

Sönnichsen A. Die Crux mit der randomisierten kontrollierten Studie. Editorial ZFA 4/2018. Z Allg Med 2018; 94: 145–146

Leserbrief von Dr. Felix S. Wicke

Die Crux mit der Signifikanz

Die Übertragung von Studienergebnissen auf individuelle Patient*Innen ist eine der großen Herausforderungen der „evidence“-basierten praktischen Allgemeinmedizin. Ich möchte daher einige Anmerkungen zu Andreas Sönnichsens kritischer Interpretation der EMPAREG-Studie machen. Die Anmerkungen sollen nicht als Widerspruch gelesen werden, sondern als hoffentlich sinnvolle Ergänzungen.

Im Editorial der ZFA 4/2018 stellt er die Frage, wer von einer Therapie mit Empagliflozin profitieren kann und analysiert dafür die Ergebnisse der Subgruppenanalysen. Bei der geschlechtsspezifischen Analyse zeigte sich: „die kardiovaskuläre Mortalität wurde nur für Männer signifikant gesenkt. Sollten also nur Männer mit manifester kardiovaskulärer Erkrankung mit Empagliflozin behandelt werden?“ Wer hier „statistisch nicht-signifikant“ mit „klinisch unwirksam“ gleichsetzt, würde die Frage mit „ja“ beantworten, bzw. schlussfolgern, dass das Medikament bei Frauen keinen Vorteil bringt. Doch hier wäre es besser p-Werte und den Begriff „signifikant“ zu vergessen. Wichtiger ist der Blick auf den Effektschätzer (hier die Hazard Ratio) und sein Konfidenzintervall. Die Hazard Ratios zeigen sowohl für Männer als auch für Frauen einen Vorteil durch Empagliflozin, das Konfidenzintervall bei Frauen ist aber deutlich

breiter. Die richtige Interpretation (unter der Annahme, dass kein systematisches Bias vorliegt) lautet: Ein Vorteil von Empagliflozin zeigt sich für Frauen und Männer, aber bei Frauen ist diese Aussage aufgrund des großen Konfidenzintervalls mit höherer Unsicherheit behaftet.

Anders sieht es bei der Subgruppenanalyse nach HbA_{1c} aus. Hier zeigt sich bezüglich des primären Endpunktes sogar ein höheres Risiko durch Empagliflozin bei einem Ausgangs-HbA_{1c} \geq 8,5 %. Dieses Ergebnis *könnte* rein zufällig bedingt sein, ein Blick auf die nur knapp überlappenden Konfidenzintervalle hinterlässt aber eher ein Fragezeichen bezüglich dieser Subgruppenanalyse.

Ein weiterer Punkt betrifft das optimale Übertragen der Subgruppenresultate auf einzelne Patienten. Es erscheint zwar wahrscheinlich, dass ein Patient, dessen Eigenschaften in die positiv ausgefallenen Subgruppenanalysen fallen, eher profitiert, dies ist aber streng genommen aus der Publikation nicht ableitbar. Andersherum: Obwohl die Subgruppe „Frauen“ und die Subgruppe „HbA_{1c} \geq 8,5 %“ jeweils „nicht-signifikante“ Subgruppenergebnisse hatten, ist es denkbar, dass ausgerechnet die Subgruppe „Frauen plus HbA_{1c} 8,5 %“ ein vorteilhaftes Ergebnis haben. Diese (Sub-)Subgruppenanalyse wurde schlicht nicht durchgeführt. Subgruppenanalysen sollten für vorteilhafte und nachteilige Ergebnisse nicht überinterpretiert werden. Insbesondere heterogene Ergebnisse sollten aber hinterfragt werden (z.B. warum waren die Ergebnisse für

HbA_{1c} \geq 8,5 % unvorteilhaft? Warum für Patient*Innen in Europa und Nordamerika?).

Einer Aussage möchte ich doch noch widersprechen und zwar, dass bei einer NNT von 31 in Konsequenz 30 Patient*Innen unnötig behandelt würden. Diese Schlussfolgerung ist aus zwei Gründen nicht korrekt. Erstens profitieren *alle* von der Risikoreduktion, schließlich kann der Endpunkt nur verhindert werden, wenn das Medikament eingenommen wird (man kann auch nur im Lotto gewinnen, wenn man ein Los kauft, und bei einer Chance von 1:30 ist das zumindest überlegenswert). Zweitens ist davon auszugehen, dass eine pharmakologische Wirkung (z.B. HbA_{1c}-Senkung, Gewichtsreduktion oder Nebenwirkungen) bei *allen* auftritt, dies wird durch die NNT für einen *dichotomen Endpunkt* nicht wiedergespiegelt.

Es bleibt eine grundsätzliche Einschränkung, dass eine sichere Vorhersage von Nutzen oder Schaden einer Therapie für den einzelnen Menschen nicht möglich ist, daran ändert auch Big Data nichts. Wir können bestenfalls aus (möglichst qualitativ hochwertigen) Studien gewonnene Wahrscheinlichkeiten anführen.

Korrespondenzadresse

Dr. med. Felix S. Wicke
Institut für Allgemeinmedizin
Goethe-Universität Frankfurt am Main
Theodor-Stern-Kai 7
60590 Frankfurt am Main
wicke@allgemeinmedizin.uni-frankfurt.de

Antwort von Prof. Dr. Andreas Sönnichsen

Mein Editorial war bewusst etwas provokant formuliert, um die Unsicherheit von Studienergebnissen generell hervorzuheben. Diese Unsicherheit ist auch durch einen Blickwechsel von „signifikant“ auf „klinisch wirksam“ bzw. von „nicht-signifikant“ auf „klinisch unwirksam“ nicht zu eliminieren. Signifikanz heißt ja eben nichts anderes, als dass die Unsicherheit der

getroffenen Aussage ein bestimmtes Maß nicht überschreitet. So können selbst signifikante Ergebnisse eben zufällig zustande gekommen sein und den Leser/Anwender in die Irre führen. Das Risiko, zufällig signifikante Ergebnisse zu erhalten, steigt mit der Anzahl der durchgeführten Subgruppenanalysen, was die Aussagekraft solcher Analysen weiter relativiert. Das Editorial sollte daher nicht als Plädoyer für Subgruppenanalysen missverstanden werden.

Bezüglich der NNT teile ich die Meinung des Leserbriefautors nicht: Im Hinblick auf den untersuchten Endpunkt werden diejenigen, bei denen dieser Endpunkt trotz Behandlung eintritt, tatsächlich „umsonst“ behandelt, ebenso wie diejenigen, die den Endpunkt ohnehin nicht erlitten hätten. Wenn man ein Los kauft und dieses Los nicht gewinnt, hat man das Geld auch umsonst ausgegeben. Hier ist zwischen dem „umsonst“ für das Individuum und dem „umsonst“ für die Gruppe zu unter-

scheiden. Das Argument, dass die NNT für einen dichotomen Endpunkt eine HbA_{1c}-Senkung nicht widerspiegelt, ist natürlich korrekt, aber irrelevant für das Individuum, das eben durch die HbA_{1c}-Senkung keinen klinischen Nutzen erzielt.

Ob es durch Big Data, Pharmakogenomics oder noch zu entwickelnde Verfahren einmal möglich sein wird,

besser vorherzusagen, wer von einer bestimmten Therapie profitiert und wer nicht, bleibt abzuwarten. Ich bin hier optimistischer als der Leserbriefautor. Im Moment ballern wir Medikamente im Schrotschussverfahren auf unsere Patienten, und nur wenige profitieren davon. Eine Verbesserung unserer Treffsicherheit wäre jedenfalls sehr wünschenswert.

Korrespondenzadresse

Prof. Dr. med. Andreas Sönnichsen
Institut für Allgemeinmedizin und
Familienmedizin
Universität Witten/Herdecke
Alfred-Herrhausen-Straße 50
58448 Witten
andreas.soennichsen@uni-wh.de