

*P. Kranke, A. Redel, Th. Smul, E.-M. Kranke¹
Klinik und Poliklinik für Anästhesiologie
(Direktor: Prof. Dr. Norbert Roewer)
¹Klinik und Poliklinik für Herz-
und Thoraxchirurgie
(Direktor: Prof. Dr. Rainer Leyh)
Universität Würzburg*

Grundlagen und Prinzipien klinischer Studien: Untersuchungen zum Nachweis der Wirksamkeit medizinischer Interventionen

Teil 1

ZUSAMMENFASSUNG

Zur Anwendung kommende Verfahren und Behandlungsmethoden in der klinischen Medizin sollten wirksam und nebenwirkungsarm sein und dem besten aktuell zur Verfügung stehenden Wissen entsprechen. Hierfür ist der Beleg der Überlegenheit gegenüber alternativen oder historischen Behandlungsverfahren erforderlich. Dies ist

ein wesentlicher Motivationsgrund für die Durchführung klinischer Studien.

Der Artikel beschreibt die grundlegenden Anforderungen an Untersuchungen zum Nachweis der Wirksamkeit medizinischer Interventionen. Insbesondere wird die Wertigkeit vergleichbarer Ausgangsbedingungen durch eine zufällige Verteilung der Patienten auf die Behandlungs- und

Kontrollgruppe (Randomisierung) erläutert. Alternativen zur Generierung von Kontrollgruppen (historische Kontrollen, nicht-randomisierte Zuteilung) werden mit ihren Vor- und Nachteilen diskutiert.

SCHLÜSSELWÖRTER

Klinische Studien, Randomisierung, randomisierte klinische Studien, nicht-rando-

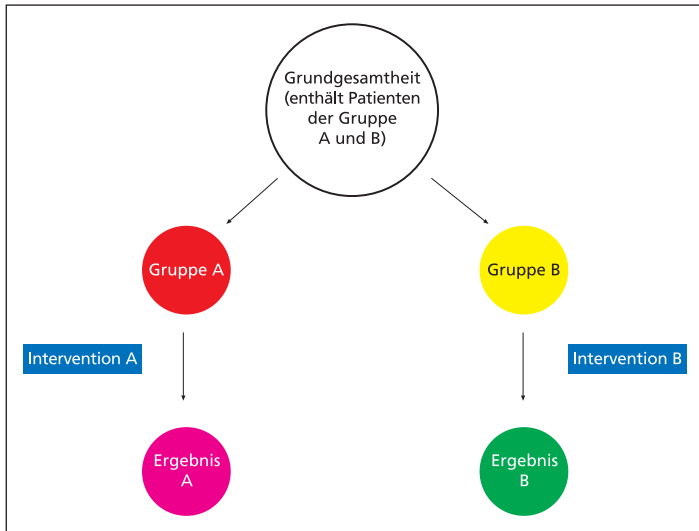


Abb. 1: Unterschiedliche Ergebnisse trotz vergleichbarer Interventionen. Die Ergebnisdifferenzen sind durch die unterschiedliche Ausgangslage (Gruppen A und B) bedingt und nicht durch eine Differenz im Behandlungseffekt.

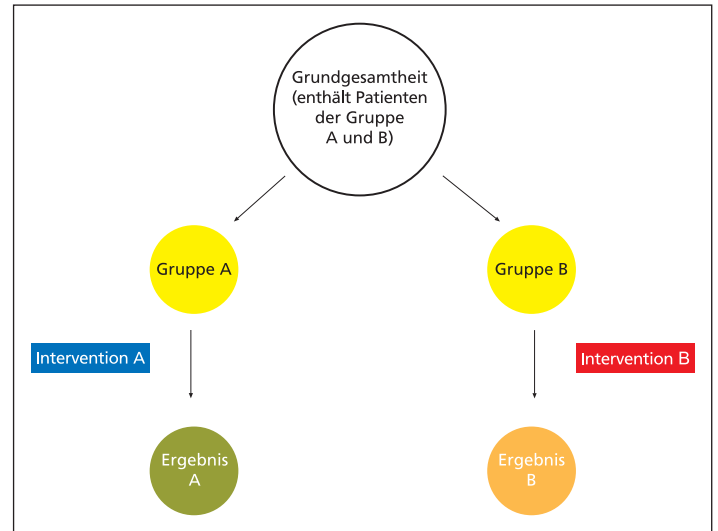


Abb. 2: Unterschiedliche Ergebnisse trotz vergleichbarer Ausgangslage. Die Ergebnisdifferenzen sind durch die unterschiedlichen Interventionen bedingt und nicht durch eine Differenz in der Gruppenzusammensetzung.

misierte Studien, historische Kontrollgruppen, Wirksamkeitsnachweis, evidenzbasierte Medizin

ABSTRACT

This first article in a series of papers to discuss the methodology of clinical research intends to introduce research methods to demonstrate the efficacy of medical interventions. Medical intervention need be effective and safe and should be based on current medical knowledge and evidence. In order to demonstrate improved efficacy and efficiency compared to established/historical medical treatment clinical trials are of utmost importance.

The article focuses on the rationale and prerequisites for clinical trials that aim to demonstrate the efficacy of medical interventions. Special emphasis is put on the importance of ensuring equal initial conditions by using randomization methods. Further, alternative methods to obtain control groups for clinical trials (historical controls, non-randomized controls) with its associated advantages and pitfalls are discussed.

KEY WORDS

Clinical trials, randomization, randomized clinical trials (RCT), non-randomized clinical trials, historical control groups, efficacy, evidence-based medicine

Einleitung

In der medizinischen Versorgung von Patienten kommen die unterschiedlichsten Verfahren zur Anwendung. Dies kann auf dem Gebiet der Kardiotechnik beispielsweise die Verwendung einer Rollerpumpe

oder einer Zentrifugalpumpe zur extrakorporalen Zirkulation sein. Das gewählte Verfahren (auch als „Intervention“ bezeichnet) sollte naturgemäß eine effektive, wirksame Behandlungsform darstellen, die mit möglichst wenigen Nebenwirkungen assoziiert ist und eine gute Kosten-Nutzen-Relation aufweist.

Bevor aber Überlegungen zur Kosten-Effizienz angestellt werden können, müssen die Wirksamkeit und das Ausmaß der Wirksamkeit (Quantifizierung), zum Beispiel gegenüber einem Vergleichsverfahren, definiert werden.

Neben „technischen“ Interventionen, wie die oben genannte Frage nach der Verwendung von Rollerpumpe oder Zentrifugalpumpe, werden auch medikamentöse Therapien oder chirurgische Maßnahmen, wie zum Beispiel modifizierte Maze-Operationen zur Behandlung des Vorhofflimmerns, unter dem Terminus „Intervention“ subsumiert.

Vielfach hat sich eine Behandlungsmethode historisch bedingt im Laufe von Jahren als „Standard“ etabliert. Werden nun auf Basis theoretischer Überlegungen oder experimenteller Versuche neue Verfahren entwickelt, so müssen sie zunächst ihre Wirksamkeit in einer Studie unter Beweis stellen. Nur wenn die neue Methode im Vergleich zur herkömmlichen effektiver oder nebenwirkungsärmer ist als das vormalig als Standard definierte Verfahren, kann empfohlen werden, die alte Methode durch die neue zu ersetzen. Andererseits kann sich ein neues Verfahren aber auch als gleichwertig oder weniger gut wirksam herausstellen als ein etabliertes Standardverfahren, so dass dem herkömmlichen Ver-

fahren der Vorzug gegenüber dem neuen gegeben werden muss.

Voraussetzungen für einen Wirksamkeitsbeweis

Zum Zweck eines Wirksamkeitsnachweises ist es erforderlich, dass

- eine Kontrollgruppe existiert (herkömmliches Verfahren),
- eine Interventionsgruppe existiert (neues Verfahren) und
- dass Kontrollgruppe und Interventionsgruppe in ihrer Zusammensetzung möglichst ähnlich sind und sich nur durch die Intervention (herkömmliches Verfahren, Intervention A, vs. neues Verfahren, Intervention B) unterscheiden.

Diese Voraussetzungen sind notwendig, da bei nicht vorhandener Gleichheit der beiden Gruppen nur schwerlich zu beweisen ist, ob ein unter Umständen beobachteter Effekt durch die Intervention oder durch andere Unterschiede in der Gruppenzusammensetzung herbeigeführt wurde. Dies soll in den Abb. 1 und 2 verdeutlicht werden.

Warum sind Kontrollgruppen meist unabdingbar?

Untersuchungen ohne Kontrollgruppen sind mit Blick auf die Wirksamkeitsbeurteilung einer Intervention naturgemäß ebenso problematisch wie Untersuchungen, in denen sich die untersuchten Gruppen substantiell in Bezug auf die Ausgangsbedingungen unterscheiden. Bei unterschiedlichen Ausgangsbedingungen ist eine Unter- wie auch Überschätzung des durch die untersuchte Intervention erzielten Effektes möglich.

Untersuchungen ohne Kontrollgruppe konstatieren oftmals einen Behandlungs-

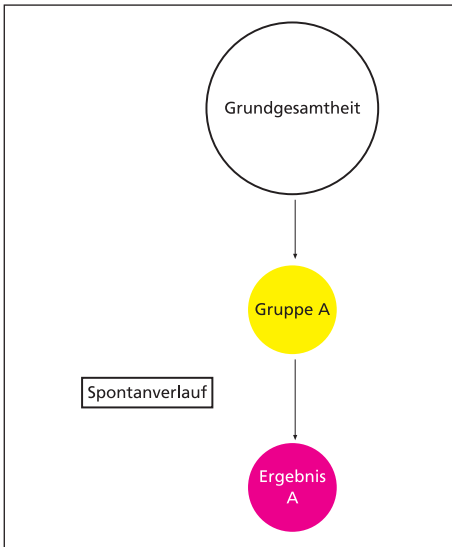


Abb. 3: Veränderung auch ohne Intervention. Die Ergebnisdifferenz im Vergleich zur Ausgangslage beruht auf einem dynamischen Spontanverlauf.

effekt bzw. überschätzen den Therapieeffekt, sofern dieser vorhanden ist. Neben einer subjektiven Färbung der Beobachtung besteht bei Untersuchungen ohne Kontrollgruppe die Schwierigkeit darin, zu unterscheiden, was nun dem Spontanverlauf im Rahmen einer Erkrankung entspricht und welcher Effekt (welche Veränderung) der Intervention zuzuschreiben ist.

Abb. 3 und 4 stellen dar, dass ein am Ende einer Beobachtungsphase zu konstatierender Effekt sowohl durch einen dynamischen Spontanverlauf als auch durch eine stabile Situation und einen „echten“ Behandlungseffekt einer Intervention herbeigeführt werden kann.

WIE KÖNNEN KONTROLLGRUPPEN ERSTELLT WERDEN?

Historische Kontrollgruppe

Die historische Kontrollgruppe ist eine naheliegende, aber auch die für Verzerrungen jeglicher Art anfälligste Variante einer Kon-

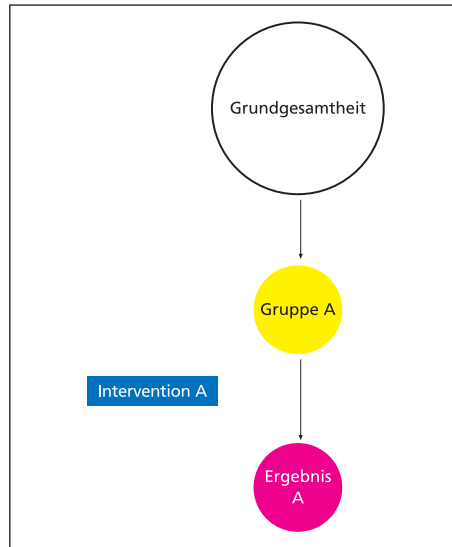


Abb. 4: Veränderung bedingt durch eine Intervention. Die Ergebnisdifferenz im Vergleich zur Ausgangslage beruht auf einem tatsächlichen Behandlungseffekt.

trollgruppe für klinische Studien. Zumeist werden Patienten als historische Kontrollgruppe herangezogen, die zu einer Zeit untersucht und behandelt worden sind, in der es die zu untersuchende Therapie noch nicht gab. Problematisch ist insbesondere die Vermischung einer retrospektiven (rückblickenden) Datenerhebung für die Kontrollgruppe und einer prospektiven (in die Zukunft blickenden) Erhebung für die zu untersuchende („neue“) Intervention. In diesem Fall ist eine einheitliche Behandlung und eine vergleichbare Beurteilung der Effekte, abgesehen von „harten“ (weniger leicht zu verfälschenden) Ergebnissparametern, eher unwahrscheinlich.

Normalerweise werden neue Interventionen im Vergleich zu historischen Kontrollen eher positiv bewertet. Eine Ursache hierfür kann sein, dass sich neben der untersuchten Intervention auch die Begleitmaßnahmen verbessern und damit das „Outcome“ (das Ergebnis) aus Patientensichtweise oder auf-

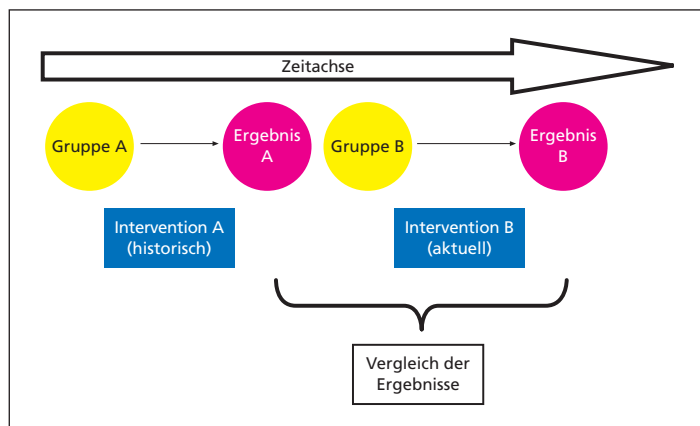


Abb. 5: Historische Kontrollgruppe zur Beurteilung einer durchgeführten Intervention. Verglichen wird die historische Therapie mit der aktuellen Therapie über einen Vergleich der in der Vergangenheit liegenden Ergebnisse mit den gegenwärtig erzielten Ergebnissen.

len jedoch einen großen Vorteil bei eher selten angewendeten Verfahren, in denen eine prospektive Untersuchung nur unter erschwerten Bedingungen und mit erheblichem Zeitaufwand möglich ist. Unter der Prämisse, dass sich die diagnostischen Kriterien und die Kriterien zur Beurteilung des Effektes der Intervention nicht geändert haben, bietet die große Anzahl der verfügbaren Fälle mitunter eine recht stabile Abschätzung von Ereignissen und Effekten. Das Grundprinzip historischer Kontrollgruppen verdeutlicht Abb. 5.

Nicht-randomisierte Auswahl einer Kontrollgruppe

Werden die Gruppen parallel zwei verschiedenen Interventionen zugeordnet, ohne dass diese Zuordnung zufällig (also per „Randomisierung“) geschieht, spricht man von nicht-zufälligem Auswahlverfahren oder eben „nicht-randomisierter“ Zuteilung der Behandlungsgruppen. Selbst wenn man dabei bemüht ist, durch bewusste Aufteilung der bekannten Risikofaktoren eine Gleichheit der Gruppen zu gewährleisten, kann hierbei nicht sicher ausgeschlossen werden, dass ein Ungleichgewicht der Gruppenzuordnung kriert wird.

Unter nicht-randomisierter Zuteilung werden all jene Verfahren subsumiert, bei denen der Arzt oder der jeweilige Untersucher bzw. Studienleiter entscheidet, wie die Gruppenzuteilung erfolgt. Unbewusste und bewusste Manipulation kann hierbei die Vergleichbarkeit der Gruppen beeinträchtigen, so dass am Ende gerade bei vergleichsweise kleinen Effekten nicht sicher gesagt werden kann, welcher Anteil des Effektes der Intervention und welcher Anteil der unterschiedlichen Gruppenzusammensetzung zuzuordnen ist.

Im Vergleich von operativem zu konservativem Vorgehen bei einer bestimmten Erkrankung könnte der die Patienten rekrutierende Arzt zum Beispiel multimorbide Patienten bewusst oder unbewusst der Gruppe mit konservativem (nicht-operativem) Behandlungsschema zuordnen, und umgekehrt für die vergleichsweise jungen und gesunden Patienten eher das operative Vorgehen avisieren. Schon wäre nicht mehr klar, ob ein besseres oder schlechteres „Abschneiden“ der einen oder anderen Gruppe der Intervention oder eben der Verschiedenheit der Gruppen zuzuordnen wäre.

Auch eine Zuteilung nach Wochentagen oder Aufnahmetagen bzw. Geburtstagen ist problematisch. Durch eine Zuteilung nach Wochentagen kann es unter Umständen zu einer Selektion bestimmter Risikoprofile

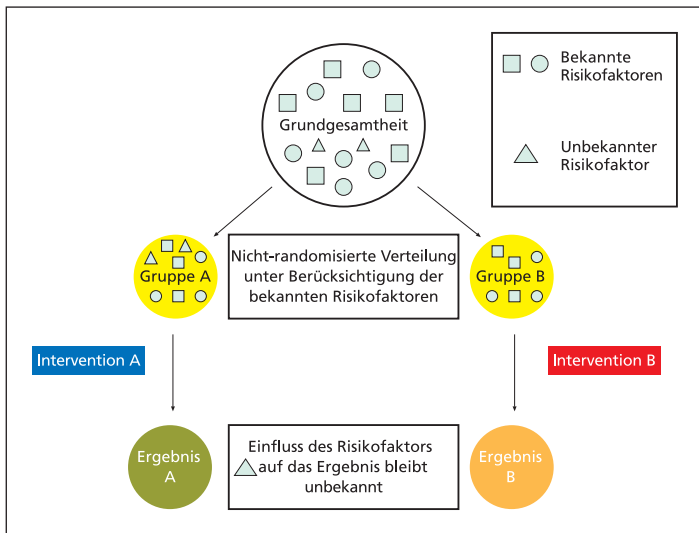


Abb. 6: Nicht-randomisierte Gruppenzuteilung. Da nur die bekannten Risikofaktoren „ausgewogen“ auf die beiden Gruppen verteilt werden können, bleibt ein Restrisiko, dass bislang unbekannte Risikofaktoren den beobachteten Behandlungseffekt bedingen und nicht die durchgeführten Interventionen.

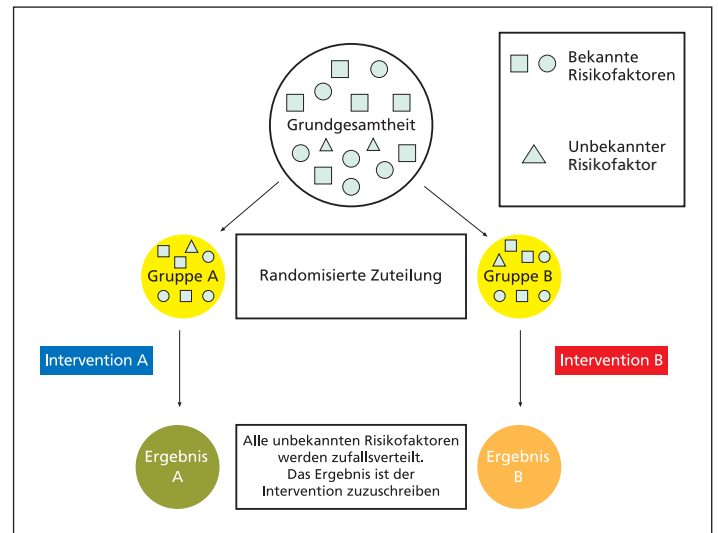


Abb. 7: Randomisierte Gruppenzuteilung. Durch das Zufallsprinzip werden alle Risikofaktoren, auch die bis dato unbekanntem beeinflussenden Faktoren, „ausgewogen“ auf die beiden Gruppen verteilt. Es besteht eine größere Sicherheit, dass die durchgeführten Interventionen den Behandlungseffekt bedingen.

kommen. Bei alternierender Zuteilung, eine in den Anfängen der Entwicklung klinischer Studien häufig geübte Praxis, kann der Arzt ebenfalls durch das Wissen, wie die Zuteilung erfolgen wird, maßgeblich die Gruppenzuteilung bestimmen, indem die Ein- und Ausschlusskriterien für die Untersuchung, und damit die Auswahl der für die Studie in Frage kommenden Patienten, verschieden interpretiert werden. Die nicht-randomisierte Gruppenzuteilung und die hiermit verbundenen Effekte zeigt Abb. 6.

Randomisierte Zuteilung zu Interventions- und Kontrollgruppe

Die zufällige („randomisierte“) Zuteilung der Teilnehmer einer Untersuchung zur Interventions- und Kontroll- bzw. Vergleichsgruppe gilt als Goldstandard, um einen Behandlungseffekt eindeutig einer Intervention zuzuschreiben.

Die Randomisierung dient dabei ganz besonders der zufälligen Aufteilung der mit Blick auf ein Ereignis unbekanntem Risikofaktoren. Wäre zum Beispiel bekannt, dass das Patientengeschlecht ein entscheidender Risikofaktor für das Auftreten von Mikroembolien wäre, und würde man die Hypothese testen, dass Zentrifugal- im Vergleich zu Rollerpumpen ein niedrigeres Embolierisiko mit sich bringen, so könnte man diesen bekannten Risikofaktor bei nicht-randomisierter Zuteilung ja bequem gleich verteilen. Es müssten nur gleich viele Männer wie Frauen in beiden Gruppen zugeteilt werden. Dieses Vorgehen findet im Rahmen klinischer Studien als sogenanntes faktorielles Design in Bezug auf Risikofaktoren auch durchaus Anwendung. Ein

solches Vorgehen funktioniert aber nur für die bekannten Risikofaktoren. Unbekannte Risikofaktoren können zwangsweise nicht gleich verteilt werden. Zudem kann durch die multiple Interaktion der Risikofaktoren eine wirkliche Gleichverteilung des Risikos als Ausgangslage für eine Studie kaum beurteilt werden. Dieses Problem wird, bei entsprechend großen Gruppen, durch die Randomisierung wirksam gelöst. Der Zufall, so das zugrunde liegende Grundprinzip, „verteilt“ die unbekanntem Risikofaktoren mit ihrem jeweiligen relativen Risiko mit Blick auf einen Studienendpunkt gleich auf die Interventions- und Kontrollgruppen.

Die ersten randomisierten Studien wurden in der Landwirtschaft bei der Erprobung von Dünge- und Saatformen eingesetzt. Erst später wurde das Prinzip von den Begründern der modernen Epidemiologie für medizinische Studien und Untersuchungen des öffentlichen Gesundheitswesens übernommen. Das Prinzip der randomisierten Zuteilung verdeutlicht Abb. 7.

ZUSAMMENFASSUNG

Randomisierte Studien gelten als Goldstandard, um den Effekt einer Intervention und dessen Ausmaß im Vergleich zu einer inaktiven Kontrolle oder einer Vergleichsintervention zu bestimmen. Randomisierte Studien ermöglichen sowohl die Festlegung eines kausalen Zusammenhangs als auch die Untersuchung mehrerer Endpunkte sowie die Erfassung der Häufigkeit des Auftretens (Inzidenz) eines Ereignisses.

Dennoch können nicht alle Fragestellungen mit vernünftigen Aufwand-Nut-

zen-Verhältnis auf diese Weise beantwortet werden, weil randomisierte Studien in der Regel zeit- und kostenintensiv sind und keine schon vorhandenen Daten (historische Kontrollen) verwendet werden können. Für seltene Endpunkte ist ihr Einsatz daher fraglich. Durch den prospektiven (in die Zukunft gerichteten) Charakter ergeben sich bei spät eintretendem Effekt mitunter lange Beobachtungs- und damit Studienzeiten.

Alternativen zur randomisierten Studie bestehen in Einzelfällen im Heranziehen historischer Kontrollen bzw. einer nicht-randomisierten Zuteilung zur Interventions- und Kontrollgruppe.

Priv.-Doz. Dr. Peter Kranke, MBA
 Oberarzt
 Klinik und Poliklinik für Anästhesiologie
 Universitätsklinikum Würzburg
 Zentrum Operative Medizin
 Oberdürrbacher Str. 6
 97080 Würzburg
 E-Mail: peter.kranke@t-online.de