Konsens zu bündeln und nach aussen zu kommunizieren, um zu moderieren und Kräfte zu sammeln, aber sie sind immer über zyklische Rückmeldeprozesse und Kontrollprozeduren nach und von „unten“ mit ihrer Basis verbunden.

Wie sieht es mit grossen Wirtschaftsorganisationen aus? Auch hier zeigt sich: die erfolgreichsten Unternehmen haben über die letzten Dekaden begonnen, Hierarchien abzuflachen und sich modulare Strukturen zu geben. Zwar gibt es natürlich CEOs, die sehr entscheidungsgewaltig sind. Aber ein guter und moderner CEO wird sich hüten,

Entscheidungen einsam zu treffen, ohne alle Elemente seines Betriebes gut zu kennen. Auch hier gibt es flache Hierarchien, Rückkopplungen mit der Basis, Prozeduren, bei denen die Einzelnen Anteil an Entscheidungen und Zugriff auf Macht haben.

Wie steht es mit anderen wichtigen Bereichen im Leben? Liebe zum Beispiel? Dient alle Liebe zu verschiedenen Menschen, zu Eltern, Freunden, Kindern, anderen Menschen nur dazu, die eine wahre Liebe des Lebens zu nähren, die zu Ihrer Frau, Ihrem Mann, Ihrem Freund oder Lebenspartner, Ihrer Freundin oder Lebenspartnerin? Offensichtlich nicht. Offenbar ist es eher so: je mehr Liebe wir in einem Bereich erleben und mobilisieren können, umso mehr haben wir auch für andere Menschen zur Verfügung. Wenn eine Liebesbeziehung gut funktioniert und sich Partner lieben, dann strahlt diese Liebe auch auf die Kinder aus und zurück. Offensichtlich ist dies ein sich zyklisch bereichernder und beeinflussender Prozess.

Oder organismische Systeme? Nehmen wir das Immunsystem. Gibt es hier eine zentrale Immunerkennungsleitstelle? Nein, klarerweise nicht. Das Immunsystem ist ein hoch vernetztes, extrem effizientes, aber dezentrales System. Wenn, dann könnte höchstens das Gehirn als eine Zentrale gelten. Es schaltet sich aber allenfalls als Koordinationsfunktion zu, wenn bewusstes Verhalten gefragt ist, also, wenn uns ein Erreger so zu schaffen macht, dass wir unsere Aktivität einschränken müssen und mit Fieber ins Bett legen sollten. Ansonsten ist das System zyklisch organisiert und eine Hierarchie im Sinne von Steuerungsprozessen, die von oben kommen, wird nur im Ernstfall aktiviert.

Wie sieht es mit dem Bewusstsein und unserem neuronalen System aus? Dies ist ein interessantes Beispiel. Denn in den Anfängen der Hirnforschung, in der Phrenologie des 19. Jahrhunderts, war man der Meinung es gebe auch hier einen hierarchischen Aufbau und irgendwo müsste all die Information, all das Bewusstsein in einem speziellen Zentrum und in spezialisierten „Bewusstseinsneuronen“ zusammenlaufen. Dann müsste es „Grossmutterneuronen“ geben, die all die Information, die mir dazu dient meine Grossmutter zu erkennen, zusammenführen und dann den bewussten Erkenntnisakt: „Das ist meine Grossmutter“ herbeiführen. Diese hierarchische Vorstellung hat die Forschung lange gelähmt. Mittlerweile wissen wir, dass auch im neuronalen System ein hohes Mass an dezentraler, modularer Organisation herrscht, und dass das Bewusstsein irgendwie überall ist und durch die koordinierte Zusammenarbeit vieler neuronaler Zentren zustande kommt.

Es ist sogar so, dass unterschiedliche Aspekte eines Wahrnehmungsgegenstandes, also das Aussehen, der Geruch, der Klang der Stimme, die Art der Bewegung, die Körpergrösse meiner Grossmuter in ganz verschiedenen Zentren des Gehirns verarbeitet werden und erst durch Koordination vermutlich zeitlicher Muster zu einer Einheit im Bewusstsein verbunden werden. Hierarchie kommt zwar schon auch ins Spiel, aber eher so, dass bewusst gewählte

Ziele, Wertvorstellungen und Wünsche „von oben“, also aus Bereichen der Grosshirnrinde, hemmend auf andere Zentren wirken und so für eine gewisse Einheit sorgen. Auch hier haben wir also wieder eine sehr filigrane Mischung von zirkulär-modularen Prozessen und moderat hierarchischen, die aber immer in Rückkopplung mit anderen Prozessen aus tiefer liegenden Zentren zusammenarbeiten.

Ich könnte die Beispiele vermehren. Wir würden sehen: Es gibt keinen Bereich der uns bekannten Wirklichkeit der rein hierarchisch aufgebaut ist, ausser vielleicht dem Krankenhaus und der katholischen Kirche und vergleichbar veralteten Institutionen, die in ihrem Festhalten an veralteten Strukturen zu erkennen geben, dass sie Kandidaten für die Sackgasse der Evolution sind. Die Vorliebe der medizinischen Methodologie scheint mir aus dieser impliziten medizinischen Hierarchie zu kommen, die unreflektiert auf die Methodik übertragen wird.

Wenn man sich die Natur und die Wirklichkeit als paradigmatisches Beispiel nimmt, dann müsste man sich eigentlich einer etwas komplexeren Metapher bedienen. Dann müsste man zyklische Rückmeldeprozesse auch in den methodologischen Erkenntnisprozess einbauen.

2. Eine zirkuläre Alternative
Das Leitbild dafür ist ein Kreis oder meinethalben eine Spirale. Sicherlich ist Erkenntnis und Einsicht ein Ergebnis vieler einzelner Schritte und hat insofern eine Art hierarchische Struktur. Denn eine Einsicht ist mehr als die einzelnen Schritte, die dazu beigetragen haben. Der springende Punkt ist aber: wir gelangen zu dieser Erkenntnis nur, wenn wir alle Schritte in eine zyklische Gestalt bringen, also in Rückbezug zueinander und zu einer vermutlichen Synthese der Information. Technisch gesprochen: Wenn wir eine vermeintlich richtige Effektgrösse aus allen vorhandenen Studien isoliert haben, dann stellt sich die Frage: Wie bewährt sich nun diese Effektgrösse, die wir aus experimentellen Studien gewonnen haben, in der Praxis?

Können wir in der Praxis, bei allen möglichen Patienten, auch bei solchen, die gar nicht in den untersuchten Studien eingeschlossen waren, damit rechnen, dass wir solche Effekte sehen (und genauso wenig oder viel Nebenwirkungen wie in den Studien)? Funktioniert, was in einer experimentellen Studie bei Patienten ohne Vorliebe für irgendein Verfahren geklappt hat, auch bei Patienten, die skeptisch sind, die verzweifelt sind, die keinen Glauben mehr an sich und die Medizin haben? Lässt sich eine Intervention, die bei Patienten mit „einfachen“ Krankheiten, ohne Mehrfachdiagnosen getestet worden ist, auch bei Patienten anwenden, die den „wirklichen“ Patienten mehr gleichen, also bei solchen, die ausser Bluthochdruck auch noch Depression, Angst und Diabetes haben?

Denn: in experimentellen Studien, die zur Testung von Arzneien und Interventionen auf Wirksamkeit dienen wird zugunsten der internen Validität, also der Sicherheit der

Schlussfolgerungen, die Generalisierbarkeit der Ergebnisse, die externe Validität reduziert. In experimentellen Studien werden Patienten sorgfältig ausgewählt. Solche mit komplizierten Erkrankungen, mit mehrfachen Diagnosen, etc. gehen gar nicht erst in die Studie ein. Nur so funktioniert das gegenwärtige hierarchische Modell. Daher muss das Ergebnis, das so gewonnen wurde, rückgekoppelt werden mit der Basis, mit der Praxis. Hier müssen aber andere Studientypen durchgeführt werden. Denn wenn man wissen will, ob andere Patienten auch von einer Intervention profitieren, muss man möglichst ohne Ausschlusskriterien operieren und evtl. einfach eine breite Praxisdokumentation machen. Oder man muss natürlich vorkommende Gruppen vergleichen, solche die so und solche die anders behandelt werden. Wenn man wissen will, ob möglicherweise die aktive Entscheidung für eine Therapie therapeutische Potenz hat, darf man gar nicht experimentell untersuchen, denn die experimentelle Zuweisung schaltet das möglicherweise wichtigste Element im Therapieprozess aus, nämlich die Entscheidung des Patienten, etwas für seine Gesundheit tun zu wollen.

Ich plädiere also für einen zyklischen Erkenntnisprozess. Wer der Meinung ist, Hierarchien müssen trotzdem sein, der kann sich eine Hierarchie zyklischer Prozesse vorstellen, die im Stile einer Spirale unsere Erkenntnis verfeinern und verbessern. Wie dies funktionieren kann, dazu mehr in einem der nächsten Beiträge. Genauer habe ich dies ausgeführt und begründet in der unten angegeben Literatur. Und interessanter Weise sehen auch führende Methodiker das in zwischen ganz ähnlich.

Literatur:

- Tinetti, M. E., & Studenski, S. A. (2011). Comparative effectiveness research and patients with multiple chronic conditions. New England Journal of Medicine, 364, 2478-2481.
- Walach, H. (2006). Zirkulär statt hierarchisch – methodische Überlegungen zur Evaluation komplexer therapeutischer Massnahmen [Circular in stead of hiearchical – methodological considerations for evaluating complex therapeutic interventions]. In W. Bödeker & J. Kreis (Eds.), Evidenzbasierung in Gesundheitsförderung und Prävention (pp. 33-56). Bremerhaven: Wirschaftsverlag NW.
- Walach, H., Falkenberg, T., Fonnebo, V., Lewith, G., & Jonas, W. (2006). Circular instead of hierarchical – Methodological principles for the evaluation of complex interventions. BMC Medical Research Methodology, 6(29).