

Interne Validität und Bias: Verzerrte Ergebnisse erkennen

Die interne Validität ist gegeben, wenn „wahre“ Messwerte nicht durch systematische Fehler verfälscht sind, sondern nur zufällig variieren. Denn Statistik kann mit zufälligen Schwankungen umgehen, nicht aber mit systematischen Abweichungen: Sie führen zu verzerrten Ergebnissen.

Roland Müller-Waldeck



Das Bias-Risiko muss von anderen Aspekten der Qualität einer Studie abgegrenzt werden:

- Die Präzision der Ergebnisse steigt mit der Fallzahl;
- die Studienqualität bemisst sich daran, wie gut sich eine Studie an die Prinzipien von Good Clinical Practice (GCP) oder Guter Epidemiologischer Praxis (GEP) hält und
- die Berichtsqualität hängt davon ab, wie vollständig die Informationen sind, die notwendig sind, um die interne Validität beurteilen zu können.

Verzerrung durch Selektion der Studienteilnehmer: Selektions-Bias

Wenn Studienteilnehmer nicht tatsächlich zufällig aus der Menge aller möglichen Teilnehmer (Grundgesamtheit) ausgewählt werden, kann sich die Studienpopulation systematisch von der Grundgesamtheit unterscheiden und die Studienergebnisse können nicht mehr auf die Grundgesamtheit übertragen werden: Eine Form von Selektions-Bias liegt vor.

Prevalence incidence Bias oder Neyman's Bias:

Fall-Kontrollstudien besonders, aber auch Querschnitts-, experimentelle und Kohortenstudien sind anfällig für Prevalence incidence Bias oder Neyman's Bias: Werden in der Zeit zwischen Exposition und Datenerhebung verstorbene Teilnehmer nicht berücksichtigt (die Gruppe der Verstorbenen wird aus der Studienpopulation ausgeschlossen), erscheint das Risiko für die Erkrankung geringer als das wahre Risiko. Werden in diesem Zeitraum geheilte Teilnehmer nicht berücksichtigt, erscheint das Risiko größer, als es tatsächlich ist.

Je mehr Zeit zwischen Exposition und Datenerhebung vergeht, desto größer ist das Risiko für diesen Bias und er ist bei lang dauernden Krankheiten wahrscheinlicher als bei kurz dauernden. Ein Beispiel: 20 Probanden werden über einen Zeitraum t_0 bis t_2 beobachtet. In diesem Zeitraum entwickeln 4 Teilnehmer eine leichte und 4 Teilnehmer eine schwere Form der Erkrankung, das Verhältnis von schwerer zu leichter Erkrankung ist 1. Für eine Querschnittsstudie werden

die Verhältnisse zum Zeitpunkt t_1 untersucht: Weil schwer erkrankte Teilnehmer kürzer überleben als leicht erkrankte, werden zum Zeitpunkt t_1 zwei leicht erkrankte und ein schwer erkrankter Teilnehmer gezählt, das Verhältnis von schwerer zu leichter Erkrankung ist 0,5.

Loss to follow-up Bias:

Er entsteht, wenn in einer klinischen Studie eine Gruppe von Teilnehmern bevorzugt für den Follow-up verloren geht. In diesem Fall verändert sich das Studienergebnis, weil sich die Studienpopulation verändert: Gesunde oder kurierte Teilnehmer scheiden z.B. eher aus einer Studie aus als Erkrankte.

Non-response Bias:

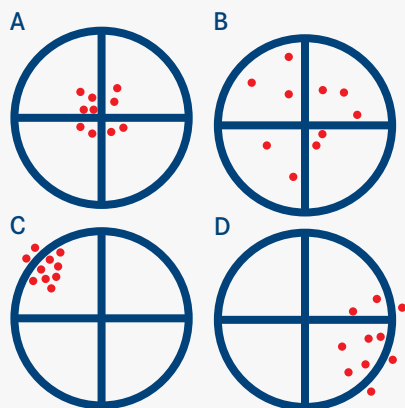
Er entsteht, wenn potenzielle Probanden mit höherem Risiko die Teilnahme überdurchschnittlich häufig ablehnen.

Verzerrung bei der Datenerhebung: Informations-Bias

Informations-Bias entsteht bei der Datenerhebung, z.B. weil fehlerhafte Messmethoden oder ungeeignete Messgeräte zu unpräzisen Ergebnissen führen.

Recall Bias:

Erkrankte Studienteilnehmer erinnern sich häufig besser an Dinge, die mit der Erkrankung zusammenhängen als gesunde Teilnehmer der Kontrollgruppe. Das kann vermieden werden, indem beispielsweise die Kontrollgruppe aus



Präzision und Bias: Der wahre Wert liegt im Zentrum der Koordinatenkreuze

- A: Die Ergebnisse streuen gering und zufällig um den wahren Wert: **hohe Präzision, kein Bias.**
- B: Die Werte streuen stark und zufällig um den wahren Wert: **geringe Präzision, kein Bias.**
- C: Die Werte streuen wenig und weichen systematisch vom wahren Wert ab: **hohe Präzision, Bias.**
- D: Die Werte streuen stark und weichen systematisch vom wahren Wert ab: **geringe Präzision, Bias.**

Erkrankten gebildet wird und indem Angaben an Hand von Krankenakten überprüft werden.

Interviewer Bias:

Interviewer bevorzugen Antworten, die ihre vorgefassten Meinungen bestätigen. So können „Erklärungen“ den Teilnehmer zu bestimmten Antworten verleiten. Um das zu vermeiden, müssen Fragen und Erklärungen sorgfältig standardisiert sein und der Interviewer sollte nicht wissen, ob der jeweils Interviewte der Kontroll- oder der Verumgruppe angehört.

Observer Bias:

Das Wissen um den Expositions-Status durch den Outcome-Gutachter kann das Ergebnis beeinflussen. Daher sollte der Outcome-Gutachter gegen den Expositions-Status verblindet sein. Mehrere Outcome-Gutachter können eingesetzt werden und wenn möglich, sollte das Outcome in „möglich“, „wahrscheinlich“ und „definitiv“ klassifiziert werden; Ein Observer Bias ist wahrscheinlich, wenn

nur Ergebnisse in der Kategorie „möglich“ auftreten.

Regression dilution Bias:

Messwerte, die wegen zufälliger Messfehler stark vom Mittelwert abweichen, nähern sich bei einer zweiten Messung häufig dem Mittelwert an. Solche zufälligen Abweichungen können eine Korrelation schwächer erscheinen lassen, als sie tatsächlich ist (Regression dilution Bias): Geht man von einer wahren 1:1 Korrelation zwischen der unabhängigen Variablen X und der abhängigen Variablen Y aus, entsteht durch wahre Messpunkte eine Gerade mit der Steigung 1. Im Falle von Messfehlern der X-Werte verschieben sich die Messwerte parallel zur X-Achse und die Kurve wird flacher (die Steigung wird geringer). Sind die Y-Werte ungenau, bleibt die Steigung der Geraden gleich.

Das Risiko solcher Verzerrungen besteht, wenn in longitudinalen Studien, die über einen längeren Zeitraum Daten erheben, eine Assoziation zwischen einer Baseline-gemessenen kontinu-

ierlichen und unabhängigen Variablen (Exposition mit einem Risikofaktor) mit dem Risiko eines Endpunktes (Erkrankung) untersucht wird. In einem solchen Fall kann die Präzision der X-Werte mit einem Durchschnittswert aus mehreren Messungen pro Patient gesteigert werden, Messgeräte und Messmethoden sollten möglichst präzise sein und das Studiendesign sollte angepasst sein.

Lead time Bias:

Die Vorlaufzeit-Verfälschung spielt eine Rolle, wenn beurteilt werden soll, ob Teilnehmer eines Screenings länger überleben als Nicht-Teilnehmer. In der Screening-Gruppe kann die Überlebenszeit verlängert scheinen, weil die Diagnose früher stattfindet als in der Vergleichsgruppe.

Um das zu vermeiden, kann das Todesfallrisiko bei allen gescreenten und nichtgescreenten Teilnehmern anstelle der kumulativen Überlebenswahrscheinlichkeit nach der Diagnose berechnet werden. ■

Literatur: Tripepi G et al. Bias in clinical research. *Kidney international* (2018)73,148-153. Tripepi G et al. Bias and information Bias in clinical research. *Nephron Clin Pract* 2010;115:c94-c99. Hutcheon J A et al. Random measurement error and regression dilution bias. *BMJ* 2010;340:c2289. Cochrane Deutschland, Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften, Institut für Medizinisches Wissensmanagement. „Bewertung des Biasrisikos (Risiko systematischer Fehler) in klinischen Studien: ein Manual für die Leitlinienerstellung“ 1. Auflage 2016.