

AMB 2017, 51, 61

Geographische Unterschiede bei Studienergebnissen im Zeitalter der globalisierten klinischen Forschung

Vor drei Jahren haben wir über die TOPCAT-Studie (1, 2) berichtet, die Spironolacton bei diastolischer Herzinsuffizienz untersuchte und (wie viele andere Arzneimittel-Studien zu dieser Erkrankung) keinen signifikanten Vorteil fand. Die multinationale Studie aus 200 Zentren in Nord- und Südamerika (51% aller Patienten) sowie in Russland und Georgien (49%) wurde vom National Heart, Lung, and Blood Institute (NHLBI) der US-amerikanischen National Institutes of Health (NIH) finanziert. In Subgruppenanalysen fiel auf, dass Patienten in Georgien und Russland deutlich niedrigere Gesamt-Ereignisraten (kardiovaskulärer Tod, Hospitalisierung wegen Herzinsuffizienz und Plötzlicher Herztod) und ein anderes Ansprechverhalten auf Spironolacton hatten als Patienten anderer Regionen. Als Ursachen für die niedrigeren Gesamt-Ereignisraten in Georgien und Russland vermutete man Unterschiede in den Ausgangscharakteristika der Patienten und/oder in einer zu großzügigen Diagnosestellung. Zudem erweckte ein Vergleich der Serum-Kaliumspiegel den Verdacht auf Unregelmäßigkeiten in der Einnahme der Studienmedikation. Dieser Verdacht wurde nun durch eine Post-hoc-Analyse eingefrorener Serumproben bestätigt (3). In Proben von 366 Patienten (206 aus USA und Kanada, 160 aus Russland) wurde der Spironolacton-Metabolit Canrenon untersucht. Bei den russischen Patienten, die bei der 12-Monats-Visite angegeben hatten, Spironolacton korrekt wie vorgegeben einzunehmen, fanden sich signifikant häufiger nicht-nachweisbare Canrenon-Spiegel als bei den nordamerikanischen (30% vs. 3%; $p < 0,0001$). Somit ist von einer fehlerhaften Durchführung der Studie in russischen und vermutlich auch in georgischen Zentren auszugehen. US-amerikanische Leitlinien haben deshalb – in einer umstrittenen Entscheidung und trotz der insgesamt negativen TOPCAT-Ergebnisse – in einem aktuellen Update Spironolacton als Klasse-IIb-Empfehlung in die Leitlinien zur Behandlung der diastolischen Herzinsuffizienz aufgenommen (4), für die ja generell nur wenige medikamentöse Therapieoptionen bestehen. Voraussetzung ist allerdings, dass keine Niereninsuffizienz \geq Stadium III und keine Hyperkaliämie besteht.

Die TOPCAT-Studie ist nicht die einzige Studie mit geographisch unterschiedlichen Ergebnissen: Deutliche Differenzen in der Wirkung von Ticagrelor (Indikation: Akutes

Koronarsyndrom) zwischen europäischen (Effekt signifikant) und amerikanischen Patienten (Effekt nicht signifikant) in der PLATO-Studie waren mit großer Wahrscheinlichkeit durch die in Nordamerika und Europa unterschiedlich üblichen ASS-Dosierungen bedingt (5). Mehr amerikanische Patienten hatten Dosierungen von ≥ 300 mg ASS/d. Umgekehrt war die Situation beim PCSK9-Inhibitor Evolocumab (Indikation: Koronare Herzkrankheit, LDL-Cholesterin ≥ 70 mg/dl und Statin-Therapie): Nordamerikanische und asiatische Patienten zeigten in der Subgruppenanalyse der FOURIER-Studie mit Abstand den größten Nutzen, während dieser bei europäischen und lateinamerikanischen Patienten nur grenzwertig bzw. nicht signifikant war (6, 7). Die Gründe dafür sind bisher unklar.

Immer mehr randomisierte kontrollierte Studien (RCT) werden nicht nur multizentrisch, sondern auch multinational durchgeführt. Dies hat verschiedene Gründe: Um relativ kleine Vorteile präventiver oder neuer therapeutischer Maßnahmen statistisch signifikant nachweisen zu können, müssen klinische Studien tausende, manchmal zehntausende Patienten einschließen, die dann oft in mehreren hundert Zentren und in mehreren verschiedenen Ländern rekrutiert werden. Manchmal müssen Studien auch infolge der Seltenheit der untersuchten Erkrankung multinational angelegt werden – dann oft mit nur wenigen Patienten pro Land. Bei Zulassungsstudien können regulatorische Bestimmungen Anlass für eine multinationale Ausrichtung sein, wenn der pharmazeutische Unternehmer anstrebt, für ein Arzneimittel in möglichst vielen Ländern eine Marktzulassung zu beantragen. Öffentlich gesponserte Studien (wie z.B. TOPCAT) versuchen, durch die Mitbeteiligung anderer Länder die Kosten auf mehrere Gesundheitsbehörden zu verteilen.

Eine aktuelle Review im Rahmen der Artikelserie „The Changing Face of Clinical Trials“ des N. Engl. J. Med. widmete sich jetzt dem immer öfter beobachteten Phänomen unerwarteter geographischer Unterschiede klinischer Studienergebnisse (8) – definiert als Unterschiede zwischen Ländern/Regionen, nicht zwischen einzelnen Studienzentren. Dies wird anhand mehrerer Beispiele illustriert. Grundsätzlich gilt es in solchen Fällen – wie bei anderen heterogenen Subgruppenanalysen auch – primär zu unterscheiden, ob dieser Effekt „zufällig“ oder „wahr“ ist. Die Review-Autoren gehen davon aus, dass ein Großteil der

geographischen Unterschiede in klinischen Studien primär auf Zufallseffekten beruht, z.B. infolge kleiner Patientenzahlen. Sind Zufallseffekte mittels statistischer Tests auf Homogenität und Interaktion sowie Plausibilitätsprüfungen weitgehend ausgeschlossen, muss von einem tatsächlich vorhandenen Unterschied ausgegangen und dann nach möglichen Ursachen gesucht werden. Es existiert eine ganze Reihe geographisch ungleich verteilter potenzieller Störfaktoren (Confounder), die klinisch relevant sind, wie z.B.:

- medikamentöse Begleittherapien,
- Qualität der (Standard-)Gesundheitsversorgung,
- Lebensstil, Ernährung und Ernährungszustand, Hygiene, Gesundheitszustand,
- bei chirurgischen/interventionellen Studien: Expertise der Operateure und Qualität der perioperativen Behandlung,
- genetische Eigenschaften der Patientenpopulation,
- bei Antiinfektiva genetische Eigenschaften der Erreger.

Außerdem gibt es Faktoren, die zwar klinisch nicht relevant sind, aber durch Beeinflussung der Durchführung der Studien zu Artefakten in den Ergebnissen führen können, z.B.:

- qualitativ unterschiedliche Angebote ambulanter und stationärer Versorgung mit Auswirkung auf Studienendpunkte, z.B. Hospitalisierungsraten,
- sozioökonomische und kulturelle Faktoren mit Auswirkungen, z.B. auf Therapieadhärenz, Wahrnehmung und Meldung von Nebenwirkungen,
- Ausbildung des Studienpersonals.

Die Wahrscheinlichkeit „unerwarteter“ regionaler Unterschiede in Studienergebnissen durch solche Störfaktoren sollte eigentlich durch eine sorgfältige Auswahl der teilnehmenden Regionen und Länder, ein überlegtes Studiendesign und eine exakte Studienplanung minimiert sein. Erkennbare und unvermeidbare potenzielle Störfaktoren sollten prospektiv erfasst werden, um die Ergebnisse anschließend beispielsweise mittels multivariater Analysen oder Matching überprüfen zu können. Dies wird allerdings umso schwieriger, je größer und je weiter geographisch verteilt (und damit sozioökonomisch und kulturell inhomogener) eine Studienpopulation und das für die Durchführung und das Monitoring der Studie verantwortliche Personal ist.

Fazit: Unerwartete geographische Unterschiede in den Ergebnissen klinischer Studien sind wahrscheinlich meist Zufallsprodukte. Sie können aber auch reale und klinisch relevante regionale Unterschiede in den positiven und/oder negativen Effekten einer präventiven oder therapeutischen Maßnahme widerspiegeln. Falls diese nicht durch bekannte und miterfasste Störfaktoren (Confounder) erklärt werden können, muss nach unbekanntem Faktoren gesucht werden. Neben mannigfachen potenziellen Ursachen für „wahre“ Effekte auf die Studienergebnisse müssen auch Artefakte durch Fehler in Studiendesign und -durchführung in Betracht gezogen werden. Diese Phänomene werden wohl zunehmen, da künftig aus mehreren Gründen die Zahl multinationaler Studien sicher weiter steigen wird.

Literatur

1. Pitt, B., et al. (TOPCAT = **T**reatment **O**f **P**reserved **C**ardiac function heart failure with an **A**ldosterone an**T**agonist): N. Engl. J. Med. 2014, **370**, 1383.
2. AMB 2014, **48**, 52.
3. de Denus, S., et al. (TOPCAT = **T**reatment **O**f **P**reserved **C**ardiac function heart failure with an **A**ldosterone antagonist): N. Engl. J. Med. 2017, **376**, 1690.
4. Yancy, C.W., et al.: J. Am. Coll. Cardiol. 2017. doi: 10.1016/j.jacc.2017.04.025.
5. Mahaffey, K.W., et al. (PLATO = **P**LATelet inhibition and patient **O**utcomes): Circulation 2011, **124**, 544.
6. Sabatine, M.S., et al. (FOURIER = **F**urther cardiovascular **O**utcomes **R**esearch with PCSK9 Inhibition in subjects with **E**levated **R**isk): N. Engl. J. Med. 2017, **376**, 1713.
7. AMB 2017, **51**, 33.
8. Yusuf, S., und Wittes, J.: N. Engl. J. Med. 2016, **375**, 2263.