Evidenzbasierte Medizin: Von der nachrangigen Bedeutung des p-Wertes

Der diesjährige Jahreskongress des Deutschen Netzwerkes Evidenzbasierte Medizin (DNEbM e. V.), der im März in Hamburg stattfand, widmete sich dem Hauptthema „Klasse statt Masse – wider die wertlose Wissenschaft“. Es war meine zweite Teilnahme an einem Kongress des Netzwerkes und mir fielen zwei Dinge besonders auf: erstens die hohe Kompetenz, mit der die Teilnehmer um das bestmögliche Wissen einer evidenzbasierten Medizin ringen, und zweitens die relativ geringe Anzahl an teilnehmenden Ärzten, die in der Krankenversorgung tätig sind. Ich hatte den Eindruck, dass evidenzbasierte Medizin und deren Errungenschaften auf der einen Seite in einem Elfenbeinturm von Methodikern zelebriert und auf höchstem Niveau weiterentwickelt wird. Auf der anderen Seite steht dagegen die praktizierende Medizin, die davon nichts mitbekommt und sich allzu oft noch in „eminenzbasierter“ oder gar „marketingbasierter“ Medizin übt.

Ich möchte nur einen Aspekt herausgreifen, der schon lange bei den Teilnehmern der EbM-Kongresse angekommen ist, nicht aber bei den meisten praktizierenden Ärzten und Wissenschaftlern: Die nachrangige Bedeutung des p-Wertes. Wir alle haben im Studium die Bedeutung des p-Wertes gelernt, aber vergessen, dass er nur eine Aussage über die Wahrscheinlichkeit eines Unterschiedes zwischen zwei Interventionen anhand des Signifikanzniveaus (in der Regel $p < 0,05$) macht, aber nichts über die Größe und Relevanz des Effekts aussagt. In einer kürzlich publizierten Studie wurde gezeigt, dass 96 % der Abstracts und Publikationen mindestens ein „statistisch signifikantes“ Resultat zeigen [1]. Der signifikante Effekt ist zur Währung des wissenschaftlichen Fortschritts geworden – aber leider wird hier mit Falschgeld gespielt: Nur ein Bruchteil dieser Effekte ist reproduzierbar und viele „signifikante“ Ergebnisse sind therapeutisch völlig irrelevant!

Um eine Aussage über die Größe und damit die Relevanz eines Effektunterschiedes zwischen zwei Interventionen machen zu können, eignet sich die Berechnung der Effektstärke, also der standardisierten Differenz der Mittelwerte nach Cohen. Mit der Berechnung von Effektstärken lassen sich Therapieeffekte in kleine (Effektstärke = 0,2), mittelgroße (0,5), große (0,8) und sehr große (1,2) einteilen und damit auch vergleichend bewerten. In anderen Worten: Während etwa alle auf dem Markt befindlichen modernen Psychopharmaka „signi-

Prof. Dr. med. Klaus Lieb

Direktor der Klinik für Psychiatrie und Psychotherapie, Universitätsmedizin Mainz
Untere Zahlbacher Str. 8, 55131 Mainz
E-Mail: klaus.lieb@unimedizin-mainz.de

fikant besser“ wirken als Placebo, so zeigen sie doch teils erhebliche Unterschiede in den Effektstärken. Das Antidepressivum Agomelatin etwa hat eine kleine Effektstärke von 0,2, während das Antipsychotikum Clozapin mit 0,9 eine große aufweist.

Jeder Arzt kann die Größe der Effektstärkenunterschiede direkt aus seiner therapeutischen Arbeit bestätigen! Nicht überraschend stimmen die Effektstärken also viel stärker mit der Erfahrung praktizierender Ärzte überein als der p-Wert. Und wer die Effektstärken der Medikamente kennt, ist auch immun gegen Beeinflussungsversuche von Pharmavertretern oder Professoren, wenn sie in Broschüren oder Vorträgen von „hochsignifikanten“ Effekten ihrer Medikamente, Medizinprodukte oder selbst entwickelten Psychotherapien sprechen. Fragen Sie doch bei der nächsten Fortbildungsveranstaltung einfach mal nach den Effektstärken!

Der nächste EbM-Kongress findet übrigens vom 8. bis 10. März 2018 an der Medizinischen Universität Graz statt und widmet sich dem Thema „Brücken bauen – von der Evidenz zum Patientenwohl“. Ein passender Titel, der ausdrückt, dass es allerhöchste Zeit wird, dass EbM wirklich in der Krankenversorgung ankommt und dass bei allem das Patientenwohl im Zentrum stehen muss. Fahren Sie mal hin!

Klaus Lieb, Mainz

Referenz:

1. Chavalarias D, Wallach JD, Li AH, Ioannidis JP. Evolution of reporting p-values in the biomedical literature, 1990-2015. JAMA 2016; 315 (11): 1141 – 8