

MEDLINEのCD-ROM版におけるフリーキーワード検索の検討

菅野育子 草田直子

愛知淑徳大学文学部図書館情報学科

MEDLINEのシソーラスであるMeSHの用語を用いるThesaurusコマンドとフリーキーワードを用いるFindコマンドで、公衆衛生学の検索事例をCD-ROM版で検索し、その検索結果を比較検討した。検索事例5件の作成、フリーキーワードの選定及び検索結果の評価は当該分野の研究者に依頼し、MeSH用語の選定は医学図書館員に依頼した。検索事例のうち、「歯科治療によるHIV感染」では、Findコマンド検索のみで得られた適合文献が67件もあった。それらの文献を分析した結果、Findコマンドによって、MeSHの関連参照の不備を補う広い概念が検索されていることが明らかになった。今後は、MeSH用語による検索とフリーキーワード検索の併用について検討が必要である。

A study of searching MEDLINE on CD-ROM by free keyword

Ikuko Sugano and Naoko Kusada

Department of Library and Information Science
Aichi Shukutoku University

Katahira 9, Nagakute-cho, Aichi Prefecture, 480-11 Japan

The 5 cases of public health were retrieved. Each case was searched by MEDLINE on CD-ROM (Silver Platter) comparing with the effectiveness of the Thesaurus command by MeSH and Find command by free keyword. The free keywords and evaluations were made by the medical specialist. The one case of searching, "HIV infection due to dental treatment" was hit 67 items only by Find command. According to this analysis, Find command mode have covered the more useful concepts than the MeSH. It is necessary to require further examination by way of combined searchings with Thesaurus command and Find command.

1. 研究目的

医学情報を提供する代表的なデータベースMEDLINEのCD-ROM版は、1986年から発売され、1988年後半から医学図書館や医薬品メーカーなどを中心に導入が始まったと報告されている[1]。現在わが国では、医学図書館の90%以上がMEDLINEのCD-ROM版のサービスを開始し、その結果オンラインでMEDLINEを検索した経験がない研究者もエンドユーザーとして利用するようになった。

一般的に、シソーラスをもつデータベースは、シソーラスを用いた検索が奨励されており、MEDLINEの検索ではMeSHの利用が不可欠であると言われてきた。しかし、MEDLINEのCD-ROM版を利用している多くの研究者は、MeSHを用いなくてフリーキーワードで検索しており、さらに、研究者はその検索結果にかなり満足していると報告されている[2]。その一方で、フリーキーワード検索が普及した理由として、“シソーラスの使い方を習得するのはかなり困難であり”[3]、MEDLINEの利用者は直接CD-ROMを検索できることに満足してしまい、フリーキーワードによる検索結果にもれやノイズがあることに気づかず、さらに進んだところにあるシソーラス検索に踏み込みにくいといった批判もある[3]。

医学研究者はなぜMEDLINEのCD-ROM版でのフリーキーワード検索の結果に満足しているのであろうか。直接CD-ROMを検索できることから、研究者は自らの検索方法に疑いをもたず、そのため検索結果についても満足しているのであろうか。

本研究は、MEDLINEのCD-ROM版でMeSHを用いた検索結果とフリーキーワードを用いた検索結果を比較し、それぞれの特徴を明らかにすることから、多くの研究者にフリーキーワード検索が普及している理由を明らかにするとともに、研究者の立場からのフリーキーワード検索を基本とした検索方法のあり方を検討することを目的とする。

2. MEDLINEのCD-ROM版に関する既往調査

2.1 医学図書館の現状[4]

四谷は、MEDLINEのCD-ROM版を導入した日本の図

書館の現状を調査し、CD-ROMの今後の改善点を明らかにしている[4]。調査方法は、質問紙法により行なわれ、調査対象は東京地区の医学図書館7館と東海地区の医学図書館5館とした。またそれとあわせて各図書館の情報検索担当者にインタビューを行なった。調査結果では、MEDLINEのCD-ROM版を検索する上で必要となるはずの、MeSHに関する知識についての説明を行っていると答えた図書館は全体の半数であった。さらに、CD-ROM検索を導入したばかりであることから、詳しいMeSHの説明はあまり行われておらず、利用者が簡単なCD-ROM検索に慣れた段階で、MeSHの教育を予定する館がそのほとんどであった。また、Silver Platter版は、Findコマンド検索による利用が簡単であるため、MeSHの使い方まで指導する必要がないと答えた図書館もあった。

2.2 MEDLINEのCD-ROM版の利用調査1

細矢は、東京慈恵会医科大学医学情報センター図書館がSilver Platter社のMEDLINEのCD-ROM版を導入し、1989年5月から一般利用を開始したのにもなって、CD-ROMの有効活用を促進することを目的に調査を行った[2]。この調査はMEDLINEのCD-ROM版の検索者を対象に、1ヵ月間アンケート調査を実施し、利用者の評価と詳しい利用状況を分析したものである。その中で、MeSHの利用状況についての結果は、「利用しない」と答えた研究者が61%と半数以上を占めており、「利用した」研究者の39%を上回っていた。この結果は、MeSHを確認しなくても、論題・抄録中の語から検索できる利点が生かされていると細矢は述べている。また検索結果の満足度に関する質問では、求める文献が「見つかった」、「かなり見つかった」とする肯定的な回答が85%であり、大半の利用者が満足していることが明らかになった。

2.3 MEDLINEのCD-ROM版の利用調査2

名古屋大学鶴舞地区キャンパス情報ネットワーク委員会は、Silver Platter社製MEDLINEのCD-ROM版を導入した医学部図書館の利用者を対象にアン

ケート調査を行った[5]。学内の研究者603人に質問紙を配布し、184人からの回答を得た(回収率31%)。そのうちの医学研究者166人について結果をまとめると、MeSHを用いて(Thesaurusコマンド)検索している研究者は28人(16.9%)、フリーキーワードのみを用いて(Findコマンド)検索している研究者は123人(74.1%)であった。

この委員会報告書[5]のデータをもとに、先の検索方法の異なるグループ(MeSH利用グループ28人とフリーキーワード利用グループ123人)で検索結果の満足度を比較した調査がある[6]。その結果によると、「充分満足している」と答えた研究者が、MeSH利用グループでは、8人(28.6%)であり、フリーキーワード利用グループでは、29人(23.6%)であった。さらに、「まずまず、普通」と答えた研究者が、MeSH利用グループでは、20人(71.4%)であり、フリーキーワード利用グループでは、88人(71.5%)、「やや不満」と答えた研究者は、MeSH利用グループでは無く、フリーキーワード利用グループでは、6人(4.9%)であった。以上の結果から、MeSH利用グループの方が、検索結果に充分満足している研究者の割合がわずかに高いことが明らかになった。しかし、フリーキーワード利用グループでも、検索結果に満足している研究者が多く存在することから、医学研究者は直接自由にMEDLINEを検索できることが利点となって、CD-ROMの利用とそのフリーキーワードを用いた検索に対して満足していると考えられる。

上記の2つの既往調査から、フリーキーワード検索を行っている医学研究者の検索結果に対する満足度が高いことが明らかになった。これは、Silver Platter社製のMEDLINEのCD-ROM版がフリーキーワードで簡単に検索できるように設計されていることと、MeSHに関する利用者教育の遅れなどが関係していると考えられるが、MeSHを用いた検索結果とフリーキーワードを用いた検索結果とには大きな相違がないとも考えられる。そこでこの点を明らかにするために、検索事例を用いた調査を行うことにした。

3 検索方法別による検索結果の比較調査

ここでは、MEDLINEやMeSHを直接利用した経験のない研究者が、MEDLINEのCD-ROM版をフリーキーワードで検索した場合に、その検索結果をどの程度評価するかについて調査した。比較のために、MeSH用語を用いた検索結果も用意した。

3.1 調査方法

公衆衛生学の検索事例をMEDLINEのSilver Platter社製CD-ROM1992年版を用いて、MeSH用語による検索(Thesaurusコマンド検索)とフリーキーワードによる検索(Findコマンド検索)を行い、その検索結果を比較検討した。検索事例は以下の5件である。

検索事例：

1. キノコによる肝毒性について
2. 小児の高血圧発症の危険因子について
3. 歯科治療によるHIV感染について
4. 老人の癌の在宅での末期看護について
5. 癌の発症に与える低コレステロール血症の影響について

これらの事例を作成するにあたって、公衆衛生学分野の研究者1名に協力を依頼した。また、Findコマンド検索に用いるキーワードの選定と検索結果の評価についても依頼した。検索結果の評価の方法は、研究者に検索結果(書誌事項、抄録、ディスクリプタ)を読んで評価することを依頼し、さらに評価の際に評価基準を記録することも依頼した。

Thesaurusコマンド検索に用いるMeSH用語の選定は、16年間Index Medicusに携わっている医学系図書館の図書館員の協力を得て行った。

3.2 調査結果

表1に検索事例ごとの検索数、研究者が評価した適合文献数、不適合文献数の内訳を示した。検索数について見ると、検索事例5を除いて、Findコマンド検索の方が多かった。研究者が評価した適合文献数については、検索事例3を除いて数件の違いはあるものの、2つのコマンド検索では結

果に差がなかった。不適合文献数については、検索事例5を除いて、Findコマンド検索の方が多かった。

表1 検索方法別による結果の一覧

	Thesaurusコマンド検索			Findコマンド検索		
	検索結果 (件)	適合文献 (件)	ノイズ (件(%))	検索結果 (件)	適合文献 (件)	ノイズ (件(%))
検索事例1	1	1	0(0%)	5	4	1(0.2%)
検索事例2	51	16	35(68.6%)	68	15	53(77.9%)
検索事例3	67	48	19(28.4%)	205	113	92(44.9%)
検索事例4	3	3	0(0%)	4	4	0(0%)
検索事例5	101	18	83(82.2%)	47	13	34(72.3%)

表2は、2つの検索の適合文献の違いを表すため、Thesaurusコマンド検索の適合文献数(A)とFindコマンド検索の適合文献数(B)、Thesaurusコマンド検索とFindコマンド検索の共通の適合文献数(C)、Thesaurusコマンド検索のみの適合文献数(A-C)、さらに、Findコマンド検索のみの適合文献数(B-C)を示したものである。また評価では、Findコマンド検索でより多くの適合文献数を得た事例3(F1)と事例1(F2)、Thesaurusコマンド検索でより多くの適合文献数を得た事例5(T1)、両方の検索で同じ程度の結果を得た事例2(S1)と事例4(S2)に分けることができた。

表2 適合文献の内訳

	適合文献数		TとF検索 の共通 適合文献 C (件)	T検索 のみの 適合文献 A-C (件)	F検索 のみの 適合文献 B-C (件)	評 価
	T検索 A (件)	F検索 B (件)				
検索事例3	48	113	46	2	67	F1
検索事例1	1	4	1	0	3	F2
検索事例5	18	13	13	5	0	T1
検索事例2	16	15	13	3	2	S1
検索事例4	3	4	1	2	3	S2

このように、5件の検索事例を比較しただけでも、Thesaurusコマンド検索の検索結果がよい場合と、Findコマンド検索の検索結果がよい場合とに

分かれる結果になった。これは、それぞれの事例ごとに、フリーキーワード検索の利点およびMeSH利用検索の利点が含まれているためと考えられる。

そこで、事例5件のうち、Findコマンド検索の結果がThesaurusコマンド検索より優れていた事例3を詳細に分析し、フリーキーワードを用いた検索の利点を検討する。

4 Findコマンド検索の特徴

4.1 著しい適合文献の違いがある事例

事例5件の中でもっともFindコマンド検索とThesaurusコマンド検索との適合文献の結果に著しい違いがあったのが事例3であった。各々の検索式をもとに、結果の違いについて分析する。

Findコマンドでの検索式(図1参照)では、「歯科治療」に対する英語として「Dental care」や「Dental treatment」などが研究者から挙げられ。また、歯科全般に関係するHIV感染についての広い検索を研究者は希望していたため、Dentという接頭語をもつ用語でも検索ができるように、「Dent#」で、前方一致検索を行った。さらに、「HIV感染」については研究者から「HIV」とそれに関係する「AIDS」が挙げられたためそれらを用いて検索した。「感染」という概念については、「HIV」および「AIDS」と歯科に関係する「DENT#」で検索すれば、自ずと、感染についての文献が検索されるという研究者の判断で、検索式には入れなかった。

1:	6456	DENT#
2:	4856	HIV
3:	3317	AIDS
4:	205	#1 and (#2 or #3)

図1 Findコマンドによる検索式

Thesaurusコマンドの検索式(図2参照)作成では、医学系図書館員の協力を得ながら、MeSH(冊子体マニュアル)を用いて「歯科治療」のMeSH用語「DENTAL CARE(歯科医療)」を見つけ、さらにその下位概念を調べた。しかし、「DENTAL CARE FOR AGED(老人歯科医療)」と「DENTAL CARE FO

R DISABLED (障害者用歯科医療)」のみがあるだけで、歯科治療の概念としては不十分であることから、「DENTAL CARE (歯科医療)」も含む上位語「DENTISTRY (歯科学)」を用いて、下位語をすべて検索する方法であるExplode検索を行った。「HIV感染」のMeSH用語として、「HIV INFECTIONS (HIV感染)」が見つかり、この語を用いてExplode検索を行った。

1:	5121	explode HIV-INFECTIONS / all subheadings
2:	3164	explode DENTISTRY / all subheadings
3:	67	#1 and #2

図2 Thesaurusコマンドによる検索式

なお、図2の「all subheadings」は、MeSHの主標目(この場合は「HIV-INFECTIONS」と「DENTISTRY」)にすべての副標目を組み合わせて検索したことを示している。副標目は医学分野の検索によく利用する概念を表す用語であり、主標目と組み合わせて検索することによって側面的な限定を与え、主標目が該当文献中で果たしている役割を明確にする働きを持っている。しかし、今回は研究者が広い範囲を検索することを希望していたため、副標目との組み合わせ検索をすることで検索限定のしすぎによる検索もれを引き起こさないように、副標目は限定せず、全ての副標目を組み合わせて検索した。

Findコマンド検索とThesaurusコマンド検索の各々の検索結果をみると、Findコマンド検索では広い範囲からの検索が行われ205件の文献が検索されたのに対し、Thesaurusコマンド検索ではExplode検索を行ったにもかかわらず67件しか検索されなかった。さらに適合文献の割合はFindコマンド検索55.1%に対してThesaurusコマンド検索の方が71.6%と多かったが、件数をみるとThesaurusコマンド検索では48件しか得られなかったが、Findコマンド検索では113件が適合文献であった。ただし、Thesaurusコマンド検索の不適合文献が19件(28.4%)に対して、Findコマンド検索では92件(検索結果の44.9%)とノイズの多い結果となった。これは常に

フリーキーワード検索での問題点として挙げられる点ではあるが、検索結果の評価を行った研究者は広い検索ができた結果、多くの適合文献を得られた点を評価していた。この点は、フリーキーワード検索を行った場合、タイトルをはじめすべてのフィールドを対象とした検索をするので、広い検索を希望する場合には有効であることから考えられる[7]。

4.2 Findコマンド検索による適合文献の分析

4.2.1 分析の手順

事例3では、Findコマンド検索で検索した方がThesaurusコマンド検索より多くの適合文献(113件)を得ることができた。検索式を比較するとFindコマンド検索での「Dent*」は、6,546件得られたのに対し、Thesaurusコマンド検索での「DENTISTRY (歯科学)」は3,164件と、半数以下の結果であった。このことから、Thesaurusコマンド検索では、「DENTISTRY (歯科学)」によって検索もれを引き起こしたことが考えられ、この点を詳細に分析するために、Findコマンド検索のみで得られた適合文献(67件)を分析した(図3参照)。

ThesaurusコマンドがMeSHフィールドのみを対象に検索するのに対して、Findコマンドでは、全てのフィールドを対象とする。事例3では、Findコマンドでの「Dent」前方一致検索で、「TI (タイトル)」、「AD (所属機関団体名)」、「AB (抄録)」、「SB (主題限定フィールド)」、「MESH (MeSHフィールド)」の各フィールドで、Dentを接頭語として持つ用語を検索した。その結果と「HIV」と「AIDS」の論理積によって得られた適合文献(67件)を対象に、その書誌事項のうち、MeSHフィールドにはどのようなMeSH用語が付与されているかを図3の「分析結果1」に示し、またタイトルや主題を示すフィールドを対象に「Dent」前方一致検索によって得られたキーワードおよび、その適合文献の件数を「分析結果2」に示した。

4.2.2 MeSHの関連語参照の不備

「分析結果1」の「DENTAL STAFF (歯科スタッ

Findコマンドで「DENT*」と「HIV」, 「AIDS」のキーワードを用いて検索した結果(205件)の評価を公衆衛生学の研究者に依頼した。



研究者による適合文献の評価

< 適合文献の判断基準 >

- ・ HIVにかかっている患者を歯科治療するにあたって、次の患者に感染させないようにする実際の報告例
- ・ 医師から治療者、治療者から治療者、治療者から医師に歯科治療によって HIVに感染した事例の文献
- ・ 歯科治療における HIV感染の予防について書かれた論文
- ・ 歯科治療に限らず、歯科関係の雑誌に掲載されている HIV感染の論文

評価結果



Findコマンド検索の適合文献113件のうち67件が、Findコマンド検索のみで得られた件数であり、Thesaurus検索では検索されなかった適合文献である。



フィールド別のMeSH用語および、キーワードの分析

分析結果 1



MeSHフィールド (MeSH用語)	
「DENTISTS」	29件
「DENTAL, STAFF」	9件
「EDUCATION, DENTAL」	6件

分析結果 2



その他のフィールド (キーワード)	
TI : タイトル	(31件)
Dental	18件
Dentists	11件
Dentistry	2件
SB : 限定フィールド	
DENTAL	49件

図3 検索事例3の分析手順と結果

「DENTISTS (歯科医)」, 「EDUCATION, DENTAL (歯科教育)」は、Findコマンド検索の「Dent*」の前方一致検索によって検索されたMeSH用語である。これらのMeSH用語は「DENTISTRY (歯科学)」の下位語でないために、Thesaurusコマンド検索では検索されなかったと考えられる。そこで、これらのMeSH用語をMeSH (冊子マニュアル)を用いて、どのトゥリー構造に属しているか、またそれぞれのMeSH用語と「DENTISTRY (歯科学)」との関係についても調べた。

その結果、「EDUCATION, DENTAL (歯科教育)」は、「I 2 : 教育」のトゥリーに属しているため、「DENTISTRY (歯科学)」の下位語ではないが、「DENTISTRY (歯科学)」への関連語参照がなされていた(図4参照)。

一方、「DENTAL STAFF (歯科スタッフ)」と「DENTISTS (歯科医)」は「M 1 : 名称・グループ」と「N 2 : 施設・人的資源とサービス」のトゥリーに属しており、「DENTISTRY (歯科学)」の下位語ではなく、さらに、関連語参照もなされて

いなかった。図5に示したようにMeSHには、「MIDWIFERY (助産学)」から「NURSE MIDWIVES (助産婦)」への関連語参照があるからにもかかわらず、「DENTISTRY (歯科学)」と「DENTISTS (歯科医)」が関連語となっていない点はMeSHの関連語参照の不備と言えよう。

DENTISTRY
歯科学(シカガク)・
E6+ G2.163+
see related DENTAL EQUIPMENT
DENTAL INSTRUMENTS
ECONOMICS, DENTAL
<u>EDUCATION, DENTAL</u>
HISTORY OF DENTISTRY
LEGISLATION, DENTAL
X ORAL MEDICINE
X STOMATOLOGY

図4 「DENTISTRY」から「EDUCATION, DENTAL」への参照 [8]

MIDWIFERY
助産学(ジョサンガク)・
G2.478.676.416
see related HOME CHILDBIRTH
<u>NURSE MIDWIVES</u>
X TRADITIONAL BIRTH ATTENDANT

図5 「MIDWIFERY」から「NURSE MIDWIVES」への参照 [8]

4.2.3 「Dent」前方一致検索の効果

「分析結果2」の「Dental」、「Dentist」、「Dentistry」はタイトルに、また「DENTAL」は主題限定のフィールドに、「Dent」を接頭語に持つ語を含む適合文献の件数を示した。さらに、これらの31件と49件の適合文献に付与されていたMeSH用語を調べたところ、タイトルと主題限定フィールドで共通して「PATIENTS (患者)」、「OCCUPATIONAL-DISEASES (職業病)」、「RISK FACTORS (危険因子)」、「MOUTH-DISEASES (口腔疾患)」などが見つかった。これらは、Thesaurusコマンド

による「DENTISTRY (歯科学)」の検索では得られない、MeSH用語であった。つまり、事例3では、Findコマンドを用いた前方一致検索によって、Thesaurusコマンドでは検索できなかった多くのキーワードを用いた検索を行うことになり、「歯科治療におけるHIV感染」についての広い範囲での検索を可能にした。

4.3 まとめ

多くの医学研究者がMEDLINEのCD-ROM版の普及によって、フリーキーワード検索によって満足した検索結果を得ているという報告 [2, 5] をもとに、その理由を明らかにするため、MeSH用語を用いたThesaurusコマンド検索とフリーキーワードを用いたFindコマンド検索の結果の比較を行った。その結果、MeSHを用いた検索と、フリーキーワード検索で得られる適合文献の内容に大きな差異はないということが明らかになった。しかし、実際にフリーキーワードのみの検索で充分であるか否かについて疑問が残った。そこで、Findコマンド検索がより多くの適合文献を得た事例3をもとに、Thesaurusコマンド検索では得られなかった適合文献67件について詳細に分析を行った。その結果、フリーキーワード検索は、広く文献を検索することを検索者が希望する場合、MeSHの関連語参照の不備を補い、効果的な検索方法となり得ることが明らかになった。

今後、さらに分析を進め、それぞれの検索方法の利点を明らかにした上で、フリーキーワード検索とMeSHを用いた検索の望ましい併用の仕方を明らかにしていくことにする。

- 1) CD-ROM調査研究委員会編. CD-ROM調査研究報告書. 東京, データベース振興センター, 1990. 146p.
- 2) 細矢敬子. CD-ROM利用者調査報告. 医学図書館. Vol. 37, No. 1, p. 47-52(1990)
- 3) 蒲池 昭, 正木達之. CD-ROM MEDLINE導入後のエンドユーザー教育. 薬学図書館. Vol. 36, No. 3・4, p. 188-191(1991)

- 4) 四谷あさみ. CD-ROMが医学図書館に与えた影響. 愛知淑徳大学文学部図書館情報学科平成2年度卒業論文, 1990, 28p.
- 5) 鶴舞地区キャンパス情報ネットワーク委員会. MEDLINE利用に関するアンケート調査結果. 愛知, 名古屋大学医学部図書館, 1993. 5p.
- 6) 草田直子. CD-ROM版MEDLINEにおけるフリーキーワード検索とMeSH利用の問題点. 愛知淑徳大学文学部図書館情報学科平成5年度卒業論文, 1993, 47p.
- 7) 児玉関. 医学中央雑誌CD-ROM版における医学用語シソーラスの有効性. 医学図書館. Vol. 40, No. 4, p. 387-392(1993)
- 8) 日本医学図書館協会編. MEDICAL SUBJECT HEADINGS 医学件名標目表 1993 英和版. 東京, 紀伊國屋書店, 1993. 1255p.