

## **Efficacy und Effectiveness [Efficacy and effectiveness]**

Jürgen Windeler, Gerd Antes

Schon vor der Ära einer evidenzbasierten Medizin ist in Zusammenhang mit der Nutzenbewertung von medizinischen Behandlungsverfahren die Diskussion um verschiedene Dimensionen wünschenswerter Effekte kontrovers geführt worden. So hat sich im Verlauf der heftigen Diskussion um die Neufassung des Arzneimittelgesetzes (AMG) in den 70er Jahren bereits Kienle zu Unterschieden zwischen Wirkung und Wirksamkeit geäußert (7). Er verwies auf die heute unverändert gültige Einschätzung, dass für die Bewertung von Therapien nur therapeutisch-relevante Effekte für die Patienten (Wirksamkeit) und nicht, oder höchstens als „Surrogatkriterium“, Veränderungen von physiologischen Merkmalen gelten können. Auch wenn die Differenzierung zwischen Wirkung und Wirksamkeit bis in die Rechtsprechung des Bundessozialgerichts, das angemessener Weise ausschließlich auf die Wirksamkeit abhebt, inzwischen weitgehend konsentiert ist, so werden doch andere Dimensionen einer Nutzenbewertung sehr unterschiedlich und missverständlich diskutiert. Angesprochen werden soll hier die Unterscheidung zwischen „Efficacy“ und „Effectiveness“.

Mit dem Versuch einer Übersetzung dieser beiden Begriffe fangen die Schwierigkeiten bereits an. Mit „Wirkung“ und „Wirksamkeit“ sind sie jedenfalls nicht gleichzusetzen, möglicherweise noch eher mit dem Begriffspaar „Wirksamkeit“ und „Nutzen“, wobei diese Übersetzung aber in keinem der einschlägigen deutschsprachigen EbM-Lehrbücher zu finden ist. Sie löst das Problem einer Definition auch nicht ohne weiteres, da durchaus nicht klar ist, was unter „Nutzen“ genau verstanden werden soll (zur „Wirksamkeit“ siehe den Definitionsversuch in 1 ). Meist werden die Begriffe daher in ihrer englischen Originalform benutzt.

Eine Recherche zur Definition und zur Unterscheidung der beiden Begriffe „Efficacy“ und „Effectiveness“ führt zunächst zu einem bemerkenswerten Ergebnis: Standardlehrbücher widersprechen sich gravierend und entsprechen zudem nicht einer andernorts zu findenden und häufig geäußerten Interpretation, dass nämlich Efficacy die Wirksamkeit einer Behandlung in der „Kunstwelt“ kontrollierter Studien (unter Idealbedingungen) beschreibe, die Effectiveness dagegen die Wirksamkeit in der routinemäßigen Anwendung (unter Alltagsbedingungen; z.B. 9). So werden in dem Lehrbuch von Kramer (8) die Begriffe Efficacy und Effectiveness z.B. mit der Durchführung bzw. Nicht-Durchführung einer Intention-to-treat-Analyse gekoppelt. Eine solche Verbindung macht die Verwirrung jedoch komplett, da sich die Sinnhaftigkeit von Intention-to-treat-Analysen auf die Aussagekraft von kontrollierten, speziell randomisierten Studien generell bezieht, nicht darauf, welches Zielkriterium oder welche Rahmenbedingungen für die Ermittlung eines positiven Effektes verwendet werden. Auch T. Greenhalgh führt in ihrem Buch (5) aus, dass eine Efficacy-Analyse eine Auswertung ohne Intention-to-treat sei. Diese Aussage steht in krassem Gegensatz zu der üblichen Interpretation (s.o.), die Efficacy mit „Wirksamkeit in klinischen Studien“ gleichsetzt. Für solche Studien muss das Intention-to-treat-Prinzip aber als State-of-the-Art angesehen werden. Es wird zwar oft nur als Auswertungskonzept verstanden, beeinflusst aber stark die Planung einer Studie.

Eine andere Sicht vertritt Feinstein in seinem Buch „Clinical Epidemiology“ (4) unter explizitem Verweis auf A. Cochrane. Danach erfordere die Ermittlung der Efficacy eine kontrollierte Studie, die Ermittlung der Effectiveness dagegen nicht, da ihre Fragestellung sich auf die Auswirkung einer Behandlungsmaßnahme „in the community“ beziehe, eine Definition, aus der sich vermutlich der Begriff Community-Effectiveness abgeleitet hat.

Ein Blick in ein Inhaltsverzeichnis der großen Medizin-Zeitschriften zeigt im übrigen, dass dieses Verständnis von Effectiveness offenbar nicht verbreitet ist, denn dieser Begriff taucht sehr häufig in der – dann ja nicht sinnvollen - Kombination mit dem Begriff „randomisiert“ auf. Auch die FDA benutzt offenbar eine ganz andere Definition von Effectiveness (3)

Das Fehlen von aussagefähigen Daten zur Effectiveness wird häufig beklagt. Nicht immer ist dabei klar, welcher Definition gefolgt wird. Oft wird allerdings mit der Klage der Ruf nach neuen „ganz anderen“ Forschungsmethoden verbunden, um das Defizit zu beseitigen. An dieser Stelle lohnt sich ein etwas differenzierterer Blick. Es ist nämlich sorgfältig zu unterscheiden zwischen der Auswirkung der Einführung einer Behandlungsmaßnahme in das Gesundheitssystem und der Wirksamkeit einer Behandlungsmaßnahme unter Alltagsbedingungen. Diese Unterscheidung ist deshalb von Bedeutung, da die beiden Zielsetzungen tatsächlich unterschiedliche Forschungsmethoden erfordern und vor allem, weil nur für die zweite Zielsetzung sinnvoll argumentiert (und ggfs. beklagt) werden kann, dass vorher durchgeführte Wirksamkeitsstudien nur unzureichende Antworten auf die Alltagssituation geben.

Bei der Frage, wie weit die in vorliegenden klinischen Studien ermittelte Wirksamkeit tatsächlich die Wirksamkeit in der Routineanwendung angibt, kommt es auf die zweimalige Benutzung des Begriffes Wirksamkeit entscheidend an. Mit dieser Verwendung wird deutlich, dass es um eine gleiche Definition, wenn auch in einer anderen Situation geht. Das so beschriebene Problem hat wieder zwei Facetten: die erste betrifft die Frage, inwieweit oder wieweit nicht Ergebnisse aus üblichen klinischen Studien (z. B. zur Zulassung von Arzneimitteln) in die Alltagspraxis übertragen werden können. Diese Frage ist erkenntnistheoretisch (statistisch-)formal mit nein, praktisch kaum und pragmatisch am ehesten so zu beantworten, dass man von einer Übertragbarkeit der Ergebnisse so lange ausgeht, wie nicht entscheidende fachliche Gegenargumente vorgebracht werden können ( 2,5 ). In der Konsequenz bedeutet dies, aus einer Studie, bei der die Wirksamkeit eines Therapieverfahrens bei 40 bis 60-jährigen Männern zuverlässig belegt ist, auf die Wirksamkeit dieser Therapie bei Frauen, bei 70-jährigen Männern oder bei 50jährigen Männern mit Begleiterkrankungen zu schließen, jedenfalls so lange, wie nicht argumentiert werden kann, dass die zu behandelnde Krankheit z.B. bei Frauen anders ist, eine andere Prognose hat, wesentliche andere Begleitumstände berücksichtigt werden müssen usw. Natürlich sind solche Gegenargumente zu einem erheblichen Teil Ermessenssache und es ist daher eben nicht sinnvoll, Studien bei hochselektierten Patientengruppen zu machen und die Übertragbarkeit der Ergebnisse „her zu argumentieren“, sondern besser, die Studien gleich möglichst nah an der geplanten Anwendungssituation durchzuführen. Grundsätzlich ist die Problematik, dass sich die Situation, in der man eine Therapie einsetzen will, oder der Patient, bei dem man die Therapie einsetzen will, sich mehr oder weniger deutlich von Studienkollektiven unterscheidet, unvermeidlich und prozedural (also z.B. als statistische Schlussfolgerung) nicht lösbar.

Die zweite Facette ist die der adäquaten Methodik, wenn denn die Wirksamkeit unter Alltagsbedingungen in einer Studie untersucht werden soll. Die gleiche Verwendung des Begriffes Wirksamkeit macht nämlich deutlich, dass hier die gleichen Forschungsmethoden anzuwenden sind wie bei jeder Wirksamkeitsprüfung, also prospektive, vergleichende, möglichst randomisierte Studien. Zahlreiche Beispiele zeigen, dass solche Studien, in möglichst schlanker Ausgestaltung, auch nahe an Routinebedingungen durchgeführt werden können. Dies entspricht der Forderung nach „large and simple trials“ von Sir Richard Peto (10). Noch einmal: das zentrale Problem der Übertragbarkeit (siehe oben) lösen auch solche Studien schlussendlich nicht, es ist aber auch durch kein anderes Studiendesign zu beseitigen.

Das andere Verständnis von Effectiveness, wie es von Feinstein ausgedrückt wird, bedeutet eine andere Fragestellung. Sie ist mit den beschriebenen Studien nicht zu prüfen und der Begriff Wirksamkeit ist für den zu beschreibenden Effekt auch gar nicht sinnvoll. Es geht also um die Frage, welche Auswirkungen eine als wirksam erkannte (!) Therapie bei ihrer Einführung in das Gesundheitswesen hat. Die Formulierung macht bereits deutlich, dass hier zahlreiche Effekte zu berücksichtigen sind, die mit der Therapie als solcher gar nichts zu tun haben. Er wird vielmehr etwa durch Spezifika eines Gesundheitssystems (Verfügbarkeit von Ressourcen, Leistungserbringung, Finanzierung), in diesem System verfügbare Alternativen (gibt es ein chirurgisches Verfahren, das durch die Einführung einer konservativen Maßnahme abgelöst werden würde?) und durch viele andere Nebenaspekte bestimmt werden. Es mag sein, dass für eine Evaluierung dieser Effekte andere als die bisher üblichen Methoden der Sozialwissenschaft notwendig sind, die dann noch zu entwickeln wären. Es sollte allerdings auch hinterfragt werden, welches eigentlich der zu beschreibende Effekt genau sein soll und ob dabei wirklich von der Effectiveness einer Methode gesprochen werden kann. Um ein Beispiel zu nennen: es können sicherlich für eine Reihe von Erkrankungen zeitliche Trends der Inzidenz oder Mortalität relativ zuverlässig dokumentiert werden. Die Zuordnung dieses möglicherweise als Effectiveness zu bezeichnenden Nutzeneffektes zur Anwendung bestimmter Maßnahmen wird aber wohl kaum oder nur unter sehr großen Schwierigkeiten gelingen.

Die Begriffe Efficacy und Effectiveness unterliegen einer ausgeprägten Interpretationsvielfalt. Das muss akzeptiert werden und ist vermutlich auch nicht zu ändern. Jeder Benutzer dieser Begriffe sollte sich aber über die Vielfalt klar sein und jedenfalls selbst genau verstehen, unter welcher Definition er/sie die Begriffe benutzt.

## Literatur

1. Abel U, Windeler J (1996) Erkenntnistheoretische Aspekte klinischer Studien. 2. Die „Wirksamkeit“ von Behandlungen. intern praxis 36: 375-390
2. Dans AL, Dans LF, Guyatt GH, Richardson S (1998) User's guide to the medical literature. XIV: How to decide on the applicability of clinical trial results to your patient. JAMA 279: 545-549
3. FDA: Guidance for Industry: Providing Clinical Evidence of Effectiveness for Human Drug and Biological Products. [www.fda.gov/cder/guidance/index.htm](http://www.fda.gov/cder/guidance/index.htm)
4. Feinstein AR. (1985) Clinical Epidemiology. The architecture of clinical research. Philadelphia: Saunders
5. Greenhalgh T. (1997) How to read a paper. The basics of evidence based medicine. London: BMJ Publishing Group
6. Guyatt G (1999) A 75-year-old man with congestive heart failure. JAMA 281: 2321-2327
7. Kienle G. (1974) Arzneimittelsicherheit und Gesellschaft - Eine kritische Untersuchung. Stuttgart, New York: Schattauer
8. Kramer MS. (1988) Clinical Epidemiology and Biostatistics. Berlin, Heidelberg: Springer
9. Kunz R, Ollenschläger G, Raspe H, Jonitz G, Kolkman FW. (2000) Lehrbuch Evidenzbasierte Medizin in Klinik und Praxis. Köln: Deutscher Ärzte-Verlag
10. Peto R, Collins R, Gray R (1995) Large-scale randomized evidence: large, simple trials and overviews of trials. J Clin Epidemiol 48: 23-40

## Korrespondenzadresse:

Prof. Dr. med. Jürgen Windeler  
Fachbereich Evidenz-basierte Medizin  
Medizinischer Dienst der Spitzenverbände der Krankenkassen e.V. (MDS)  
Lützowstraße 53  
D – 45141 Essen;

Tel.: 0201 - 8327-125;  
Fax: 0201 – 8327-3154;  
E-mail: [j.windeler@mds-ev.de](mailto:j.windeler@mds-ev.de)

### What does work

Effectiveness (Wirksamkeit unter Alltagsbedingungen)

Die unter alltäglichen Bedingungen erreichbare Wirksamkeit einer medizinischen Intervention für den Patienten

### What can work

Efficacy (Wirksamkeit unter Idealbedingungen)

Die unter Idealbedingungen erreichbare Wirksamkeit einer medizinischen Intervention für den Patienten

**Abb. 1:** What can work: Efficacy und Effectiveness (nach 9)