

Die Grenzen der Literatur-Datenbanken

Wie gut oder schlecht das Ergebnis einer Literatur-Recherche ist, hängt auch von der Qualifikation des Suchers ab

Aktuelle und verlässliche medizinischen Informationen sind in referierten Zeitschriften zu finden. Wenn auch oft kritisiert und bei weitem nicht perfekt, sorgt dabei das Peer-Review-System für ein gewisses Qualitätsniveau der Veröffentlichungen. Da man aber nur ein begrenztes Spektrum von Fachzeitschriften beziehen und lesen kann, ist es unmöglich, auf jedem Gebiet hundertprozentig informiert zu sein. Angesichts der zunehmenden Interdisziplinarität des medizinischen Wissens wird derjenige Teil der Medizin immer kleiner, der überblickt werden kann. Bereits im letzten Jahrhundert wurde dieses Problem erkannt und Übersichts- bzw. Referatedienste wie z.B. der Index Medicus geschaffen. Diese verzeichneten nicht nur die erschienene Literatur, sondern strukturierten sie durch formale und sachliche Erschließung (Indexierung) und machten sie dadurch leichter auffindbar. Wer heute die neueste Literatur zu aktuellen Fragestellungen sucht, führt in der Regel eine Recherche in MEDLINE durch, dem Nachfolger des Index Medicus.¹ Dabei vertraut man auf den guten Ruf dieser zentralen medizinischen Literaturdatenbank und nimmt stillschweigend an, dass MEDLINE die gesamte oder zumindest den wichtigsten Teil der medizinischen Zeitschriftenliteratur verzeichnet. Dieser Artikel geht im folgenden der Frage nach, ob diese Annahme gerechtfertigt ist. Dazu wird untersucht, welche medizinischen Zeitschriften MEDLINE indexiert und ob es andere Datenbanken gibt, welche die medizinische Literatur besser abdecken.

Umfang der Datenbanken

Die umfassendste Informationsquelle für Zeitschriften weltweit ist das *Ulrich's International Periodicals Directory*². Der *Ulrich's* enthält detaillierte Angaben zu 236.000 Zeitschriften und deren Verzeichnung in MEDLINE, Current Contents und weitere 1.500 Literaturdatenbanken. Mit diesen Angaben kann die Frage beantwortet werden, welches medizinische Fachgebiet von welcher Datenbank am besten abgedeckt wird. Für diese Studie wurden nicht nur die Zeitschriften eines einzigen Fachgebiet betrachtet,³ sondern mit insgesamt 45 Fachgebieten das gesamte Fächerspektrum der Medizin. Dabei wurde sich auf MEDLINE und sieben weitere große biomedizinische Literaturdatenbanken beschränkt (Abb.1). Der *Ulrich's* weist auf den untersuchten 45 Fachgebieten 13.426 laufende Zeitschriften aus, von denen lediglich 6.146 (45,8%) überhaupt nur von einer der 1.500 Literaturdatenbanken indexiert werden. Die überwiegende Mehrheit dieser indexierten Zeitschriften (79,6%) wird von mindestens einer der acht untersuchten Datenbanken indexiert, was deren zentrale Stellung unterstreicht.

In Abb. 1 sehen Sie die Zahl der indexierten Titel pro Datenbank. Es ist dabei zu beachten, dass mit MEDLINE, EMBASE und CINAHL nur drei dieser Datenbanken einen ausschließlich medizinischen Bezug haben und daher zahlenmässig vergleichbar sind. Current Contents, Chemical Abstracts und der Science Citation Index sind fachübergreifend ausgerichtet, was die Zahl der indexierten Titel in die Höhe treibt.

¹ MEDLINE = MEDlars OnLINE, die größte und meistbenutzte medizinische Literaturdatenbank der Welt, wird von der National Library of Medicine der USA hergestellt. MEDLINE besteht aus Index Medicus, Index to Dental Literature und International Nursing Index und enthält 11 Mio. Artikelnachweise.

² <http://www.ulrichsweb.com>

³ S. McDonald, L. Taylor, C. Adams "Searching the right database. A comparison of four databases for psychiatry journals" *Health Libraries Review* 16(3):151-156 (1999)

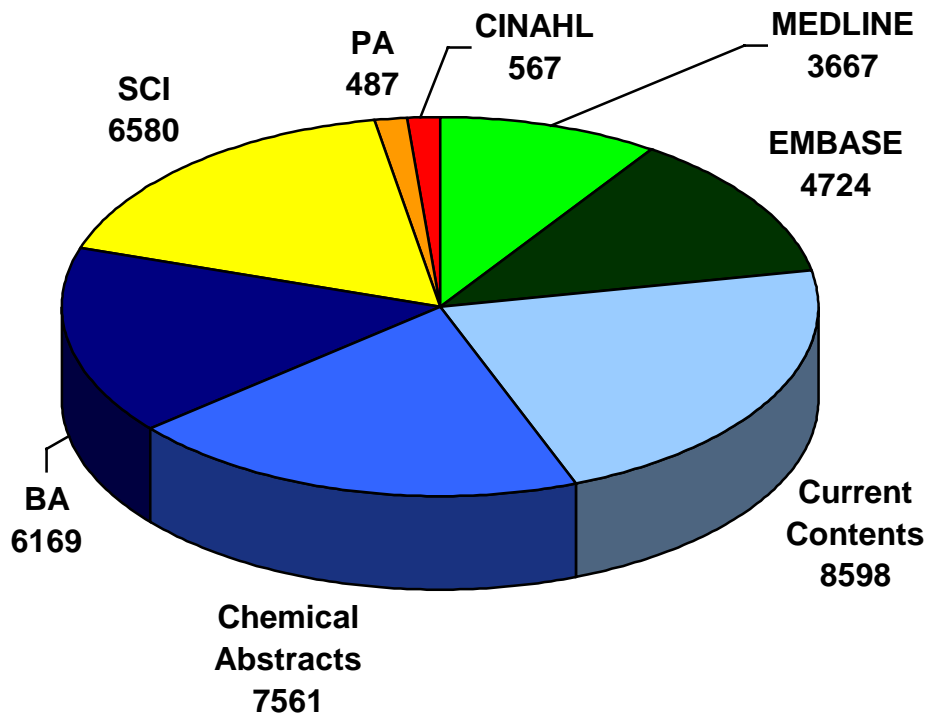


Abbildung 1: MEDLINE enthält nicht die meisten Zeitschriften, besitzt aber die beste fachliche Abdeckung in der Medizin.

[EMBASE = Excerpta Medica dataBASE, CINAHL = Cumulative Index to Nursing & Allied Health Literature, BA = Biological Abstracts, CC= Current Contents (alle Editionen), SCI = Science und Social Science Citation Index, PA = Psychological Abstracts, CA = Chemical Abstracts]

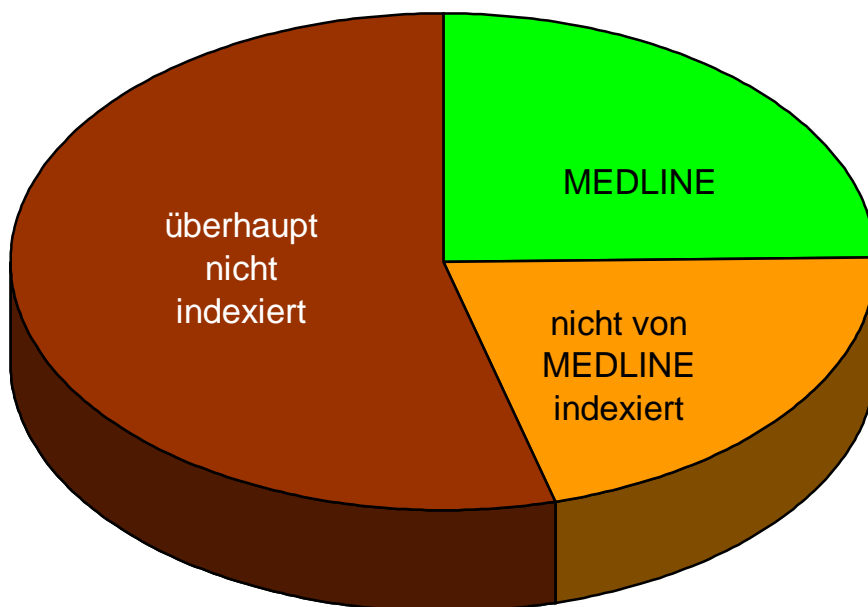


Abbildung 2: MEDLINE indexiert knapp ein Viertel aller 13.426 medizinischen Zeitschriften

Stärken und Schwächen der großen Datenbanken

MEDLINE deckt durchschnittlich 24,7% der Zeitschriften in den 45 betrachteten Fachgebieten ab. In den einzelnen Fachgebieten kann der Abdeckungsgrad jedoch weitaus geringer sein, wie z.B. bei Public Health mit knapp 12%, in der Gerontologie (10%) oder der Chiropraxis (4%). In

Randgebieten wie der Alternativmedizin, der Arbeitsmedizin und dem Public Health-Bereich ist MEDLINE - wie übrigens auch alle anderen untersuchten Datenbanken – ebenfalls schlecht bestückt. Dagegen liegt die Stärke von MEDLINE in den klinischen Fachbereichen und bei den Grundlagenfächern wie Anatomie, Biochemie, Genetik, usw. Wie in Tabelle 1 zu sehen, reicht die Spannweite des fachlichen Abdeckungsgrades von 3,6% bis 45,8%.

Tabelle 1

FACHGEBIETE	Zeitschriften weltweit	In Datenbanken nachgewiesene Zeitschriften	Von MEDLINE nachgewiesen (in % der aktiven Titel)	Von EMBASE nachgewiesen (in % der aktiven Titel)	Überlappung (EMBASE-Titel, die auch von MEDLINE nachgewiesen werden)
Allgemeinmedizin	3095	1293	19,5%	19,6%	70,0%
Alternativmedizin	185	32	6,5%	4,3%	37,5%
Anästhesiologie	110	55	26,4%	30,0%	81,8%
Anatomie, Histologie und Zytologie	201	147	45,8%	37,8%	86,8%
Arbeitsmedizin und Arbeitssicherheit	405	124	7,7%	8,4%	70,6%
Atemwegserkrankungen	144	62	24,3%	22,2%	71,9%
Augenheilkunde und Optometrie	287	113	23,3%	25,8%	70,3%
Biochemie	458	326	37,6%	32,8%	79,3%
Chiropraktik, Homöopathie, Osteopathie	101	23	4,0%	5,9%	50,0%
Chirurgie	549	290	32,8%	31,3%	77,9%
Dermatologie und Venerologie	163	93	30,7%	30,7%	76,0%
Drogensucht und Alkoholismus	205	70	12,2%	13,2%	66,7%
Endokrinologie	218	123	30,3%	28,4%	80,6%
Ernährung und Diätetik	394	155	9,4%	9,6%	65,8%
Gastroenterologie	169	95	32,0%	35,5%	71,7%
Geburts- und Frauenheilkunde	389	174	23,4%	20,1%	80,8%
Genetik	254	171	33,5%	29,9%	84,2%
Gerichtsmedizin	77	41	22,1%	26,0%	70,0%
Gerontologie und Geriatrie	418	118	9,8%	10,5%	61,4%
Gesundheitszustand und Hygiene	558	126	3,6%	4,3%	66,7%
Hals-, Nasen-, Ohrenheilkunde	140	82	36,4%	39,3%	80,0%
Hämatologie	125	69	36,8%	35,2%	84,1%
Immunologie, Allergologie	221	134	39,8%	35,7%	83,5%
Infektiologie	311	154	31,2%	25,1%	79,5%
Innere Medizin	72	39	31,9%	27,8%	70,0%
Kardiologie	369	165	24,1%	25,5%	76,6%
Kinderheilkunde	365	182	27,7%	25,2%	70,7%
Klinische Chemie und Labormedizin	228	121	22,8%	22,4%	74,5%
Krankenhäuser, Verwaltung, Management	574	129	5,2%	10,3%	37,3%
Krankenpflege	528	255	31,4%	3,2%	94,1%
Medizinische Informatik	70	33	24,3%	12,9%	100,0%
Mikrobiologie	299	210	37,5%	33,1%	76,8%
Öffentliche Gesundheitspflege (Public Health)	830	299	11,9%	11,8%	61,2%
Onkologie	389	205	32,6%	26,5%	82,5%
Orthopädie und Traumatologie	346	134	20,5%	20,8%	72,2%
Pharmazie und Pharmakologie	1294	587	16,0%	20,5%	63,4%
Physikalische Therapie und Rehabilitation	177	65	10,7%	11,9%	47,6%
Physiologie	204	133	42,2%	36,3%	90,5%
Psychiatrie und Neurologie	1087	617	29,3%	30,4%	72,4%
Radiologie und Nuklearmedizin	278	140	30,9%	27,3%	80,3%
Rheumatologie	96	42	24,0%	18,8%	83,3%
Sportmedizin	107	61	15,9%	20,6%	54,5%
Umweltmedizin und Toxikologie	175	101	31,4%	36,0%	79,4%
Urologie und Nephrologie	185	86	25,4%	22,2%	78,0%
Zahnheilkunde	531	246	37,3%	8,5%	95,6%
<i>Summe/Durchschnitt</i>	<i>13.426</i>	<i>6.146</i>	<i>24,7%</i>	<i>22,5%</i>	<i>71,9%</i>

In obiger Tabelle wurde zugleich das Fachprofil von MEDLINE dem seines größten Konkurrenten EMBASE gegenübergestellt; der jeweils bessere Abdeckungsgrad wurde bei den Prozentzahlen fett hervorgehoben. Der in der letzten Spalte ausgewiesene Überlappungsgrad gibt den Prozentsatz der in EMBASE enthaltenen Titel wieder, die auch von MEDLINE indiziert werden. EMBASE enthält nur sehr wenige Pflege- und Zahnmedizin-Titel, was auch am Überlappungsgrad von 100 bzw. 95,6% sichtbar wird. In den klinischen Fächern schenken sich beide Datenbanken nichts, während bei den Grundlagenfächern MEDLINE die Nase vorne hat. 14,3% aller in EMBASE verzeichneten Zeitschriften stammen aus deutschsprachigen Ländern, aber nur 11,5% der in MEDLINE verzeichneten. Obwohl EMBASE damit rund 250 deutschsprachige Titel mehr enthält als MEDLINE, sind doch viele deutschsprachige Zeitschriften in den untersuchten Datenbanken überhaupt nicht nachgewiesen.⁴

Die kardiologischen Zeitschriften

Wie sieht es in der Kardiologie aus? Auf dem Gebiet der Kardiologie existieren laut *Ulrich's* 369 Zeitschriften weltweit. Doch davon sind lediglich 165 in irgendeiner Literaturdatenbank indiziert. Der Prozentsatz der potentiell nicht zu findenden Literatur wird noch größer, wenn man sich bei der Suche auf eine Datenbank wie z.B. MEDLINE beschränkt. In MEDLINE sind von diesen 369 Titeln nur 89 nachgewiesen (24,1%). Bei einer MEDLINE-Recherche kann man also bestenfalls die Artikel dieser 89 Zeitschriften finden, nicht jedoch die Arbeiten der übrigen 280 Zeitschriften (Tab. 2).

Tabelle 2: Nicht in MEDLINE enthaltene deutschsprachige kardiologische Zeitschriften (in Auswahl)

Titel	Nachweisdatenbank
Cardio News	-
Das Neurophysiologie-Labor (war: Das EEG-Labor)	EMBASE
Gefäßchirurgie : Zeitschrift für vaskuläre und endovaskuläre Chirurgie	EMBASE
Herz + Gefäße	EMBASE
Herz-Kreislauf	EMBASE
Herzschrittmachertherapie & Elektrophysiologie	EMBASE
Kardiologie Assistenz	-
Nieren- und Hochdruckkrankheiten	Biological Abstracts, EMBASE
Perfusion : Durchblutungsstörungen und Arteriosklerose in Klinik und Praxis	Current Contents
Vasomed : die Fachzeitschrift für Gefässerkrankungen	EMBASE
Zeitschrift für Herz-, Thorax- und Gefäßchirurgie	EMBASE

Andere Datenbanken

Sind in den anderen untersuchten Datenbanken mehr Zeitschriften nachgewiesen als in MEDLINE? In manchen Fachgebieten schon, aber im Durchschnitt aller Fächer ist MEDLINE mit 24,7% der klare Spitzenreiter. Der Konkurrent EMBASE folgt mit 22,5% auf dem Fuß, vier weitere Literaturdatenbanken knapp dahinter. Die beiden fachspezifischen Datenbanken Psychological Abstracts (PA) und CINAHL belegen mit 3,3 und 0,9% die letzten beiden Plätze. MEDLINE dominiert in insgesamt 19 der 45 Fachgebiete, EMBASE in 17. Damit sind diese beide Datenbanken erste Wahl in mehr als drei Viertel aller Fachgebiete. Current Contents enthält in sechs Fachgebieten die meisten Zeitschriften, Chemical Abstracts in vier und Biological Abstracts in zwei. Die übrigen drei Datenbanken (CINAHL, PA, SCI) gehen leer aus. Current Contents und Biological Abstracts sind durch ihre gute Abdeckung der Life Sciences bei

⁴ Um die Artikel dieser Titel zu finden, muss man auf spezielle Anbieter ausweichen wie den Springer-Verlag (<http://link.springer.de/search.htm>) oder die ZB Med in Köln mit ihrem „Current Contents Medizin deutscher und deutschsprachiger Zeitschriften“ (<http://medsun.zbmed.uni-koeln.de/webOPAC/>).

den vorklinischen Zeitschriften besonders stark, während sie in den anderen Fachgebieten meistens nur eine untergeordnete Rolle spielen. Wie nicht weiter verwunderlich, stellt EMBASE die beste Ergänzung zu MEDLINE dar, weil es in den meisten Fachgebieten den maximalen Zugewinn an Literaturzitationen verspricht (Tab. 3). Besonders erfolgversprechend ist eine EMBASE-Recherche in den klinisch-theoretischen und gesundheitsorientierten Fachgebieten mit zusätzlichen Literaturhinweisen von plus 32% bzw. plus 49%. In den Grundlagenfächern lohnt sich eine Zusatzrecherche dagegen vor allem in Biological Abstracts bzw. Current Contents und in der Krankenpflege natürlich in CINAHL. Lediglich in der Zahnmedizin und der Medizinischen Informatik bringt eine zusätzliche Recherche fast keinen Zugewinn.

Tabelle 3

Fachgebiet	Beste Ergänzung zu MEDLINE	Zugewinn bei Recherche in ‚besten Ergänzung‘	Zugewinn bei Recherche in allen sieben weiteren DBs
Allgemeinmedizin	EMBASE	30%	70%
Alternativmedizin	EMBASE	42%	50%
Anästhesiologie	CC	24%	62%
Anatomie, Histologie und Zytologie	BA	29%	52%
Arbeitsmedizin und Arbeitssicherheit	EMBASE	32%	106%
Atemwegserkrankungen	BA	40%	60%
Augenheilkunde und Optometrie	EMBASE	33%	46%
Biochemie	CA	59%	76%
Chiropraktik, Homöopathie, Osteopathie	EMBASE/CC	75%	200%
Chirurgie	EMBASE	21%	37%
Dermatologie und Venerologie	EMBASE	24%	44%
Drogensucht und Alkoholismus	CC	44%	96%
Endokrinologie	CC	39%	64%
Ernährung und Diätetik	BA	73%	165%
Gastroenterologie	EMBASE	31%	48%
Geburts- und Frauenheilkunde	CC	20%	46%
Genetik	BA	39%	75%
Gerichtsmedizin	EMBASE	35%	76%
Gerontologie und Geriatrie	CC	51%	98%
Gesundheitszustand und Hygiene	EMBASE/CIN	40%	160%
Hals-, Nasen-, Ohrenheilkunde	EMBASE	22%	29%
Hämatologie	CC	20%	39%
Immunologie, Allergologie	CC	22%	57%
Infektiologie	EMBASE	16%	33%
Innere Medizin	EMBASE	26%	41%
Kardiologie	CC	31%	52%
Kinderheilkunde	EMBASE	27%	51%
Klinische Chemie und Labormedizin	CA	50%	73%
Krankenhäuser, Verwaltung, Management	EMBASE	123%	177%
Krankenpflege	CINAHL	36%	37%
Medizinische Informatik	SCI	12%	18%
Mikrobiologie	BA	29%	63%
Öffentliche Gesundheitspflege	EMBASE	38%	88%
Onkologie	CC	18%	39%
Orthopädie und Traumatologie	EMBASE	28%	54%
Pharmazie und Pharmakologie	CA	84%	109%
Physikalische Therapie und Rehabilitation	EMBASE	58%	111%
Physiologie	BA	22%	47%
Psychiatrie und Neurologie	SCI	33%	64%
Radiologie und Nuklearmedizin	EMBASE	17%	40%
Rheumatologie	CC	35%	48%
Sportmedizin	CC	71%	153%
Umweltmedizin und Toxikologie	CA	40%	47%
Urologie und Nephrologie	CC	30%	57%
Zahnheilkunde	CC	5%	10%
<i>Durchschnitt</i>		37%	70%

Qualität kardiologischer Literaturrecherchen

Obwohl weniger als ein Viertel der weltweit vorhandenen Zeitschriften von MEDLINE indiziert wird, stellt sie dennoch – nicht nur aufgrund ihrer ubiquitären und kostenfreien Zugänglichkeit, sondern wegen ihrer überragenden fachlichen Abdeckung – die erste Wahl für biomedizinische Literaturrecherchen dar. Durch zusätzliche Recherchen in EMBASE und weiteren Datenbanken kann die Suche erheblich vervollständigt werden, auch wenn nie 100% zu erreichen sind. Eine qualitative Aussage über die recherchierbaren Zeitschriften ist mit dem *Ulrich's* jedoch nicht möglich. Ein – wenn auch umstrittener Indikator für die Qualität von Zeitschriften – ist der Journal Impact Faktor des Science Citation Index. Dieser zeichnet ein ganz anderes Bild von der Vollständigkeit MEDLINE's: Von den 63 im SCI enthaltenen kardiologischen Titeln sind die besten 61 in MEDLINE nachgewiesen. Diese vereinigen 99,9% aller Zitierungen in der Kardiologie auf sich (EMBASE kommt auf 93,1%). Dies zeigt deutlich, dass die - aus angloamerikanischer Sicht - wichtigsten kardiologischen Zeitschriften in MEDLINE nachgewiesen werden.

Auch wenn eine Datenbank 100 Artikel zu Diabetes mellitus enthalten sollte, kann selbst durch die allerbeste Suchstrategie nur ein bestimmter Prozentsatz davon gefunden werden. Dieser Prozentsatz wird Recall genannt. Selbst Experten erreichen nur einen Recall von gut 50%, Ungeübte einen von durchschnittlich 28%, abhängig vom Retrievalsystem.^{56,7,8} Von den 25% der Literatur, die durchschnittlich in MEDLINE enthalten ist, können Datenbankneulinge wiederum also höchstens 28% wiederfinden, dies sind 7% der gesamten Literatur; mit anderen Worten: *Von vierzehn relevanten Artikeln weltweit kann über eine MEDLINE-Recherche lediglich ein einziger gefunden worden.* Der niedrige Recall von Gelegenheitsbenutzern macht es wenig effizient auch die allerletzte Datenbank, die vielleicht noch ein Artikelchen zum Thema enthalten könnte, mühsam zu durchsuchen. Weist MEDLINE eine vernünftige Abdeckung in diesem Gebiet auf, so kann eine genau so große Abschöpfung der Literatur stattfinden, wenn stattdessen durch intensive Schulungen der Recall verbessert wird. Auch alternative Wege der Literatursuche wie z.B. ein Verfolgen von Referenzen erhöhen die Vollständigkeit erheblich. Allerdings dürfte die unter diesen Umständen eigentlich sinnvollste Alternative zugleich am seltensten benutzt werden: die kostenpflichtige Recherche durch Bibliothekare z.B. einer Universitäts- oder Pharmabibliothek. Nur sehr geübte und geschulte Endnutzer erreichen ähnlich hohe Trefferzahlen wie diese professionellen Informationsvermittler.

Dr. Oliver Obst

⁵ W. Hersh et al. "Evaluation of SAPHIRE: an automated approach to indexing and retrieving medical literature" Proceedings of the Annual Symposium on Computer Application in Medical Care 808-12 (1991)

⁶ A.A. Verhoeven, E.J. Boerma, B. Meyboom-de Jong "Which literature retrieval method is most effective for GPs?" Family Practice 17(1):30-5 (2000)

⁷ K.A. McKibbin et al. "How good are clinical MEDLINE searches? A comparison of clinical end-user and librarian searches" Computer in Biomedical Research 23(6):583-593 (1990)

⁸ C.H. Fenichel: "Online searching: measures that discriminate among users with different types of experience" In: Journal of the American Society for Information Science 32:23-32 (1981)

Publication Bias

Wie kommen die Zeitschriften in MEDLINE? Der Direktor der National Library of Medicine entscheidet auf Antrag darüber, ob ein Titel Aufnahme in MEDLINE findet (<http://www.nlm.nih.gov/pubs/factsheets/jsel.html>). Die Chance, in diesen heiligen Gral des medizinischen Literaturnachweises aufgenommen zu werden, sinkt erheblich, wenn eine Zeitschrift nicht auf Englisch publiziert, aus einem Entwicklungsland stammt oder über exotische Themen und Therapieformen berichtet. Die WHO beklagte unlängst den angloamerikanischen Hochmut gegenüber wichtigen tropenmedizinischen Zeitschriften aus der dritten Welt. Diese Art von Blindheit führt zu einem Publication und Language Bias*: Viele gute Artikel und durchdachte klinische Studien bekommen dadurch keine Chance, gelesen, zitiert und somit Eingang in die wissenschaftliche Rezeption zu finden.

* Handbuch der Cochrane Collaboration, Kap. 5.0 [Online: http://www.cochrane.dk/cochrane/handbook/hbook50_Locating_and_selecting_studie.htm]