

# PubMed と MEDLINE とその他のデータベースの比較

小河 邦 雄\*

【抄録】 PubMed は医学関連文献を無料で検索できるため、企業の研究所においても使用する人は多い。しかし、その特徴を理解して使うべきであり、セキュリティの面にも注意する必要がある。今回、PubMed の主な情報源である MEDLINE を薬学関連の調査で定評のある EMBASE と比較することにより、収録雑誌や索引の違いによって、大きく検索される件数が異なることを確認した。また、PubMed の検索システムや位置づけなどを考察することにより、企業の研究部門として、どのように対応していくかを述べる。

【キーワード】 MEDLINE, EMBASE, 情報検索, エンドユーザー検索, 情報セキュリティ

## 1. はじめに

PubMed は MEDLINE を検索するための Web ツールとして高機能化しており、無料で操作性も良いため企業の研究者も多く利用している。しかし、インターネット特有のセキュリティ面での心配があるにもかかわらず、情報部門として利用状況をつかんでいないのが実態である。さらに、PubMed は使用や秘密保持に関する契約もなく、自己責任で使用されている。これらの現状を踏まえ、ユーザーが適切にデータベース(以下 DB)を使い分けられるように、その利用注意点について喚起する必要から、今回、PubMed とそのデータソースである MEDLINE について他の DB と比較し、その特徴を調べた。

## 2. 薬剤の検索

MEDLINE は、本来医学関連 DB であり、索引語付与数も決まっているため、文献中のすべての薬剤が索引されるわけではない。また、統制語シソーラスの MeSH も全体量がある程度決まっ

ているため、新しい薬剤名が次々に追加されることはない。このことから、薬剤の検索でどの程度の再現性があるかを他の DB と比較し、その原因を調べた。

### 2.1. 日本製薬情報協議会

今回の調査は、日本製薬情報協議会 (Pharmaceutical Information Association of Japan : 以下 PIAJ と略す) の関東支部勉強会において、参加各社 (13 社 17 名) で分担して調査を行ったものを基に、筆者が必要な追加調査を行いまとめた。以前の勉強会で治験薬の各段階での採録状況を見るために Derwent Drug File, MEDLINE, EMBASE, CA の比較を行い<sup>1)</sup>、EMBASE の薬剤検索における優位性が実証されていたので、MEDLINE で検索されずに EMBASE でのみ検索されるケースの原因について検証した。

### 2.2. 調査の方法

文献が DB で検索されるためには、①収録雑誌であること②論文が採録基準を満たしていること③タイトル、抄録、キーワードの中に検索語が含まれることの 3 点が必要である。これらの点について MEDLINE と EMBASE の検索結果を比較した。検索する薬剤は、Pharmaprojects で 2005 年以降に臨床試験に移行した 491 薬剤を STN の MEDLINE, EMBASE で一般名を用いて検索し、その件数差が 10 件以内の 63 薬剤を抽出し

\* Kunio OGAWA  
大正製薬(株)総合研究所研究システム部  
(日本製薬情報協議会会長)  
〒331-9530 さいたま市北区吉野町1-403  
E-mail: piai21@ybb.ne.jp

一般名	PubMed件数 EMBASE 独自論文 件数 Pharmaprojects 掲載時期と現在の Phaseの時期	PubMed Details部分の コピー	DB種 M:PubMed E:EMBASE	タイトル, 雑誌名 (改行する時は、Altキーを押しながらenterキーを押す)	コメント	MED 採録誌 照雑誌	MED 採録論 文	MED 索引の 有無
zanolimumab	全(重複なし)- 5 件 重複 - 1件 PubMed- 1件 EMBASEとのみ - 4 件 開業 2005 phase-3 1998 収載	zanolimumab[ All Fields]	E	Human antibodies from transgenic animals. Nature Biotechnology, (2005) Vol. 23, No. 9,	Med収録論文であるがインデキシングが無し EMBASEはControlled Term(CT)中に該当薬剤の記載	○	○	×
			E	Monoclonal antibody successes in the clinic. Nature Biotechnology, (2005) Vol. 23, No. 9,	Med収録論文であるがインデキシングが無し EMBASEはCT中に該当薬剤の記載本文を確認したところ 臨床ページにある薬剤のTable中に記載されているのみ	○	○	×
			E	Licensing highlights. IDrugs, (2005) Vol. 8, No. 12, pp. 1015-1019.	MEDLINE採録誌であるが該当論文の掲載なし EMBASEはCT中に該当薬剤の記載 Idrugsの中身を確認できなかったためMEDLINEの採用 非採用の他向は確認できず。	○	×	-
			E	Recommended International Nonproprietary Names: List 53. WHO Drug Information, (2005) Vol. 19, No. 1,	MEDLINE採録誌 EMBASEはCT中に該当薬剤の記載	×	-	-
			E	Gateways to Clinical Trials: June 2005. Methods and Findings in Experimental and Clinical Pharmacology, (2005) Vol. 27, No. 5, pp. 331-372.	タイトル欄記が Medlineと異なったため Dup Remコマンド では除けなかった	○	○	○
			M	Gateways to clinical trials. Methods and findings in experimental and clinical pharmacology, (2005 Jun) 27 (5) 331-72.				

図1 調査結果記入シート記入例

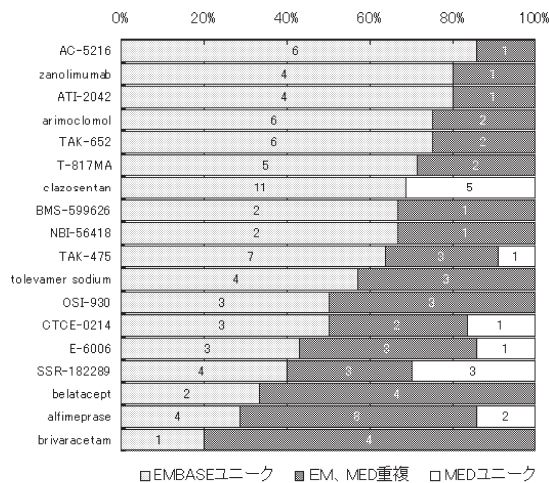


図2 MEDLINEとEMBASEでの治験薬検索数と重複状況

た。これは、臨床試験の薬剤の収録状況に興味を持たれることと、この段階の方が検索漏れの影響が大きいと考えられるためである。差を10件以内としたのは、解析する負担を調整するためである。次に検索ノイズの可能性のある単純な名称を除いて18薬剤に絞り、エンドユーザー検索を想定して、一般名のみでPubMedとSTNのMEDLINE, EMBASEを検索した。そしてMEDLINE検索分を除いたEMBASEデータを出力し、所定の表にしたがって結果をまとめ、考察を記入した(図1)。

### 2.3. 臨床試験移行治験薬の検索結果

STNの検索の結果、個々の薬剤において、

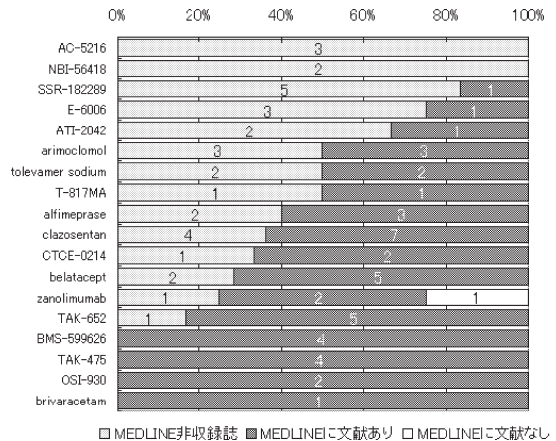


図3 EMBASEのみで検索された文献の内容解析

EMBASEだけで検索されたユニークな文献数が半分を超えるものが15薬剤(83%)あり、EMBASEの優位性が確認された(図2)。一方、MEDLINEだけで検索された文献も6薬剤(33%)あったが、EMBASEの独自性に比べMEDLINEの独自性の割合は少なかった。また、EMBASEとMEDLINEの両方で検索された文献の占める割合は、平均35%であり、収録雑誌の重複率4割より低かった<sup>2)</sup>。この結果を単純化すると、ある薬剤について10件の文献がある場合、EMBASE独自が6件、EMBASEとMEDLINEの重複が3件、MEDLINE独自が1件という関係になる。さらに、EMBASE独自の6件の文献のうち、2件がMEDLINEの収録誌ではなく、残りの4件はMEDLINEに採録されては

いるが、索引がないため検索できなかったことがわかった (図 3)。非収録誌では、Drugs of the Future (Prous Science) などが目立った。これは新規薬剤の研究開発情報誌で、基礎から臨床までの治験薬調査に有効な雑誌であるが、MEDLINE では収録していない。また、PubMed 独自の論文が採録された理由は、Automatic Term Mapping によって化合物の IUPAC 名が付与されて検索されたものや、巻号頁の付いていない Pre-MEDLINE 段階の Article in Press であった。これらの結果を踏まえて、さらに多くのサンプルから傾向を確認するために、市販薬について調査を行った。

#### 2.4. 市販薬での件数の比較

市販薬は Pharmaprojects を使用し、Novelty が Leading Compound、特許優先権が 1990 年以降、さらに Market Rating が US\$2,001 million 以上の 52 薬剤をピックアップした。これらの一般名から、ノイズを含みそうな短い名称を除いて検索リストを作成した。MEDLINE で予備検索

を行い、化学物質名フィールド CN に索引がある薬剤で、件数的に 100 件から 1,000 件のものを選択し、最終的に残った 19 薬剤を比較対象とした。

#### 2.5. 複数のデータベースでの比較の結果

これらの薬剤について MEDLINE, EMBASE, Cplus で一般名検索を行った (表 1)。ここでは、化学、生化学、医薬等の分野を含む 9,500 誌を収録対象とし、2,400 万件の論文や特許を採録している CA を加えた。CA は、本来、化合物を検索する場合は CAS 番号で検索すべきであるが、同義語の収録が多く、比較のために加えた。しかしながら、臨床論文が多い市販薬では、CA の採録件数は低かった。MEDLINE は CN フィールドにあるキーワードを使用したので比較的妥当な件数だと思われるが、EMBASE については MEDLINE よりも件数が少ないものがあり、最適なキーワードでない可能性がある。しかし、全般的に、EMBASE は MEDLINE の 2~3 倍検索され、全体では 2.4 倍と市販薬でも

表 1 3種のDBによる市販薬の一般名検索件数

No.	Drug Name	MEDLINE	EMBASE	Cplus
1	TERIPARATIDE	876	182	116
2	BOSENTAN	961	1,804	795
3	MONTELUKAST	805	2,252	603
4	TIROFIBAN	706	2,317	565
5	EPTIFIBATIDE	642	2,040	444
6	VERTEPORFIN	584	466	245
7	BORTEZOMIB	600	1,282	468
8	ZANAMIVIR	470	1,265	335
9	OSELTAMIVIR	313	1,223	276
10	EZETIMIBE	421	1,025	353
11	ERLOTINIB	434	1,738	344
12	GLIMEPIRIDE	355	1,202	587
13	REPAGLINIDE	327	1,080	448
14	TADALAFIL	325	855	270
15	VARDENAFIL	306	813	274
16	GEMIFLOXACIN	277	709	366
17	BEXAROTENE	186	436	103
18	PARECOXIB	134	545	307
合計		8,908	21,670	7,002

太字：最も多いもの

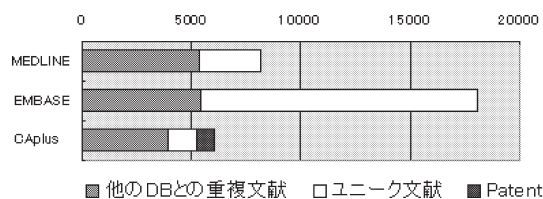


図4 市販薬検索結果全体の重複状況

EMBASEの優位性が確認できた。これらの結果から、市販薬もMEDLINEだけで検索を行うことは、多くの文献を見逃す結果となることが確認できた。また、これらの検索結果全体の重複状況を調べたのが図4である。他のDBと重複する文献は5,000件程度あるが、DB独自のユニークな文献は、EMBASEが圧倒的に多く約1.3万件であった。しかし、MEDLINEも独自の文献を3割ほど持っているため、EMBASEの検索だけでも網羅性を確保できない。ただし、STNは電子ジャーナルと印刷体の間では重複除去できないことがあるので、実際のユニークな文献はもう少し少なくなる可能性がある。また、市販薬については、毒性、副作用情報やEBMのシステムティックレビューなどで漏れの少ない検索をしたい場合もあるが、件数が多い場合は逆に目的の文献に絞り込む機能が重要となる。

## 2.6. 市販薬がMEDLINEで検索できなかった理由

次に、EMBASEのみで検索された文献がMEDLINEの収録誌かどうかを調べた。収録誌であれば、索引に薬剤名が付与されていないことが検索漏れの原因と考えられる。MEDLINEの2006年収録誌リストを使用し<sup>3)</sup>、上位35件の中で、MEDLINE収録誌は23件(66%)と多く、薬剤名が索引されていないことが原因とわかった(表2)。

## 3. 薬理メカニズムでの比較

適応疾病や薬理作用などのキーワードは比較的規則性を持って詳細に索引されているが、薬理メカニズムは新しい概念も多く、統制されたキーワードがなかったり、自由語での検索も略語やキーワード表記のバリエーションが多かったりなど、

検索式を組み立てることが難しい。しかし、創薬の研究においては、活性の作用点である受容体タンパク質や酵素の発見とその阻害剤などのリガンドの情報は非常に重要である。新規の創薬テーマ企画に関する調査においても、薬理メカニズム関連の調査は重要で、それらを調査するためのデータベースとしてMEDLINEが適当であるか検証した。

### 3.1. 調査方法と結果

Pharmaprojectsで前臨床段階の薬剤3,867件の中から薬理メカニズム351種を抽出し、その中から、その他の市販、臨床中、中止などの段階にもあるメカニズムを除いた105件の集合を作成した。この集合は、比較的新規な薬理メカニズムと考えられ、これらがDBでどのように検索されるかを検証した。検索式としては、Pharmaprojectsの薬理メカニズムが一般的記載で文献中の記載に近いと想定し、そのまま検索した場合と、近接演算子を使用して語順の違いを含めた漏れの少ない検索式を組み立てて検索した場合との両方で調べた。105件のメカニズムをSTNで予備検索して、件数が非常に多い11件と件数の少ない64件を除いて30件とし、これらをMEDLINE, EMBASE, C.Aplusで本検索した(表3)。以前の勉強会でリガンドを薬理作用の対象として調査した際、CAでは独自の文献が多く見られたので<sup>1)</sup>、ここでもCAを加えた。近接演算子を使用した時の個々の件数の増加率は、MEDLINE, EMBASE, C.Aplusでそのままの検索と比べてそれぞれ平均29.0, 26.5, 16.3倍であり、近接演算子の使用が有効であったが、これらの中には主題と直接関係のないノイズ文献もある程度含まれてしまう。3種類のDBの中で件数が多いものを網掛けにしたが、ほとんどがCAで、次がEMBASEで、MEDLINEはどれも少なかった。論文中でのメカニズムの記載は、…Inhibitorなどの記述がInhibitor of…などと倒置されたり、途中の語句の順番が前後したりすることも多く、近接演算子を使えないとかなりの漏れが生じる。STN MEDLINEでは近接演算子ができるが、PubMedでは使えないので、検索結果が少なくなる可能性がある。さらに、PubMedには



表2 EMBASE のみで検索された論文の収録雑誌 (市販薬)

	雑誌名	論文数
1	Deutsche Apotheker Zeitung	217
2	Pharmaceutical Journal	145
3	<b>Expert Opinion on Investigational Drugs</b>	<b>100</b>
4	<b>Nature Reviews Drug Discovery</b>	<b>100</b>
5	<b>New England Journal of Medicine</b>	<b>99</b>
6	Pharmazeutische Zeitung	97
7	<b>Expert Opinion on Pharmacotherapy</b>	<b>93</b>
8	<b>IDrugs</b>	<b>89</b>
9	<b>American Heart Journal</b>	<b>85</b>
10	Drugs of the Future	85
11	<b>American Journal of Cardiology</b>	<b>83</b>
12	<b>Seminars in Oncology</b>	<b>82</b>
13	Pharmazeutische Industrie	80
14	<b>Journal of the American College of Cardiology</b>	<b>80</b>
15	<b>Journal of Invasive Cardiology</b>	<b>78</b>
16	<b>Lancet</b>	<b>77</b>
17	<b>Circulation</b>	<b>76</b>
18	Hospital Pharmacy	75
19	Pharmaceutisch Weekblad	75
20	<b>Chest</b>	<b>66</b>
21	<b>Journal of Allergy and Clinical Immunology</b>	<b>66</b>
22	<b>European Heart Journal</b>	<b>66</b>
23	Formulary	62
24	British Journal of Cardiology	59
25	<b>MMW-Fortschritte der Medizin</b>	<b>58</b>
26	<b>Drugs(e)</b>	<b>56</b>
27	<b>American Journal of Respiratory and Critical Care Medicine</b>	<b>54</b>
28	Expert Opinion on Therapeutic Patents	53
29	<b>Current Pharmaceutical Design</b>	<b>52</b>
30	<b>Clinical Lung Cancer</b>	<b>52</b>
31	Drugs and Therapy Perspectives(e)	51
32	<b>Journal of Antimicrobial Chemotherapy</b>	<b>49</b>
33	<b>Annals of Oncology</b>	<b>49</b>
34	European Heart Journal, Supplement	49
35	<b>Journal of the American Medical Association</b>	<b>48</b>

太字：MEDLINE 収録対象誌

MeSH の Pharmacologic Action で下位に該当する薬剤を一括して検索する機能もないため、同じ MEDLINE をデータソースとしていても、STN MEDLINE と比較して、この面でも不利である。

### 3. 2. 薬理メカニズム検索と PubMed 収録雑誌

MEDLINE と EMBASE を近接演算子を使用して検索した全 6,620 件の中で、MEDLINE のみが 893 件、重複が 3,518 件、EMBASE のみが 2,209 件であった。この EMBASE のみで検索さ

表3 3種のDBによる薬理メカニズム検索件数

No.	検索式	A:そのまま検索			B:近接演算子(6A)を使用		
		MED	EM	CA	MED	EM	CA
1	THROMBOXANE A2 ANTAGONIST?	103	107	<b>375</b>	837	815	<b>1,182</b>
2	PHOSPHODIESTERASE V INHIBITOR?	22	<b>960</b>	173	123	<b>1,063</b>	401
3	PROSTAGLANDIN AGONIST?	16	17	<b>64</b>	752	732	<b>956</b>
4	SODIUM ANTAGONIST?	8	11	<b>17</b>	584	601	<b>978</b>
5	HISTAMINE RELEASE INHIBITOR?	26	<b>136</b>	100	483	590	<b>796</b>
6	POTASSIUM ANTAGONIST?	5	5	<b>28</b>	407	435	<b>786</b>
7	GASTRIN INHIBITOR?	9	11	<b>55</b>	434	439	<b>582</b>
8	SELECTIN ANTAGONIST?	56	<b>136</b>	126	98	175	<b>218</b>
9	TESTOSTERONE AGONIST?	1	1	<b>9</b>	201	189	<b>349</b>
10	ENDOTHELIN B RECEPTOR AGONIST?	25	<b>208</b>	24	98	<b>295</b>	72
11	GLUTAMATE 3 RECEPTOR ANTAGONIST?	0	0	0	155	159	<b>191</b>
12	INTERLEUKIN-10 ANTAGONIST?	0	1	<b>11</b>	124	124	<b>206</b>
13	FACTOR XIA INHIBITOR?	5	4	<b>31</b>	66	59	<b>111</b>
14	INTERFERON AGONIST?	0	0	<b>11</b>	38	34	<b>176</b>
15	TOLL-LIKE RECEPTOR ANTAGONIST?	2	7	7	17	26	<b>46</b>
16	UROTENSIN II RECEPTOR AGONIST?	2	<b>6</b>	4	8	9	<b>15</b>
17	PHOSPHOFRUCTOKINASE 2 INHIBITOR?	0	0	<b>1</b>	12	6	<b>14</b>
18	PANTOTHENATE KINASE INHIBITOR?	1	2	<b>5</b>	4	2	<b>9</b>
19	SERINE RACEMASE INHIBITOR?	2	2	2	4	3	<b>8</b>
20	BMX TYROSINE KINASE INHIBITOR?	0	0	<b>4</b>	0	0	<b>11</b>
21	LEUKOTRIENE D4 ANTAGONIST?	1	2	2	1	<b>5</b>	4
22	TIE-1 TYROSINE KINASE INHIBITOR?	0	0	0	0	0	<b>15</b>
23	PROLACTIN RELEASE STIMULANT?	0	0	0	5	5	3
24	EPHRIN B2 INHIBITOR?	0	0	0	3	3	<b>6</b>
25	FIBROBLAST GROWTH FACTOR RECEPTOR 2 ANTAGONIST?	0	0	0	1	3	8
26	HYALURONAN SYNTHASE INHIBITOR?	1	1	<b>2</b>	2	1	<b>5</b>
27	HYALURONIC ACID AGONIST?	0	0	0	1	1	<b>10</b>
28	INTERLEUKIN 22 RECEPTOR ANTAGONIST?	0	0	0	2	2	<b>7</b>
29	INTERLEUKIN 31 ANTAGONIST?	0	0	0	4	3	<b>4</b>
30	PROTEASE-ACTIVATED RECEPTOR-2 ANTAGONIST?	1	0	<b>2</b>	2	1	<b>5</b>

太字:最も多いもの

れた論文を雑誌別に集計した(表4)。上位35件中29件がMEDLINE収録誌で、抄録は著者抄録でEMBASEと同じと考え、薬剤名の索引と同様、メカニズム索引もEMBASEの方が豊富と推測された。

#### 4. MEDLINEの雑誌別論文採録数

ここまでの調査で検索数の差は索引が主な原因と思われたが、論文単位で未採録の場合もあり、データベースによる採録の差を確かめてみた。ここでは、収録対象誌の選定基準に引用分析を含み、その結果によって科学技術、生物医学分野の

表4 EMBASE のみで検索された論文の収録雑誌 (薬理メカニズム)

No.	雑誌名	論文数
1	<b>The American Journal of Physiology</b>	53
2	<b>British Journal of Pharmacology</b>	50
3	<b>International Journal of Impotence Research</b>	46
4	<b>Journal of Cardiovascular Pharmacology</b>	43
5	<b>European Journal of Pharmacology</b>	41
6	<b>Journal of Sexual Medicine</b>	39
7	<b>American Journal of Physiology—Heart and Circulatory Physiology</b>	31
8	<b>European Urology</b>	29
9	Drugs of the Future	25
10	<b>Bioorganic and Medicinal Chemistry Letters</b>	25
11	<b>Journal of Urology</b>	24
12	<b>BJU International</b>	22
13	<b>American Journal of Cardiology</b>	18
14	<b>Agents and Actions.</b>	17
15	<b>Journal of Pharmacology and Experimental Therapeutics</b>	16
16	<b>American Journal of Physiology—Gastrointestinal and Liver Physiology</b>	16
17	<b>Hypertension</b>	16
18	<b>American Journal of Physiology—Regulatory, Integrative and Comparative Physiology</b>	16
19	<b>Journal of Sexual Medicine (ep)</b>	15
20	<b>Urology</b>	14
21	<b>Urologe-Ausgabe A</b>	13
22	<b>Circulation</b>	12
23	Drug Development Research	12
24	<b>International Journal of Clinical Practice</b>	12
25	<b>Journal of Medicinal Chemistry</b>	11
26	Journal of Labelled Compounds and Radiopharmaceuticals	11
27	Japanese Pharmacology and Therapeutics	11
28	Expert Opinion on Therapeutic Patents	11
29	<b>Urologic Clinics of North America</b>	10
30	<b>American Journal of Physiology—Renal Physiology</b>	10
31	<b>Journal of the American College of Cardiology</b>	10
32	<b>Nature Clinical Practice Urology</b>	10
33	<b>Life Sciences</b>	9
34	Tetrahedron Letters	9
35	<b>European Journal of Medicinal Chemistry</b>	9

太字：MEDLINE 収録対象誌

最も重要な雑誌約 4,500 誌を収録する SciSearch も加えた。比較した雑誌は、SciSearch の引用分析から作成される JCR (Journal Citation Reports) の IF (インパクトファクター) 上位 30 誌である (表 5)。ISSN (印刷体と電子体) で 2005 年発行の雑誌を検索したが、著名な雑誌

においても、DB によってかなり採録数が違った。採録数が多いのは CA, SciSearch で、MEDLINE, EMBASE は少なかったが、これは同じ雑誌であっても、MEDLINE や EMBASE は医学を中心に収録しているためと推測される。IF の上位誌はレビュー誌が多く、また、コラム

表5 4種のDBによるIF上位雑誌の論文収録数

No.	MED	EM	CA	SCI	JOURNAL NAME	発行国	IF
1	29	29	<b>33</b>	29	ANNUAL REVIEW OF IMMUNOLOGY	USA	52.4
2	22	33	0	<b>48</b>	CA-A CANCER JOURNAL FOR CLINICIANS	USA	44.5
3	1,467	1,202	<b>2,100</b>	1,798	NEW ENGLAND JOURNAL OF MEDICINE	USA	38.6
4	83	156	<b>256</b>	211	NATURE REVIEWS CANCER	UK	36.6
5	33	<b>34</b>	30	<b>34</b>	PHYSIOLOGICAL REVIEWS	USA	33.9
6	82	137	<b>235</b>	194	NATURE REVIEWS MOLECULAR CELL BIOLOGY	UK	33.2
7	0	0	27	<b>33</b>	REVIEWS OF MODERN PHYSICS	USA	32.8
8	78	131	<b>228</b>	187	NATURE REVIEWS IMMUNOLOGY	UK	32.7
9	2,534	2,407	<b>3,546</b>	2,808	NATURE	UK	32.2
10	2,409	2,314	<b>4,373</b>	2,698	SCIENCE	USA	31.9
11	28	28	<b>30</b>	28	ANNUAL REVIEW OF BIOCHEMISTRY	USA	31.5
12	383	398	373	<b>456</b>	NATURE MEDICINE	USA	31.2
13	449	462	<b>507</b>	462	CELL	USA	28.4
14	209	224	222	272	NATURE IMMUNOLOGY	USA	27.6
15	1,028	920	<b>2,262</b>	1,387	JAMA-JOURNAL OF THE AMERICAN MEDICAL ASSOCIATION	USA	24.8
16	284	306	271	<b>350</b>	NATURE GENETICS	USA	24.7
17	19	19	<b>22</b>	19	ANNUAL REVIEW OF NEUROSCIENCE	USA	23.1
18	29	<b>30</b>	28	<b>30</b>	PHARMACOLOGICAL REVIEWS	USA	22.8
19	399	345	<b>513</b>	471	NATURE BIOTECHNOLOGY	USA	22.4
20	1,581	1,407	160	<b>1,738</b>	LANCET	UK	21.7
21	0	0	<b>14</b>	12	SURFACE SCIENCE REPORTS	NL	21.4
22	84	121	<b>261</b>	205	NATURE REVIEWS GENETICS	UK	21.3
23	78	122	<b>244</b>	207	NATURE REVIEWS NEUROSCIENCE	UK	21.2
24	29	29	<b>33</b>	29	ANNUAL REVIEW OF PHARMACOLGY AND TOXICOLOGY	USA	21.1
25	223	231	<b>264</b>	239	NATURE CELL BIOLOGY	UK	20.6
26	135	0	<b>140</b>	<b>140</b>	CHEMICAL REVIEWS	USA	20.2
27	169	<b>339</b>	306	258	NATURE REVIEWS DRUG DISCOVERY	UK	19.6
28	0	0	9	<b>21</b>	ANNUAL REIEW OF ASTRONOMY AND ASTROPHYSICS	USA	18.8
29	40	43	40	<b>55</b>	ENDOCRINE REVIEWS	USA	18.8
30	121	127	<b>150</b>	128	CANCER CELL	USA	18.1

太字：最も多いもの

のような記事もあり、DBによって論文採録基準が大きく異なることがわかった。ただ、MEDLINEにまったく収録されていない雑誌は3誌で、いかにIF上位に医学やライフサイエンス関連誌が多いかがわかる。特に30誌中12誌がnature系雑誌であり、ブランド力を感じた。結論として、雑誌が収録対象誌であっても、採録さ

れる論文の数はDBによって大きく異なり、それが検索数の差の原因の1つと考えられた。

##### 5. MEDLINEの収録対象雑誌のISSN番号

2006年版リストにある4,884誌のISSN表示は冊子・電子版の併記が1,220誌(25%)、冊子3,564誌(73%)、電子版100誌(2%)である



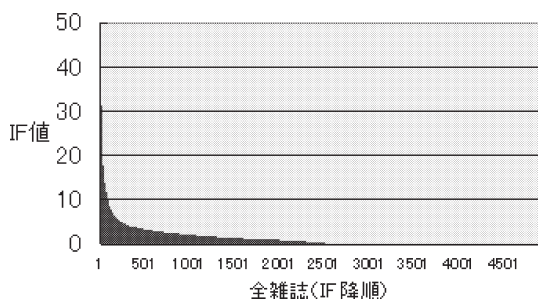


図5 MEDLINE 収録雑誌と IF 値

が、STN の MEDLINE の書誌には一方の ISSN のみ付与されている。例えば nature は、2002 年までは冊子の ISSN で、2003 年は冊子か電子版のどちらか一方で、2004 年からは電子版のものが付与されていた。ISSN で検索する場合の検索漏れやマルチファイル検索での重複除去には注意が必要である。MEDLINE の収録対象全雑誌と IF の関係が図 5 である。2,500 件 (51%) と約半分の雑誌が JCR の収録誌であるが、IF はその収録されている雑誌分しかわからない。

## 6. MEDLINE で収録していない高 IF 雑誌

MEDLINE の収録対象誌はメジャーなものに関しては、地域に関係なく収録されているようである。表 6 は EMBASE の収録対象誌であるが MEDLINE では収録対象でない雑誌の中で、IF 値が 1 以上のもの 75 誌である。JCR 全体の中で IF 値が 1 以上の雑誌は 49% あり、この 75 誌は半分以上のランクにある著名な雑誌といえる。これらを見ていくと、62 件 (83%) は米国以外の雑誌であるが、内容的には、化学系など医学以外の雑誌が多い。しかし、創薬の構造活性相関に関係する著名な雑誌も含まれているので、PubMed を検索する場合は、著明な雑誌でも収録されていない雑誌がある点も考慮して検索する必要がある。

## 7. PubMed の検索エンジン

PubMed の利用は、1 日 200 万件以上あるとのことであるが、検索スピードは非常に速く、瞬時に検索件数が表示される。これは、検索エンジン

をそのように設計しているからであるが、そのため、文書中の単語の位置情報は保持していないのか、フレーズ検索に有効な近接演算子が使えない。そのため、単語では全文を検索できるが、フレーズに関しては、登録されたものしか検索できない。フレーズが登録されるには、そのフレーズが複数回用いられている必要があるため、比較的新しい複数の語句で表される用語は登録されていない場合が多い。そのため、複数の語句を and 検索しなければならないので、多量のノイズを含むこともあり、あまり適当とは言えない。情報担当者にとっては、近接演算子が使えないのは、ある意味致命的ではあるが、その分、スピードを重視したと理解できる。しかし、全文が検索できる場合は、エンドユーザーがたまたま入力したフレーズでヒットし、それで検索を止めてしまう方が問題な場合もあるので、目的によって使い分けるべきであろう。

## 8. PubMed の位置づけ

PubMed は、NML (米国国立医学図書館) 内の NCBI (国立バイオテクノロジー情報センター) が作成している Entrez (アントレ) という統合型分子生物学 DB の一部として公開されている。MEDLINE のデータを中心に作られているが、本来の目的は、Entrez 中にある、核酸 DB、タンパク DB などのファクトデータと論文を結ぶ「要」の役割にある。もちろん、米国の税金で作成している DB なので、医学情報を広く公開するという役割も持っているが、個人での利用を前提としており、企業の組織的使用は想定外と言われても仕方がない。また、無料で多くの人がアクセスするために、PubMed 収録雑誌としてリストアップされようとする出版社や学会からのマーケティング対象ともなり、情報のブランドカタログ的な側面も見られる。また、大量のダウンロードを行うとサーバーに負担がかかるため、その利用者のサーバーからは利用禁止になることもある。実際に今年の 6 月には京都大学で一時使用できなくなった。こうなると組織のコンプライアンスの面からも、情報部門はその利用に関して何らかの対応をとる必要がある。

表6 EMBASE にしか収録されていない高 IF 雑誌リスト

	雑誌名	IP	出版国
1	Biochimica et Biophysica Acta—Reviews on Cancer	16.12	NETHERLANDS
2	European Journal of Cancer, Supplement	9.02	ENGLAND
3	Current Opinion in Coloid and Interface Science	5.27	ENGLAND
4	Biochimica et Biophysica Acta—Molecular and Cell Biology of Lipids	5.00	NETHERLANDS
5	TrAC—Trends in Analytical Chemistry	3.89	NETHERLANDS
6	Mutation Reseach—Reviews in Mutation Research	3.67	NETHERLANDS
7	<b>Journal of Biomedical Materials Research—Part A</b>	<b>3.65</b>	<b>UNITED STATES</b>
8	<b>Journal of Biomedical Materials Research—Part B Applied Biomaterials</b>	<b>3.65</b>	<b>UNITED STATES</b>
9	<b>Critical Reviews in Oral Biology and Medicine</b>	<b>3.64</b>	<b>UNITED STATES</b>
10	Biochimica et Biophysica Acta—Bioenergetics	3.50	NETHERLANDS
11	Biochimica et Biophysica Acta—Molecular Cell Research	3.48	NETHERLANDS
12	Biochimica et Biophysica Acta—Biomembranes	3.44	NETHERLANDS
13	Biochimica et Biophysica Acta—General Subjects	3.37	NETHERLANDS
14	Biochimica et Biophysica Acta—Molecular Basis of Disease	3.05	NETHERLANDS
15	Current Organic Chemistry	2.78	NETHERLANDS
16	Cognitive Neuropsychology	2.75	ENGLAND
17	Synlett	2.74	GERMANY
18	New Journal of Chemistry	2.74	ENGLAND
19	European Urology, Supplements	2.70	NETHERLANDS
20	Tetrahedron	2.64	ENGLAND
21	Analytica Chimica Acta	2.59	NETHERLANDS
22	Atmospheric Environment	2.56	ENGLAND
23	Tetrahedron Letters	2.48	ENGLAND
24	Tetrahedron Asymmetry	2.39	ENGLAND
25	Synthesis	2.20	GERMANY
26	Psychiatry Research—Neuroimaging	2.15	NETHERLANDS
27	Biochimica et Biophysica Acta—Proteins and Proteomics	2.11	NETHERLANDS
28	Journal of Membrane Science	2.11	NETHERLANDS
29	Journal of Atmospheric Chemistry	2.05	NETHERLANDS
30	Biochimica et Biophysica Acta—Gene Structure and Expression	2.05	NETHERLANDS
31	Mutation Research—Genetic Toxicology and Environmental Mutagenesis	2.02	NETHERLANDS
32	Chemometrics and Intelligent Laboratory Systems	1.90	NETHERLADS
33	QSAR and Combinatorial Science	1.88	GERMANY
34	Journal of Aerosol Science	1.86	ENGLAND
35	Environmental Conservation	1.84	ENGLAND
36	Helvetica Chimica Acta	1.83	SWITZERLAND
37	<b>Fluoride</b>	<b>1.78</b>	<b>UNITED STATES</b>
38	<b>Enzyme and Microbial Technology</b>	<b>1.76</b>	<b>UNITED STATES</b>
39	<b>Clinical and Diagnostic Laboratory Immunology</b>	<b>1.72</b>	<b>UNITED STATES</b>
40	Molecular Brain Research	1.71	NETHERLANDS
41	<b>Critical Reviews in Environmental Science and Technology</b>	<b>1.68</b>	<b>UNITED STATES</b>
42	Journal of Microbiology and Biotechnology	1.66	SOUTH KOREA
43	Biochemical Engineering Journal	1.62	SWITZERLAND

表 6 EMBASE にしか収録されていない高 IF 雑誌リスト (続き)

	雑誌名	IP	出版国
44	Journal of Molecular Catalysis B: Enzymatic	1.55	NETHERLANDS
45	Food Chemistry	1.54	ENGLAND
46	Colloids and Surfaces A: Physicochemical and Engineering Aspects	1.51	NETHERLANDS
47	Expert Opinion on Therapeutic Patents	1.46	ENGLAND
48	Bulletin of the Chemical Society of Japan	1.45	JAPAN
<b>49</b>	<b>Aerosol Science and Technology</b>	<b>1.43</b>	<b>UNITED STATES</b>
<b>50</b>	<b>Journal of Neurolinguistics</b>	<b>1.43</b>	<b>UNITED STATES</b>
51	Neuropsychiatre	1.41	GERMANY
<b>52</b>	<b>Pediatric Exercise Science</b>	<b>1.38</b>	<b>UNITED STATES</b>
<b>53</b>	<b>Journal of the Air and Waste Management Association</b>	<b>1.36</b>	<b>UNITED STATES</b>
54	Basic and Clinical Pharmacology and Toxicology, Supplement	1.34	DENMARK
55	European Heart Journal, Supplement	1.32	ENGLAND
56	Nutrition Research Reviews	1.31	ENGLAND
57	Advances in Environmental Research	1.27	ENGLAND
58	Arthropod Structure and Development	1.27	ENGLAND
59	Environmental Toxicology and Pharmacology	1.25	NETHERLANDS
60	Journal of Biological Physics	1.23	NETHERLANDS
61	Hepatology Research	1.17	IRELAND
62	Clinical Neuropsychologist	1.16	NETHERLANDS
63	Chromatographia	1.15	GERMANY
64	Disability and Rehabilitation(ep)	1.14	ENGLAND
65	Colloid and Polymer Science	1.11	GERMANY
66	Phlebologie	1.10	GERMANY
<b>67</b>	<b>Clinical Neuroscience Research</b>	<b>1.08</b>	<b>UNITED STATES</b>
68	International Journal of Life Cycle Assessment	1.07	GERMANY
69	Water, Air, and Soil Pollution	1.06	NETHERLANDS
70	Canadian Journal of Chemistry	1.06	CANADA
71	Biocatalysis and Biotransformation	1.05	ENGLAND
72	Journal of Labelled Compounds and Radiopharmaceuticals	1.05	ENGLAND
73	Comparative and Functional Genomics	1.04	ENGLAND
<b>74</b>	<b>Combustion and Flame</b>	<b>1.02</b>	<b>UNITED STATES</b>
75	Aphasiology	1.01	ENGLAND

太字：米国出版

## 9. PubMed のセキュリティ

インターネットは公道であり、ネットワーク途中での一般的な情報漏洩の可能性は考えられる。また、プライバシーポリシーでは、利用者のドメインや PC の OS、アクセス記録やクッキーの利用、そしてその分析のための第三者機関への情報提供も示唆されている<sup>4)</sup>。確かに、論文の引用情

報が研究トレンドの調査に使われるように、検索結果からフルテキストへのアクセス情報などは、注目される研究のデータとして利用可能である。このプライバシーポリシーは米国内を対象としたものと思われ、海外については紳士協定的に期待するものであり、内容の修正があっても当然連絡はなく、常にモニターすることも難しい。無料で契約がないのはありがたいが、半面、契約ができ

ないことでもある。このような点から、戦略に係る調査や特許性がからむ場合は、慎重に使うべきである。

## 10. PubMed とコンピュータ解析

膨大な医学関連情報を含む MEDLINE のデータについては、それ自体をファクトデータとしてゲノムとその機能、疾病との関連性を探索するためのバイオインフォマティクス分野での使用が進んでいる。自然言語で書かれた抄録から、キーワードの揺れを吸収しながら特定の概念に集約していく作業をコンピュータで行うものであるが、これは正にインデキシングの作業に似ている。実際に、MeSH のインデキシングも本文、抄録等をコンピュータで解析し、同義語、関連語から、適切な MeSH タームのリストを提示する索引支援システムを使用している<sup>9)</sup>。このため、論文執筆の際は、本文はもちろん抄録においても、それらのコンピュータ処理を想定した記述を行うことがその後大きな影響を与えることがわかる。特に近年の検索システムの開発は、従来の図書館情報的なアプローチとは異なる研究分野から、情報資源の多様な活用を目的として、言語学、コーパス、オントロジー、タクソノミー、フォークソノミーなどの概念を組み合わせ、目的とする情報や知識を検索する方法が試みられている<sup>9)</sup>。これらは従来、特定分野に特化して検索技術を深めてきた専門サーチャーをシステム化しようという部分も感じられて興味深い。しかし、その解析コンテンツが、例えば MEDLINE だけでいいのかという点に関しては、長年それらと格闘してきた検索担当者としても気になる部分であろう。

## 11. おわりに

従来、企業の情報担当者によるエンドユーザー

教育は、自分たちが用意した情報源について、その使い方等を利用者に教えることで済んでいた。現代は、インターネット情報の充実で、我々の知らない所で社内ユーザーはいろいろな情報源にアクセスしている。PubMed についても、今回の調査を行う前は知識が少なかったが、随分勉強になった。PIAJ では、関東・関西の両支部でこれから新しい情報源についての評価や講演会などの勉強会を積極的に行っている。関心のある企業情報担当者のさらなる参加を期待したい<sup>7)</sup>。

## 参考文献

- 1) 野田利章. 医学薬学データベースの比較. 薬学図書館. 46 (3), 2001, 215-221.
- 2) エルゼビアの EMBASE.com の説明画面 (EMBASE と MEDLINE を同時に検索できる). <[http://japan.elsevier.com/products/embase\\_com/index.html](http://japan.elsevier.com/products/embase_com/index.html)>, (参照 2006-08-07).
- 3) List of Journals Indexed for MEDLINE, 2006. <[http://www.nlm.nih.gov/tsd/serials/terms\\_cond.html](http://www.nlm.nih.gov/tsd/serials/terms_cond.html)>, (accessed 2006-08-07).
- 4) NLM Privacy Policy. <<http://www.nlm.nih.gov/privacy.html>>, (accessed 2006-08-07).
- 5) 阿部信一ほか. 図解 PubMed の使い方 インターネットで医学文献を探す 第3版. 東京, 日本医学図書館協会, 2003, 91 p. (ISBN 4-931222-12-9)
- 6) Morville, P., 浅野紀予訳. アンビエント・ファインダビリティ. 東京, オーム社, 2006, 241 p. (ISBN 4-87311-283-4)
- 7) 日本製薬情報協議会 関連サイト <<http://piaj.sub.jp/ring/>>, (参照 2006-08-07).

日本製薬情報協議会 勉強会検索参加者：  
塩野邦彦：グラクソ・スミスクライン(株)，石川靖子，木村美佐子：三共(株)，奥田葉子，浜崎泰嗣，前田清人：ゼリア新薬(株)，光吉久美子：(株)ツムラ，松浦智佳子：協和発酵工業(株)，本田直樹：(株)三和化学研究所，加藤和夫：持田製薬(株)，作山幸由，中野敦子：大塚製薬(株)，中村規子：大日本住友製薬(株)，三輪 保：第一製薬(株)，村上敏信：帝人ファーマ(株)，関 英雄：田辺製薬(株)，中村文胤：日本新薬(株) (順不同)

(原稿受け：2006.8.10)