

Folge 20: Wie kann man Behandlungseffekte aus klinischen Studien bewerten? II: Dichotome Outcomes*

R. D. Herbert

Die vorhergehende Folge (Herbert 2000, deutsch 2004) behandelte die Frage, wie Artikel über klinische Studien genutzt werden können, um unverzerrte Schätzungen der Größe des Behandlungseffekts zu gewinnen. Der Artikel diskutierte, wie Leser klinischer Studien einfache Schätzungen der Größe des Behandlungseffekts erhalten können, wenn die Studienergebnisse kontinuierlich gemessen wurden. Der dabei vorgetragene Ansatz integriert klinische Intuition und Patientenpräferenzen in den klinischen Entscheidungsprozess. Es wurde die Auffassung vertreten, dass ein optimaler Entscheidungsprozess bei der Wahl

* Ursprünglich erschienen als: Herbert RD 2000. How to estimate treatment effects from reports of clinical trials. II: Dichotomous outcomes. Australian Journal of Physiotherapy, Vol. 46, 309-313. Übersetzung aus dem Englischen: Erwin Scherfer

1 Zwischen zwei beliebigen Werten einer kontinuierlichen Variable sind immer unendlich viele weitere Werte denkbar und theoretisch auch messbar, indem man weitere Dezimalstellen hinter dem Komma „misst“. Die Grenze der Genauigkeit bei Messungen von kontinuierlichen Variablen liegt also einzig in der Präzision des Messgeräts (Anm. d. Übers.).

2 Dichotom bedeutet wörtlich: „In zwei schneidend“. Eine dichotome Variable hat somit per definitionem nur zwei Ausprägungen. Diskrete Variablen sind „eindeutig abzählbar“, d.h. zwischen zwei Werten einer Variable, z. B. drei und vier, existieren keine weiteren Merkmalsausprägungen. Dichotome Variablen mit nur zwei Ausgängen sind ein Spezialfall einer diskreten Variable (Anm. d. Übers.).

Zusammenfassung

Diese Folge setzt die vorhergehende fort, indem sie diskutiert, wie Leser einfache Schätzungen der Größe von Behandlungseffekten aus den Angaben in Artikeln über klinische Studien errechnen können, wenn die Studienergebnisse dichotom gemessen wurden. Es wird eine Methode dargestellt, mit der man den Grad der Unsicherheit, mit dem solche Schätzungen behaftet sind, quantifizieren kann. Schätzungen der Größe des Behandlungseffekts können auf Basis des Ausgangsrisikos angepasst werden, um die Wahrscheinlichkeit, dass eine Behandlung einem bestimmten Patienten helfen wird, zu bestimmen. Die Wahrscheinlichkeit, dass die Behandlung hilft, sollte gegenüber den Kosten und Risiken, die mit der Behandlung einhergehen, abgewogen werden.

Schlüsselworte: Konfidenzintervall, Entscheidungsprozess, Evidenzbasierte Medizin

der Therapie für einzelne Patienten die Modifikation der Schätzung des Behandlungseffekts auf der Basis der Eigenschaften des einzelnen Patienten und den Vergleich des zu erwartenden Behandlungseffekts mit dem „kleinsten klinisch lohnenswerten Effekt“ beinhaltet. In dieser Folge wird die gleiche Vorgehensweise auf klinische Studien angewendet, deren Ergebnisse dichotom gemessen wurden.

Was sind dichotome Outcomes?

Die Beispiele der vorhergehenden Folge stammten aus klinischen Studien, deren Ergebnisse als kontinuierliche Variablen gemessen wurden. Kontinuierliche Variablen sind Variablen, die zwischen ihrem untersten und obersten Extremwert jeglichen Wert aus einer (theoretisch) unend-

lich großen Menge von Werten annehmen können¹. Ödemumfänge, Schmerzangaben auf einer VAS und Lungenfunktionstests sind Beispiele kontinuierlicher Variablen. Es gibt aber auch Outcomes, die als „dichotome“ Variablen gemessen werden. Dichotome Outcomes sind diskrete Ereignisse – sie treten ein oder sie treten nicht ein -, wie z. B. Tod, Verletzung oder „Behandlungszufriedenheit.“²

Wie werden dichotome Outcomes quantifiziert?

Wir quantifizieren dichotome Therapieergebnisse, indem wir sie als den prozentualen Anteil der Patienten, die das interessierende Ereignis erfahren, ausdrücken, und dies üblicherweise in einem näher spezifizierten Zeitraum. Hierdurch erhalten wir Informationen über das Risiko, die

Eintrittswahrscheinlichkeit des Ereignisses für Individuen aus der entsprechenden Population. Ein gutes Beispiel hierfür stellt die vor kurzem durchgeführte³ Studie zur präventiven Wirkung prophylaktischer Atemphysiotherapie in Bezug auf respiratorische Komplikationen nach großen Bauchoperationen dar (Olsen et al. 1997). In dieser Studie war das Auftreten von respiratorischen Komplikationen das interessierende Ereignis. 52 von 192 Probanden in der Kontrollgruppe erlitten respiratorische Komplikationen innerhalb von sechs Tagen nach der Operation. Das Risiko, eine respiratorische Komplikation zu entwickeln, betrug somit für diese Probanden 0,27 bzw. 27 %.

Das Behandlungsrisiko abschätzen

Bei klinischen Studien mit dichotomen Ergebnissen interessiert uns, ob die Behandlung das Risiko des interessierenden Ereignisses herabsetzt. Was wir also bestimmen müssen ist, ob sich das Risiko der Behandlungs- von dem der Kontrollgruppe unterscheidet. Die Größe der Risikoreduktion, die eine Aussage hinsichtlich des Grades der Wirksamkeit der Behandlung ermöglicht, kann auf verschiedene Arten und Weisen ausgedrückt werden (Guyatt et al., Sackett et al. 1998). Drei gebräuchliche Maße sind: Die absolute Risikoreduktion, die „numbers needed to treat“⁴ und die relative Risikoreduktion.

Die absolute Risikoreduktion

Unter der absoluten Risikoreduktion versteht man einfach die Differenz des Risikos von Behandlungs- und

Kontrollgruppe. In der Studie von Olsen et al. erfuhr nur ein relativ kleiner Anteil der Probanden der Behandlungsgruppe respiratorische Komplikationen (10 von 172 bzw. $10/172 = 0,06$ bzw. 6 %). Das Risiko, eine respiratorische Komplikation zu entwickeln, war also für die Probanden dieser Gruppe relativ gering im Vergleich zu dem Risiko in Höhe von 27 %, das die Probanden der Kontrollgruppe aufwiesen. Die absolute Risikoreduktion beträgt $0,27 - 0,06 = 0,21$ oder 21 %. Dies bedeutet, dass die behandelten Probanden gegenüber den Probanden der Kontrollgruppe ein um 21 % geringeres Risiko hatten, innerhalb von sechs Tagen nach der Operation respiratorische Komplikationen zu entwickeln.

☞ Große absolute Risikoreduktionen zeigen, dass eine Behandlung sehr effektiv ist.

☞ Negative absolute Risikoreduktionen zeigen, dass das Risiko in der Behandlungsgruppe größer ist als in der Kontrollgruppe und somit, dass die Behandlung schädlich ist (eine Ausnahme von dieser Regel ist gegeben, wenn es sich bei dem interessierenden Ereignis um ein positives Ergebnis handelt, wie z.B. die Wiederherstellung der Arbeitsfähigkeit, und nicht um ein negatives Ereignis).

„Number needed to treat“

Verständlicherweise haben viele Menschen Schwierigkeiten, die Größe einer angegebenen absoluten Risikoreduktion richtig einzuschätzen bzw. zu interpretieren. Eine Folge hiervon ist, dass es oft schwierig ist, den kleinsten klinisch lohnenswerten Behandlungseffekt ausgedrückt in absoluter Risikoreduktion zu spezifizieren (be-

sonders wenn die absolute Risikoreduktion klein ist). Wie groß ist eine 21-%ige Reduktion des absoluten Risikos? Ist eine Reduktion des absoluten Risikos um 21 % klinisch lohnenswert? Ein zweites Maß der Risikoreduktion, die „number needed to treat“ (die „Anzahl Patienten, die behandelt werden müssen“) macht die Größe einer absoluten Risikoreduktion expliziter.

☞ Die „number needed to treat“ (NNT) erhält man, indem man den Kehrwert der absoluten Risikoreduktion bestimmt.

☞ In unserem Beispiel beträgt die absolute Risikoreduktion 0,21. Also ist die NNT 1 dividiert durch 0,21 bzw. $1/0,21 \approx 5$, also ungefähr 5. Dies ist die Anzahl der Patienten, die behandelt werden müssen, um im Durchschnitt einmal das interessierende Ereignis durch Prävention zu vermeiden.

In unserem Beispiel kommt im Durchschnitt auf fünf behandelte Patienten einer, bei dem eine respiratorische Komplikation präventiv vermieden wurde. Für die anderen vier machte die Behandlung keinen Unterschied. Einige hätten sowieso keine respiratorische Komplikation entwickelt und die anderen erleiden eine respiratorische Komplikation trotz Behandlung.

☞ Eine kleine NNT (wie z. B. fünf) ist besser als eine große NNT (wie z. B. 100), weil sie zeigt, dass nur eine kleine Zahl von Patienten behandelt werden muss, bis die Behandlung für einen von ihnen einen Unterschied ausmacht.

³ Diese Angabe bezieht sich auf den Zeitpunkt der Erst- bzw. Originalveröffentlichung dieses Artikels (Anm. d. Übers.).

⁴ Auch in der deutschsprachigen methodologischen bzw. klinischen Literatur wird der englische Begriff verwendet (Anm. d. Übers.).

Die NNT ist sehr hilfreich

Sie macht es relativ leicht, einen kleinsten klinisch lohnenswerten Effekt zu spezifizieren. Mit der NNT können wir leichter den Nutzen einer Behandlung, nämlich inwiefern sie bei einem Patienten das Eintreten eines Ereignisses präventiv verhindert, abwägen gegenüber den mit ihrer Anwendung verbundenen Kosten und Risiken (insbesondere in Bezug auf die Kosten und Risiken für die Patienten, bei denen die Therapie keine Wirkung haben wird). In unserem Beispiel würden wohl die meisten zustimmen, dass eine NNT von 10 bedeutete, dass die Behandlung klinisch lohnenswert ist, denn die Prävention einer respiratorischen Komplikation ist eine sehr wünschenswerte Angelegenheit, und die Kosten und Risiken der hier zur Diskussion stehenden Behandlung sind minimal. Es macht also nicht viel aus, wenn neun von zehn Patienten „umsonst“ behandelt werden. Die meisten würden aber wohl auch zustimmen, dass eine NNT von 100 zu groß wäre, um die Therapie als klinisch lohnenswert gelten zu lassen. Es mag zwar wenig Risiko in der Therapie liegen, aber wahrscheinlich ist es zu teuer, 99 Patienten zu behandeln ohne eine Wirkung zu erzielen, um einen Fall von respiratorischer Komplikation zu vermeiden.

Wo ist die Grenze?

Wie groß ist aber die größte NNT für prophylaktische Atemphysiotherapie, um sie noch als klinisch lohnenswerte Therapie zu akzeptieren (d.h. wie groß ist der kleinste klinisch lohnenswerte Effekt?). Als der Autor einige erfahrene Atemphysiotherapeuten

befragte, gaben diese an, dass sie die Therapie nicht veranlassen würden, wenn mehr als 20 Patienten behandeln müssten, um eine respiratorische Komplikation präventiv zu verhindern. Sie nannten also eine NNT von 20 als den kleinsten klinisch lohnenswerten Effekt. Dies entspricht einer absoluten Risikoreduktion von 5 %. Es wäre interessant, eine Befragung von Patienten durchzuführen, die vor einer großen Bauchoperation stehen, um zu ermitteln, was sie für den kleinsten klinisch lohnenswerten Effekt halten. Die Größe des Behandlungseffekts, den die Studie von *Olsen* ergab (NNT = 5), ist viel größer als die meisten Therapeuten als kleinsten klinisch lohnenswerten Effekt ansehen würden (NNT \oplus 20; erinnern wir uns, dass eine kleine NNT für einen großen Behandlungseffekt steht).

Die Gefährlichkeit der Folgen entscheidet

Natürlich gibt es nicht *den einen* NNT-Wert, den man als den kleinsten klinisch lohnenswerten Effekt ansehen kann. Die Größe des kleinsten klinisch lohnenswerten Effekts wird immer abhängen von der Ernsthaftigkeit des Ereignisses und den Kosten und Risiken der Behandlung. So könnte der kleinste klinisch lohnenswerte Effekt eines dreimonatigen Übungsprogramms nur bei zwei oder drei liegen, wenn es sich bei dem zu vermeidenden Ereignis um eine gelegentliche Instabilität des Knies handelt, wohingegen der kleinste klinisch lohnenswerte Effekt für die Anwendung einer Atemtrainingsmethode wie der SMI (mit einem incentive spirometer) in der unmittelbaren postoperativen Phase nach einer Thoraxoperation bei einer NNT von

100 liegen könnte, wenn das zu vermeidende Ereignis der Tod ist. Bei Langzeittherapien sollte die NNT, genau wie die absolute Risikoreduktion, im Verhältnis zur Dauer der Interventionsperiode gesehen werden. Eine NNT von 10 für eine drei Monate lang durchgeführte Therapie mit dem Ziel, respiratorische Komplikationen bei Kindern mit cystischer Fibrose, zu vermeiden, ist in ihrer Effektgröße vergleichbar mit einer anderen Behandlung, die eine NNT von 5 bei einer sechs Monate dauernden Intervention hat.

Relative Risikoreduktion

Eine üblichere, aber nicht so unmittelbar hilfreiche Weise, die Risikoreduktion auszudrücken ist, sie als Prozentsatz des Risikos der nicht behandelten Patienten anzugeben. Dies wird die relative Risikoreduktion genannt.

Man erhält sie, indem man die absolute Risikoreduktion durch das Risiko der Kontrollgruppe teilt.

So beträgt die relative Risikoreduktion, die durch die präventive Atemphysiotherapie erreicht wird $0,21/0,27$ und damit 0,78 oder 78 %. Mit anderen Worten: Präventive Atemphysiotherapie verminderte das Risiko respiratorischer Komplikationen um 78 % des Risikos der unbehandelten Patienten. Deutlich ist, dass die relative Risikoreduktion mit 78 % viel größer ausfällt als die absolute Risikoreduktion mit 21 %.

Tatsächlich ist die relative Risikoreduktion immer größer als die absolute Risikoreduktion, weil man sie erhält, indem man die absolute Risikoreduktion durch eine Zahl teilt, die immer kleiner als 1 ist.

Absolut oder relativ – das ist die Frage

Was ist nun das beste Maß für die Größe eines Behandlungseffekts. Sollten wir die absolute Risikoreduktion, ihren Kehrwert (die NNT) oder die relative Risikoreduktion heranziehen?

Die relative Risikoreduktion hat einige Eigenschaften, die sie hilfreich für den Vergleich von Ergebnissen verschiedener Studien machen, aber sie kann in die Irre führen, wenn sie für klinische Entscheidungsprozesse genutzt wird.

Zum Beispiel zeigten *Lauritzen et al.* (1993), dass die Versorgung der Bewohner von Pflegeheimen mit Hüftschutz-Pads (Polstern) eine große relative Risikoreduktion in Bezug auf Hüftfrakturen bewirkte (relative Risikoreduktion von 56 %).

Dies klingt nach einem sehr großen Effekt, und auf Basis dieser Statistik ist man versucht zu schlussfolgern, dass Hüft-Pads klinisch lohnenswert sind. Die Inzidenz von Hüftfrakturen in der Population von Pflegeheimbewohnern beträgt jedoch ungefähr 5 % pro Jahr (*Lauritzen et al.* 1993), sodass die absolute Risikoreduktion durch das Tragen von Hüft-Pads in Bezug auf Hüftfrakturen innerhalb der betrachteten Population 56 % von 5 % und somit weniger als 3 % beträgt. Drücken wir dies als NNT aus, so wird deutlich, dass 36 Personen ein Jahr lang Hüft-Pads tragen müssen, um eine Hüftfraktur zu verhindern. Wenn die Risikoreduktion als absolute Risikoreduktion oder besser noch als NNT ausgedrückt wird, erscheint die Größe des Behandlungseffekts klein, und vielleicht zu klein, um klinisch lohnenswert zu sein.

Das Beispiel illustriert, dass es wahrscheinlich besser ist, die Größe von Behandlungseffekten als absolute Risikoreduktion oder in NNT auszudrücken anstatt als relative Risikoreduktion.

Die Bedeutung des Ausgangsrisikos

Ganz allgemein gilt, dass sogar die besten Behandlungen (die mit großen relativen Risikoreduktionen) nur kleine absolute Risikoreduktionen erzeugen, wenn das Risiko des Ereignisses, also seine Eintrittswahrscheinlichkeit, bei unbehandelten Patienten gering ist, wenn also das „baseline“- bzw. Ausgangsrisiko niedrig ist. Dies ist vielleicht intuitiv einleuchtend. Wenn nur wenige Personen Gefahr laufen, das Ereignis zu erfahren, dann ist auch nicht möglich, das Ereignis häufig durch Prävention zu vermeiden. Das hat zwei *Implikationen für die Praxis*:

Erstens: Sogar die besten Behandlungen werden wahrscheinlich keine klinisch lohnenswerten Effekte haben, wenn das zu vermeidende Ereignis unwahrscheinlich ist. Und umgekehrt wird eine Behandlung sich eher als klinisch lohnenswert erweisen, wenn sie das Risiko eines häufig auftretenden Ereignisses vermindert (eine besonders klare Diskussion dieses Punktes leisten *Glazou und Irwig* 1995).

Zweitens: Da die Größe des Behandlungseffekts häufig stark abhängt von dem Risiko, dem die unbehandelten Patienten ausgesetzt sind, sind die Resultate nur mit Vorsicht auf einen individuellen Patienten anzuwenden, wenn das Risiko der Patienten der Kontrollgruppe aus

der Studie sich deutlich von dem Risiko, dem der Patient, für den die Therapie erwogen wird, unterscheidet. Wenn das Risiko der Probanden der Kontrollgruppe viel höher ist, als das des Patienten, um den es geht, so besteht die Gefahr, dazu zu tendieren, den Behandlungseffekt zu überschätzen (die aus den Studiendaten errechnete absolute Risikoreduktion wird zu groß sein, die NNT wird zu niedrig sein).

Lösung des Problems: der Faktor f

Es gibt eine einfache Lösung für das Problem, die Ergebnisse einer Studie auf einen Patienten mit höherem oder niedrigerem Risiko anzuwenden. Der hier beschriebene Ansatz geht auf eine Methode von *Straus und Sackett* (1999, s. a. *McAlister* 2000) zurück. Die absolute Risikoreduktion wird genauso wie oben beschrieben aus den Daten der Studie berechnet, wird aber dann um einen Faktor f korrigiert, der angibt, um wie viel größer das Risiko der unbehandelten Patienten der Kontrollgruppe ist.

Ein f größer 1 wird verwendet, wenn die Patienten, auf die die Resultate angewendet werden sollen, einem größeren Risiko ausgesetzt sind als die unbehandelten Patienten der Kontrollgruppe, und ein f kleiner 1 wird verwendet, wenn die Patienten, auf die die Resultate angewendet werden sollen, ein kleineres Risiko tragen als die Patienten der Kontrollgruppe. Die absolute Risikoreduktion wird angepasst, indem mit f multipliziert wird, und die NNT wird angepasst, indem durch f dividiert wird.

Beispiel

Ein Therapeut, der einen krankhaft adipösen Patienten behandelt, der vor einer großen Bauchoperation steht, könnte schätzen, dass der Patient ein doppelt so hohes Risiko einer respiratorischen Komplikation trägt wie die Patienten in der Studie von *Olsen et al.* Um eine angemessene Schätzung des zu erwartenden Therapieeffekts unter Einbezug des größeren Ausgangsrisikos dieses Patienten im Vergleich zu den Patienten in der Studie zu erhalten, könnte die NNT (sie beträgt 5, wie wir oben ausgerechnet haben) durch 2 dividiert werden. Hieraus ergibt sich eine NNT von 2,5 (aufgerundet zu 3) für krankhaft übergewichtige Personen. Somit können wir einen noch größeren Behandlungseffekt prophylaktischer Atemphysiotherapie bei Hochrisikopatienten erwarten. Dieser Ansatz kann verwendet werden, um Schätzungen des wahrscheinlichen Behandlungseffekts für jeden individuellen Patienten auf der Basis der therapeutischen Einschätzung des individuellen Patientenrisikos zu berechnen.

Konfidenzintervalle berechnen für Behandlungseffekte bei dichotomen Outcomes

Wie auch Studien, die kontinuierliche Outcomevariablen messen, geben viele Studien mit dichotomen Outcomes keine Konfidenzintervalle zur relativen Risikoreduktion, absoluten Risikoreduktion oder zur NNT an. Die meisten enthalten aber genug Daten, sodass man diese selbst ausrechnen kann. Ein grobes 95 %-Konfidenzintervall für die absolute Risikoreduktion kann man einfach errechnen, wenn die durchschnittliche Gruppen-

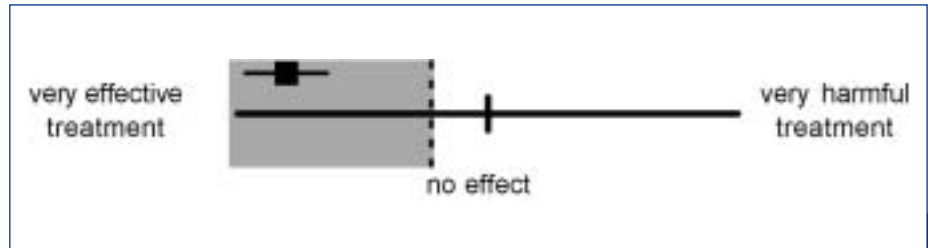


Abb. 1:

Ein Baumdiagramm der Größe des Behandlungseffekts, das die Ergebnisse der Studie von *Olsen et al.* (1997) illustriert. Das Baumdiagramm besteht aus einer horizontalen Linie, die den Behandlungseffekt darstellt. An den beiden Enden der Linie finden sich sehr effektive bzw. sehr schädliche Behandlungen. Der kleinste klinisch lohnenswerte Behandlungseffekt wird durch die unterbrochene Linie dargestellt. Der grau schattierte Bereich links von ihr repräsentiert klinisch lohnenswerte Behandlungen. In diesem Beispiel zur absoluten Risikoreduktion in Bezug auf respiratorische Komplikationen wurde der kleinste klinisch lohnenswerte Effekt auf 5 % festgelegt. Die beste Schätzung für die Größe des Behandlungseffekts (21 %) wird durch das kleine Quadrat wiedergegeben, und das 95 %-Konfidenzintervall zu dieser Schätzung (14 % bis 28 %) durch die horizontale Linie durch das Quadrat. Der Behandlungseffekt ist deutlich größer als der kleinste klinisch lohnenswerte Effekt.

größe (n) der Experimental- und der Kontrollgruppe gegeben sind:

$$95\% \text{ KI} = \text{Differenz des Risikos} \pm 1,96/n$$

(zur Ableitung dieser Gleichung s. Anhang). Diese Annäherung funktioniert recht gut (ihre Resultate sind recht nahe an denen komplexerer Formeln), wenn das durchschnittliche Risiko der Ereignisse in der behandelten und der Kontrollgruppe größer ist als $\oplus 10\%$ und kleiner ist als $\oplus 90\%$.

Um die Berechnung von Konfidenzintervallen für dichotome Daten zu illustrieren, erinnern wir uns, dass in der Studie von *Olsen et al.* (1997) das Risiko für die Probanden der Kontrollgruppe 0,27 und für die der Experimentalgruppe 0,06 betrug, während die durchschnittliche Gruppengröße 182 war. Also:

$$95\% \text{ KI} = 0,27 - 0,06 \pm 1,96/182$$

$$95\% \text{ KI} = 0,21 \pm 0,07$$

Somit beträgt die beste Schätzung der absoluten Risikoreduktion 21 % bei einem 95 %-Konfidenzintervall von 14 % bis 28 %. Rechnet man das 95 %-

Konfidenzintervall der absoluten Risikoreduktion um in ein Konfidenzintervall für die NNT, so reicht dies von 4 bis 7.

Dieses Resultat ist als Baumdiagramm dargestellt (Abb. 1)

Die Logik dieses Diagramms ist genau die gleiche wie bei den drei Graphiken, die in der vorhergehenden Folge für kontinuierliche Outcomevariablen präsentiert wurden (*Herbert 2000, 2004*), außer dass hier das linke Ende der Linie den positiven Effekt (eine Risikoreduktion) kennzeichnet. Wieder zeichnen wir den kleinsten klinisch lohnenswerten Effekt (eine absolute Risikoreduktion von 5 % die einer NNT von 20 entspricht), die Größe des Behandlungseffekts (eine absolute Risikoreduktion von 21 %) und deren Konfidenzintervall (von 14 % bis 28 %) ein. In diesem Beispiel sind die geschätzte absolute Risikoreduktion und ihr Konfidenzintervall eindeutig größer als der kleinste klinisch lohnenswerte Effekt, sodass wir mit Sicherheit schlussfolgern können,

dass diese Behandlung klinisch lohnenswert ist. Für krankhaft adipöse Patienten (für die wir die absolute Risikoreduktion mit $f = 2$ multiplizieren konnten, um ihr größeres Risiko bei Nichtbehandlung zu berücksichtigen) ist die Therapie sogar noch lohnenswerter.

Dieser Artikel und der vorhergehende haben einen Prozess beschrieben, mit dem Informationen aus klinischen Studien in den klinischen bzw. praktischen Entscheidungsprozess integriert werden können. Gut durchgeführte klinische Studien liefern akzeptabel unverzerrte Schätzungen der Größen von Behandlungseffekten. Diese Schätzungen können auf Grundlage der individuellen Charakteristika von Patienten modifiziert und dann mit dem kleinsten klinisch lohnenswerten Effekt verglichen werden, um zu bestimmen, ob eine Therapie wahrscheinlich mehr Nutzen als Schaden wird.

Literatur:

1. Gardner MJ, Altman DG 1989. Statistics with Confidence – Confidence Intervals and Statistical Guidelines. London. British Medical Journal, 20-33.
2. Glasziou PP, Irwig LM 1995. An evidence based approach to individualizing treatment. British Medical Journal, 311, 1356-1359.
3. Guyatt GH, Sackett DL, Cook DJ 1994. User's guide to the medical literature: II. How to use an article about therapy or prevention: What were the results and will they help me in caring for my patients? Journal of the American Medical Association, 271; 59-63.
4. Herbert RD 2000. How to estimate treatment effects from reports of clinical trials. I: Continuous Outcomes. Australian Journal of Physiotherapy, 46, 229-235
5. Herbert RD 2004. Wie kann man Behandlungseffekte aus klinischen Studien bewerten? I: Kontinuierliche Outcomes. Zeitschrift für Physiotherapeuten; 56;
6. Newcombe RG 1999. Interval Estimation for the difference between independent proportions: Comparison of eleven methods. Statistics in Medicine, 17, 873-890
7. Newcombe RG 1999. Improved confidence intervals for the difference between binomial proportions based on paired data. Statistics in Medicine, 17. 2635-2650

Anhang:

Angenähertes 95 % Konfidenzintervall für die absolute Risikoreduktion

Die übliche Gleichung für das Konfidenzintervall zur absoluten Risikoreduktion (d. h. die Differenz zwischen zwei prozentualen Anteilen) ist:

$$95\%KI = \text{Risikodifferenz} \pm z_{(1-\alpha/2)} \times \sqrt{\frac{R_c(1-R_c)}{n_c} + \frac{R_t(1-R_t)}{n_t}}$$

wobei $z_{(1-\alpha/2)}$ der entsprechende Wert einer z-Verteilung ist, R_c das Risiko der Probanden in der Kontrollgruppe (c für controls), R_t das Risiko der Probanden in der Behandlungsgruppe (t für treatment = Behandlung) und n_c und n_t die Größe der Kontroll- bzw. der Behandlungsgruppe ist (Gardner und Altman, 1989). Unter bestimmten Bedingungen, insbesondere wenn R_c und R_t sich 0% oder 100% annähern, ist diese Gleichung nicht sehr genau (Newcombe 1998, 1999), aber für die meisten klinischen Zwecke ist sie hinreichend präzise. In randomisierten klinischen Studien sind die Gruppengrößen in der Regel annähernd gleich groß (d. h. $n_t \approx n_c$), so dass wir n_c und n_t durch n_{av} , das Mittel beider Gruppengrößen, ersetzen können. Da ferner $z_{(1-\alpha/2)} \approx 2$, vereinfacht sich der Ausdruck zu:

$$95\%KI = \text{Risikodifferenz} \pm 2 \times \sqrt{\frac{R_c(1-R_c) + R_t(1-R_t)}{n}}$$

Piantadosi (1997) hat herausgestellt, dass $R[1-R]$ relativ wenig mit R variiert, zumindest im Bereich von $0,1 < R < 0,9$. Um die Gleichung zu vereinfachen, können wir den Wert 1 für den Ausdruck $2 \times \sqrt{R_c[1-R_c] + R_t[1-R_t]}$ setzen, wodurch sich die Gleichung vereinfacht zu:

$$95\%KI = \text{Risikodifferenz} \pm 1/\sqrt{n}$$

Die mit dieser Annäherung verbundene Fehleranalyse ist komplex, aber die Annäherung scheint mit hinreichender Genauigkeit anwendbar zu sein. Sogar bei kleinen Stichproben (z.B. $n=30$ beträgt der relative Fehler (das ist der Fehler ausgedrückt als Prozentanteil an der Weite des „wahren“ Konfidenzintervalls) normalerweise weniger als 20 %, was wohl für klinische Entscheidungsprozesse ausreichend ist.

Kleines Wörterbuch zum Verständnis der Grafiken:

very harmful effect – sehr schädliche Wirkung

no effect – keine Wirkung

very effective treatment – sehr wirksame Behandlung

8. Olsen MF, Hahn I, Nordgren S, Lonroth H, Lundholm K 1997. Randomized controlled trial of prophylactic chest physiotherapy in major abdominal surgery. *British Journal of Surgery* 84, 1535-1538.
9. Piantadosi S 1997. *Clinical Trials: A methodological Perspective*. New York. Wiley
10. Sackett DL, Richardson WS, Rosenberg W, Haynes RB 1998. *Evidence-based Medicine. How to practice and teach EBM*. Edinburgh; Churchill Livingstone; 91-96
11. Straus SE, Sackett DL 1999. Applying evidence to the individual patient. *Annals of Oncology*, 10, 29-32.



R. D. HERBERT

- *Senior Lecturer der Physiotherapieschule an der Universität Sydney, Australien*
- *Unterrichtung von undergraduate und postgraduate Studenten in Evidenz basierter Praxis*
- *Forschungsfelder: Muskelmechanik und klinische Untersuchung der Effekte physiotherapeutischer Interventionen*
- *Direktor des Zentrums für Evidenz basierte Physiotherapie und Wissenschaftlicher Redakteur des australischen Journal of Physiotherapy*

■ Korrespondenzadressen:

Dr. Robert D. Herbert
 School of Physiotherapy
 The University of Sydney
 PO Box 170
 Lidcombe NSW 1825, Australia
 r.herbert@fhs.usyd.edu.au

Dr. Erwin Scherfer
 Bildungswerk Physio-Akademie
 des ZVK gGmbH
 Wremer Specken 4, 27638 Wremen
 E-Mail: e.scherfer@physio-akademie.de