

## Externe Validität

# „Für welche Patienten gelten die Ergebnisse dieser Studie?“

Die externe Validität einer Studie zeigt auf, ob sich deren Ergebnisse auf andere Zielpopulationen übertragen lassen. Welche Kriterien dabei erfüllt sein müssen.

*Roland Müller-Waldeck*

■ Klinische Studien untersuchen an einer Studienpopulation, wie Therapien wirken. Die interne Validität setzt voraus, dass die Ergebnisse nicht durch systematische Fehler verfälscht sind und nur zufällig variieren.

Die interne Validität ist Voraussetzung für die externe Validität: Externe Validität untersucht, ob die Studienergebnisse auf eine andere Gruppe als die Studienpopulation (auf eine Zielpopulation) anwendbar sind.

Einige Autoren unterscheiden externe Validität (Patienten in genau der Behandlungssituation wie in der Studie) und Anwendbarkeit (Patienten in nicht genau der Behandlungssituation wie in der Studie). Wenn beispielsweise ein chirurgischer Eingriff in einem darauf spezialisierten Krankenhaus sehr gute Ergebnisse erzielt (Studienergebnis), wird das auch für Patienten gelten, die sich in diesem Krankenhaus unter den gleichen Bedingungen dem Eingriff unterziehen (externe Validität). In einem nicht darauf spezialisierten Krankenhaus kann das Ergebnis ganz anders sein (Anwendbarkeit).

Allgemeine Aussagen über die externe Validität lassen sich nicht treffen, denn die Zielpopulation von Ärzten kann sich sehr stark in ihren Eigenschaften unterscheiden. Selbst wenn die Zielpopulation exakt nach den (häufig nur unzu-

reichend) beschriebenen Einschlusskriterien ausgewählt würde, so würde sie sich doch unterscheiden in geografischen, ethnischen und zeitlichen Merkmalen.

Dekkers et al. schlagen vor, die externe Validität anhand von 3 Fragekomplexen zu beurteilen:

## 1. Weicht die Studienpopulation in ihren Eigenschaften von der durch die Einschlusskriterien definierten Population ab, die das Studienergebnis ändern können?

Studienteilnehmer und mit ihnen auch Eigenschaften der Erkrankung, Behandlung und Diagnose können auch auf andere Weise als durch die Einschlusskriterien selektiert werden. Das ist umso wahrscheinlicher, je weniger geeignete Patienten tatsächlich eingeschlossen werden.

### Selektion vor Studieneinschluss:

- Die Patientengruppe, die ein Studienzentrum (z. B. Krankenhaus) oder einen Studienarzt aufsucht, kann sich in wesentlichen Eigenschaften von der Patientengruppe in der Primärversorgung unterscheiden.
- Patienten, die teilnehmen wollen, können sich von den Patienten unterscheiden, die das nicht tun (z. B. Schwere der Erkrankung, Stadium).
- Bestimmte Patientengruppen (Frauen,

ältere Patienten, komorbide Patienten) können ausgeschlossen sein.

### Andere Selektionsmechanismen für Patienten:

- Häufig bevorzugen Studienärzte jüngere Patienten, Patienten mit weniger Komorbiditäten und geringerer Mortalität. Das Verhältnis von Patienten, welche die Einschlusskriterien erfüllen, zu Patienten, die tatsächlich rekrutiert wurden, kann ein Hinweis darauf sein.
- In einer Run-In-Periode können Patienten mit schlechter Compliance oder Nebenwirkungen sowie Patienten, bei denen die Intervention nicht wirkt, ausgeschlossen werden.

### Anreicherungsstrategien

### Selektion durch Auswahl der Studienzentren:

- Patienten in stark spezialisierten Kliniken werden häufig anders behandelt und diagnostiziert als Patienten der Primärversorgung.
- Teilnehmende Ärzte können nach ihrer Erfolgsbilanz ausgewählt werden.

## 2. Beeinflussen die zeitlichen, ethnischen und geografischen Unterschiede zwischen Studienpopulation und Zielpopulation die Übertragbarkeit der Ergebnisse?

Studienergebnisse sind nicht immer von Land zu Land und auch nicht unabhängig von der Zeit übertragbar.

### Geografische Einflussfaktoren:

- Patienten in unterschiedlichen Gesundheitssystemen können sich in wichtigen Merkmalen (z. B. Diagnostik, Behandlung) unterscheiden.
- Krankheitserreger können sich geografisch unterscheiden.
- Behandlungen außerhalb der Studie können in verschiedenen Ländern differieren.

### Ethnische Einflussfaktoren:

- Ethnische Unterschiede können bei Krankheitsanfälligkeit, Krankheitsverlauf, Ernährung, Alkoholkonsum etc. bestehen.

### Zeitliche Einflussfaktoren:

- Wenn sich nach der Studie ein Einflussfaktor (Diagnose, Schweregradeinteilung, Vorbehandlung, Resistenzen) geändert hat, ist das Studienergebnis nicht in die Gegenwart übertragbar.

### 3. Lassen sich die Studienergebnisse auf eine Zielpopulation übertragen, die nicht in allen Punkten den Einschlusskriterien der Studie entspricht?

Der Arzt muss sich ein Urteil darüber bilden, ob die Ergebnisse einer Studie auf seine Patienten übertragbar sind. Dabei können bereits vorhandene Erkenntnisse oder das Prinzip der biologischen Plausibilität helfen.

- Das Outcome von Patienten mit gleichen oder scheinbar irrelevant verschiedenen Ausgangscharakteristika kann sich unterscheiden. Wichtige Faktoren sind z. B. die zugrundeliegende Pathologie, Krankheitsstadium, Begleiterkrankungen und das absolute Risiko für ein ungünstiges Outcome ohne Therapie.

- Um die Anwendbarkeit zu beurteilen, muss das Behandlungssetting in der Studie mit dem Routine-Behandlungssetting verglichen werden. Hier spielen Häufigkeit und Parameter von Kontrolluntersuchungen oder diagnostische Gegebenheiten eine Rolle.

**Kommentar:** Die Fragekomplexe von Dekkers et al. lassen wichtige Themen unbeachtet. Einige davon sind:

- Eine Studie kann nur dann extern valide sein, wenn die Endpunkte klinisch relevant sind. Surrogatendpunkte sind häufig fragwürdig oder irreführend. Die größte Aussagekraft haben einfache klinische Endpunkte, die auch die Meinung der Patienten wiedergeben. Bei kombinierten Endpunkten kann der für den Patienten wichtigste Endpunkt durch die Behandlung anders beeinflusst werden als der kombinierte Endpunkt. Wenn in kombinierten Endpunkten Ereignisse unterschiedlicher Schweregrade zusammenfasst sind, können die Effekte für den geringsten Schweregrad am stärksten sein, weil er am häufigsten vorkommt.
- Häufig sind Behandlungsdauer und Nachbeobachtung zu kurz für Langzeiteffekte: Initiales Ansprechen ist kein Indikator für Langzeitnutzen.
- Nebenwirkungen werden häufig nur unzureichend dargestellt. Die Rate von Studienabbrüchen kann ein Hinweis auf Nebenwirkungen sein. Grundsätzlich gilt, dass durch Patientenselektion, Begrenzung auf Spezialzentren und intensive Patientenüberwachung die Risiken in der Studie oft geringer sind als im klinischen Alltag. ■

---

**Literatur:** Dekkers OM et al. How to assess the external validity of therapeutic trials: a conceptual approach. *International Journal of epidemiology* 2010;39:89–94. Rothwell PM. To whom do the results of this trial apply? *Lancet* 2005; 365:82–93; deutsche Fassung: Externe Validität randomisierter kontrollierter Studien: „Auf wen sind die Ergebnisse dieser Studie anwendbar?“ *Z.ärztl.Fortbild.Qual.Gesundh.wes.* 2006; 100:3003–317. Rothwell PM. Factors that can affect the external validity of randomized controlled trials. *PLOS Clin. Trials* 1(1): e9.