

AUS DER REDAKTION

Den Referenzen Reverenz erweisen

Korrektes Zitieren ist eine wissenschaftliche Tugend, aber auch in begutachteten Artikeln nicht selbstverständlich.

Stephan Mertens, Christopher Baethge

Wissenschaftliche Aussagen müssen belegt werden. In Fachartikeln wird dazu üblicherweise das Literaturzitat benutzt, die Referenz. Sie sollte eine Behauptung präzise untermauern und den aktuellen Forschungsstand repräsentieren oder aber im Falle einer Außenseitermeinung als solche kenntlich gemacht sein. Der Verweis auf eine Literaturquelle ermöglicht dem Leser eine vertiefende Auseinandersetzung mit dem Thema und versetzt ihn in die Lage, die zugrundeliegende Evidenz zu beurteilen. Den Autor verpflichtet der Gebrauch von Zitaten, präzise zu formulieren, um die Quelle genau wieder zu geben.

Die Konvention des Zitierens ist ein konstituierendes Element wissenschaftlicher Texte, grenzt sie von journalistischen Arbeiten und anderen Textformen ab und verleiht ihnen besondere inhaltliche Autorität. Ein Großteil der Leser dürfte einer referenzierten Aussage mehr Vertrauen entgegenbringen als einer ohne Literaturverweis, auch ohne sie geprüft zu haben. Wahl und Platzierung eines Zitats sind jedoch komplexe, fehlerbehaftete Vorgänge. Folgende Fehler können auftreten:

- Die Quelle untermauert nicht oder nur teilweise die Aussage.
- Der Verweis ist so platziert, dass nicht deutlich wird, auf welche Aussage er sich bezieht.
- Die bibliografischen Angaben wie Autoren, Titel, Bandnummer, Seitenzahl oder Name der Zeitschrift sind unvollständig oder falsch. Dies kann dazu führen, dass es schwierig oder unmöglich ist, die Quelle zu finden.

Insbesondere inhaltliche Fehler sind problematisch, denn sie sind schwer zu durchschauen: Der Leser kann die Genauigkeit eines Verweises häufig nur dann beurteilen, wenn er über einschlägige Fachkenntnisse verfügt oder sich die Mühe macht, die zitierte Quelle zu recherchieren. Eine andere Form von Referenzierungsfehler, die im vorliegenden Text allerdings nicht thematisiert wird, stellen Plagiate dar: Hier wird die Informationsquelle bewusst verschwiegen und fälschlich als eigene intellektuelle Leistung präsentiert (1).

Inhaltliche Fehler

In einer systematischen Übersicht werteten Wager und Middleton 66 Studien aus, in denen insgesamt 3 836 Referenzen aus 74 biomedizinischen Fachzeitschriften untersucht wurden (2). Neben Zitaten in allgemeinmedizinischen Journalen wie The Lancet oder dem British

Medical Journal wurden auch Untersuchungen von Zeitschriften für Chirurgie, Pädiatrie und andere Fächer berücksichtigt. Die durchschnittliche Quote großer und kleiner inhaltlicher Fehler betrug in dieser Cochrane-Analyse insgesamt 20 %, mit einer Spannweite von 0 bis 50 % (2). In der systematischen Übersicht wurden Arbeiten bis einschließlich 2007 erfasst.

Eine eigene Literaturrecherche in Medline beschränkte sich auf Arbeiten ab 2008, um keine Veröffentlichungen doppelt zu zählen. Diese Publikationen sind in der *Tabelle* zusammengefasst. Ab 2008 konnten fünf Veröffentlichungen identifiziert werden, in denen insgesamt 1 521 Zitate aus 16 Zeitschriften analysiert wurden. Hier identifizierte man 299 Fehler, was einer Quote von 19,7 % entspricht und die zuvor ermittelten Werte bestätigt. In den Veröffentlichungen gab es keine deutlichen Hinweise darauf, ob die Fehlerquote beispielsweise mit dem Impact-Faktor der untersuchten Zeitschrift oder einer bestimmten medizinischen Disziplin zusammenhängt. Darüber hinaus scheint die Fehlerquote von 1977, als die ersten Studien veröffentlicht wurden, bis 2011 konstant zu bleiben, wie eine eigene Analyse der Daten zeigt.

Deutsches Ärzteblatt

Wie häufig inhaltliche Zitationsfehler sind, wurde in den vergangenen Jahren mehrfach untersucht, allerdings noch nicht in deutschsprachigen Fachzeitschriften. Aus diesem Grund haben die Autoren eine kleine Pilotstudie von Referenzen im Deutschen Ärzteblatt durchgeführt.

Um einen Eindruck von der Korrektheit der Zitate im wissenschaftlichen Teil des Deutschen Ärzteblatts zu erhalten, wurden 50 zufällig ausgewählte Zitate aus Artikeln des Jahres 2010 ausgewertet. Die Übereinstimmung von Aussage und Zitat beurteilten S. M. und C. B. unabhängig voneinander anhand der Zusammenfassung oder des Volltextes der Referenz. Jedes Zitat wurde entweder als richtig oder als falsch bewertet, wobei in kleine und substanzielle Fehler unterteilt wurde. Als kleine Fehler wurden beispielsweise Zitate definiert, in denen geringe, die Aussage aber dennoch stützende Abweichungen bei Zahlenangaben bestanden. Unterschiede in der Beurteilung wurden gemeinsam diskutiert und bewertet.

Bei neun der fünfzig Aussagen fand man substanzielle inhaltliche Fehler (18 %). Die Häufigkeit klei-

TABELLE

Studien, in denen inhaltliche Fehler bei der Referenzierung festgestellt wurden

Studie	Gepüfzte Referenzen (N)	Fehlerhafte Referenzierung* ¹ (%)	Untersuchte Zeitschriften (N)
Wager, Middleton 2008, Cochrane Review (2)	3 836	20, Spannweite: 0 bis 50	74 (verschiedene Fächer)
Reddy et al. 2008 (3)	258	7,8, Spannweite: 3,5 bis 10,3	Chirurgie (4)
Al-Benna et al. 2009 (6)	117	13,7; 95%-KI: 8,6–19,5	Verbrennungsmedizin (2)
Singh, Chaudhary 2009 (7)	46	43,5; 95%-KI: 29,2–57,8	Dermatologie (1)
Davis et al. 2010 (8)	200	38; 95%-KI: 30,1–47	Orthopädie (4)
Awrey et al. 2011 (5)	900	18,5, Spannweite: 10,5 bis 22,0	Chirurgie (5)
Eigene Pilotstudie* ²	50	Schwere Fehler: 18, 95%-KI: 9,2–30,5 Kleine Fehler: 14, 95%-KI: 6,3–25,7	Medizin, allgemein (Deutsches Ärzteblatt)

Bei der Studie von Wager und Middleton handelt es sich um eine systematische Übersicht (Cochrane), die Veröffentlichungen bis Ende 2007 berücksichtigt hat. Um keine Arbeiten doppelt zu bewerten, wurden bei den nachfolgend aufgeführten Veröffentlichungen nur Arbeiten ab 2008 erfasst.

*¹ Referenz untermauert nicht oder nicht vollständig die im Artikel dargelegte Aussage.

*² In der Pilotstudie wurden Referenzen nicht gewertet, die sich auf Bücher bezogen.

Die Einschätzung der Bewerter stimmte befriedigend überein (gewichteter Cohens Kappa = 0,583).

ner Fehler betrug 14 %. Obwohl diese kleine Stichprobe nur eine ungenaue Abschätzung der Fehlerrate ermöglicht (95%-Konfidenzintervall substanzieller inhaltlicher Fehler: 9,2 bis 30,5 %), passt das Ergebnis gut zu den in der Literatur gefundenen Häufigkeiten.

Ein direkter Vergleich dieser Daten mit der Literatur ist dennoch schwierig, weil die Definition von Fehlern uneinheitlich ist. Bereits Wager und Middleton konnten aus diesem Grund einige Arbeiten in der Cochrane-Publikation nicht berücksichtigen (2). Darüber hinaus ist das Verhältnis von großen zu kleinen Fehlern variabel. So wird der Anteil großer Fehler an der Gesamtfehlerquote mit bis zu 80 % angegeben (3). Vor diesem Hintergrund scheinen die im Deutschen Ärzteblatt gefundenen Fehlerraten mit denen anderer Titel vergleichbar zu sein (*Grafik*).

Bibliografische Angaben

In der Cochrane-Publikation werteten Wager und Middleton 66 Arbeiten aus, in denen auch die bibliografischen Angaben überprüft wurden. Von den mehr als 27 000 Referenzen aus über 100 Zeitschriften waren durchschnittlich 38 % fehlerhaft (2). Je nach Zeitschrift bestanden Fehlerquoten zwischen 4 % bis 67 %. Die Cochrane-Autoren haben Hinweise gefunden, dass eine redaktionelle Bearbeitung mit einer geringeren Fehlerquote verbunden ist.

In einer eigenen Medline-Recherche ab 2008 konnten sechs weitere Arbeiten identifiziert werden, in denen man die bibliografischen Angaben überprüft hatte. Von den insgesamt 1 928 Zitationen aus 18 Zeitschriften waren durchschnittlich 16,4 % fehlerhaft (Spannweite: 7,5 % bis 59,6 %) (3–8). Große Fehler, durch die die Originalquelle nicht identifiziert werden

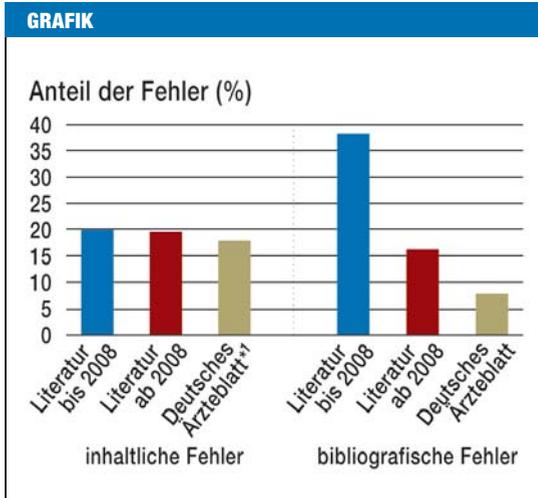
konnte, waren selten. In der Gesamtschau zeigte sich, dass die Quote bibliografischer Fehler in den Publikationen seit 2008 deutlich niedriger ist als in den Jahren davor. Dies steht im Gegensatz zur inhaltlichen Korrektheit, bei der sich eine über die Zeit konstante Fehlerrate abzuzeichnen scheint. Die niedrigere Rate bibliografischer Fehler könnte auf die verstärkte Nutzung von elektronischer Referenzierungssoftware zurückzuführen sein.

Deutsches Ärzteblatt

Da es offenbar keine Veröffentlichungen gibt, die die bibliografischen Angaben in deutschsprachigen Journalen untersucht haben, werteten die Autoren 50 zufällig ausgewählte Referenzen aus 50 Artikeln des Deutschen Ärzteblatts aus dem Jahr 2010 aus. Hierbei wurden andere Referenzen betrachtet als im ersten Teil. Vier der Zitate waren fehlerhaft (8 %, 95%-KI: 2,6 % bis 18,2 %). Allerdings handelte es sich ausschließlich um kleine Fehler, wie die verkehrte Schreibweise eines Autors. Das Auffinden des Artikels war in keinem Fall beeinträchtigt. Diese Quote liegt deutlich unter den in der Literatur gefundenen Werten (*Grafik*) und könnte auf der umfangreichen redaktionellen Bearbeitung der Manuskripte beim Deutschen Ärzteblatt beruhen.

Schlussfolgerung

Referenzen in wissenschaftlichen Arbeiten sind häufig nicht korrekt. Publierte und eigene Daten weisen auf eine Fehlerquote von etwa 20 % hin. Dies ist eine eher konservative Schätzung, denn in den Untersuchungen wurde lediglich geprüft, ob die Aussage mit der Quelle übereinstimmt. Interessanter aber auch ungleich aufwendiger wäre es zu prüfen, ob die gewähl-



Darstellung der inhaltlichen und bibliografischen Fehler.
 Die Erfassung der Literatur bis 2008 bezieht sich auf (2), ab 2008 auf eine eigene Literaturrecherche und die Angaben für das Deutsche Ärzteblatt auf die Pilotstudie. ^{*1}Darstellung der substanziiellen Fehler

te Zitierung wissenschaftlich korrekt und für den Stand der Forschung repräsentativ ist und nicht eher vom hypothesen- oder interessengetriebenen Wissenschaftsbild des Autors beeinflusst wird.

Wenn 20 % der Zitate falsch sind, bedeutet dies, dass auch der Impact-Faktor mit einer entsprechend hohen Fehlerrate behaftet ist. Zur Erinnerung: Der Impact-Faktor eines Journals erfasst, wie oft Artikel einer Zeitschrift in einem bestimmten Zeitraum zitiert werden und dient der Evaluation von Zeitschriften und Autoren (9). Bei der Berechnung wird davon ausgegangen, dass es keine inhaltlichen Zitationsfehler gibt. Wenn die Ungenauigkeit systematischer Natur wäre, sich also auf alle Artikel gleich auswirken würde, wären zwar die absoluten Impact-Faktor-Werte fehlerhaft, aber zumindest das Verhältnis zwischen den Zeitschriften bliebe gleich. Es ist aber nicht sehr wahrscheinlich, dass eine solche systematische Verzerrung vorliegt. So ist vorstellbar, dass populäre Arbeiten überproportional häufig falsch zitiert werden, alleine schon deshalb, weil sie vielen Wissenschaftlern bekannt sind. Dies könnte zu einer Verzerrung des Impact-Faktors zugunsten viel zitierter Arbeiten führen.

Der vergleichsweise hohe Anteil nicht korrekter Zitate könnte aber nicht nur ein Hinweis auf Nachlässigkeit seitens der Autoren sein. Er könnte auch bedeuten, dass die Forderung nach Belegen für jede Behauptung vielfach eine Überforderung für Wissenschaftler darstellt. Schließlich haben die Verfasser wissenschaftlicher Arbeiten vor allem die Darstellung ihrer Ergebnisse oder – im Falle einer Übersichtsarbeit – ihrer Ansicht eines klinischen Problems im Sinn. In diesem Rahmen kann der lückenlose Literaturnachweis fehlerhaft sein, ohne dass die wesentli-

chen Aussagen eines Artikels falsch sein müssen. Allerdings ist unklar, ob Zitierfehler eher periphere oder zentrale Aspekte berühren.

Um mehr Transparenz bei der Auswahl der Quellen zu schaffen, wäre es hilfreich, nicht nur bei systematischen Reviews, sondern idealerweise auch bei Originalarbeiten und bei narrativen Übersichtsartikeln zu beschreiben, wie die Autoren die verwendete Literatur ausgewählt haben: Gibt es Ein- und Ausschlusskriterien für die herangezogenen Publikationen? In welchen Datenbanken und in welchem Zeitraum wurde recherchiert? Basieren die Zitierungen nur auf Originalarbeiten oder auch auf Übersichtsartikeln? Dem kritischen Leser könnten diese Angaben ermöglichen, die Auswahlkriterien der zitierten Literatur zu prüfen und nachzuvollziehen, um so die Ausgewogenheit der Zitate besser zu beurteilen. Darüber hinaus könnte dies Autoren motivieren, intensiver zu recherchieren und weitere, vorher nicht bekannte relevante Arbeiten zu erwähnen.

Interessenkonflikt

Dr. Mertens ist Redakteur und Prof. Baethge Schriftleiter in der medizinisch-wissenschaftlichen Redaktion des Deutschen Ärzteblatts.

LITERATUR

- Mertens S: Spotlight on plagiarism. Dtsch Arztebl Int 2010; 107: 863–5.
- Wager E, Middleton P: Technical editing of research reports in biomedical journals. Cochrane Database Syst Rev 2008 (4): MR000002.
- Reddy MS, Srinivas S, Sabanayagam N, Balasubramanian SP: Accuracy of references in general surgical journals—an old problem revisited. Surgeon 2008; 6: 71–5.
- Unver B, Senduran M, Unver Kocak F, Gunal I, Karatosun V: Reference accuracy in four rehabilitation journals. Clin Rehabil 2009; 23: 741–5.
- Awrey J, Inaba K, Barnparas G, et al.: Reference accuracy in the general surgery literature. World J Surg 2011; 35: 475–9.
- Al-Benna S, Rajgarhia P, Ahmed S, Sheikh Z: Accuracy of references in burns journals. Burns 2009; 35: 677–80.
- Singh S, Chaudhary R: Accuracy of references cited in articles published in Indian Journal of Dermatology, Venereology and Leprology: a pilot study. Indian J Dermatol Venereol Leprol 2009; 75: 488–91.
- Davis JR, Weigl DM, Edmonds JP, Blackhurst DW: Reference accuracy in peer-reviewed pediatric orthopaedic literature. J Bone Joint Surg Am 2010; 92: 1155–61.
- Baethge C: A good start: Deutsches Ärzteblatt gets its first impact factor. Dtsch Arztebl Int 2010; 107: 693.

Anschrift für die Verfasser

Dr. sc. nat. Stephan Mertens
 Medizinisch-wissenschaftliche Redaktion
 Deutsches Ärzteblatt, Ottostraße 12, 50859 Köln
 Mertens@aerzteblatt.de

Zitierweise

Mertens S, Baethge C: The virtues of correct citation—careful referencing is important but often neglected even in peer reviewed articles. Dtsch Arztebl Int 2011; 108(33): 550–2. DOI: 10.3238/arztebl.2011.0550

The English version of this article is available online: www.aerzteblatt-international.de