

Skepsis und wachsendes Unbehagen – eine Methodenkritik an Großstudien

Scepticism and Growing Unease – Thinking About the Big Trials

Philipp Conradi

Ich bin niedergelassener Allgemeinarzt und lese von Zeit zu Zeit einige der größeren Medikamentenstudien. Ich habe weder wissenschaftliche Ambitionen noch habe ich berufspolitische Absichten und beschäftige mich mit diesen Studien lediglich, um die Medikamente genauer zu kennen, die ich tagtäglich verschreibe. Meine Facharztausbildung absolvierte ich in Großbritannien. Dort machte ich meine ersten grundlegend positiven Erfahrungen mit der Evidenzbasierten Medizin (EbM). Ich arbeitete bis 2003 als Hausarzt in Birmingham/UK, bevor ich aus familiären Gründen wieder nach Deutschland zog.

In diese Zeit fällt die Veröffentlichung der Heart Protection Study (HPS) [2–4]. Die HPS ist die bislang größte Studie mit einem cholesterinsenkenden Medikament und beschreibt einen großen Nutzen von Simvastatin für eine Vielzahl von verschiedenen Patientengruppen.

Die damalige Euphorie hatte mich verwundert. Ich begann mit dem intensiven Lesen dieser und anderer Studien. Meine ersten Irritationen waren zaghaft und betrachteten vor allem die alltagsferne Patientenrekrutierung. Doch je genauer ich mich in diese Studien einzulesen begann, umso stärker wuchsen meine Skepsis und mein Unbehagen.

Ermuntert durch die aktuelle Diskussion um „EbM: a movement in crisis“ des BMJ, die jüngst im Deutschen Ärzteblatt aufgenommen wurde [1], möchte ich exemplarisch anhand einiger großer, häufig zitierter Studien meine Zweifel an deren Interpretation erläutern.

Heart Protection Study (HPS)

Die HPS testete die Therapie mit 40 mg Simvastatin gegenüber Placebo an Patienten in der kardiovaskulären Primär- und Sekundärprävention und wurde von 1994 bis 2001 in Großbritannien durchgeführt.

Nach durchschnittlich 5 Jahren wurde die Gesamterblichkeit im Interventionsarm um 1,8 Absolutprozent gesenkt – dies entspricht einer NNT von 55, die Differenz eines zusammengesetzten Endpunktes verschiedener vaskulärer Ereignisse betrug darüber hinaus 5,4 % – entsprechend einer NNT von 18. Die Studie wirft jedoch viele Fragen auf, die mit der Patientenrekrutierung beginnen.

Patientenrekrutierung

Gut 63.000 von ca. 130.000 angeschriebenen Patienten folgten einer Einladung für eine Vorauswahl in eines der 69 Studienzentren. Hier verließen alle Teilnehmer die Studie, die wesentliche Begleitkrankheiten hatten, deren weitere Adhärenz fraglich war oder die nicht weiter an der Studie teilnehmen wollten.

Es verblieben 32.145 Patienten, die in einer Vorlaufphase vier Wochen lang Placebo und danach sechs Wochen Simvastatin erhielten. Nach diesem Erstkontakt verabschiedeten sich weitere 11.609 Patienten, sodass schließlich 20.536 Patienten in die Studie eingeschlossen wurden. Dieses Vorgehen ist fragwürdig, denn durch den Simvastatin-Erstkontakt im Vorlauf wurden alle Patienten mit Nebenwirkungen aus der Studie im Vorfeld exkludiert, und die Autoren kamen da-

her irrigerweise zu der Aussage, dass die Nebenwirkungen im Placebo- und Simvastatin-Kollektiv bis auf die zweite Kommastelle nahezu identisch seien. Im Übrigen wurde die Studie dadurch entblindet und war somit anfällig für eine Ungleichbehandlung.

Begleitmedikation

Die kardioprotektive Begleitmedikation der 20.536 eingeschlossenen Patienten war lückenhaft, zu Beginn der Studie nahmen zum Beispiel nur 57 % aller Patienten mit einer peripheren Verschlusskrankheit Acetylsalicylsäure oder ein Antikoagulans [2] und nur 31 % der 5.400 Patienten mit Typ-II-Diabetes mellitus Metformin ein [4]. Es findet sich übrigens keine Aussage über die Gleichverteilung von Metformin zu Beginn der Studie. Dies ist unbefriedigend, denn Metformin ist bekanntermaßen das einzige Antidiabetikum mit einem gesicherten makrovaskulären Nutzen.

Zur Relativierung der spärlichen Begleitmedikation wurde von den HPS-Autoren auf die Subgruppenanalysen verwiesen, die einen proportionalen Nutzen der Simvastatin-Gruppe mit und ohne jeweiliges Begleitmedikament zeigten. Dieser annähernd parallele Nutzen über alle Subgruppen hinweg macht mich misstrauisch, weil er meinem Verständnis von komplexen Therapiesituationen widerspricht. Zur Illustration meiner Skepsis verweise ich auf die Abbildung 8 in [2] mit der perschnurartigen Verteilung der odds ratio an der 0,8-Linie für alle Subgruppen zugunsten von Simvastatin.

Die Studie trifft die Aussage, dass der Gebrauch der Begleitmedikation

während der gesamten Studiendauer im Allgemeinen gleich blieb („... *such use generally continued throughout the study ...*“) – konkrete Zahlen werden jedoch nicht genannt. Diese spärliche Therapie ist für den einzelnen Patienten weder wünschenswert, noch ist es für mich glaubhaft, dass Hausarzt (und Studienzentrum) über die gesamte Studiendauer keine Therapieverbesserungen veranlasst haben sollten. Ich nehme darüber hinaus an, dass Simvastatin-Patienten und Placebogruppe unterschiedlich von einer verbesserten Begleitmedikation profitierten. Die Placebogruppe zeigte nämlich eine deutlich schlechtere Therapieadhärenz und ein höheres Ausscherverhalten. Die Ausscherer gingen dennoch im Sinne der *intention to treat analysis* in die Ergebniskalkulation ein, weil das Studienzentrum sich telefonisch

mit etablierten kardiovaskulären Erkrankungen als auch Hochrisikopatienten (Primärprävention) und wurde zwischen 1997 und 2001 durchgeführt. Die Patientenvorauswahl war ähnlich streng wie in der HPS, insgesamt wurden ca. 23.700 Patienten auf eine mögliche Teilnahme hin untersucht und nur ein gutes Viertel schließlich in die Studie hineingenommen. Die Patienten durchliefen zu Beginn eine vierwöchige Placebophase, danach erfolgte die Blockrandomisierung pro Studienort in ausbalancierten Vierer-Gruppen. Diese Blockrandomisierungen nach Placebobe-ginn sind ein übliches Verfahren, die LIPID-Studie [7] und 4S Study [8] wurden nach dem gleichen Modus durchgeführt. Es besteht dadurch jedoch die Gefahr der Entblindung und Ungleichbehandlung, insbesondere,

Zahlen dazu sind in der Studie nicht gesondert ausgewiesen worden und müssen aus einer unübersichtlichen Tabelle (Figure 5) entnommen und berechnet werden.

PROSPER schließt mit einer aus meiner Sicht angreifbaren Ergebnisdiskussion, denn die Autoren „verdünnen“ die erhöhte Krebsmorbidity in einer fehlerhaften Metaanalyse und vermeiden die klare Benennung unerwünschter Ergebnisse. Unklar bleibt ebenfalls, warum die Studie anscheinend vorzeitig beendet wurde.

Metaanalyse von Afilalo

Metaanalysen rechnen die Ergebnisse gleichartiger Studien zusammen und treffen Aussagen anhand großer Patientengruppen. Metaanalysen gelten daher als Goldstandard der Evidenzbasierten Medizin. Verzerrungen treten auf, wenn die ausgewählten Studien von vornherein fehlerhaft sind oder auch wenn die einzelnen Studienergebnisse unkorrekt zusammengefasst werden. Außerdem sind Metaanalysen – genauso wie die eingegangenen Großstudien – anfällig für die handwerklichen Fehler der Autoren. Das folgende Beispiel soll dies erläutern.

2008 publizierten Afilalo et al. eine Metaanalyse zur kardiovaskulären Sekundärprophylaxe mit Statinen bei älteren Patienten [9]. Afilalo und Mitarbeiter suchten in verschiedenen Datenbanken nach Statinstudien, die zwischen 1966 und 2007 publiziert wurden und Patientengruppen im Alter zwischen 65 und 82 Jahren untersuchten. Die Autoren inkludierten 19.569 Patienten aus neun verschiedenen Versuchen inklusive der oben erwähnten HPS, PROSPER, LIPID und 4S Study.

Der erste systematische Fehler in dieser Zusammenfassung von Studiendaten ist meines Erachtens die über große Altersspreizung der untersuchten Patientengruppe zwischen 65 und 82 Jahren. Ein 65-jähriger Patient hat annehmbar weniger häufig chronische Krankheiten, weniger Dauermedikamente und eine höhere verbleibende Lebenserwartung als ein 82-jähriger Patient. Diese Altersgruppen sind nicht miteinander vergleichbar.

Altersgruppe	Anzahl der Patienten
< 65 Jahre	9839
65–70 Jahre	4891
70–75 Jahre	4543
75–80 Jahre	1263
Gesamt (65–80 Jahre)	10.697

Tabelle 1 Altersverteilung der 10.697 Patienten, älter als 65 Jahre in HPS [2, 3]

bei den Hausärzten über erreichte Endpunkte informierte.

Ich schlage daher eine andere Deutung der Studie als die Autoren vor: Die Subgruppenanalysen zeigen einen proportionalen und annähernd parallelen Nutzen des Interventionsarmes, weil das Simvastatin-Kollektiv etwas besser gestellt war und besser betreut wurde. Die Entblindung zu Studienbeginn hat dabei geholfen.

PROSPER

PROSPER ist eine irisch-britisch-niederländische Studie und untersuchte die Wirkung von 40 mg Pravastatin gegenüber Placebo an 5.804 Patienten im Alter von 70 bis 82 Jahren [5, 6]. Die Studie inkludierte sowohl Patienten

wenn der eigene Hausarzt des Patienten – wie in PROSPER – als Mitbetreuer fungiert.

Die Ergebnisse von PROSPER sind unspektakulär, der primäre Endpunkt sank um 2,1 Absolutprozent zugunsten von Pravastatin. Dieses scheinbar günstige Ergebnis ist vor allem auf das bessere Abschneiden bei nicht tödlichen Herzinfarkten und TIAs zurückzuführen. Die koronare Sterblichkeit war im Interventionsarm um 0,9 % günstiger, demgegenüber stand ein kleiner Zuwachs an Krebsmortalität. Die Gesamtsterblichkeit war dadurch zwischen beiden Gruppen ausgeglichen. Der Pravastatinarm der Studie zeigte darüber hinaus eine erhöhte Krebsmorbidity gegenüber Placebopatienten (8,4 % gegenüber 6,8 % bei einer Hazard ratio größer als 1,25). Die

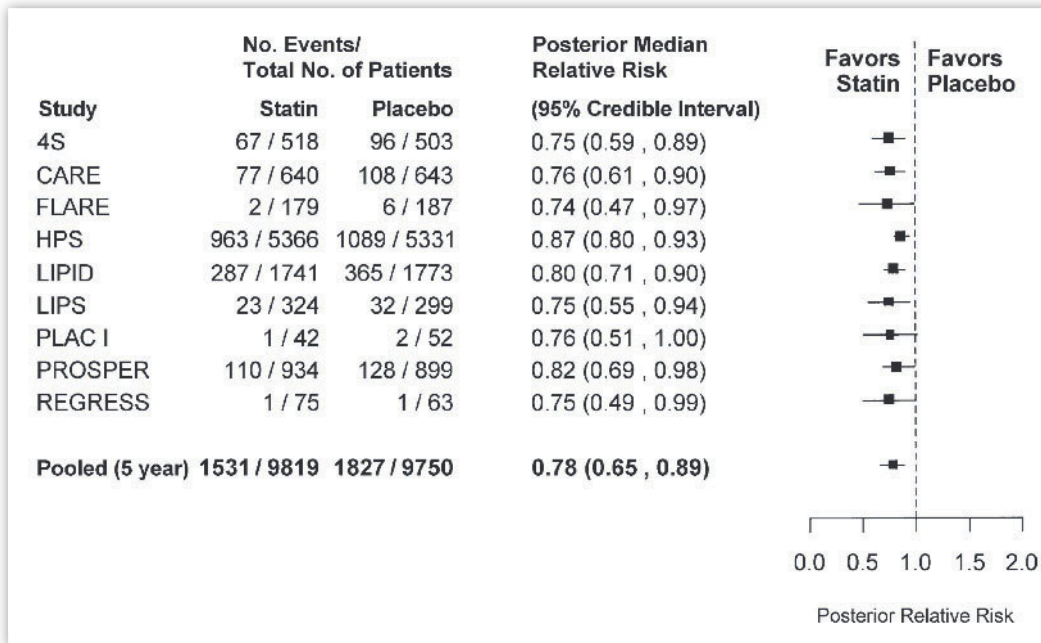


Abbildung 1
Gesamt mortalität
Placebogruppe /
Statingruppe der
einzelnen Studien –
nur Sekundärpro-
phylaxe [9]

Die Heart Protection Study stellte mit 10.697 Patienten mehr als die Hälfte der zusammengefassten Patienten dieser Metaanalyse. Diese Zusammenstellung ist bei genauerer Betrachtung unkorrekt, wie Tabelle 1 auf S. 174 zeigt.

Die Patientenzahl wurde durch das Hinzufügen der „jüngeren“ alten Patienten unnötig aufgebläht. Ein Altersmittelwert (oder Median) der einzelnen Altersgruppen wird in der HPS nicht angegeben. Bei Annahme des ungünstigsten Szenarios beträgt das Durchschnittsalter aller dieser Patienten zusammen 68 Jahre und ist damit nicht aussagetauglich für die oben erwähnte Fragestellung in Bezug auf „ältere“ Patienten.

Abbildung 1 benennt die Gesamt mortalität und zählt die Ereignisse pro Patientenzahl der einzelnen Studien jeweils für Placebo- und Statinkollektiv. Die ausgewiesene Gesamt mortalität für HPS und PROSPER ist in der Zusammenstellung von Afilalo deutlich besser als in den Originalstudien. Afilalo begründet dies mit der Hinzunahme von bisher unveröffentlichten Einzelergebnissen zur Sekundärprophylaxe, die er von den Studienautoren erhalten habe.

Dieses Vorgehen ist für mich nicht nachvollziehbar. Es wird Gründe gegeben haben, dass die Autoren der beiden Studien die Einzelergebnisse für die Sekundärprophylaxe zurückhielten. Wir erfahren beispielsweise nicht,

ob die einzelnen Gruppen ausgeglichen und die Ergebnisse robust waren. Weiterhin ist festzustellen, dass bei der Subtraktion der günstigen Ergebnisse der Sekundärprophylaxe vom Gesamtergebnis die Werte für die Primärprophylaxe deutlich schlechter ausfallen würden und damit der Therapienutzen für riesige Patientenzahlen infrage gestellt wäre. In PROSPER beträgt die errechnete Differenz der Gesamt mortalität bei Patienten der Primärprophylaxe 3,3 % zuungunsten von Pravastatin, gleichbedeutend einer *number needed to kill* von 30!

Resümee

Nach sorgfältigem Lesen habe ich große Zweifel an der Korrektheit und Seriosität vieler Großstudien und Metaanalysen insbesondere von Studien zum Einsatz von Statinen. Das Studiendesign der Großstudien ermöglicht die Entblindung zu Studienbeginn. Diese Studien entsprechen somit nicht den Anforderungen an eine randomisierte kontrollierte doppelblinde Studie, weil die Möglichkeit der (unbewussten) Betterbehandlung des Interventionsarmes angenommen werden muss. Die Ergebnisdarstellung von PROSPER und HPS ist zumindest unvollständig, Nebenwirkungen werden heruntergespielt und der Therapienutzen vergrößert dargestellt. Das Studi-

endesign der beiden vorgestellten Großstudien ist darüber hinaus eine Mischkalkulation aus Patientengruppen mit einem höheren Nutzen durch eine Statintherapie und Patientengruppen, die eher wenig von dieser Therapie profitieren. Auf diese Weise werden angebliche Behandlungsnutzen für eine Vielzahl von Patientengruppen nachgewiesen. Das Beispiel von PROSPER zeigt jedoch, dass beispielsweise eine Primärprävention für Patienten ab 70 Jahren mit Pravastatin nicht zu empfehlen ist, da diese Therapie mehr geschadet als genutzt hat.

Mit Skepsis betrachte ich ebenfalls die derzeit modernen multinationalen Studien, deren Nutzen im untersten einstelligen Prozentbereich durch das auffallend gute Abschneiden des getesteten Medikamentes in einigen Schwellenländern herrührt, wie zum Beispiel PLATO [10] oder Re-LY [11].

Die *Zeit* veröffentlichte am 14. März 2014 einen Artikel [12] über Fälschungen bei klinischer Auftragsforschung in Russland. Dort und in anderen osteuropäischen Ländern lassen internationale Pharmafirmen zunehmend neue Medikamente testen. Der Artikel beschreibt systematische Fälschungen einschließlich der Durchführung von Phantomstudien, „Patientenverdoppelungen“ sowie die Nichtdokumentation von Nebenwirkungen. In einem besonderen Falle sei ein Studienzentrum unter Druck ge-

Dr. med. Philipp Conradi ...



... Jahrgang 1963, Medizinstudium in Rostock 1986–1992,
Hausarzt in Dresden

setzt worden, den unerwarteten Tod eines 60-jährigen Patienten als natürliche Todesursache zu werten und nicht als Reaktion auf das getestete Arzneimittel. Die Kontrolle zur Einhaltung von Forschungsstandards ist hinfällig, da die Überprüfung durch dieselben Subunternehmen erfolgte, die auch die Studienaufträge an die einzelnen Zentren vergeben hatten.

Evidenzbasierte Medizin ist bekanntermaßen der umsichtige und behutsame Gebrauch der derzeit besten Beweise bei der Behandlung eines Patienten. Evidenzbasierte Medizin ist nicht die Aneinanderreihung von Wunschannahmen außerhalb des logischen Zusammenhangs.

Die Diskussionen in der Ärzteschaft über eine Korrektur der derzeitigen EbM haben begonnen. Am 9. Januar des Jahres referierte der Direktor des Deutschen Cochrane Zentrums, Gerd Antes, in Köln zum Thema „Wie glaubwürdig ist die Evidenz – von verzerrter Publikation bis zur Fälschung“ [13]. Er beschreibt systematische Verzerrungen in der EbM und benennt als „Achse des Bösen“ nicht nur die Industrie, sondern eine Vielzahl von Institutionen und Gruppen, einschließlich der Ärzteschaft, der Forscher und Wissenschaftler, Leitlinienkommissionen, Ministerien, Regierungen, Fakultäten, Ethikkommissionen, Zeitschriften, der Patienten und anderen mehr.

Die große Anzahl der aufgeführten „Mitspieler“ mit ihren unterschiedlichen Interessen und Empfindlichkeiten verweist auf die Schwierigkeiten, die sich bei der nötigen Korrektur von falschen Evidenzen ergeben werden.

Und wer soll die Evidenzbasierte Medizin vom Kopf wieder auf die Füße stellen? Jedes Essay lebt von der freundlichen Zuspitzung – die Korrektur von fehlerhaften Nutznachweisen, die die allgemeinärztliche Tätigkeit betreffen, ist eine Angelegenheit der akademischen Allgemeinmedizin. Bitte!

Interessenkonflikte: keine angegeben.

Korrespondenzadresse

Dr. med. Philipp Conradi
Facharzt für Allgemeinmedizin
Otto-Dix-Ring 98
01219 Dresden
Tel.: 0351 2752457
philippconradi@yahoo.de

Literatur

1. Leiß O. Evidenzbasierte Medizin: Kein L'art pour l'art, sondern zum Nutzen der Patienten. Dtsch Arztebl 2015; 112: A-130–132
2. Heart Protection Study Collaborative Group. MRC / BHF Heart Protection Study of cholesterol lowering with simvastatin in 20,536 high risk individuals: a randomised placebo controlled trial. Lancet 2002; 360: 7–22
3. MRC/BHF Heart Protection Study Collaborative Group. MRCBHF Heart Protection Study of cholesterol lowering therapy and of antioxidant vitamin supplementation in a wide range of patients at increased risk of coronary heart disease death: early safety and efficacy experience. Eur Heart J 1998; 20: 725–741
4. Heart Protection Study Collaborative Group. MRC / BHF Heart Protection Study of cholesterol lowering with simvastatin in 5,963 people with diabetes: a randomised placebo controlled trial. Lancet 2003; 361: 2005–2016
5. Shepherd J, Blauw G, Murphy M, et al. Prospective study of pravastatin in the elderly at risk (PROSPER). Lancet 2002; 360: 1623–1630
6. Shepherd J, Blauw G, Murphy M, et al. The design of a prospective study of pravastatin in the elderly at risk (PROSER). Amer J Cardiol 1999; 84: 1192–1197
7. The Long-Term Intervention with Pravastatin in Ischaemic Disease (LIPID) Study Group. Prevention of cardiovascular events and death with pravastatin in patients with coronary heart disease and a broad range of initial cholesterol levels. N Engl J Med 1998; 339: 1349–1357
8. Scandinavian Simvastatin Survival Study Group. Randomised trial of cholesterol lowering in 4444 patients with coronary heart disease: the Scandinavian Simvastatin Survival Study (4S). Lancet 1994; 344: 1383–89
9. Afilalo J, Duque G, Steele R, et al. Statins for secondary prevention in elderly patients. JACC 2008; 51: 37–45
10. Walentin L, Becker R, Budaj A, et al. Ticagrelor versus clopidogrel in patients with acute coronary syndromes – PLATElets inhibition an patients outcomes (PLATO) trial. N Engl J Med 2009; 361: 1045–1057
11. Connolly S, Ezekowitz M, Yusuf S, et al. Dabigatran versus warfarin in patients with atrial fibrillation. N Engl J Med 2009; 361: 1139–51
12. Ahr N, Hawranek C. Eine Überdosis Risiko. DIE ZEIT vom 14. März 2014
13. www.adka-arznei.info/index.php/programm-main/programm-uebersicht.html