Thieme

Konfundierungen und Störfaktoren in rehabilitationswissenschaftlichen Beobachtungs- und Interventionsstudien

Confounding and Biasing Variables in Observational and Intervention Studies in Rehabilitation

Autor
Markus Antonius Wirtz

Institut
Abteilung für Forschungsmethoden, Institut für Psychologie, Pädagogische Hochschule Freiburg

Schlüsselwörter
Konfundierungen, Störfaktoren, Kausalität, Evidenz

Key words
confounders, biasing factors, causation, evidence

Bibliografie
DOI https://doi.org/10.1055/s-0043-120911
Online-Publikation: 7.12.2017
Rehabilitation 2018; 57: 381–387
© Georg Thieme Verlag KG Stuttgart · New York
ISSN 0034-3536

Kontaktadresse
Prof. Dr. Markus Antonius Wirtz
Abteilung für Forschungsmethoden, Institut für Psychologie
Pädagogische Hochschule Freiburg, Kartäuserstraße 47,
79102 Freiburg
markus.wirtz@ph-freiburg.de

ZUSAMMENFASSUNG

Um Effekte von Rehabilitationsmaßnahmen prüfen zu können, muss das Studiendesign so gewählt werden, dass Veränderungen der Zielvariablen (z. B. Gesundheitszustand) möglichst eindeutig auf die Interventionsmaßnahme zurückgeführt werden können. Bei Beobachtungsstudien muss davon ausgegangen werden, dass konfundierende Merkmale den Einfluss der Maßnahme auf die Zielvariable systematisch überlagern und Effektschätzungen verzerren. Die Präferenz für eine Behandlungsalternative, die Behandlungsmotivation sowie das Ausmaß der initialen Beeinträchtigung der Rehabilitanden sind typische Konfundierungen, die insbesondere bei freiwilliger Behandlungsauswahl oder bei Teilnahme nach Behandlungsempfehlung berücksichtigt werden müssen (programmexterne Störfaktoren). Bei gut kontrollierten Interventionsstudien (insb. randomisiert kontrollierte Studien, RCT), bei denen die Inanspruchnahme der Behandlung systematisch kontrolliert erfolgt, können diese Fehlerquellen vermieden werden. Aber auch bei gut kontrollierten Interventionsstudien kann die Studiendurchführung (z. B. aufgrund der Stichprobenauswahl, des Einsatzes von Messinstrumenten, des Austauschs zwischen Vergleichsgruppen; Designeffekte, Treatmentkontaminationen) die eindeutige Interpretierbarkeit der Studienbefunde beeinträchtigen. In diesem Beitrag wird gezeigt, welche Konfundierungen und Störeinflüsse typischerweise die Validität der Schlussfolgerungen aus empirischen Zusammenhängen verringern können. Es wird verdeutlicht, wie potenziell verzerrende Konfundierungen und Effekte von Störfaktoren durch die Gestaltung der Untersuchung vermieden oder bestmöglich kontrolliert werden können.

ABSTRACT

To prove and determine effects of rehabilitation treatments, appropriate study designs have to be applied which allow inferring that differences in the outcome variables (e .g. health state) are causally determined by rehabilitation treatments. In observational studies confounding variables may affect or distort the association between treatment and outcome variables. Preference for a treatment, motivation for treatment or level of impairment before rehabilitation are typical confounders in rehabilitational intervention studies. These potential confounders are especially important if patients chose treatment voluntarily or the treatment has been suggested because of specific impairment characteristics (external program effects). Appropriately controlled intervention studies (especially randomized control trials, RCTs) allow controlling for such biasing effects. But also for appropriate controlled intervention studies, biasing effects may arise due to the implementation of the study design (e. g. selection of participants, applying assessment instruments, communication of participants of different study groups; design effects, treatment contaminations) and may deteriorate the unambiguous interpretability of empirical results. It is shown which confounding and distortional effects have to be regarded that may affect the interpretability of associations. Recommendations for planning, conducting and analyzing empirical studies are given to avoid such biasing effects in the best possible way.

Dieses Dokument wurde zum persönlichen Gebrauch heruntergeladen. Vervielfältigung nur mit Zustimmung des Verlages.

Einleitung

In der rehabilitativen Versorgung werden insbesondere Maßnahmen angewendet, die den Funktions- und Gesundheitszustand der Rehabilitanden positiv beeinflussen sollen. Um Informationen zu systematischen Zusammenhängen zwischen Einflussgrößen bzw. Maßnahmen und Zielkriterien zu erhalten, werden v. a. Beobachtungs- und Interventionsstudien durchgeführt. Gut kontrollierte Interventionsstudien liefern eine höhere Evidenz für die Wirksamkeit von Maßnahmen als Beobachtungsstudien [1], weil wichtige Störeinflüsse, die die eindeutige Interpretierbarkeit der gewonnenen Ergebnisse gefährden, kontrolliert werden.

Ausgehend von der Klärung wichtiger Fehler- und Störquellen wird verdeutlicht, welche Aspekte des Studiendesigns berücksichtigt werden müssen, um die Belastbarkeit von Befunden im Sinne von Wirkungsindikatoren fundiert einschätzen zu können. Die Darstellung kann zudem für die Planung rehabilitationswissenschaftlicher Studien genutzt werden, um durch die Wahl geeigneter Designelemente die Gefahr verfälschter Befunde zu minimieren.

Verzerrungen im Rahmen von Beobachtungsstudien – Konfundierungen und der Einfluss programmexterner Störfaktoren[1]

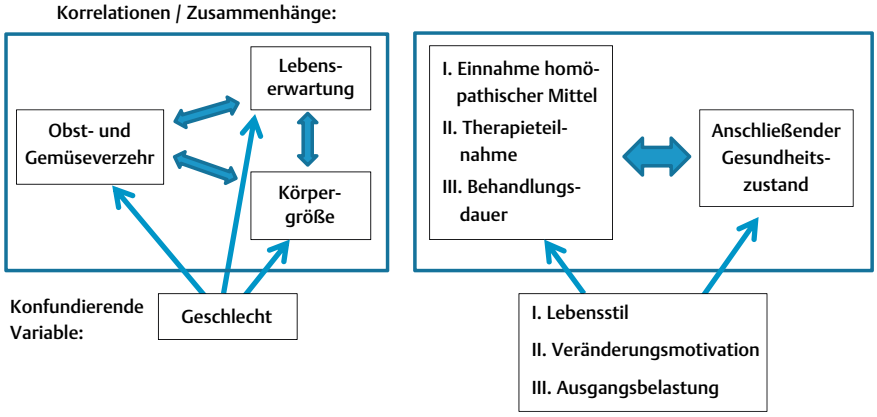
In Beobachtungsstudien werden empirische Sachverhalte und Merkmalsstrukturen erfasst, ohne dass die Untersucher das Geschehen gezielt beeinflussen bzw. intervenierend in empirische Strukturen und Abläufe eingreifen. Neben Querschnittstudien stellen Kohortenstudien und Fall-Kontroll-Studien wichtige Formen epidemiologischer Beobachtungsstudien dar [2, 3]. Bei Beobachtungsstudien ist der Rückschluss von empirischen Zusammenhängen auf potenziell zugrunde liegende Wirkprozesse stets spekulativ, da verzerrende Einflüsse konfundierender Variablen nicht ausgeschlossen werden können.

Was versteht man unter einer Konfundierung?

Stehen 2 Merkmale in einem Zusammenhang, so sind Menschen geneigt anzunehmen, dass diese Merkmale ursächlich miteinander gekoppelt sind, d. h. dass die Ausprägung eines Merkmals diejenige des anderen Merkmals ursächlich bewirkt [4, 5]. Konfundierende Merkmale können jedoch dazu führen, dass Zusammenhänge nachweisbar sind, obwohl zwischen den assoziierten Merkmalen keine wechselseitige Beeinflussung (verursachender, kausaler Wirkeffekt) vorliegt. Im folgenden Beispiel wird die Unzulässigkeit der Ableitung eines ursächlichen Wirkeffekts aus einem Merkmalszusammenhang unmittelbar klar: Angenommen, in einer epidemiologischen Studie wurde festgestellt, dass Menschen, die in höherem Maße Gemüse und Obst verzehren, länger leben und eine geringere Körpergröße aufweisen. Während die Assoziation des Gemüse- und Obstverzehrs mit der Lebenserwartung plausibel erscheint, ist der Zusammenhang mit der Körpergröße überraschend. Das Beispiel wurde so gewählt, dass beide Zusammenhänge durch das konfundierende Merkmal „Geschlecht" bedingt sein könnten. Männer sind durchschnittlich größer als Frauen und haben eine geringere Lebenserwartung. Wenn nun Männer in geringerem Maße Gemüse und Obst verzehren, so entsteht hierdurch indirekt ein Zusammenhang des Obst- und Gemüsekonsums sowohl mit der Lebenserwartung als auch mit der Körpergröße. Aufgrund dieser Zusammenhänge folgt sogar die skurril erscheinende Assoziation, dass mit höherer Körpergröße eine geringere Lebenserwartung korrespondiert (▶ **Abb. 1**, links).

Die folgenden Beispielszenarien verdeutlichen, dass die in diesem Beispiel geschilderte Grundproblematik der Konfundierung für rehabilitationswissenschaftliche Beobachtungsstudien von hoher Bedeutung sein kann.

Szenario 1 – Problem der Unterschiedlichkeit von Vergleichsgruppen Angenommen, es zeige sich, dass Menschen, die homöopathische Mittel einnehmen, generell gesünder sind als Men-



▶ **Abb. 1** Schematische Darstellung, gemäß der die Merkmale im oberen Kasten miteinander assoziiert sind (angedeutet durch Doppelpfeile). Diese Assoziation resultiert gemäß der Erläuterung im Text jedoch lediglich daraus, dass die Ausprägungen dieser 3 Merkmale systematisch mit den konfundierenden Variablen in Zusammenhang stehen.

[1] Der Begriff Programm stammt aus der Evaluationsforschung. Ein Programm kann als Maßnahmenbündel verstanden werden.

Dieses Dokument wurde zum persönlichen Gebrauch heruntergeladen. Vervielfältigung nur mit Zustimmung des Verlages.

schen, die dies nicht tun. Die Schlussfolgerung, dass die Einnahme homöopathischer Mittel die bessere Gesundheit ursächlich bedingt, ist hieraus nicht ableitbar, da sich der allgemeine Lebensstil (z. B. Ernährung, Bewegung) zwischen den beiden Vergleichsgruppen ebenfalls unterscheiden könnte. Aspekte des Lebensstils wären dann die potenziell konfundierenden Variablen: Menschen, die homöopathische Mittel nehmen, legen ggf. allgemein größeren Wert auf einen gesunden Lebensstil, sodass dieser eine alternative Erklärung für den besseren Gesundheitszustand liefern könnte. Nehmen jüngere Menschen eher homöopathische Mittel und sind jüngere Menschen tendenziell gesünder, so würde indirekt ein Zusammenhang aufgrund der konfundierenden Variable „Alter" resultieren.

Szenario 2 – Problem der freiwilligen Teilnahmebereitschaft Menschen, die freiwillig an einem Therapieangebot teilnehmen, seien im Anschluss an die Therapie gesünder als Menschen, die sich bewusst gegen die Therapie entschieden haben. Hieraus kann die Therapiewirkung nicht eindeutig abgeleitet werden. Ggf. sind die freiwillig an der Therapie Teilnehmenden grundsätzlich motivierter, ihren Gesundheitszustand verbessern zu wollen, sodass diese Einstellung und Bereitwilligkeit wesentlich für die positive Veränderung sein kann. Die Veränderungsbereitschaft wäre demnach eine konfundierende Variable: Wer veränderungsbereiter ist, nimmt eher an einem Therapieangebot teil und verändert sich unabhängig von der Therapieteilnahme auch positiver. Aufgrund der Konfundierung lässt sich aus dem Zusammenhang nicht eindeutig ableiten, ob sowohl die Therapieteilnahme als auch die Veränderungsbereitschaft oder nur eines der beiden Merkmale für die positive Veränderung ausschlaggebend sind bzw. ist.

Szenario 3 – Problem der initialen Beeinträchtigung Angenommen, Personen, die sich lange in einer Reha befinden, sind am Ende der Reha nicht gesünder als Personen, die kürzer behandelt wurden. Hieraus lässt sich nicht ableiten, dass die Reha unwirksam ist: Werden Menschen mit starken Beeinträchtigungen lange behandelt und Menschen mit weniger starken Beeinträchtigungen kürzer behandelt, so ist zu erwarten, dass die zweite Gruppe am Ende der Behandlung trotz der geringeren Behandlungsdauer gesünder ist, da Personen dieser Gruppe ja schon günstigere Ausgangsbedingungen aufwiesen. Hier wäre die Beeinträchtigung zu Behandlungsbeginn als konfundierende Variable zu betrachten: denn diese determiniert zum einen die als angemessen erachtete Behandlungsdauer und zum anderen den Gesundheitszustand der Rehabilitanden unabhängig vom Behandlungseffekt.

In den 3 Szenarien liegt dieselbe Problematik vor: 2 Merkmale bzw. Variablen stehen in Zusammenhang (Merkmal 1: Einnahme homöopathischer Mittel, Therapieteilnahme bzw. Behandlungsdauer; Merkmal 2: Gesundheitszustand; ▶ **Abb. 1** rechts). Bei isolierter Betrachtung der Zusammenhänge scheint es naheliegend anzunehmen, dass der Gesundheitszustand durch das jeweils erstgenannte Merkmal determiniert wird. Die Zusammenhänge könnten jedoch durch ein drittes, konfundierendes Merkmal (Lebensstil, Teilnahmebereitschaft, initiale Belastung) bedingt oder mitbestimmt sein.

Das wesentliche Problem von Beobachtungsstudien resultiert also daraus, dass Merkmalsausprägungen dokumentiert und analysiert werden, ohne dass die zugrunde liegenden Wirkprozesse durch ein Untersuchungsdesign kontrolliert werden. Weder die Zuordnung der Rehabilitanden zu den Vergleichsbedingungen noch die Applikation der Maßnahme erfolgt kontrolliert. Möchte man feststellen, welche Prozesse die Zusammenhänge bestimmen, so ist dies mit einer hohen Unsicherheit verbunden: In der Regel bleibt es intransparent, weshalb sich Rehabilitanden für eine Maßnahme entschieden haben. Je unkontrollierter der Studienablauf erfolgt, desto schwerwiegender ist in der Regel die Problematik von Konfundierungen und Störeinflüssen zu werten.

Konfundierende Variablen sind sowohl mit der vermuteten Einflussvariable (z. B. Behandlungsmaßnahme) als auch mit dem Zielkriterium (z. B. Gesundheitszustand) assoziiert. Aus Beobachtungsstudien lässt sich nicht ableiten, ob und in welchem Maße eine vermutete Einflussvariable ein Zielkriterium ursächlich beeinflusst (mangelnde interne Validität), weil konfundierende Variablen die Zusammenhänge vollständig bedingen oder die Zusammenhangsstärke mitbestimmen können. In Beobachtungsstudien ist ein verzerrender Einfluss konfundierender Effekte grundsätzlich nicht auszuschließen, weil nicht sichergestellt werden kann, dass alle konfundierenden Variablen bekannt sind oder angemessen erfasst werden können.

Programmexterne Störfaktoren

Ziel von Wirksamkeitsstudien ist es, den Wirkeffekt einer Behandlung unabhängig von Veränderungen, die auch ohne die Behandlung eingetreten wären, zu bestimmen. Verbessert sich die Funktionsfähigkeit der Rehabilitanden im Verlauf einer 4-wöchigen Rehamaßnahme durchschnittlich um 10 Punkte (Bruttoeffekt), so könnte der tatsächliche Wirkeffekt der Rehabilitation mit einer Verbesserung um z. B. lediglich 5 Punkte (Nettoeffekt) geringer ausfallen, weil sich die Funktionsfähigkeit auch aufgrund des natürlichen Krankheitsverlaufs (z. B. Spontanremission) verbessert. Veränderungen können nur eindeutig auf die Intervention zurückgeführt werden, wenn ausgeschlossen werden kann, dass Veränderungsanteile durch überlagernde Störeinflüsse bedingt sind. Nur direkt durch die Behandlung verursachte Veränderungen dürfen als Wirkung bezeichnet werden.

Rossi et al. [6] unterscheiden verschiedene Arten von Störfaktoren, wobei die Gruppe der programmexternen Störfaktoren besonders bei Beobachtungsstudien beachtet werden muss. Als programmexterne Störfaktoren werden Einflussfaktoren bezeichnet, die unabhängig von der empirischen Untersuchung eines Phänomens wirksam werden.

- **Endogener Wandel:** Verändert sich ein Untersuchungsmerkmal aufgrund eines inhärenten Entwicklungsverlaufs unabhängig von äußeren Umständen, so spricht man von endogenem Wandel. Bewirken die natürlichen Heilungskräfte des Körpers (z. B. Immunreaktionen) einen Regenerationsprozess oder eine Spontanremission, so bessert sich die Symptomatik im Zeitverlauf unabhängig von der Behandlung. Verläuft eine Erkrankung in Schüben, so kann die Messung des Gesundheitszustands zu einem bestimmten Zeitpunkt erheblich durch die aktuelle Erkrankungsphase bestimmt sein.

Dieses Dokument wurde zum persönlichen Gebrauch heruntergeladen. Vervielfältigung nur mit Zustimmung des Verlages.

- **Reife- und Degenerationsprozesse** sind bei der Analyse langfristiger Veränderungen bedeutsam. Ein Programm, das dazu führt, dass die Gedächtnisleistung vom 70. bis zum 80. Lebensjahr unverändert bleibt (Bruttoeffekt = 0), wäre als hoch effektiv zu werten, da natürliche Degenerationsprozesse in dieser Zeitspanne eine deutliche Abnahme der Gedächtnisleistung erwarten ließen. Bei Jugendlichen unterliegt der körperliche Zustand einer systematischen Entwicklung, sodass diese Entwicklung bei der Veränderung der körperlichen Fähigkeiten im Zeitverlauf in Rechnung gestellt werden muss.
- **Exogener Wandel** liegt vor, wenn die Zielvariable von Rahmenbedingungen beeinflusst wird, die selbst einer Veränderung durch natürliche Prozesse, strukturelle Trends oder langfristige Entwicklungen unterliegen. Das psychische Befinden wird z. B. von jahreszeitlichen Veränderungen beeinflusst: Depressionsindikatoren sind z. B. im Winter höher ausgeprägt als im Sommer. Verschlechtert sich die soziale Integration eines Rehabilitanden aufgrund veränderter Umstände im Rahmen der beruflichen Tätigkeit (z. B. Verringerung der Sozialkontakte, Verlust der Mobilität), so muss dies berücksichtigt werden, wenn behandlungsbedingte Veränderungen der sozialen Integration bewertet werden sollen.
- **Akute Ereignisse:** Während exogener Wandel kontinuierliche, eher langsame Veränderungsprozesse bezeichnet, treten akute Ereignisse plötzlich ein und korrespondieren mit abrupten Veränderungen. Der Wechsel des Behandlungspersonals oder der Stationszugehörigkeit kann den Zustand des Rehabilitanden beeinflussen. Die Auskunft, dass die zukünftige Wiederaufnahme der Berufstätigkeit in der bisherigen Form unrealistisch erscheint, kann ebenfalls die interventionsbedingte Veränderung des Gesundheitszustands des Rehabilitanden systematisch überlagern.
- **Unkontrollierte Auswahl:** Erfolgt die Teilnahme an einer Behandlung freiwillig oder aufgrund einer Behandlerempfehlung, so repräsentieren die Behandelten die Population aller Erkrankten nicht angemessen. Nehmen nur Motivierte und Veränderungswillige an einer Maßnahme teil (Selbstselektion), so wird die Behandlungswirkung in der Regel systematisch überschätzt. Scheiden behandelte Patienten im Verlauf aus (Drop-Out), weil sie z. B. frühzeitig feststellen, dass für sie keine Besserung zu erwarten ist, so resultiert in der Regel eine Überschätzung der Effekte, wenn der systematische Drop-Out bei der Auswertung nicht angemessen berücksichtigt wird.
- **Regression zur Mitte** bezeichnet einen besonderen stochastischen Effekt: Die Werte von Personengruppen mit extremen Merkmalsausprägungen tendieren dazu, bei einer Folgemessung weniger extrem zu liegen. Rehabilitanden, die extrem belastet sind, werden in Referenz zu der Population aller Rehabilitanden bei einer zweiten Messung weniger extrem belastet erscheinen. Dies führt in der Regel zu einer Überschätzung von Interventionseffekten, wenn keine adäquate Kontrollgruppe vergleichend betrachtet wird [7].

Kontrollierte Interventionsstudien – Designeffekte und Treatmentkontaminationen

Kontrollierte Studiendesigns vergleichen die Veränderung eines Zielkriteriums (abhängige Variable; z. B. Gesundheitszustand) zwischen Gruppen (unabhängige Variable), die die Behandlungsmaßnahme in Anspruch genommen haben (Interventionsgruppe) oder unbehandelt blieben bzw. an einer alternativen Maßnahme teilgenommen haben (Kontrollgruppe). Bei Interventionsstudien wird die zu untersuchende Interventionsmaßnahme im Rahmen der Studie kontrolliert appliziert. Bei gut kontrollierten Interventionsstudien wird sichergestellt, dass sich die Studienteilnehmer der Interventions- und der Kontrollgruppe zu Studienbeginn nicht systematisch voneinander unterscheiden. Dies kann optimal gelingen, wenn Studienteilnehmer den Vergleichsgruppen zufallsgesteuert zugewiesen werden (Experimentaldesign bzw. randomisiert kontrollierte Studie; RCT). Dadurch, dass sich die Vergleichsgruppen initial nicht systematisch unterscheiden und auch im Behandlungsverlauf gleichen externen Einflüssen unterliegen, wird ein unverzerrter Vergleich sichergestellt.

> Unterscheiden sich Interventions- und Kontrollgruppe zu Studienbeginn hinsichtlich keines Merkmals systematisch, kann eine Konfundierung von Drittvariablen mit den nachweisbaren Veränderungen im Behandlungsverlauf ausgeschlossen werden. Werden die Vergleichsgruppen zudem unter den gleichen Rahmenbedingungen behandelt, kann der Einfluss von im Behandlungsverlauf auftretenden programmexternen Störfaktoren kontrolliert ausgeglichen werden. Entwicklungen in der Interventionsgruppe, die in der Kontrollgruppe nicht auftreten, können somit nur durch die kontrolliert applizierte Intervention kausal bedingt sein. Die Untersuchung hinreichend großer Stichproben stellt sicher, dass Effekte, die lediglich individuell wirksam werden, statistisch kontrolliert werden können.

Erfolgt die Zuweisung zu den Vergleichsgruppen nicht per Zufall, handelt es sich um ein quasi-experimentelles Studiendesign. Gelingt es im Rahmen einer quasi-experimentellen Studie nicht, eine gute Vergleichbarkeit der Ausgangsbedingungen zwischen den Vergleichsgruppen sicherzustellen, ist die Studienqualität hinsichtlich der Gefahr verzerrter Befunde aufgrund von Konfundierungen ähnlich einzuschätzen wie bei Beobachtungsstudien. Gelingt jedoch ein gut begründeter Ausgleich potenziell konfundierende Variablen mittels Techniken wie z. B. Matching oder Parallelisierung oder können Gruppenunterschiede adäquat statistisch kontrolliert werden, so kann ein experimenteller Standard ggf. gut angenähert werden. Trotz der Elimination programmexterner Störfaktoren in gut kontrollierten Interventionsstudien können aber auch bei experimentellen und quasi-experimentellen Designs Störfaktoren verzerrend wirken. Rossi et al. [6] und Shadish, Cook und Campbell [5] unterscheiden Designeffekte und Treatmentkontaminationen.

Dieses Dokument wurde zum persönlichen Gebrauch heruntergeladen. Vervielfältigung nur mit Zustimmung des Verlages.

Designeffekte – verzerrende Effekte, die durch die Studiendurchführung bedingt sind

Rossi et al. [6] bezeichnen verzerrende Effekte in Interventionsstudien, die durch die Anwendung eines Untersuchungsdesigns bedingt sind, als Designeffekte. Ein Studiendesign führt immer zu einer Veränderung der Behandlungspraxis: Es werden Messungen durchgeführt, die adäquate Durchführung der Intervention muss kontrolliert erfolgen und Studienteilnehmer sind sich bewusst, dass sie an einer Untersuchung teilnehmen [6, 8]. Erwartungseffekte können verzerrend wirken, wenn die Untersuchungsleiter, die Behandelnden oder die Rehabilitanden davon überzeugt sind, dass eine Maßnahme besonders wirksam oder überlegen ist. Insbesondere positivere Einstellungen oder größeres Engagement in der Interventionsgruppe sowie den Erwartungen angepasstes Interaktionsverhalten von Behandlern und Rehabilitanden können diesen Effekt bedingen (Pygmalion-Effekt, Selbsterfüllende Prophezeihung; [9]). Der aus der klinischen Forschung bekannte Placebo-Effekt ist ein spezieller Erwartungseffekt. Die Vermeidung durch Doppelblind-Studien ist in rehabilitationswissenschaftlichen Interventionsstudien in natürlichen Behandlungssettings nur in Ausnahmefällen realisierbar.

Beim Hawthorne-Effekt wird das Verhalten dadurch beeinflusst, dass sich Studienteilnehmer bewusst sind, an einer Studie teilzunehmen. Werden z. B. Behandler in einer Rehabilitationsklinik untersucht, so empfinden sie dies ggf. als Leistungskontrolle und vermuten, dass aus den Studiendaten auf ihre individuelle Leistungsfähigkeit zurückgeschlossen wird. Die Aufklärung der Teilnehmer über die tatsächlichen Studieninhalte und die Verwertung der Studienbefunde kann erheblich dazu beitragen, solche Effekte zu vermeiden. Im Rahmen des Studienmonitorings sollten zudem Hinweise gesammelt werden, in welchem Maße die Problematik von Hawthorne-Effekten einflussreich sein kann.

Rossi et al. [6] nennen „inadäquate Ergebnisindikatoren", „mangelnde Zuverlässigkeit/Reliabilität" und „mangelnde Gültigkeit/Validität" als weitere wichtige Störfaktoren, die durch das Erhebungsdesign bedingt sein können. Nicht nur bei Interventionsstudien, sondern auch bei Beobachtungsstudien kommt der zuverlässigen und aussagekräftigen Erfassung von Merkmalen und Outcomeparametern eine besondere Bedeutung zu, damit Merkmalszusammenhänge und Indikatoren für die Wirksamkeit einer Maßnahme nicht verzerrt abgebildet werden. Hier ist auch zu berücksichtigen, dass das Antwortverhalten durch die Bearbeitung von Testmaterialien beeinflusst werden kann (Testungseffekte). Zur Sicherstellung der Aussagekraft und Generalisierbarkeit von Studienergebnissen müssen zudem repräsentative und hinreichend große [10] Stichproben analysiert werden (Störfaktoren gemäß Rossi et al. [6]: Stochastische Effekte, Stichprobeneffekte). Im Falle von Fehlwerten müssen Standards der Missing-Data/-Analyse und –Korrektur berücksichtigt werden [11, 12]. Im Falle von Behandlungsabbrüchen sollten Intention-to-treat-Analysen erfolgen [13].

Treatmentkontaminationen – Implementationsbedingte Verfälschungen der Treatmenteigenschaften

Für die Bewertung von Interventionseffekten ist es wesentlich, dass die Intervention in der Praxis angemessen umgesetzt wird (Treatmentintegrität). Shadish et al. [4] weisen auf die Bedeutung von Veränderungen von Treatmenteigenschaften hin, die dadurch entstehen, dass parallel eine Kontrollbedingung eingeführt wird und die Studienteilnehmer und Behandler ihr Verhalten ggf. entsprechend anpassen.

- **Mangelnde Treatmentintegrität:** Wird eine Maßnahme in der rehabilitativen Versorgung umgesetzt, so ist sicherzustellen, dass durch die lokalen Rahmenbedingungen und die für die Umsetzung verantwortlichen Personen die zentralen Prinzipien und Kernmerkmale der Intervention nicht verändert werden. Die für die Umsetzung Verantwortlichen müssen genau über den Studieninhalt aufgeklärt und ggf. entsprechend geschult werden. Die erforderlichen zeitlichen und materiellen Ressourcen müssen verfügbar sein. Verändern sich die Wahrnehmung und die Gestaltung des Treatments durch die Kenntnis der Existenz der Kontrollgruppe, so ist die Treatmentintegrität ebenfalls gefährdet.
- **Treatmentdiffusion** liegt vor, wenn das Wissen über die in der Vergleichsgruppe angewandten Maßnahmen dazu führt, dass sich die in den Interventions- und/oder Kontrollgruppe angewandten Maßnahmen aneinander angleichen. Erfahren z. B. Mitglieder der Kontrollgruppe, in der die etablierte Routineversorgung umgesetzt wird, dass in der Intervention interessante oder als wirksamer erscheinende Neuerungen eingeführt werden, so werden ggf. Elemente dieser Neuerungen abweichend vom Studienkonzept imitiert oder übernommen. Somit verändert sich das Verhalten in der Treatmentgruppe und/oder Kontrollgruppe durch das Wissen, was in der jeweiligen Vergleichsgruppe erfolgt. Auf Behandlerebene ist diese Gefahr insbesondere dann gegeben, wenn die Behandler für die Umsetzung der Maßnahmen sowohl in der Interventions- als auch in der Kontrollgruppe zuständig sind, oder aber, wenn die Behandler in einem Team zusammenarbeiten bzw. sich miteinander austauschen. Auf Ebene der Rehabilitanden ist die Gefahr einer Treatmentdiffusion besonders dann gegeben, wenn die Interventions- und Kontrollbedingung in der gleichen Einrichtung oder gar auf derselben Station bei unterschiedlichen Rehabilitanden umgesetzt werden, da der Austausch über die Behandlungsinhalte wahrscheinlich ist.
- **Kompensatorischer Ausgleich** liegt vor, wenn die Behandler den Eindruck gewinnen, dass Rehabilitanden in der Kontrollgruppe benachteiligt werden. Werden dann Maßnahmen empfohlen oder angewendet, die die befürchtete Benachteiligung ausgleichen sollen, so wird der empirisch bestimmbare Unterschied zwischen Interventions- und Kontrollgruppe ggf. verfälscht.
- **Kompensatorischer Wettstreit** bezeichnet das Phänomen, dass Behandler der Kontrollgruppe befürchten, als die schlechteren Therapeuten evaluiert zu werden und als Reaktion darauf im Verlauf der Studie besonders engagiert handeln. Ähnlich wie beim Hawthorne-Effekt ist dieses Phänomen besonders dann zu befürchten, wenn die Studienteilnehmer die Evaluation so verstehen, dass das eigene Handeln und nicht die Wirksamkeit der Intervention bewertet werden soll.
- **Empörte Demoralisierung** bezeichnet das Gegenteil kompensatorischen Wettstreits. Fühlen sich Behandler oder Rehabilitanden als Mitglieder der Kontrollgruppe benachtei-

Dieses Dokument wurde zum persönlichen Gebrauch heruntergeladen. Vervielfältigung nur mit Zustimmung des Verlages.

ligt, so verlieren sie ggf. die Motivation, überhaupt aktiv an den in der Kontrollgruppe applizierten Maßnahmen teilzunehmen oder diese engagiert umzusetzen. Hierbei besteht die Gefahr, dass allein aufgrund der erwarteten Unterlegenheit der Kontrollgruppe die Interventionsgruppe fälschlicherweise oder in zu starkem Maße als überlegen nachgewiesen wird.

Designeffekte und Treatmentkontaminationen können durch die Wahl einer Kontrollgruppe in der Regel nicht angemessen kontrolliert werden. Vielmehr bedarf es einer Aufklärung über die Studienhintergründe und einer sauberen Implementierung des zu untersuchenden Treatments und der Kontrollbedingung. Um die Angemessenheit und Qualität der tatsächlichen Umsetzung von Maßnahmen gewährleisten und ggf. korrigierend eingreifen zu können, ist ein Monitoring des Implementationsprozesses erforderlich [14]. Das Monitoring sollte den Umsetzungsprozess detailliert dokumentieren, sowie Barrieren und Erfolgsfaktoren identifizieren. Die Gefahr von Treatmentkontaminationen kann bei Studien, die in denselben Behandlungseinrichtungen Treatment- und Kontrollbedingungen parallel anbieten, besonders hoch sein. Bei groß angelegten, multizentrischen Studien bietet hier die Randomisierung von Versorgungseinheiten (z. B. Stationen) im Rahmen cluster-randomisierten Studien eine günstigere Designvariante [15].

Anwendungsbeispiel: Förderung der Nachhaltigkeit von Rehaeffekten durch ein internetgestütztes Nachsorgeangebot

Es soll untersucht werden, ob ein internetgestütztes Nachsorgeangebot die Nachhaltigkeit von Rehaeffekten fördert. Das Nachsorgeangebot stellt im Internet Informationen und Übungsmaterialien zur Verfügung und nutzt soziale Netzwerke für die Aufrechterhaltung der Kommunikation der Rehabilitanden untereinander und mit den Behandlern nach Rehaende. Von einer Beobachtungsstudie würde man hier z. B. sprechen, wenn im Rahmen der Rehabilitation lediglich auf solche Angebote als eine von mehreren Maßnahmen hingewiesen wird und die Rehabilitanden sich freiwillig und nach Bedarf für dessen Nutzung entscheiden können. Ein Jahr nach Rehaende wird erhoben, ob bzw. welche internetgestützten Angebote genutzt wurden und wie sich der Funktions- und Gesundheitszustand entwickelt hat. Da es sich nicht um eine gut kontrollierte Interventionsstudie handelt, erschweren potenzielle Konfundierungen und programmexterne Störfaktoren die Wirksamkeitseinschätzung: Es wäre plausibel anzunehmen, dass jüngere Menschen und Menschen mit höherem Bildungsgrad das Internet und damit auch das Nachsorgeprogramm eher nutzen (Selektionsprozesse). Wenn bei diesen auch grundsätzlich eine bessere Gesundung erwartet werden kann, liegt eine systematische Konfundierung vor. Zudem ist zu beachten, dass sich der Gesundheitszustand im Verlauf einer 1-Jahres-Katamnese ohnehin verändert (i. d. R. Abschwächung der Rehaeffekte; endogener Wandel). Der Gesundheitszustand wird zudem ggf. durch parallele Alterungsprozesse (Degenerationsprozesse) oder durch die (modifizierte) Wiederaufnahme bzw. Nichtwiederaufnahme einer beruflichen Tätigkeit beeinflusst. Möglicherweise etablieren sich alternative Gesundheitsförderungsmaßnahmen durch den Gesetzgeber oder die Krankenkassen (exogener Wandel, akute Ereignisse). Nehmen nur hochbelastete das Zusatzangebot in Anspruch, so wird zudem der Effekt der Regression zur Mitte bedeutsam.

Durch eine kontrollierte Interventionsstudie können diese Effekte vermieden werden: Eine Interventionsgruppe wird explizit in der Nutzung des internetgestützten Angebots geschult und erklärt sich bereit, dieses nach Beendigung der Reha systematisch zu nutzen. Die Kontrollgruppe bleibt ohne Schulung. Würden die Rehabilitanden selbst entscheiden, ob sie das Internet nutzen oder nicht (schwach kontrolliertes quasi-experimentelles Design), müsste hier ebenfalls durch Selektionseffekte bedingt mit erheblichen Verzerrungen gerechnet werden. Deswegen sollte eine randomisierte Zuweisung zu den Vergleichsgruppen angestrebt werden (experimentelles Design; RCT). Dadurch, dass die Studie nun aufwendiger strukturiert und evaluiert wird, tritt der Studiencharakter der Intervention deutlicher in den Vordergrund. Alle im Text genannten Designeffekte, die Verhaltensänderungen aller an der Maßnahme Beteiligten hervorrufen können (insb. Erwartungseffekte, Hawthorne-Effekt), müssen hier berücksichtigt werden. Es sollte eine explizite Aufklärung aller Beteiligten über Aspekte, die für die adäquate Studiendurchführung erforderlich sind, erfolgen. Im Rahmen des Studienmonitoring sollte Aufschluss über Effekte und Phänomene gewonnen werden, die Hinweise auf potenzielle Verzerrungseffekte liefern.

Es besteht zudem die Gefahr, dass es zu Veränderungen des Praxishandelns kommt, wenn Behandler und Rehabilitanden Kenntnis über die Inhalte des Behandlungsprogramms in der Vergleichsgruppe haben. Würde die Schulung vs. Nicht-Schulung bei Rehabilitanden, die auf derselben Station versorgt werden, erfolgen, so bestünde eine hohe Gefahr von Treatmentkontaminationen. Behandler und Rehabilitanden der Kontrollgruppe wären über das Informationsangebot in der Schulungsgruppe informiert: Die Übernahme von Schulungsinformationen in der Kontrollgruppe (z. B. Austausch über interessanter Internetquellen und –angebote; Zusammenfassung der wichtigsten Maßnahmenelemente in einer verschriftlichten Handreichung; Aufnahme von Kontrollgruppenmitgliedern in die sozialen Netzwerke aufgrund persönlicher Kontakte; Treatmentdiffusion), das Vermitteln alternativer Unterstützungsangebote in der Kontrollgruppe (z. B. Information über alternative Online-Angebote; individualisierte Betreuung aufgrund größerer Zeitressourcen; kompensatorischer Ausgleich) oder die Intensivierung der üblichen Behandlungselemente (z. B. kompensatorischer Wettstreit) wären nicht auszuschließen. Ebenfalls könnte in der Kontrollgruppe der Eindruck entstehen, schlechter informiert und betreut zu werden, sodass deren Behandlungsadhärenz verringert würde (empörte Demoralisierung). Aus diesen Gründen sollte die Durchführung einer cluster-randomisierten Studie, bei der Stationen den Vergleichsgruppen zufallsgesteuert zugewiesen werden, in Erwägung gezogen werden.

Dieses Dokument wurde zum persönlichen Gebrauch heruntergeladen. Vervielfältigung nur mit Zustimmung des Verlages.

KERNBOTSCHAFT

Die Aussagekraft empirischer Studienbefunde wird maßgeblich durch das Studiendesign festgelegt. Beobachtungsstudien liefern Hinweise auf mögliche Ursache-Wirkungsbeziehungen. Sie sind jedoch nicht geeignet, diese zu belegen (unzureichende interne Validität, geringe Evidenz), da mögliche Konfundierungen und programmexterne Störfaktoren die eindeutige Interpretierbarkeit von Befunden erschweren. Ein Merkmal ist konfundiert, wenn dieses sowohl mit der Inanspruchnahme der Behandlung als auch dem abhängigen Merkmal in Zusammenhang steht. Als programmexterne Störfaktoren werden Einflussmerkmale bezeichnet, die unabhängig von der empirischen Untersuchung wirksam werden. Der Nachweis einer hohen Evidenz für die Wirksamkeit der Behandlung setzt voraus, dass sich im Rahmen einer kontrollierten Interventionsstudie die Interventions- und Kontrollgruppe zu Behandlungsbeginn nicht systematisch voneinander unterscheiden und beide Gruppen im Untersuchungsverlauf gleichen Rahmenbedingungen ausgesetzt sind. So können Verzerrungen durch Konfundierungen und programmexterne Störfaktoren ausgeschlossen werden. Aber auch hier muss sichergestellt werden, dass durch die Durchführung der wissenschaftlichen Studie keine Designeffekte die Aussagekraft der Daten beeinträchtigen. Zudem müssen sowohl in der Interventions- als auch in der Kontrollgruppe die Behandlungsmaßnahmen adäquat realisiert werden (Vermeidung von Treatmentkontaminationen). Die zufallsgesteuerte Zuweisung der Studienteilnehmer zu den Vergleichsgruppen im Rahmen eines RCT und ein systematisches Studienmonitoring dienen der Vermeidung von Verzerrungen durch Konfundierung und Störfaktoren.

Koordinatoren der Reihe „Methoden in der Rehabilitationsforschung“

Prof. Dr. Dr. Hermann Faller, Würzburg; Prof. Dr. Thomas Kohlmann, Greifswald; Prof. Dr. Markus Antonius Wirtz, Freiburg; Prof. Dr. Dr. Christian Zwingmann, Siegburg. Interessenten, die einen Beitrag zur Reihe beisteuern möchten, werden gebeten, vorab Kontakt aufzunehmen: h.faller@uni-wuerzburg. de.

Interessenkonflikt

Der Autor gibt an, dass kein Interessenkonflikt besteht.

Literatur

[1] Kunz R, Ollenschläger G, Raspe H et al., (Hrsg). Lehrbuch Evidenzbasierte Medizin in Klinik und Praxis. Köln: Deutscher Ärzte-Verlag; 2007

[2] Hammer GP, Prel J-B, Blettner M. Vermeidung verzerrter Ergebnisse in Beobachtungsstudien. Deutsches Ärzteblatt 2009; 106: 664–668

[3] Machin D, Campbell MJ. Design of studies for medical research. Chichester: Wiley; 2005

[4] Campbell DT, Stanley JC. Experimental and quasi-experimental designs for generalized causal inference. Boston: Hughton; 1963

[5] Shadish WR, Cook TD, Campbell DT. Experimental and quasi-experimental designs for generalized causal inference. Boston: Houghton Mifflin; 2002

[6] Rossi PH, Lipsey MW, Freeman HE. Evaluation: A systematic approach. Thousand Oaks: Sage; 2003

[7] Zwingmann C, Wirtz M. Regression zur Mitte. Rehabilitation 2005; 44: 244–251

[8] Döring N, Bortz J. Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften. Berlin: Springer; 2016

[9] Watzlawick P, Beavin JH, Jackson DD. Menschliche Kommunikation. Bern: Huber; 2011

[10] Kutschmann M, Bender R, Grouven U et al. Aspekte der Fallzahlkalkulation und Powerberechnung anhand von Beispielen aus der rehabilitationswissenschaftlichen Forschung. Rehabilitation 2006; 45: 377–384

[11] Cho A, Leonhart R. Lösungsansätze für das Problem fehlender Werte: Mechanismen erkennen und adäquat behandeln. Rehabilitation 2013; 52: 273–279

[12] Wirtz M. Über das Problem fehlender Werte. Rehabilitation 2004; 43: 109–115

[13] Faller H. Intention-to-treat. Rehabilitation 2004; 43: 52–55

[14] Christiansen G. Evaluation – Ein Instrument zur Qualitätssicherung in der Gesundheitsförderung. Köln: BZgA; 1999

[15] Campbell MK, Piaggio G, Elbourne DR et al. Consort 2010 statement: Extension to cluster randomised trials. BMJ 2012, doi:10.1136/bmj.e5661

Dieses Dokument wurde zum persönlichen Gebrauch heruntergeladen. Vervielfältigung nur mit Zustimmung des Verlages.