

Klinische Studien zu Arzneimitteln – Wo ist der Haken?

Verloren, aber nicht unersetzlich? – Vom Umgang mit fehlenden Daten

Dr. med. Natascha Einhart, Prof. Dr. Tim Mathes

Nachdruck aus „Arzneiverordnung in der Praxis“, Band 52, Heft 2, September 2025. Im Internet unter www.akdae.de oder via Kurzlink <https://t1p.de/h3wyi>

Wer evidenzbasiert argumentiert, bezieht sich auf den aktuellen Wissensstand aus Studien, nicht nur auf persönliche Erfahrungen oder auf die Meinung von Experten. Das bloße Zitat einer Studie ist aber noch keine evidenzbasierte Argumentation. Studien bieten keine unumstößlichen Wahrheiten, sondern Ergebnisse statistischer Analysen. Jeder Studientyp hat dabei spezifische Stärken und Schwächen. Diese Artikelreihe in AVP konzentriert sich auf klinische Studien, die experimentell Wirksamkeit und Verträglichkeit von Arzneimitteln prüfen. In kurzen Beiträgen möchten wir Sie mit dem nötigen „Werkzeug“ ausstatten, um klinische Studien zu Arzneimitteln kritisch zu lesen und sich Ihre eigene, evidenzbasierte Meinung zu bilden. Im HÄBL 12/2025 findet sich ein weiterer Teil dieser Serie, Link: <https://t1p.de/5pnp6> – der QR-Code führt direkt dorthin.



Rückblick: Per-protokoll, As-treated oder Intention-to-treat: Was machen wir mit den „Abtrünnigen“? (siehe HÄBL 12/2025, Link siehe oben)

In Studien weichen Teilnehmer häufig von den geplanten Therapieschemata ab. Diese Abweichungen können durch Patientenmerkmale bedingt sein, die gleichzeitig den Krankheitsverlauf oder die Wirksamkeit und Verträglichkeit des Arzneimittels beeinflussen (z. B. vorbestehende Allergien). Bei einer Intention-to-treat(ITT)-Analyse werden Studienteilnehmer auch dann entsprechend ihrer ursprünglichen Zuteilung ausgewertet, wenn sie die Studienmedikation nicht wie vorgesehen einnahmen. Dadurch bleibt die durch die Randomisierung erzeugte Strukturgleichheit von Kontroll- und Interventionsgruppe erhalten. Bei As-treated(AT)- oder Per-protocol(PP)-Analysen ist dagegen nicht gewährleistet, dass die ausgewerteten Gruppen strukturgleich sind. Die Aussagesicherheit von AT- und PP-Analysen ist deshalb geringer als die Aussagesicherheit von ITT-Analysen.

Im Beispiel unserer Artikelserie möchten Sie herausfinden, ob die medizinische Handcreme Supersoft das Risiko für Hand-

ekzeme bei Pflegepersonal stärker reduziert als in Drogerien erhältliche Handcremes. Dafür führen Sie eine kontrollierte randomisierte Studie durch, in der sie die Teilnehmer zufällig einer von zwei Behandlungen zuteilen¹: Die Studienteilnehmer tragen dreimal täglich entweder Supersoft (Interventionsgruppe) oder ihre bisher genutzte Handcreme (Kontrollgruppe) auf. Durch die Randomisierung haben Sie (tendenziell) strukturgleiche Kontroll- und Vergleichsgruppen erreicht, beispielsweise hinsichtlich der Häufigkeit vorbestehender Allergien. Damit diese Strukturgleichheit erhalten bleibt, wollen Sie eine Intention-to-treat(ITT)-Analyse durchführen: Sie möchten alle Studienteilnehmer entsprechend ihrer ursprünglichen Zuteilung auswerten („once randomized, always analyzed“ bzw. „analysiert wie randomisiert“). Sie laden deshalb alle 16 Teilnehmer ein Jahr nach Studienbeginn zur Abschlussvisite ein, auch diejenigen, die ihre zugeteilte Handcreme nicht wie vorgesehen benutzen.

Das Ergebnis der Abschlussvisite ist in Abbildung 1 dargestellt und scheint eindeutig für die Handcreme Supersoft zu sprechen: In der Abschlussvisite weist kein Patient der Interventionsgruppe, aber fast

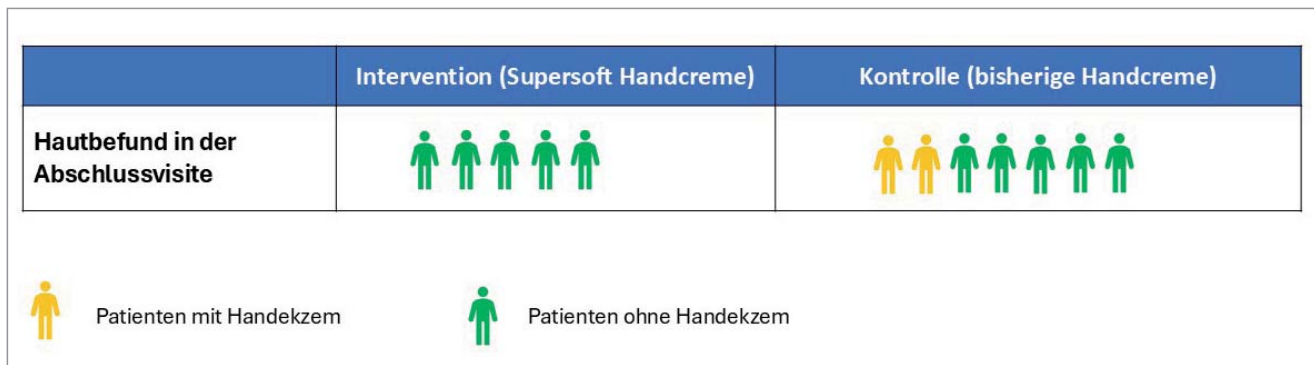


Abb. 1: Zur Abschlussvisite erscheinen zwölf Studienteilnehmer. Bei der Untersuchung wird festgestellt, dass in der Interventionsgruppe kein Teilnehmer ein Handekzem aufweist, während in der Kontrollgruppe bei zwei Teilnehmern ein Handekzem besteht.

¹ Bei diesem Gedankenexperiment klammern wir statistische Fragen zur Errechnung der geeigneten Stichprobengröße ebenso aus wie die notwendige Bewilligung Ihrer Studie durch die Ethikkommission.

	Intervention (Supersoft Handcreme)	Kontrolle (bisherige Handcreme)
Hautbefund zu Studienende		
Erscheinen zur Visite		
	Patienten mit schwerem Handekzem	Patienten mit leichtem Handekzem Patienten ohne Handekzem

Abb. 2: Jeweils drei Teilnehmer der Interventions- und der Kontrollgruppe haben ein Handekzem entwickelt. Bei den Teilnehmern aus der Interventionsgruppe war das Handekzem so stark ausgeprägt, dass sie bei Studienende ihren Arbeitsplatz gewechselt hatten und nicht zur Abschlussvisite erschienen. In der Kontrollgruppe bestanden nur leichte Handekzeme. Lediglich ein Teilnehmer aus der Kontrollgruppe fehlte bei der Abschlussuntersuchung.

30 % der Kontrollgruppe ein Handekzem auf. Wo könnte hier der Haken sein?

Offensichtlich sind nicht alle Teilnehmer ihrer Einladung zur Abschlussvisite gefolgt. Bei 25 % der Studienteilnehmer wissen Sie deshalb nicht, ob bei ihnen zum Studienende ein Handekzem bestand. Es sind viele Gründe denkbar, warum Untersuchungsdaten fehlen. Denkbar sind beispielsweise folgende Szenarien:

1. Der Studienassistent war erkrankt und geplante Untersuchungen sind ausgefallen.
2. Studienteilnehmer haben aufgrund allergischer Reaktionen nicht nur die Studienmedikation abgesetzt, sondern ihre Teilnahme gänzlich abgebrochen.
3. Studienteilnehmer konnten nicht kontaktiert werden, weil sie den Arbeitsplatz gewechselt haben.

Fehlende Daten verzerren nur dann die Beurteilung der untersuchten Endpunkte, wenn 1. bestimmte Merkmale der Studienteilnehmer zu ihrem Fehlen führten und wenn 2. diese Merkmale durch die Wirksamkeit oder Verträglichkeit der Studienmedikation beeinflusst wurden. In Szenario 1 (Erkrankung des Studienassistenten) kann davon ausgegangen werden, dass die Eigenschaften der Studienteilnehmer keinen Einfluss darauf hatten, ob die Abschlussuntersuchung stattfand oder nicht; in Szenario 2 (Studienabbruch aufgrund unerwünschter Ereignisse) erscheint es plausibel, dass das Auftreten allergischer

Reaktionen durch die Studienmedikation verursacht wurde. Komplizierter ist dagegen die Beurteilung von Szenario 3: Ein Arbeitsplatzwechsel kann durch gänzlich von der Studienmedikation unabhängige Gründe motiviert sein (beispielsweise durch Konflikte mit der Stationsleitung), aber auch auf eine unzureichende Wirksamkeit der Studienmedikation zurückzuführen sein (höheres Risiko für das Auftreten von Handekzemen). Die Beurteilung, ob Daten „zufällig“ (d. h. unabhängig von der untersuchten Intervention) fehlen, basiert somit immer auf mehr oder weniger sicheren Annahmen.

Abbildung 2 stellt ein Szenario dar, in dem mehrere Teilnehmer ihren Arbeitsplatz aufgrund eines Handekzems gewechselt haben. In beiden Behandlungsgruppen erkrankten im Studienverlauf drei von acht Teilnehmern an einem Handekzem. Unter Supersoft war das Handekzem so stark ausgeprägt, dass nach einem Jahr alle drei Pflegekräfte eine andere Tätigkeit ausübten und deshalb die Einladung zur Abschlussvisite nicht erhielten.

In der Kontrollgruppe war das Handekzem deutlich leichter ausgeprägt. Nur ein Studienteilnehmer hatte aufgrund des Handekzems seinen Arbeitsplatz gewechselt, die beiden anderen Teilnehmer waren weiterhin in ihrer Station tätig und nahmen an der Abschlussuntersuchung teil. Die erfassten Daten suggerieren in diesem Szenario fälschlicherweise, dass unter Supersoft der Entwicklung von Handekze-

men deutlich besser vorgebeugt wird als unter einer handelsüblichen Handcreme.

Wie kann eine solche Verzerrung durch fehlende Daten vermieden werden? Im besten Fall gelingt es, das Fehlen von Daten weitestgehend zu vermeiden. Anstatt die Studienteilnehmer nur über die dienstliche E-Mail anzuschreiben, könnte beispielsweise auch ein Brief an die Meldeadresse verschickt werden – dann würde die Einladung zur Abschlussvisite auch diejenigen Studienteilnehmer erreichen, die zwischenzeitlich ihre Arbeitsstelle gewechselt haben.

Insbesondere bei langen Studiendauern ist es allerdings unvermeidlich, dass Untersuchungsdaten fehlen, weil Patienten aufgrund gesundheitlicher oder persönlicher Gründe nicht mehr an den Erhebungen teilnehmen können bzw. wollen. Es gibt keinen allgemein anerkannten Grenzwert, bis zu welchem die Aussagesicherheit der Studie nicht gefährdet ist. Im Risk-of-Bias (RoB) Tool von Cochrane [1] wird für kontinuierliche Endpunkte eine (potenziell) reduzierte Aussagesicherheit bereits bei mehr als 5 % Datenverlust angenommen und für dichotome Endpunkte² auf eine Festlegung verzichtet (die Anzahl beobachteter Ereignisse sollte hier laut Cochrane „deutlich größer“ sein als die Anzahl der Teilnehmer mit fehlenden Ereignisdaten, siehe auch [2]).

Fehlende Daten werden in der Statistik als „missing completely at random“, „missing at random“ oder „missing not at random“

² Kontinuierliche Endpunkte sind Ergebnisse, die auf einer kontinuierlichen Werteskala gemessen werden, wie zum Beispiel die Höhe des Blutdrucks. Dichotome Endpunkte sind Ereignisse, die für einen Patienten nur entweder eintreten oder nicht eintreten können (beispielsweise haben Teilnehmer entweder einen Herzinfarkt oder sie haben keinen Herzinfarkt) [3].

bezeichnet, je nachdem, ob die Daten zufällig fehlen und ob ihr Fehlen durch vorhandene Informationen erklärt werden kann. Bei einem vollständig zufälligen Fehlen (**missing completely at random**) beeinflussen die fehlenden Daten nicht das Studienergebnis, da die nicht zur Abschlussvisite erschienenen Patienten sich nicht systematisch von den untersuchten Patienten unterscheiden (Szenario 1: Erkrankung des Studienassistenten). Falls die „missing completely at random“-Annahme adäquat erscheint, kann deshalb eine Complete Case Analyse durchgeführt werden, bei der nur diejenigen Patienten analysiert werden, für die vollständige Daten vorliegen.

Wenn Daten nicht zufällig fehlen, würde eine Complete Case Analyse das Studienergebnis verzerren (siehe Abb. 2). Eine Lösungsstrategie ist die sogenannte **Imputation**: Hierbei werden fehlende Daten durch plausibel erscheinende Werte ersetzt, indem man Informationen aus den vorhandenen Daten verwendet. Die Schätzung für die Imputation ist umso genauer, je mehr man über die Studienteilnehmer weiß, bevor sie „verloren“ gehen. Im Alltag gehen wir ähnlich vor, wenn wir

das zukünftige Verhalten einer bestimmten Person vorhersagen wollen. Vor Ihrem mündlichen Staatsexamen hatten Sie beispielsweise ein großes Interesse daran, möglichst viel über Ihnen zukünftigen Prüfer zu erfahren. Sobald Sie die „Baseline“-Daten Ihres Prüfers kannten, also seinen Fachbereich und seine Forschungsschwerpunkte, konnten Sie abschätzen, welche Themen er bei Ihrer Prüfung mit einiger Wahrscheinlichkeit abfragen würde. Noch besser ließ sich eine Vorhersage treffen, wenn Sie wussten, welche Fragen er in der Vergangenheit bei Prüfungen gestellt hatte – weshalb Prüfungsprotokolle bei Medizinstudenten von Jahrgang zu Jahrgang weitergegeben werden. Aus dem gleichen Grund ist es in klinischen Studien wichtig, dass sowohl umfangreiche Baseline-Charakteristika erhoben werden als auch regelmäßige Zwischenerhebungen erfolgen.

Bei der Auswertung klinischer Studien ist die Nutzung von Imputationsverfahren nur zulässig, wenn davon ausgegangen wird, dass die nicht zufällig fehlenden Daten vollständig mit Hilfe der vorhandenen Daten erklärt werden können. Unter dieser Annahme der Ersetzbarkeit werden

nicht zufällig fehlende Daten (verwirrenderweise) von Statistikern als **missing at random** (zufällig fehlend) beschrieben. Dagegen bezeichnen Statistiker Daten als **missing not at random** (nicht zufällig fehlend), wenn sie vermuten, dass unbekannte Faktoren die Datenerhebung beeinflussten, die mit den vorhandenen Daten nicht aufgeklärt werden können.³ In Szenario 3 wären fehlende Abschlussdaten somit „missing at random“, wenn der Arbeitsplatzwechsel und vorausgegangene Krankschreibungen dem Studienleiter bekannt wären (z. B. über eine Abfrage bei der Personalverwaltung) und sie wären „missing not at random“, wenn diese Informationen nicht für die Ersetzung der fehlenden Werte verwendet werden könnten.

Patienten teilen im Regelfall nicht detailliert mit, warum sie eine Studie abbrechen bzw. an einzelnen Datenerhebungen nicht teilnehmen. Studienleiter und Statistiker können somit oft nur mutmaßen, ob Daten zufällig fehlen und ob in ihre Imputation alle wesentlichen Faktoren einfließen. Zur Überprüfung der „missing completely at random“- und „missing at random“-Annahme sollten deshalb Sensitivitätsanaly-

Tabelle 1: Wo gibt es Informationen zum Umgang mit fehlenden Daten?

Methodische Frage*	Fundort in der Primärpublikation
Unterscheidet sich der Anteil fehlender Daten zwischen den Armen?	Der Anteil fehlender Daten sollte pro Endpunkt und Erhebungszeitpunkt angegeben werden. Sofern publiziert, finden sich diese Informationen am ehesten in der tabellarischen Darstellung der Ergebnisse (Achtung: in der Kopfzeile der tabellarischen Endpunktdarstellung wird in der Regel die Anzahl der randomisierten, nicht die Anzahl der in der Auswertung berücksichtigten Patienten genannt.) Sofern keine endpunktbezogenen Informationen verfügbar sind, kann orientierend die Anzahl der Studienabbrecher pro Arm verglichen werden. Diese Information findet sich in der Regel in der Grafik „Flow of patients“ und/oder im Fließtext zu Beginn des Ergebnisteils („Results“). Insbesondere bei Patient-Reported Outcomes (PRO), z. B. zur Lebensqualität, kann sich die Antwortrate jedoch deutlich von der Anzahl der Studienabbrecher unterscheiden.
Unterscheiden sich die Gründe für die Studienabbrüche zwischen den Armen und könnte dies mit der Studienintervention in Zusammenhang stehen?	Die Gründe für den Studienabbruch werden zumeist in der Grafik „Flow of patients“ aufgeführt. Interessant ist insbesondere, wie viele Patienten die Studie aufgrund von Ereignissen im Zusammenhang mit der Studienintervention abgebrochen haben, z. B. aufgrund unzureichender Wirksamkeit („lack of efficacy“) oder aufgrund unerwünschter Ereignisse („adverse event“). Leider sind die genannten Gründe aber oft nicht hilfreich für die Beurteilung, ob ein Zusammenhang zur Studienintervention besteht („consent withdrawn“, „physician decision“, „other“).

* Diese Fragen orientieren sich am Risk of Bias (RoB) 2 Tool von Cochrane [10]. Das RoB 2 Tool wird in einem späteren Artikel dieser Reihe ausführlich vorgestellt.

³ Diese Begriffe werden praxisnah in dem Artikel „What is the difference between missing completely at random and missing at random?“ [4] anhand eines fiktiven Gesprächs zwischen einem Kliniker und einem Statistiker erläutert.

sen durchgeführt werden. Sensitivitätsanalysen prüfen, ob sich das Ergebnis der primären Analyse (Complete case oder Imputation) verändert, wenn bei den fehlenden Daten z. B. der schlechtestmögliche Fall („worst case scenario“), der bestmögliche Fall („best case scenario“) oder ein als plausibel eingeschätztes Szenario („base case“) angenommen wird.⁴ Abb. 2 illustriert annähernd ein „worst case scenario“: In der Interventionsgruppe hatten alle Personen, die nicht zur Abschlussvisite erschienen sind, ein Handekzem entwickelt, während in der Kontrollgruppe die meisten Personen mit Handekzem an der Untersuchung teilnahmen.

Die Auswahl des Imputationsverfahrens kann das Studienergebnis erheblich beeinflussen [6]. Die CONSORT-Leitlinie [7] gibt deshalb Empfehlungen, wie der Umgang mit fehlenden Daten bei randomisierten Studien berichtet werden soll. Leider finden sich bislang in Studienpublikationen häufig nur äußerst rudimentäre Informationen zum Umgang mit fehlenden Werten (z. B. keine Diskussion der Plausibilität der „missing at random“-Annahme). Oftmals bleibt sogar unklar, ob überhaupt eine Imputation oder Sensitivitätsanalyse angewandt wurde [8, 9]. Hinweise auf ein niedriges Verzerrungspotenzial durch fehlende Daten ergeben sich dann

Artikelserie aus Arzneiverordnung in der Praxis

In der AVP-Serie „Klinische Studien – wo ist der Haken?“ zu Klinischen Studien in der Kategorie „Evidenzbasierte Medizin“ sind folgende Beiträge erschienen:

- AVP 2/2024: Studientypen – Ohne Kontrolle geht nichts (Kurzlink: <https://tinyurl.com/39txa8hs>)
- AVP 4/2024: Randomisierung – Der reine Zufall (Kurzlink: <https://tinyurl.com/bdh7hdjy>)

- AVP 1/2025: Per-protokoll, As-treated oder Intention-to-treat: Was machen wir mit den „Abtrünnigen“? (Kurzlink: <https://t1p.de/15o0b>) HÄBL 12/25
- AVP 2/2025: Verloren, aber nicht unersetzlich? – Vom Umgang mit fehlenden Daten (Kurzlink: <https://t1p.de/h3wyyi>) HÄBL aktuelle Ausgabe

lediglich dadurch, dass der Datenverlust in den Studien armen ähnlich hoch ist und die Gründe für Studienabbrüche zwischen den Armen gleich verteilt sind (siehe Tabelle 1).

Zusammenfassung

Fehlende Daten können das Studienergebnis verzerren, wenn ihr Fehlen durch die Verträglichkeit oder Wirksamkeit der Studienmedikation beeinflusst wurde. Mit Hilfe von Imputationsverfahren können fehlende Daten durch plausibel erscheinende Werte ersetzt werden. Die Aussagesicherheit der Imputationsverfahren sollte durch Sensitivitätsanalysen über-

prüft werden. Leider wird in Publikationen oftmals nur rudimentär geschildert, wie mit fehlenden Daten umgegangen wurde. Hinweise auf ein hohes Verzerrungsrisiko sind ein unterschiedlicher Anteil fehlender Daten sowie unterschiedliche Gründe für den Studienabbruch zwischen Interventions- und Kontrollgruppe.

Dr. med. Natascha Einhart, Berlin
E-Mail: natascha.einhart@baek.de

Prof. Dr. Tim Mathes, Köln

Die Literaturhinweise finden sich online auf der Website www.laekh.de unter der aktuellen Ausgabe.

⁴ Informationen sowohl zu Imputationsverfahren als auch zu Sensitivitätsanalysen finden sich beispielsweise bei [5]. Allgemeinverständlicher führt [2] in die Thematik ein.

Ärztekammer

Ukrainehilfe: Schnelle Hilfe auch im Einzelfall

Wie in der vergangenen Ausgabe des Hessischen Ärzteblattes berichtet, gibt es jenseits der größeren Projekte in unserer Ukrainehilfe auch Gelegenheit für andere Hilfen: Anfang März erreichte uns über Vermittlung unseres Partners Hope for Ukraine e. V. (Non Profit Organisation, NPO) eine dringliche ärztliche Bitte um Unterstützung. Es wurde konkret das Zytostatikum Cytarabin in definierter Konzentration zur Herstellung mehrerer Infusionslösungen für die anstehende Behandlung eines neunjährigen Jungen mit einer akuten lymphatischen Leukämie aus Ternopil in der Westukraine erbeten. Auch diese vermeintlich kleine, individuelle Medikamentenausfuhr musste trotz der

Eile einige bürokratische Hürden überwinden: Zunächst musste das ukrainische Rezept zur eindeutigen Identifikation von Behandler, Patient und Medikation zeitnah fachkundig übersetzt werden. Danach konnte über eine geeignete Apotheke das Cytarabin in der richtigen Konzentration beschafft werden. Mit Einbindung der Rechtsabteilung unserer Kammer (meinen herzlichen Dank an Justitiar Manuel Maier und die Syndikusrechtsanwältin Katja Jörg!) wurde die beabsichtigte Ausfuhr dem zuständigen Gesundheitsamt angezeigt und um Gestattung gebeten. Abschließend musste der rasche Transport des Zytostatikums (der Autor wohnt in Düsseldorf) in die Wege geleitet werden.

Was einfach aufgezählt wurde, kann ein längerer bürokratischer Weg sein; erfreulicherweise jedoch nicht in diesem Fall. Das Übersetzungsbüro, die Apotheke, die Rechtsabteilung, das Gesundheitsamt und die NPO Hope for Ukraine e. V. haben sich alle „in die Kurve gelegt“, um das Kind zu versorgen: Meinen herzlichen Dank dafür!

Dr. med. Alexander Marković
Ehrenamtlicher Beauftragter für die
Zivil-Militärische Zusammenarbeit/
Ukrainehilfe
der Landesärztekammer Hessen