

Indirekte Vergleiche – zwischen Ergebnissicherheit und Realität

Prof. Dr. Ralph Tunder und Belinda Martschinke, EBS Business School, Health Care Management Institute

Veröffentlicht in *market access & health policy* 06/2015 (November)

Um den Nutzen bzw. den Zusatznutzen eines neuen Arzneimittels innerhalb des AMNOG Prozesses gem. § 35a SGB V (A) zu bewerten, muss dieses mit einem bereits etablierten Therapiestandard (B) verglichen werden. Idealerweise werden hierzu Therapie A und B direkt miteinander verglichen. Aufgrund verschiedener Umstände liegen allerdings nicht immer die geforderten direkten Vergleiche zwischen A und B vor, wodurch in solchen Fällen vermehrt auf Analysen mittels indirekten Vergleichen ausgewichen wird. Bei einem derartigen indirekten Vergleich wird Therapie A gegen eine Therapie C (bspw. Placebo) getestet; Therapie B ebenfalls gegen C. Mit verschiedenen Verfahren kann daraufhin mittels des indirekten Komparators C der Effekt, den A verglichen mit B haben müsste, über die relativen Effekte von AC und BC berechnet werden. So einfach wie dies klingt, so sehr sitzt der Teufel im Detail, wenn es um die Umsetzung und Ergebnissicherheit indirekter Vergleiche geht. Bereits 2009 konnte im Auftrag des DIMDI nachgewiesen werden, dass indirekte Vergleiche in vielen Fällen empirisch verlässliche Ergebnisse vorweisen, auch wenn Head-to-head-Studien fehlen. Dies gilt selbst für ein gewisses Maß an Heterogenität zwischen den Studien, wobei die Frage, wieviel Heterogenität hinnehmbar ist, nicht abschließend beantwortet wurde[1]. Diese Unklarheit drückt sich auch noch gegenwärtig vielerorts aus, wie in den G-BA-Beschlüssen zu den indirekten Vergleichen der Hersteller [2].

Gründe für fehlende direkte Vergleiche

Neue Arzneimittel werden in den Zulassungsstudien regelmäßig gegen Placebo oder die Standardtherapie getestet. Allerdings besteht oft keinerlei Incentive, einen direkten head-to-head Vergleich mit anderen Arzneimitteln unter kontrollierten Studienbedingungen durchzuführen. Allerdings ist es oft nicht einmal die explizite Entscheidung des pharmazeutischen Unternehmers, einen solchen Vergleich nicht durchführen zu *wollen*. Vielmehr besteht im Vorfeld keine Klarheit darüber, welche zweckmäßige Vergleichstherapie in Deutschland nach der Zulassung in der Frühen Nutzenbewertung eines Tages zugrunde gelegt wird. Es kann sogar passieren, dass sich zum Zeitpunkt der Nutzenbewertung eine neue Evidenz oder neue Therapieoptionen ergeben haben, die zu Zeiten des Studiendesigns nicht absehbar waren. Des Weiteren werden in anderen Ländern andere Vergleichstherapien oder Standardtherapien herangezogen[2], was auch aus ökonomischer sowie zeitlicher Sicht eine zusätzliche direkte Studie, speziell konzipiert für die deutsche Nutzenbewertung, unverhältnismäßig wenn nicht gar in manchen Fällen unmöglich macht. All dies sind als Gründe anzuführen für fehlende direkte Vergleiche, wodurch die Notwendigkeit indirekter Vergleiche entstehen kann.

Anspruch der evidenzbasierten Medizin

„Gold-Standard zur Ableitung von Nutzenbelegen ist in der Regel nach wie vor die Synthese direkter Evidenz aus ergebnissicheren Studien.“[3] Selbst wenn indirekte Vergleiche in manchen Settings durchgeführt wurden und dabei nützliche Ergebnisse erzielt haben, gibt es zahlreiche Limitationen im Rahmen der Evidenzbasierung, da Ergebnisse abgeleitet oder extrapoliert werden. So sind Institutionen nicht nur in Deutschland, sondern beispielsweise auch in Kanada, skeptisch, inwiefern valide Ergebnisse generierbar sind. Angriffsfläche für Diskussionen sind dabei eventuell bestehende Unterschiede in Patientenpopulationen, die zu Heterogenität zwischen den Studien führen und somit den Effekt verfälschen können. Zudem können unterschiedliche Follow-up Zeiträume, Messung der Endpunkte oder der diagnostischen Kriterien zusätzlich die Validität infrage stellen. Weitere Bedenken beziehen sich auf veraltete Studien, die evtl. mit heutigen, etablierten Vorgehen bei klinischen Studien nicht mehr konform sind[4]. Um vergleichbare Studienbedingungen annehmen zu können, müssen daher a) biographische Patientendaten, b) Ein- und Ausschlusskriterien, sowie c)

Definition und Dokumentation von Wirksamkeits- und Sicherheitsendpunkten vor der Durchführung eines indirekten Vergleichs untersucht werden[5].

Das IQWiG schließt die Verwendung indirekter Vergleiche in seinen allgemeinen Methoden keineswegs aus, rät aber grundlegend davon ab und weist auf die hohen methodischen Anforderungen hin, damit Ergebnisse solcher Vergleiche in der Bewertung Berücksichtigung finden können (Kapitel 8.3.9)[6]. Werden Effekte geschätzt, was insb. auf indirekte Vergleiche zutrifft, werden die Angabe von Konfidenzintervallen, Szenario- sowie Sensitivitätsanalysen empfohlen (Kapitel 4.3.4).

Studienlage in der Realität

Da die Studien in indirekten Vergleichen meist nicht für diesen indirekten Vergleich konzipiert wurden, kann eine Verzerrung durch Effektmodifikatoren (bspw. Patientencharakteristika, Endpunktdefinition und -erhebung, unterschiedliche Studienprotokolle, die z.B. unterschiedliche Komorbiditäten zulassen, unterschiedliche Follow-up Zeiträume etc.) entstehen [2]. Doch selbst wenn potentielle Effektmodifikatoren, wie ungleiche Studienbedingungen, identifiziert werden, bedeutet dies nicht zwingend, dass ein indirekter Vergleich als Evidenz abgelehnt werden sollte. In methodischen Diskussionen sollte die Frage gestellt werden, wieviel Effekt der Modifikator in welche Richtung haben könnte, um dies anschließend in den Berechnungen zu berücksichtigen. *Ergebnissicherheit* kann es letztlich niemals geben; die Frage, die sich daher jedes Gesundheitssystem beantworten muss, ist, wieviel *Ergebniswahrscheinlichkeit* zwingend bzw. ausreichend ist. Oder wie es Jansen et al. innerhalb der ISPOR Task Force ausdrücken, wäre zu fragen, wie ein nicht-modifizierter Effekt aussähe und ob dies zu einer anderen Bewertung oder Entscheidung führen würde, anstatt die Ergebnisse indirekter Vergleiche mit dem Argument nicht vergleichbarer Studien abzulehnen[2].

Ein weiteres Argument spricht für eine Spezialform des indirekten Vergleichs, sogenannte Mixed-Treatment-Comparisons (MTC, die Kombination direkter und indirekter Vergleiche). Während für Zulassungsstudien meist möglichst homogene Studienpopulationen gewählt werden, um Confounder auszuschließen, spiegeln diese Populationen nicht die Heterogenität des Versorgungsalltags wieder („efficacy“ vs. „effectiveness“), was die externe Validität der Daten verringert. Aus Sicht von Entscheidern sollte es daher für die Beurteilung der effectiveness und der Kosten-Nutzen-Relation förderlich sein, leichte Heterogenität in Populationen als Spiegel des Versorgungsalltags zu betrachten, was gleichzeitig die externe Validität des Vergleichs erhöhen kann[2]. Zugleich können MTCs auch die interne Validität erhöhen, wenn im Rahmen der Konsistenzprüfung die berechneten indirekten Effekte mit den direkten Effekten gegeneinander abgeglichen werden[7].

Anspruch	Herausforderungen
Konsistenzprüfung: vergleichbare Effektschätzungen (bei Vorliegen von direkten <u>und</u> indirekten Vergleichen)[3, 5-7]	Konsistenz der Effekte zwischen direkten und indirekten Vergleichen in meisten Fällen gegeben[1]. Sinnvolles Ausmaß, inwieweit Konsistenz erforderlich, ist festzulegen.
Ähnlichkeitsprüfung hinsichtlich der Vergleichbarkeit der Effektmodifikatoren (vergleichbare Studienbedingungen, Homogenität der Daten)[3, 5]	Da Studien meist nicht für indirekten Vergleich konzipiert wurden, kann eine Verzerrung durch Effektmodifikatoren entstehen[2]; komplette Vergleichbarkeit daher oft nicht gegeben. Sinnvolles Ausmaß, inwieweit Ähnlichkeit erforderlich, ist festzulegen

Nicht adjustierte indirekte Vergleiche nicht zugelassen (da Randomisierung erhalten bleiben soll [7])	Wenn die ZVT eine nicht medikamentöse Therapie ist (bspw. Krankengymnastik), ist es kaum möglich, eine Randomisierung von Krankengymnastik versus Placebo durchzuführen, wodurch auch kein adjustierter indirekter Vergleich möglich ist[7]
Qualitative indirekte Vergleiche nicht zugelassen[7]	Bsp. mikrobielle Collagenese aus Clostridium histolyticum: indirekter Vergleich nicht möglich, daher deskriptive Gegenüberstellung von Studien (→ G-BA: kein Nutzenbeleg mangels Brückenkomparators)[7]
Zusätzliche indirekte Vergleiche können die Evidenz vorliegender direkter Vergleiche erhöhen (= Mixed Treatment Comparison (MTC))[2]	MTCs international stetig etablierter[7], wenn auch noch methodische Fragen offen bleiben[6]
Studien sollten möglichst als RCT durchgeführt werden[2]	Durchführung einer RCT teils nicht ratsam oder möglich (z.B. unethisch, falls Schaden durch eine Alternativtherapie vermutet wird; Onkologie: Phase II oft nur ein Studienarm; zu geringe Zeit aufgrund sich schnell entwickelnder Technologie; begrenzte Ressourcen)[2]

Abb. 1: Gegenüberstellung: Ansprüche an indirekte Vergleiche versus Herausforderungen in der Realität

Ausblick

In zahlreichen Veröffentlichungen wird subsumiert, dass derzeit noch methodische Probleme bei den verschiedenen Formen indirekter Vergleiche bestehen, einhergehend mit der Forderung nach methodischer Weiterentwicklung in der Zukunft. Es ist jedoch nicht erkennbar, dass auf diesem Gebiet in den letzten Jahren allzu große Sprünge gemacht wurden. Offen bleibt bei den Forderungen, wer sich dieser methodischen Entwicklung annehmen wird. Zwar hat sich die Pharmaindustrie in den vergangenen Jahren, insb. durch die Erfordernisse des AMNOG vermehrt diesen Methoden zugewandt, nur blieb die erwünschte Anerkennung bisher aus. Die Argumente für die Ablehnung indirekter Vergleiche, falls überhaupt Begründungen abgegeben werden, bleiben die Gleichen[7]. Das Dilemma bleibt bislang bestehen: indirekte Vergleiche sollen zeitaufwendige, ressourcenintensive direkte Vergleiche ersetzen, wenn solche aufgrund verschiedener Ursachen (siehe oben) nicht vorliegen oder nicht durchgeführt werden können. Kernproblem bleibt, dass meist jene Studien, die zu indirekten Vergleichen herangezogen werden, nicht für diesen indirekten Vergleich designiert wurden. Somit ist es nur in wenigen Fällen möglich, die bislang sehr eng gefassten Anforderungen an die Ergebnissicherheit indirekter Vergleiche zu erfüllen. Zugleich lassen sich indirekte Vergleiche einschließlich Mixed-Treatment-Comparisons als Chance begreifen, einen übergreifenden Vergleich zwischen einer Vielzahl an Therapieoptionen abzubilden, wie es innerhalb einer einzelnen klinischen Studie kaum umsetzbar wäre. Auch unter diesem Gesichtspunkt scheint eine weitergehende Diskussion, inwieweit *Ergebnissicherheit* gewährleistet werden kann und welches Ausmaß an Effektmotifikatoren ggf. zuzulassen ist, eine Bereicherung der Gesundheitssysteme, ohne die Evidenzbasierung aus dem Blickwinkel zu verlieren. Je umfassender eine Vielzahl an Therapien miteinander verglichen werden können, desto bessere Rückschlüsse lassen sich für die Versorgung der Patienten ziehen.

Literatur

1. Schöttker, B., D. Lühmann, and H. Raspe, *Indirekte Vergleiche von Therapieverfahren - Eine methodische Bestandsaufnahme und Validitätsprüfung (DIMDI-finanzierter HTA Bericht „Indirekte Vergleiche von Therapieverfahren“)*, in *Workshop „indirekte Vergleiche“* 2009: Berlin 12.11.09.
2. Jansen, J.P., et al., *Interpreting Indirect Treatment Comparisons and Network Meta-Analysis for Health-Care Decision Making: Report of the ISPOR Task Force on Indirect Treatment Comparisons Good Research Practices: Part 1*. *Value in Health*, 2011. 14: p. 417-428.
3. Bender, R., et al., *Stellenwert von Ergebnissen aus indirekten Vergleichen - Gemeinsame Stellungnahme von IQWiG, GMDB und IBS-DR*. 2012.
4. Wells, G.A., et al., *Indirect Evidence: Indirect Treatment Comparisons in Meta-Analysis*. 2009, Canadian Agency for Drugs and Technologies in Health.
5. Harenberg, J., et al., *Clinical trials with new direct oral anticoagulants - Additive value of indirect comparisons (network meta-analyses)*. *Phlebologie*, 2013. 2013(42): p. 139-148.
6. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG), *Allgemeine Methoden - Version 4.2 vom 22.04.2015*. 2015: Köln.
7. Moeser, G. and C. Ecker, *Indirekte Vergleiche in der frühen Nutzenbewertung in Deutschland - eine Bestandsaufnahme*. *Gesundh ökon Qual manag*, 2013. 2013(18): p. 235-253.

Autoren: Prof. Dr. Ralph Tunder, Belinda Martschinke (beide: Health Care Management Institute, EBS Business School)