Prof. Dr. Karl Wegscheider im Titelinterview:

„Politischer Wille zur Transparenz ist nötig“

Die größten Hindernisse sieht Prof. Dr. Karl Wegscheider, Direktor des Instituts für Medizinische Biometrie und Epidemiologie am Universitätsklinikum Hamburg Eppendorf und u.a. Vorsitzender des Wissenschaftlichen Beirates des Bundesversicherungsamtes (BVA) zur Evaluation der Disease-Management-Programme, nicht in der Frage der Studiendesigns, sondern in der weit verbreiteten Verfolgung von Partikularinteressen im Gesundheitswesen. Die Verbesserung der Versorgung steht nach Meinung Wegscheiders an zweiter Stelle hinter den eigenen wirtschaftlichen oder politischen Interessen der handelnden Parteien, die entgegen vielen Lippenbekenntnissen oft gar kein Interesse an Transparenz hätten. Sein Vorwurf: „Überlebte Strukturen werden abgeschottet und verteidigt, als wären sie nicht Mittel zum Zweck, sondern selbst der Zweck des Gesundheitswesens.“ Entsprechend sei Forschung häufig nicht wirklich gewollt und Ergebnisse, die zu Konsequenzen führen müssten, würden ausgesessen.

>> Sehr geehrter Herr Professor Wegscheider, Sie sind nicht nur Forscher, sondern durch Ihre Kinderlähmung auch Betroffener. Was haben Sie eigentlich als Patient bisher von der Versorgungsforschung gehabt? Und welche Wünsche haben Sie für die Zukunft?

Die intensivierte Versorgungsforschung trägt dazu bei, dass sich in Deutschland ein wachsendes Bewusstsein für Qualität in der medizinischen Versorgung entwickelt. Während früher Forschung und Öffentlichkeit fast ausschließlich auf die spektakulären Möglichkeiten moderner Medizin schauten, wird heute auch immer häufiger die Frage gestellt, ob und wie diese Medizin zum Patienten kommt und wer davon überhaupt profitieren kann. Ich merke, dass einige (leider nicht alle) Ärzte, die mich behandeln, offener geworden sind z.B. für die Frage, was einem multimorbiden, mit Einschränkungen lebenden Patienten wie mir wirklich nutzt und was nicht. Was ich jedoch oft vermisse, ist eine Bereitschaft, sich selbst an der Forschung zu beteiligen und aktiv an der Verbesserung von Strukturen mitzuarbeiten.

Goldstandard der klinischen Prüfung von Arzneimitteln ist die randomisierte kontrollierte Studie (RCT). Gleichzeitig gibt es das Phasenmodell mit der stufenweisen Abklärung von Sicherheit und Wirksamkeit von der präklinischen Forschung bis zur sog. Phase-III-Studie, in der die klinische Wirksamkeit bei breiter Anwendung in zahlreichen Zentren nachgewiesen werden soll. Die Verantwortung für die sachgerechte und methodisch korrekte Durchführung der klinischen Prüfung liegt dabei bis zur Zulassung laut Gesetz bei den Herstellern. Dann endet diese Verpflichtung und es beginnt die Versorgungsforschung. Wie geht die Versorgungsforschung dabei am besten vor? Wie ist der heutige Standard? Wo sehen Sie Lücken und Probleme?

Versorgungsforschung gehört in der Tat bis heute nicht zu den Aufgaben des Herstellers. Entsprechend wird über die Versorgung zu wenig geforscht. Wenn geforscht wird, werden zudem häufig wegen der finanziellen Begrenzungen methodische Kompromisse eingegangen, die zu suboptimalen Forschungsergebnissen führen. Die billige und deshalb verbreitete Form der Phase-IV-Studie ist die Anwendungsbeobachtung, die zumeist reinen Marketingzwecken dient und dann kaum ernsthaft als Forschung bezeichnet werden kann. Es gibt aber auch die seriöse Versorgungsforschung, die mit ausgeklügelten Designs versucht, realitätsnahe und dennoch methodisch saubere Studien durchzuführen. Diese Art der Forschung wird an Bedeutung gewinnen, wenn wir stärker an die Erforschung des Nutzens von Therapien herangehen. Sie ist allerdings nicht billig zu bekommen und verlangt erfahrene methodische Begleitung, die unabhängig von Interessen sein sollte.

Versorgungsforschung macht eigentlich Phase-IV-Studien. Daraus hat sich ein neuer Studientyp entwickelt, die sogenannten Comparative Effectiveness Trials. Sehen Sie in diesen CET eine Brücke zu den im AMNOG geforderten „Versorgungsstudien“?

Der Streit um die Randomisation in Versorgungsstudien ist ein Streit um eine falsche Alternative. Man kann und sollte im Rahmen eines Studienprogramms unterschiedliche Designs realisieren. Randomisierte Studien weisen häufig eine bessere interne Validität auf, nicht-randomisierte Studien schneiden meist besser bei der externen Validität ab. Randomisierte Versorgungsstudien können als Substudien von nicht-randomisierten Studien durchgeführt werden. Studien mit unterschiedlichen Designs können aus einem umfassenden Register als Basispopulation erwachsen. Solche integrierten Modelle lassen zusätzliche Erkenntnisse zu. Comparative Effectiveness Trials kann man als Kompromisslösungen in einer Studie und damit als Brücken zwischen den anderen Formen ansehen. Sie können aber andere spezialisiertere Designs nicht grundsätzlich ersetzen. Die Wahl des richtigen Designs ist eine Frage der Zweckmäßigkeit, die möglichst unideologisch von Fachleuten geklärt werden und nicht vom Gesetzgeber geregelt werden sollte. Der deutsche Gesetzgeber hält sich im Übrigen durch die Wahl des nicht genau definierten Begriffes „Versorgungsstudie“ aus der Methodenfrage heraus und überlässt es ausdrücklich dem Gemeinsamen Bundesausschuss zu entscheiden, welche Studien er als Versorgungsstudien akzeptiert und welche nicht. Es gibt für mich keinen Zweifel, dass nur methodisch hochwertige Studien Sinn machen und akzeptabel sind. Die Qualität einer Studie wird aber nicht allein vom Studientyp bestimmt, sondern von vielen Faktoren.

„Die Qualität einer Studie wird nicht allein vom Studientyp bestimmt, sondern von vielen Faktoren.“

Sie haben am Beispiel der Akupunktur herausgearbeitet, wie sich die klassischen Wirksamkeitsstudien (Efficacy Trials) von diesen Comparative Effectiveness Trials unterscheiden. Wo liegen diese Unterschiede und was bedeuten sie? Welche unterschiedlichen Folgerungen kann man jeweils ziehen?

Wirksamkeitsstudien belegen die prinzipielle Wirksamkeit in dafür besonders geeigneten Kollektiven unter standardisierten Bedingungen im Hinblick auf einen relevanten klinischen Endpunkt mit einer bis auf die Studienintervention möglichst ähnlichen Kontrollgruppe. Sie sind produktorientiert. Comparative Effectiveness Trials orientieren sich an Strategie-Entscheidungen, die bei der Behandlung von Patienten unter Routinebedingungen bei Vorliegen einer bestimmten Diagnose zu treffen sind. Sie versuchen die Frage zu beantworten, was für verschiedene Patientengruppen in der Praxis unterm Strich simultan in verschiedenen patientennahen Endpunkten heraus-

kommt, wenn die Entscheidung so oder so ausfällt. Dabei wird das Produkt nicht allein, sondern in seinem Anwendungskontext bewertet, Kontrollgruppen sind meist Usual-care-Gruppen mit der lokalen Routinebehandlung. Anders als in der Phase-III-Wirksamkeitsstudie kann ein an sich wirksames Produkt im Comparative Effectiveness Trial zum Beispiel daran scheitern, dass es in der Praxis zu spät oder nur bei Patienten eingesetzt wird, die ohnehin ein geringes Risiko haben und es deshalb gar nicht benötigen.



Noch gibt es dazu keinen Konsens. Wo könnte eine Einigung liegen?

Efficacy Trials haben ihren festen Platz in den Phasen IIb und III vor der Zulassung. Spätestens nach der Zulassung sollten Nutzenaspekte eine stärkere Berücksichtigung finden. Generell sind im übrigen Efficacy und Effectiveness Trials keine scharfen Gegensätze. Es gibt fließende Übergänge. Moderne Phase-III-Studien weisen häufig schon zahlreiche Merkmale auf, die bereits der Methodik der Effectiveness Trials entlehnt sind. Umgekehrt wird kein Comparative Effectiveness Trial die Hauptwirkung einer Intervention außer Acht lassen.

Gibt es in anderen Ländern schon weitere Erfahrungen mit den Comparative Effectiveness Trials? Welche Probleme sind damit bisher untersucht worden? Und welche könnte man noch untersuchen?

Einzelne Studien, die wir heute als Comparative Effectiveness Trials bezeichnen würden, gibt es schon seit mehr als 50 Jahren. Außerhalb des Bereiches der Arzneimittelstudien gibt es viel Erfahrung mit derartigen Studien, etwa im Bereich der Medizinprodukte, der operativen Versorgung oder des Einsatzes diagnostischer Verfahren. Auch waren die frühen Arzneimittelstudien nach meinem Erleben noch weniger fokussiert auf das Nahziel der Zulassung als heute, praxisnäher und weniger standardisiert. Comparative Effectiveness Trials sind generell überall dort sinnvoll, wo Arzneimittel oder andere Interventionen in unterschiedlichen Patientengruppen oder im Rahmen von unterschiedlichen Versorgungsmodellen zur Anwendung kommen. Das ist die große Mehrheit der Arzneimittelverordnungen und klinischen Interventionen.

Die Denkweise orientiert sich hier an relativ kleinen Eingriffen, nämlich der Gabe von Arzneimitteln oder konkreten Behandlungen. Wie kann man komplexe Interventionen untersuchen, wie sie etwa die Einführung von Disease-Management-Programmen oder integrierten Versorgungsmanagement-Ansätzen darstellen? Welche passenden Methoden hat die Biometrie dazu in ihrem Werkzeugkasten?

Das Studium komplexer Interventionen läuft nicht prinzipiell nach anderen Regeln ab als klassische klinische Prüfungen. Allerdings sind die Anwendungsbedingungen vielgestaltiger, die Abläufe heterogener und die Zahl der beteiligten Institutionen und Personen ist größer. Es durchdringen sich mehrere Handlungsebenen, die im Design und im Auswertungsmodell einer Studie auseinanderdividiert werden müssen, um Fehlschlüsse zu vermeiden. Die Biometrie ist dafür gut gerüstet. In den letzten zwanzig Jahren sind eine Fülle moderner, teilweise raffinierter Designs erprobt worden, und es gibt hervorragende Software zur Analyse komplexer Designs. Hierzu sind z.B. clusterrandomisierte Studien zu zählen, bei denen nicht Patienten, sondern klinische Einrichtungen oder Arztpraxen randomisiert werden, bayesianische Studienansätze, die das vorhandene Vorwissen explizit einbauen, adaptive Designs mit explizit vorgesehenen Anpassungsmöglichkeiten während der laufenden Studie, Longitudinal-Modelle zur Analyse von Langzeitverläufen und vieles mehr.

In der regionalen Unterschiedlichkeit liegt auch eine Chance für die Versorgungsforschung. Wie könnte man denn zum Beispiel mit Stepped Wedged Design vorgehen, um sich der klassischen RCT zu nähern? Könnten nicht solche Studien in Deutschland häufiger durchgeführt werden? Welche Fragen könnten damit beantwortet werden? Wo sehen Sie praktische Möglichkeiten?

Im Allgemeinen werden in Deutschland Innovationen nicht wie damals die Disease-Management-Programme per Gesetz flächendeckend eingeführt, sondern Region für Region erst erprobt und dann schrittweise nach Kapazität eingeführt. Dabei werden Anpassungen an regionale Strukturen vorgenommen, und es wird aus vorliegenden Erfahrungen gelernt. Diesen Prozess kann man auch im Rahmen eines Studienprojektes optimieren und den Lernprozess modellieren. Das kann sogar unter Randomisierung der Einführungsreihenfolge geschehen. Mit solchen Verfahren kann man den Einführungsprozess transparent machen und Aussagen zum Nutzen eines Programms machen, die uns zum Beispiel bei den Disease-Management-Programmen mangels parallel dokumentierter Kontrollgruppen bitter fehlen. Solche Studien setzen allerdings eine zentrale Studienleitung voraus, an die die regionalen Entscheidungsträger bestimmte Kompetenzen bei der Gestaltung des Prozesses abgeben müssten. Hier könnten sich die föderale Struktur und das gegliederte Versorgungssystem als spezifisch deutsche Hürden erweisen.

„Solche Studien setzen eine zentrale Studienleitung voraus, an die die regionalen Entscheidungsträger bestimmte Kompetenzen bei der Gestaltung des Prozesses abgeben müssten.“

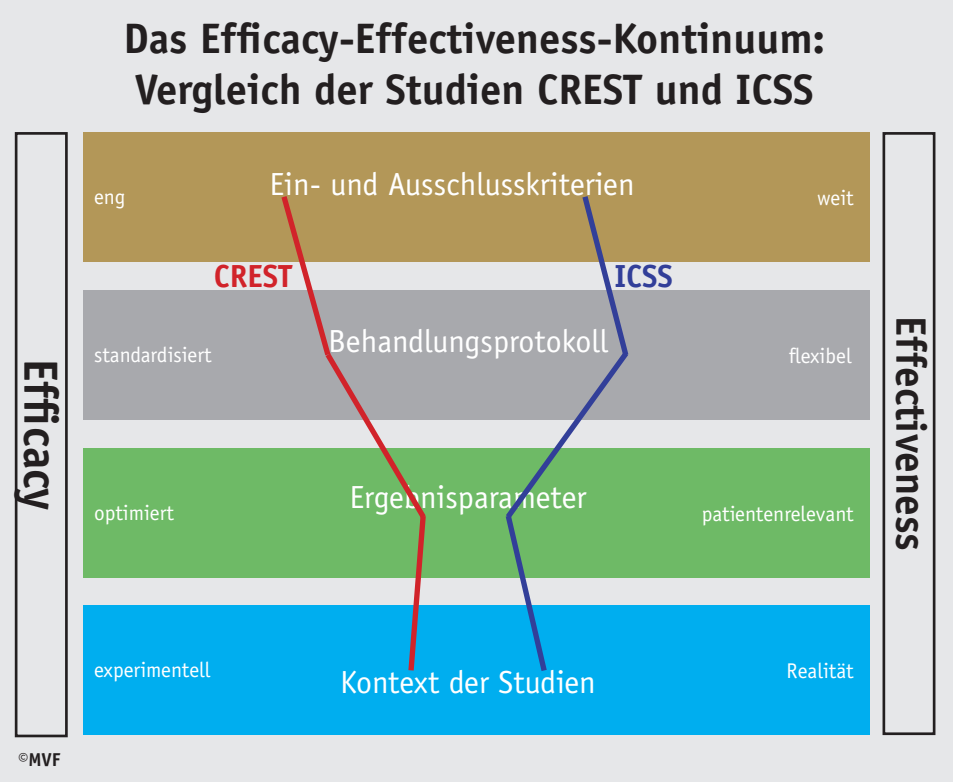


Abb. 1: Reale klinische Studien bewegen sich in ihren Merkmalen in einem Kontinuum zwischen Efficacy- und Effectiveness-Orientierung. Verschieden platzierte Studien ergänzen sich dabei und geben häufig erst gemeinsam ein zutreffendes Bild eines klinischen Sachverhalts, wie hier am Beispiel zweier Studien zum Vergleich einer chirurgischen und einer interventionellen Therapie der Carotisstenose. Unter Idealbedingungen erwies sich die interventionelle Therapie als ebenbürtig. Unter Alltagsbedingungen fiel sie aber deutlich hinter die chrirurgische Therapie zurück. Im Detail: http://www.aerzteblatt.de/archiv/113438.

Was haben Sie an Studien-Material oder auch Daten-Qualität beim BVA vorgefunden, um die Evaluation durchführen zu können?

Die Datenqualität kann ich persönlich ohne Zugang zu den Quellen nicht wirklich beurteilen, hierzu müsste man die Evaluatoren befragen. Nach anfänglich erheblichem Widerstand erlebe ich aber zumindest die Kassen als kooperativ und an Ergebnissen interessiert. Wie weit die Dokumentation vor Ort mit optimistischer Brille vorgenommen wurde, ist schwer einzuschätzen. Die dokumentierten Unterschiede lassen jedoch den Schluss zu, dass hier nicht flächendeckend Daten geschönt wurden. Die Auswertungsmethodik entspricht dem internationalen Standards für den Vergleich von Einrichtungen. Die Zukunft wird vom politischen Willen zur Transparenz abhängen. In jedem Fall muss und wird sich das Evaluationsprogramm weiter entwickeln. Es sollte schlanker und nutzerfreundlicher werden. Wir können aus den ersten zehn Jahren lernen und uns stärker auf die aussagekräftigen Teile fokussieren, die sich bewährt haben. Und das System muss sich technisch weiter entwickeln, um Aufwand zu reduzieren.



„Wir müssen zu einem entspannteren Alltagsverständnis von einer den praktischen Alltag durchdringenden allgegenwärtigen Forschung finden."

Versorgungsforschung ist das eine, die Übertragung der Erkenntnisse in den Versorgungsalltag das andere. Welche Erfolge wurden inzwischen erreicht? Wo liegen die größten Hindernisse? Wo sehen Sie die wichtigsten Aufgaben für die Zukunft?

Die größten Hindernisse sehe ich in der weit verbreiteten Verfolgung von Partikularinteressen im Gesundheitswesen. Die Verbesserung der Versorgung steht häufig an zweiter Stelle hinter eigenen wirtschaftlichen oder politischen Interessen. Entgegen vielen Lippenbekenntnissen haben Personen und Institutionen oft kein Interesse an Transparenz. Überlebte Strukturen werden abgeschottet und verteidigt, als wären sie nicht Mittel zum Zweck, sondern selbst der Zweck des Gesundheitswesens. Entsprechend wird Forschung häufig nicht wirklich gewollt, Ergebnisse, die zu Konsequenzen führen müssten, werden ausgesessen. Hier ist ein Mentalitätswandel erforderlich. Wir müssen begreifen, dass es in unser aller Interesse liegt, Mängel transparent zu machen und Alternativen zu erproben. Wir müssen zu einem entspannteren Alltagsverständnis von einer den praktischen Alltag durchdringenden allgegenwärtigen Forschung finden. Täglich Fragen zu stellen, dann etwas auszuprobieren und schließlich Antworten zu finden macht Spaß, erhöht den Lebensgenuss und hält uns lange gesund. Viele von uns haben das vergessen.

Das Gespräch führten MVF-Herausgeber Prof. Dr. Reinhold Roski und MVF-Chefredakteur Peter Stegmaier. <<