Interview mit Prof. Dr. Jürgen Windeler, Leiter des Instituts für Qualität und Wirtschaftlichkeit im Gesundheitswesen

„Register sind für Nutzenfragen nicht geeignet“

Prof. Dr. med. Jürgen Windeler wurde am 1. September 2010 als Nachfolger von Prof. Dr. med. Peter Sawicki (MVF 05/2009) zum Leiter des Instituts für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) berufen, vor allem – so wird vermutet – weil der durchaus streitlustige Arzt Sawicki wohl der Pharmaindustrie ein Dorn im Auge war. Das indes ist Windeler ebenso, wenn auch ein vielleicht weniger emotionaler, dafür aber weit methodischerer – manche sagen gar einer mit eher dogmatischen Zügen. Dabei klingt das, was er im Interview mit „Monitor Versorgungsforschung“ ausführlich darlegt, ganz anders als die verkürzten Aussagen seines Statements beim Pro-Contra-Forum „Register“ auf dem Deutschen Kongress für Versorgungsforschung (MVF 06/17).

>> In Ihrem Statement beim Pro & Contra-Forum „Register“ auf dem Deutschen Kongress für Versorgungsforschung erklärten Sie …

… gerne führe ich das gleich weiter aus. Aber lassen Sie mich vorab ein paar Worte zu dieser Session verlieren. Ich habe mich schon ein wenig gewundert über die Wellen, die mein Statement ausgelöst hat. Wenn man einen Redner, dessen Position man gut kennt, zu einem zudem zeitlich äußerst limitierten Contra-Statement einlädt, muss man sich nicht wundern, wenn dieser die ihm gestellte Aufgabe dann auch in gebotener Klarheit und Deutlichkeit wahrnimmt. Und ich neige dazu, solche Aufgaben ernst zu nehmen. Inhaltlich habe ich den Eindruck, dass die von mir vertretene Position und jene vieler Versorgungsforscher gar nicht so weit auseinander liegen. Insofern frage ich mich, warum es ein so großes Bemühen gibt, Grenzen und Gräben zu betonen, u. a. mit Äußerungen, die einfach nicht richtig sind. Damit meine ich unter anderem eine Äußerung von Prof. Neugebauer, dass das IQWiG den Begriff Register aus dem aktuellen Methodenpapier gestrichen habe – das ist einfach nicht wahr. Ebenso wenig stimmt es, dass in der frühen Nutzenbewertung keine patientenrelevanten oder patientenberichteten Outcomes vorliegen und die auch nicht bewertet würden. Es gibt in vielen Dossiers sehr wohl Daten und Informationen zur Lebensqualität oder Outcomes wie „Schmerz“ und „Symptome“ – natürlich nicht in jedem Dossier, aber sie sind vorhanden – z. B. bei Indikationen wie Rheuma oder Psoriasis.

Die Sie dann auch in der Bewertung berücksichtigen?

Es kommt darauf an. Denn hier kommen wir schon zur Methodik und zwar zur Frage der Aussagekraft von Studien und zur Unterscheidung zwischen interner und externer Validität.

Und auch zur Gretchenfrage, ob randomisierte Studien unter Idealbedingungen stattfinden …

… der nächste Punkt, der nicht stimmt. Randomisierung erfordert keine „Idealbedingungen“. Selbst Zulassungsstudien werden nicht unter „idealen“ Bedingungen durchgeführt. Ich weiß nicht, was man sich unter „Idealbedingungen“ vorstellt.

Dann gehen wir doch auf die Methodik ein und zwar gleich zur „Table 0“, in der alle Ein- und Ausschlusskriterien eines RCT zu finden sind. Beschreiben diese nicht schon alleine durch die Zahl der Ein- und Ausschlusskriterien ein künstliches Setting?

Sagen wir es einmal so: Es gibt durchaus Studien, die über Ein- und Ausschlusskriterien ein spezielles Setting oder eine relativ eingeschränkte Patientengruppe beschreiben, und die das auch mit einer bestimmten Intention so tun, nämlich um die „Empfindlichkeit“ der Studie zu erhöhen. Das ist aber natürlich nicht konstitutiv für RCT. Denn RCT kann man ja auch durchführen, wenn man – wie jeder praktisch tätige Arzt – eine Indikation stellt, aber eben keine Alters- und Geschlechtsbeschränkungen einführt oder Multimorbidität ausschließt, also ganz allgemeine, breite Einschlusskriterien vorgibt. Es gibt solche Studien durchaus. Insofern sind Aussagen, die RCT mit „Idealbedingungen“ gleichsetzen, einfach falsch.

Ob das viel schlauer ist, breite Einschlusskriterien vorzugeben?

Ich glaube nicht, dass es sehr schlau ist, sie ganz eng zu machen. Aber auch ganz breite Einschlusskriterien – der Versuch, den sogenannten „Alltag“ abzubilden – können für die Interpretation ganz erhebliche Probleme machen.

Sie werden doch konzedieren müssen, dass die meisten RCT, die Sie in der Nutzenbewertung zu sehen bekommen, mit Sicherheit ziemlich eng gefasst sind.

„Randomisierung erfordert keine Idealbedingungen.“

Aber doch nur, weil sich die meisten der von uns durchgeführten Nutzenbewertungen auf die Arzneimittelzulassung beziehen. Bei diesen sind die Ein- und Ausschlusskriterien relativ eng gefasst – das ist schon richtig. Dennoch würde ich enge Ein- und Ausschlusskriterien nicht mit „Idealbedingungen“ übersetzen wollen, denn eine Idealbedingung ist etwas völlig anderes. Andererseits sollten sich Wissenschaft und Praxis – hier gibt es unterschiedliche und ungeklärte Positionen – sehr wohl die Frage stellen, ob es grundsätzlich negativ sein muss, wenn die Einschlusskriterien eng sind. Heißt das denn, dass die Ergebnisse nur für die damit beschriebene Population gültig sind und für keine andere?

Ergänzend die Frage, ob die so gefundenen Erkenntnisse übertragbar sind. Und wenn ja, wie?

Richtig. Ist das vielleicht gar nicht so schlimm, wenn die Einschlusskriterien eng sind, die Ergebnisse aber trotzdem für eine andere, eine breitere Population nutzbar sind? Das wäre eine Frage, über die es sich intensiv zu diskutieren lohnt.

Die aber nicht diskutiert wird. Warum eigentlich? Weil man sich vielleicht scheut, hinter den RCT ein Fragezeichen zu stellen?

Diese Frage ist doch eher ein Ausrufezeichen für RCT! Man scheint offenbar der Auffassung zu sein, dass das, was man in der RCT-Po-

pulation sieht, so auch für andere, breitere Patientengruppen gilt oder anwendbar ist. Das wäre ein Ausrufezeichen für RCT, weil damit die praktische Gültigkeit von RCT über die an sich enge Population hinaus beschrieben würde. Wenn man dagegen Gründe hat, anzunehmen, dass die Effekte in anderen Populationen anders – viel kleiner oder auch viel größer – sind, dann ist das auch kein Fragezeichen an RCT, sondern ein guter Grund, für diese anderen Populationen die offene Frage mit RCT zu beantworten. Nur inwieweit die Ergebnisse von RCT verallgemeinerbar sind und wann man weitere braucht, da gehen die Auffassungen weit auseinander – wenn auch vermutlich jeder zustimmen würde, dass es für die Übertragbarkeit von in RCT erzeugten Ergebnissen irgendwo eine Grenze gibt.

Doch warum wird denn diese Diskussion nicht geführt?

Sie wird ja geführt, aber es existieren kaum Kriterien und viel zu wenig Daten dazu. Es gibt zwar einzelne Studien, die zeigen, dass der Unterschied vermutlich nicht groß ist, doch mehr als solche sporadischen Daten gibt es nicht. Man kann aber den Eindruck bekommen, dass jeder annimmt, die Effekte in der breiten Versorgung – also in den Situationen und Populationen, die vielleicht nicht in RCT betrachtet werden – seien kleiner als in RCT. Denn wenn man davon spricht, dass RCT unter „Idealbedingungen" stattfinden, bedeutet das automatisch, dass die hier gefundenen, vielleicht relativ großen Effekte „draußen" kleiner sein werden. Das aber scheint niemanden wirklich zu beschäftigen, jedenfalls hat es keine praktischen Konsequenzen.

Weil man das vielleicht gar nicht so genau wissen will?

Sollte man aber, jedenfalls diejenigen, die Arzneimittel verordnen. Dass die Industrie das vielleicht nicht wissen will, mag man ja verstehen. Doch ist die Industrie bei RCT sowieso auf einem ganz anderen Trip, was das Ganze noch unübersichtlicher macht.

Auf welchem?

Die Industrie meint, wir benötigten unbedingt mehr Real-World-Evidence – also mehr Daten aus dem sogenannten Alltag. Man darf wohl vermuten, dass sich die Industrie davon etwas verspricht, sonst würde sie in diesem Gebiet nicht so drängen.

Wie lautet Ihrer Meinung nach der Beweggrund?

Man kann doch vermuten, was sicher nicht so abwegig ist, dass sich die Industrie davon irgendetwas verspricht, sonst würde sie in dieses Gebiet auch nichts investieren.

Eben die Vermutung, dass zum Beispiel die Effekte im Alltag größer sein könnten als in RCT?

Das passt doch irgendwie nicht zusammen – oder?

Sie meinen, dass das eine falsche Hoffnung ist?

Zumindest eine relativ substanzlose Hoffnung oder besser: offensichtlich paradoxe Vermutungen und sehr wenig Wissen.

Ist das eine Frage, der sich originär das IQWiG annehmen soll oder muss?

Mit unseren Herbstsymposien nehmen wir uns seit Jahren aktuellen Problemen an. So haben wir in diesem Jahr genau dieses Thema aufgegriffen, nachdem es der Gesetzgeber mit dem sogenannten Evidenztransfer oder der Extrapolation in das Gesetz geschrieben hat. Wir diskutieren hier die Frage: Wie weit und mit welchen Methoden können wir Ergebnisse aus Studien über die untersuchte Situation und Population hinaus extrapolieren? Wir hoffen, dass wir die Diskussion damit voran bringen und versachlichen.*



„Extrapolation ist ein methodisch sehr schwer greifbares Problem."

Ist das denn Ihre Aufgabe?

Sicher! Es gehört zu unseren Aufgaben, sich solchen Themen zu widmen, weil unser Institut immer mit derartigen Fragen und Situationen konfrontiert sein wird, wenn wir Nutzenbewertungen für was auch immer machen sollen.

Könnten Sie nicht auch – wie bei CA und AHP geschehen – einen Piloten aufsetzen?

Daraus ein Projekt zu machen, wird sehr schwierig. Darum müssen wir uns im Moment damit abfinden, zumindest aktiv eine Diskussion anzustoßen.

Warum kann ein solcher Pilot nicht stattfinden?

Weil es die entsprechende Methodik, die man in einem Piloten ausprobieren könnte, gar nicht gibt. Zwar gibt es erste Überlegungen der EMA** zur Übertragbarkeit von bei Erwachsenen erforschten Arzneimitteln auf Kinder, aber das war es dann auch schon. Ich befürchte allerdings, wenn man die dort beschriebene Extrapolation wirklich umsetzen möchte, wird man erstaunt sein über die Höhe der Anforderungen.

Woran liegt es denn eigentlich, dass es bei uns in Deutschland keine Institution gibt, die sich um so grundlegende Fragen der Methodik kümmern könnte oder würde?

— — — —

* Anmerkung der Red.: Den Bericht über das IQWiG-Herbstsymposium 2017 lesen Sie auf Seite 18-21.

** Reflection paper on extrapolation of efficacy and safety in paediatric medicine development; 2016

Zunächst: Extrapolation ist ein methodisch sehr schwer greifbares Problem. Es ist jedenfalls in wesentlichen Teilen eigentlich ein Bewertungsprozess mit weiten Ermessensspielräumen. Ich bezweifle, dass es methodisch im engeren Sinne zu lösen ist. Aber man kann sicher versuchen, eine Systematik zu entwickeln. Wichtiger und erfolgversprechender ist es allerdings, Prozesse zu beschreiben, in deren Rahmen sich die Bewertung der Übertragbarkeit bzw. die Extrapolation bewegen soll. Akademische Institutionen, Fachgesellschaften und auch das Netzwerk Versorgungsforschung sind aufgerufen, solche Fragen in wissenschaftlichen Teams zu bearbeiten. Doch spielt das Thema in der Debatte außerhalb von etwas kritischeren akademischen Kreisen so gut wie überhaupt keine Rolle. Was verwundert, weil sich diese Frage doch eigentlich viele praktische Behandler stellen müssen: Wie kann ich Erkenntnisse aus Zulassungsstudien auf meine Patienten übertragen? Das müsste Ärzte doch umtreiben!

Was es aber anscheinend nicht tut, weil sie laut einer Studie der TK die in der Nutzenbewertung gewonnenen Daten eh nur zu 11 % kennen.

Ich stelle mir vor, dass die meisten Ärzte durchaus wissen, welche vergleichsweise beschränkten Patientenpopulationen in den jeweiligen Zulassungsstudien der Arzneimittel, die sie einsetzen, betrachtet werden. Meine Hypothese ist, dass die meisten davon ausgehen, dass die Erkenntnisse schon etwas für ihre Patienten bedeuten werden.



Worin ist diese Studiengläubigkeit begründet? Etwa im Einfluss der EbM, die – so zumindest die These, die Dr. Konrad Schily im Titelinterview in MVF 05/17 vertrat – bessere Ergebnisse erzielte, wenn sie sich mehr auf die ärztliche Erfahrung stützen und aus dieser heraus eine rationale Medizin weiterentwickeln würde?

„Studiengläubigkeit" klingt mir zu negativ. Vielleicht machen die beschriebenen Ärzte es ja genau richtig, indem sie die Studien kennen und ernst nehmen und daraus das Beste für ihre Patienten ableiten. Ich würde andererseits zustimmen, dass wahrscheinlich bei vielen Anwendern die Frage, woher die Evidenz ganz genau kommt, wie sie entstanden ist und ob sie überhaupt übertragbar ist, wahrscheinlich eine ziemlich untergeordnete Rolle spielt. Dass Ärzte AMNOG-Nutzenbewertungen nicht intensiv zur Kenntnis nehmen, hat meines Erachtens aber nichts mit EbM als solcher, sondern mit anderen Regularien und Entscheidungsprozessen im Rahmen des SGB V zu tun.

„RCT sind nicht die Grundlage von EbM, sondern sie sind für die Frage nach dem Nutzen die fehlerärmsten Instrumente."

Wäre es denn sinnvoll?

Evidenzbasierte Medizin ist eine notwendige Grundlage; sie ist nicht alles, wohlgemerkt, aber eine sinnvolle Grundlage. Dr. Schily hat gesagt, dass er noch kein Beispiel gefunden habe, „in dem der randomisierte Doppelblindversuch etwas Sicheres zu Tage gefördert" hätte. Ich weiß nicht, was Herr Kollege Schily in den letzten Jahrzehnten wahrgenommen hat, aber die Wahrheit ist eine andere und Beispiele für hoch relevante praktische Erkenntnisse gibt es zuhauf. Darum würde ich mir wünschen, dass in medizinischen Entscheidungen und Beratungsprozessen sowie in Gesprächen mit den Patienten Evidenz eine Rolle spielt. Auf der Grundlage vorhandener Evidenz können dann Ärzte zusammen mit den Patienten überlegen und aushandeln, was deren Präferenzen und Ziele sind und welche Therapie im individuellen Einzelfall die beste ist. Das ist ein hoch sinnvolles und ein für alle breiten Spielraum lassendes Vorgehen. Nur muss man halt die dafür nötigen Grundlagen kennen.

Womit wir zurückkehren zum RCT als Grundlage für EbM, wobei das Stichwort RCT ein breites Bündel an Instrumenten beschreibt, die man einsetzen kann – von Basket- und Umbrella-Designs, adaptiven Designs bis hin zu clusterrandomisierten und pragmatischen Studien.

RCT sind nicht die Grundlage von EbM, sondern sie sind für die Frage nach dem Nutzen die fehlerärmsten Instrumente – nicht mehr und nicht weniger. Es mag jetzt meinem Alter oder auch meiner Biografie geschuldet sein, aber das ist doch alles nichts Neues. Adaptive und pragmatische Designs diskutieren und machen wir seit mehr als 20 Jahren, clusterrandomisierte Studien seit mindestens 30. Wenn man neuere Entwicklungen wie Basket- und Umbrella-Trials einmal außen vor lässt, die in den letzten Jahren aus den molekulargenetischen Entwicklungen entstanden sind, sind andere Variationen von RCT seit langem bekannt. Wir müssen dies aber anscheinend immer wieder neu erklären und da kann man dann auch schon einmal ungeduldig werden. Dabei gibt es viele Beispiele aus der Literatur der letzten Dekaden, in denen alles theoretisch und methodisch begründet – und auch vermittelt und konkret gemacht worden ist.

Wurden denn in AMNOG-Dossiers beispielsweise pragmatische Studien aufgenommen, die in die Nutzenbewertung eingeflossen sind?

Das hängt davon ab, wie man pragmatische Studien definiert. Es ist ja nicht so, dass ein Studientyp vom anderen scharf abgegrenzt wäre. Es gibt daher in den Nutzenbewertungen viele Studien, die pragmatische Komponenten haben. Wenn es zum Beispiel in der Onkologie gleichwertig angesehene Therapiealternativen gibt und es den Ärzten überlassen wird, welche Therapiealternative sie aus einem



Im Zentrum des Gesundheitswesens? Der Patient natürlich. Wer sonst?

Darum gibt es unsere elektronische Gesundheitsakte. Das vitabook-Gesundheitskonto verschafft Bürgern die volle Kontrolle über ihre Daten. Denn vitabook ist unabhängig, genauso wie unsere Nutzer.

Alle Daten eines Patienten gespeichert an einem Ort. Unter seiner alleinigen Regie. Die Basis besserer und zeitgemäßer Gesundheitsversorgung.

www.vitabook.de

vitabook

bestimmten Spektrum auswählen dürfen – ohne dies vorzuschreiben –, würde ich das als pragmatischen Aspekt betrachten.

„Mein Nein oder Contra bei der anfangs erwähnten Veranstaltung bezieht sich darauf, aus Registern Nutzenaussagen ableiten zu wollen."

Das könnte aber auch ein Basket-Ansatz sein.

Meinetwegen. Pragmatische Studien arbeiten oft mit möglichst wenig Einschlusskriterien, um ein möglichst breites Anwendungsspektrum abdecken zu können. Das kann sehr sinnvoll sein. Solche Studien können aber sehr problematisch werden, z. B. weil Ärzte und Patienten sich nicht an die ursprüngliche Randomisierung halten, alles fröhlich durcheinander geht und am Ende die Aussagekraft gegen Null geht.

Ist das nicht der Preis für Pragmatismus?

Ja, das kann man durchaus so sagen, vielleicht ein zu hoher Preis. Man wird darum auch in pragmatischen Studien gewisse Vorgaben definieren und Leitplanken einbauen müssen. Bei Studien nach einem „Anything-Goes-Prinzip" wird man am Ende mit leeren Händen dastehen.

Nur: Wie eng müssen die Grenzen denn sein?

So eng wie nötig, dass Aussicht auf die Beantwortung einer Frage besteht. Der zweite Punkt ist natürlich der, dass die Grenzen auch in den relativ engen Zulassungsstudien nicht eins zu eins eingehalten werden, womit wir wieder beim Stichwort „Idealbedingungen" wären. Es passiert eben schon einmal, dass ein Patient mit 75 Jahren eingeschlossen wird, obwohl der Cutoff mit 72 angesetzt war. Auch werden Patienten ihre Medikamente nicht immer zuverlässig einnehmen – obwohl ihnen natürlich nahegelegt wird, dass das speziell für die Zulassung wichtig wäre. Darum haben eben solche Studien mit Idealbedingungen nichts zu tun. Eine scharfe Trennung zwischen „Ideal" und „Alltag" existiert nicht, genauso wenig wie der dauernd betonte Unterschied zwischen Efficacy und Effectiveness. Ich halte das einfach für Quatsch.

Wie das?

Die Begriffe sind schlecht definiert, sie sind eigentlich inhaltsleer und insbesondere eben nicht klar unterscheidbar. Es gibt in Studien eher pragmatische Komponenten und mehr oder weniger ideale Komponenten. Eine ideale Komponente könnte sein, dass alle Patienten nach genau vier Wochen wieder zur Nachuntersuchung kommen. Aber: Passiert das auch? Oft eben nicht! Keine Studie ist „Ideal" oder „Alltag". Jede Studie hat ihre pragmatischen und ihre nicht so pragmatischen Komponenten, womit jede Studie etwas über Efficacy und auch über Effectiveness aussagt, oder über eine Mischung von beidem. Die Begriffe und die Trennung machen daher definitorisch und auch praktisch überhaupt keinen Sinn. Und sie führen auch noch dazu, dass man dann auf die Idee kommt, Effectiveness, den Nutzen im sogenannten Alltag neu definieren und abgrenzen zu müssen, um ihn – das ist die regelhaft abgeleitete Konsequenz – mit „neuen" Methoden untersuchen zu müssen. Dabei weiß ich schon nicht, was genau dieser „Alltag" sein soll.

Weil einfach niemand ganz genau weiß, wie sich der Versorgungsalltag mit all seinen Determinanten von Arzt-Patienten-Interaktion bis hin zum Versorgungssetting auswirkt?

Nein, weil es doch nicht einen „Alltag" gibt, sondern viele „Alltage". Wer sagt denn, dass eine Studie, die einen „Alltag" abzubilden versucht, nicht viele andere „Alltage" ignoriert? Ich möchte gerne wissen, was in der Versorgung eigentlich genau passiert. Doch: Das ist etwas völlig anderes als Nutzen in irgendwelchen „Alltagen".

Wenn Sie das ausführen würden?

Nehmen wir eine Arzneimittelzulassung. Da steht ein Anwendungsgebiet drin, Dosierungen, möglicherweise die Dauer der Therapie, notwendige Kontrolluntersuchungen usw. Versorgungsforschung im besten Sinne würde sich darum bemühen herauszufinden, ob das eigentlich alles so passiert, welche Abweichungen es vielleicht gibt und welche Gründe es dafür gibt. Hierzu gibt es verschiedene geeignete Forschungsmethoden inklusive qualitativer Forschung, die hinterfragt: Warum passiert etwas eigentlich nicht oder nicht so wie geplant?

Das wird dann an Innovationshemmnissen, Sozialisierungen und vielleicht auch dem Glauben liegen, dass das, was man bisher gemacht hat, so falsch nicht gewesen sein kann.

… und dazu werden sicher auch Struktur- und Schnittstellenprobleme bzw. -erkenntnisse kommen. Es wird für vieles sogar gute und auch vernünftige Gründe geben. Zusammen zeigt das: Wir wissen viel zu wenig darüber, was in der Versorgung wirklich passiert und warum. Wir brauchen daher eine Beschreibung der Versorgung, eine Analyse der Versorgung, allgemein eine Beforschung der Versorgung.

Ein Instrument, das hier eingesetzt werden könnte, wären gute Register. Dazu werden Sie doch nicht Nein sagen können.

Zu dieser Zielsetzung von Registern sage ich sogar ausdrücklich „Ja". Mein „Nein" oder „Contra" bei der anfangs erwähnten Veranstaltung bezieht sich darauf, aus Registern Nutzenaussagen ableiten zu wollen.

Warum?

Weil man weiß, dass das nicht funktioniert. Das ist doch der Punkt.

Darum noch einmal nachgefragt: Was weiß man, dass was nicht geht?

Dass Register für Nutzenbewertungen nicht das richtige Instrument sind.

Nachgehakt: Warum ist das Ihrer Meinung nach so?

Register sind erst einmal und überhaupt nur Datensammlungen und selbst keine Studien. Nun weiß man aus rund 50 Jahren epidemiologischer Forschung, dass Studien, die auf Registerdaten aufsetzen und die nicht randomisiert sind, u. a. in der Anforderung der Strukturgleichheit – die für Nutzenaussagen nun einmal von ausschlaggebender Bedeutung ist – die Datenqualität von RCT nicht erreichen. Das ist epidemiologisches Basiswissen.

Das würde ich Ihnen absolut zustimmen. Doch können solche Registerstudien durchaus einen Hinweis geben, um weitere Evidenz schaffen zu können.

Das könnte man ja diskutieren, um dem Leitsatz der Beobachtungsforschung gerecht zu werden, dass solche Studien Hypothesengenerierend sind. Was heißt: Nachdem die Daten eines Registers mit

DocMorris
Die Apotheke.

GEMEINSAM ZU MEHR THERAPIEERFOLG

indikationsbezogene Patienten-Begleitprogramme

Zugang zum Patienten



Stärkung der Adhärenz

Verbesserung der Lebensqualität

Werden Sie Partner!

Mehr Infos unter DocMorris.de oder per E-Mail an patientenprogramme@docmorris.de



großem Aufwand ausgewertet und qualitätsgesichert sind, muss das Ergebnis – man hat z. B. einen Vorteil von A gegenüber B gefunden – erst noch mit einem RCT veri- oder falsifiziert werden. Das kann doch nicht Ihr Ernst sein oder?

Wie soll man sonst Hypothesen generieren?

Die Hypothese muss doch gar nicht generiert werden, es gibt sie ja schon: Diese Intervention hat in dieser Population einen Nutzen oder keinen Nutzen. Und diese Hypothese prüft man, indem man mit einem RCT beginnt!

Gleiches gilt doch auch für eine Registerstudie. Die zu erforschende Frage steht doch von Anfang an fest und wird nicht erst im Zuge der Registerauswertung überlegt.

Genau. Die Frage oder die Hypothese steht bereits am Anfang fest: Gibt es einen Nutzen? Die muss das Register weder generieren noch kann es sie beantworten.

Nehmen wir die Frage des Zusatznutzens von Therapie A gegen B. Nach Ergebnis der Zulassungsstudien gibt es einen Zusatznutzen von A und die Hypothese ist nun, dass dieser im sogenannten Alltag anders ist.

Dazu sagt mir die Registerstudie, dass es tatsächlich so aussieht, als ob Therapie A besser ist als Therapie B. Das ist das Ergebnis für die Hypothese, die das Register generieren kann, die Registerforscher mit großem Aufwand generieren können. Und dann schreiben die gleichen Registerforscher, dass sie dieses Ergebnis aber nicht glauben, und dass es nötig ist, dieses Ergebnis, in dessen Erarbeitung ich bereits einen sehr hohen Aufwand gesteckt habe, noch einmal in einem RCT zu verifizieren?

Sie würden darum lieber gleich ein RCT aufsetzen und die Hypothese selbst generieren?

Nein, eben nicht generieren. Ich will sie prüfen! Und zwar mit der bestmöglichen Aussicht auf Antwort. Die Hypothese ist doch schon generiert. Die Hypothese lautet: Ich vermute, dass Therapie A besser ist als B. Meines Erachtens geht man mit Registerforschung in Nutzenfragen einen unnötigen Umweg mit fragwürdigen Verzögerungen. Man hat eine Frage und man hat ein besonders geeignetes Instrument. Im Alltag ist klar, was man macht.

Das müssen Sie erklären.

Ganz einfach: Jemand will einen Nagel in ein Brett einschlagen. Jeder weiß, dass das mit einem Hammer am besten geht. Dieser jemand hat zwar einen Hammer, nimmt aber nicht den Hammer, sondern eine Gartenschere, weil er meint, das müsse auch funktionieren. Nach einigen Schlägen ist die Gartenschere kaputt und der Nagel krumm, aber nur ein paar Millimeter eingeschlagen. Der Jemand nimmt schließlich den Hammer und haut den Nagel ins Holz.

Die Erkenntnis lautet demnach: Um einen Nagel in ein Brett zu schlagen, nimm den Hammer, weil dafür die Gartenschere nicht so gut geeignet ist.

Richtig, tolle Erkenntnis. Und ich habe eine kaputte Gartenschere und viel länger gebraucht, bis ich den Nagel schließlich da habe, wo er hin soll. Das ist für mich Ressourcenverschwendung. Gleiches gilt entsprechend für den Versuch, mit Registerforschung Nutzenfragen zu beantworten. Aber jetzt bitte nicht missverstehen: Damit sage ich nichts gegen Register als solche. Mit guten Registern kann man Fragen zur Adhärenz der Patienten, zum Verschreibungsverhalten der Ärzte, zu Kontrolluntersuchungen und vielem anderen mehr beantworten: Was passiert in der Versorgung wirklich? Wie ist die Einstufung der Patienten? Wie sind die Prozesse? Hier sind Register gut geeignet – für Nutzenfragen aber eben nicht.

Herr Prof. Windeler, danke für das Gespräch. <<

Das Interview führte MVF-Chefredakteur Peter Stegmaier.

Zitationshinweis

Stegmaier, P.: „Register sind für Nutzenfragen nicht geeignet“, in „Monitor Versorgungsforschung“ (01/18), S. 6-12; doi: 10.24945/MVF.01.18.1866-0533.2055

Prof. Dr. med. Jürgen Windeler

ist seit 1. September 2010 Leiter des Instituts für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) in Köln. Davor war er von 1999 bis 2004 Leiter des Fachbereichs Evidenzbasierte Medizin beim Medizinischen Dienst des Spitzenverbandes Bund der Krankenkassen e.V. (MDS) in Essen und von 2004 bis 2010 Stellv. Geschäftsführer und Leitender Arzt beim Medizinischen Dienst des Spitzenverbandes Bund der Krankenkassen e.V. in Essen. Zudem hat er seit 2001 eine außerplanmäßige Professur für Medizinische Biometrie und Klinische Epidemiologie an der Ruhr-Universität Bochum.