

東邦大学学術リポジトリ

Toho University Academic Repository

タイトル	米国国立医学図書館分類法(NLMC)日本語版データベースを紹介します
別タイトル	Introduction of Japanese Version of National Library of Medicine Classification Database
作成者(著者)	橋本, 郷史
公開者	日本医学図書館協会
発行日	2019.09.01
ISSN	04452429
掲載情報	医学図書館. 66(3). p.212 216.
資料種別	学術雑誌論文
内容記述	JMLA活動報告
著者版フラグ	publisher
メタデータのURL	https://mylibrary.toho.u.ac.jp/webopac/TD32122614

米国国立医学図書館分類法 (NLMC) 日本語版データベースを紹介します

橋本 郷史*

NLMC-DB WG, 東邦大学医学メディアセンター

I. はじめに

日本医学図書館協会（以下、JMLA）は、2019年6月10日に米国国立医学図書館分類法（NLMC）日本語版データベース（以下、NLMC-DB）（<http://jmla1927.org/>）のサービスを開始した（図1）。本稿では、NLMC-DBの作成を担当したワーキンググループ（以下、本WG）の立場から、サービスの概要と使い方のコツ、その構造を紹介する。またサービス開始後の利用実態を報告し、今後の課題と展望についても記す。

なお、NLMC-DBは2019年9月時点では、JMLA会員のみが利用可能である。利用のためのIDとパスワードはJMLA Webサイトの会員専用ページに掲載している。

II. NLMC-DBの概要と検索方法

最初に、NLMC-DBの概要と検索方法について紹介する。

1. NLMC-DBとは

NLMC-DBとは、米国国立医学図書館によって作成された医学・医療分野に特化した分類法（National Library of Medicine Classification : NLMC）を日本語に翻訳したものを、インターネットを介して利用できるサービスで、NLMCの2016年版を翻訳し2017年6月に出版した冊子「米国国立医学図書館分類法：2016年版 日本語版」¹⁾（以下、冊子版）をデータベース化したものである。

NLMC-DBでは、冊子版の分類表と索引とをそれぞれ検索して利用することができる。

なお、冊子版の翻訳とNLMC-DBの作成は、いずれもJMLAの受託事業として行われており、本WGは、翻訳作業を行ったワーキンググループ（以下、翻訳WG）のメンバーの一部で構成されている。翻訳WGの活動及び作業内容については、翻訳WGグループ長だった小嶋が報告している^{2) 3)}。興味のある方はそちらを参照されたい。



図1. NLMC-DB TOPページ

*Satoshi HASHIMOTO : ヘルスサイエンス情報専門員 (中級)
〒143-8540 東京都大田区大森西5-21-16. Tel.03-3762-4151 (2444)
jmla-nlmc@umin.ac.jp, hashimoto@mnc.toho-u.ac.jp
(2019年7月17日 受理)

2. 分類検索

分類検索では、QZ, WAなどの分類記号と分類番号を掛け合わせて検索する。該当する組み合わせがある場合、分類表の中の該当箇所が表示され、対応する見出し語を確認することができる。前後の分類番号と見出し語も表示され、冊子版の分類表と同じように自由に前後を確認することができる。該当する組み合わせがない場合は、検索に使用した分類記号の分類表の最上部が表示される。表示される分類表は、QZ, WAなどの分類記号ごとである。分類記号をまたいだ移動はできない。他の分類記号に属する分類を調べる場合は再検索する必要がある。

3. 索引検索

1) 索引検索の概要

索引検索では、任意のキーワードで検索すると、そのキーワードを含む索引語が50音順で表示され、あわせてその索引語が対応する分類や関連語も表示される。その分類から分類表の該当箇所に飛んだり、関連語で索引を再検索したりすることができる。

索引語は日英併記されており、日本語でも英語でも検索することができる。日本語の名称は主に「医学用語シソーラス第8版」⁴⁾及び「日本医学会 医学用語辞典WEB版」⁵⁾に則っている。なお、英語で検索した場合も、結果の表示は訳語の50音順となる。

索引中の分類も検索対象になっている。分類で検索すると、その分類に対応する索引語を調べることができる。分類で調べる場合は分類記号と分類番号の間にスペースを入れる必要がある。

2) 索引検索の留意事項

検索キーワードをスペースで区切って入力すると、AND検索となる。ORやANDなどの論理演算子、アスタリスクやエクスクラメーションマーク、クォーターションマークなどの検索記号は使用できない。また、ヨミによる検索はできない。

検索対象となる索引語の階層は索引見出し語(=上位層。例:腫瘍)レベルである。索引見出し語の細目(=下位層。例:腫瘍→薬物療法)は検索の対象になっていない。

また、シソーラスのように同義語などの多様な語彙を備えているわけではない。加えて、翻訳という性質上、必ずしも利用者が想定している訳語が索引語に採用されているとは限らない。

3) 索引検索のコツ

論文検索データベースで行うように検索をしてしま

い、思いどおりの結果が得られないこともあると思われる。そこで、索引検索を使用する際のコツと便利な機能を紹介する。

例えば、「災害看護」というキーワードの分類を知りたいとする。検索したところ、残念ながら、この言葉そのままでは見出し語にはなっていないが、「看護」もしくは「災害」の見出し語の細目に出現している可能性がある。再検索したところ、災害の細目に、災害→看護 WY154.2が見つかった。このように、あくまで索引であるということ忘れずに、検索語を適宜修正しながら検索することで、よい結果が得られるものと思う。

索引検索は初期状態では部分一致で検索するようになっているが、検索キーワードの入力窓の右横で前方一致や後方一致、完全一致に検索方法を変更することができる。例えば、「糖尿病」を語中に含む語(前糖尿病状態など)を避けて、糖尿病から始まる語(糖尿病-1型など)のみを検索したいときは、前方一致に切り替えるとよい。また、「症候群」で終わる語を調べたいときなどは、後方一致に切り替えるとよい。

検索キーワードの入力窓にキーワードを入れると、サジェスト機能により、前方一致で索引語の事前検索が行われ、最大10語の候補が表示される。例えば「再生医療」というキーワードの分類番号を知りたいとする。入力窓に「再生」まで入力すると、候補として「再生医学」が出てくるが、「再生医療」は出てこない。再生医療は索引に登録されていない。Regenerative Medicineの訳語には医中誌シソーラスでは統制語として「再生医学」が使用されており、NLMC-DBでもそれを採用しているためである。このように自分の想定する訳語が登録されていない場合でも、事前検索の候補を見ることで、索引中の訳語を予測し、適切な検索を行うことができる。

3. NLMC-DBの画面構成

NLMC-DBは、機器に応じて最適な画面構成で利用できるようになっている。このため、パソコンだけでなく、タブレットやスマートフォンからでも快適に検索が可能である。パソコンユーザーでも、使用しているブラウザの表示幅を狭くすることで、スマホ上での画面構成を再現できるため、どのように見えるのか気になる方は試してみてください。

Ⅲ. NLMC-DBのデータと構造

この章では、データベース作成の側面からNLMC-DBについて説明する。

1. NLMC-DBのデータ

データベース化にあたっては、冊子版作成時の作業データを使用した。

冊子版作成時はエクセルで作業しており、このデータはデータベース化の際に必要な基本的な構成要素を備えていた。しかし、冊子版作成の作業当初からデータベース化が念頭にあったわけではなく、あくまで印刷物を作るためのデータだったため、そのままでは使用できない箇所も多くあり、修正が必要だった。また、冊子版はNLMCの2016年版を翻訳したものである。このため、現状のNLMC-DBには2017年以降の改訂は反映されていない。

今回のサービス開始に修正の作業が間に合わなかった箇所も残っている。加えて、NLMCの2017年以降の改訂に対応するかどうかという問題がある。これらの点については、V. で述べる。

2. NLMC-DBの構造

作成を担当したWGは時限的であり、長期のサポートが難しい。検索のエラーや管理の手間をなくすため、なるべくシンプルな構成でメンテナンスが不要なデータベースの構築を目指した。

NLMC-DBはMySQLを用いたりレーショナルデータベースとして構築した。分類表を格納するclassテーブルと、索引を格納するindexテーブルの2テーブルからなる。

図2に簡略化して抜粋したindexテーブルの構造を示す。冊子版作成時には灰色の部分にデータがなかった。

dbno ID	index_j 索引見出し語	subhead_j 索引細目	refer_number 参照先分類	refer_index_j 参照先索引語	see_j を(も)見よ	result 翻訳結果	sort ソート用ヨミ
27957	糖尿病		WK 810-850			糖尿病 (Diabetes Mellitus) WK810-850	トウニョウビョウ 10631
27958	糖尿病	妊娠中	WQ 248			妊娠中 WQ248	トウニョウビョウ 10631@10632
27959	糖尿病	看護	WY 155			看護 WY155	トウニョウビョウ 10631@10633
27960	糖尿病	治療	WK 815-825			治療 WK815-825	トウニョウビョウ 10631@10634
27961	糖尿病			尿崩症	をも見よ	尿崩症をも見よ	トウニョウビョウ 10631@10635
27962	糖尿病			食事-糖尿病	をも見よ	食事-糖尿病をも見よ	トウニョウビョウ 10631@10636
27963	糖尿病			前糖尿病状態	をも見よ	前糖尿病状態をも見よ	トウニョウビョウ 10631@10637

図2. indexテーブル簡略構造

冊子版の索引の順番は、カタカナと番号で生成されたソート用ヨミ列 (sort) で制御されていた。冊子版ではこれでよかったが、データベースとしては、このままの構造では、検索時に見出し語と細目をどう紐付けるかが問題となった。対処方法としていくつかの方法を考えたが、最終的には図にある通り、細目にも見出し語の情報を持たせるようにデータを修正した。これにより、索引見出し語列 (index_j) を検索し、HITした行の結果列 (result) をID (dbno) 順で表示するだけでデータベースが成立するようになった。

この検索処理は、次のような最も簡単な検索用SQLで記述できる。このため、検索で生じるイレギュラーなエラーは少なく、検索速度も早いものとなった。

```
select result from index where index_j like '%キーワード%' order by dbno;6)
```

IV. NLMC-DBの利用実態

サービス開始から1ヶ月経った7月9日までの利用ログを集計したところ、以下ようになった。

30日間のトップページの訪問数は612回である。平均すると1日あたり約20回の訪問となるが、利用の96%は月曜日から金曜日の平日に集中しており、平日の訪問数の平均は約27回だった。

検索は計2,372回行われ、内、分類検索が1,101回、索引検索が1,271回であった。

複数のIPアドレスから100回を超える検索が行われており、すでに業務の中でNLMC-DBを利用し始めている機関があることをうかがわせた。

一方で索引の検索結果からは、少々悲しい実態もわかった。索引検索の内、約3割に当たる359回の検索がHIT数0となっていた。Ⅱ. 3. 2) で説明した索引検索特有の事情が影響していると思われる。

索引検索の検索対象は冊子版の索引語で、シソーラスのような語彙の豊富さや、一般的なWeb検索のような自由度はない。冊子版の翻訳作業時には、4万語を超える索引語の翻訳(と訳語の統一、調整)を行っており、作業者にとっては膨大な作業で気が遠くなったが、データベースとしては物足りない語数であることを思い知らされる。

V. 課題と展望

NLMC-DBの修正と更新、今後の展望について述べる。

1. NLMC-DBの修正

NLMC-DBには以下のような修正の余地があると考ええる。

1) 誤字・脱字・不体裁の修正

冊子版は、データ作成後、印刷して校正作業を行った。今回使用したデータにはその校正が反映されていないため、誤字・脱字が残っている箇所がある。また、データベースとして構築した際に、見出し語のレベルがずれるなど不体裁になっている箇所があり、これらは今後修正したいと考えている。

2) 索引語の修正

冊子版作成時に、原本の索引にある語だけでなく、翻訳WG独自に、分類表にある語や、同義語を補ったものがある。これらは、冊子版に最低限必要なデータ(result, sort列など)しかなく、データベースとして必要なデータが欠けていたり、間違っていたりする箇所がある。検索時にエラーやノイズとなっているため、1)と同様に、今後修正を検討したい。

検索の利便性を向上させるために索引の同義語などを補うということも考えられるが、採用する語の判断が難しく、また、継続して作業できる人員がおらず、この作業の予定はない。

3) ヨミの整備

Ⅱ. 3. 2) で記したように、ヨミでの検索ができない。冊子版作成時の作業データのヨミ情報は並び順のソートのために適宜編集されており、ヨミが完全ではなく数字や記号混じりとなっている。このため、ヨミ情報はデータベース上に存在しているものの、そのままでは検索に使用できない状態である。ヨミでの検索を実装するためには一度情報を整備する必要があるが、作業量が大きく、現状では整備の予定はない。

4) インターフェースの修正

インターフェースについて、大規模な修正は難しいが、小規模であれば可能なものもあると考えている。例えば、NLMCに不慣れな利用者にとっては、QZ, WAと言った分類記号が対応する分類がわからないため分類検索が単独では使いづらい。この点は、分類検索の分類記号窓に分類名を併記するなどの方法で改善できる余地があると考えている。

2. NLMC-DBの更新

Ⅲ. 1. で触れたように、NLMC-DBはNLMCの2016年版を元にしており、その後の改訂は反映されていない。毎年の改訂分も翻訳できればそれに越したことはないが、改訂の内容によってはデータベース構造が変わる箇所が出てくることも予想され、対応するにはそれなりの体制を整える必要がある。このため、残念ながら、改訂分の更新についても、現状は予定がない。

3. NLMC-DBの今後の展望

今後の展望としては以下の2点がある。

1) 一般公開

現在NLMC-DBはJMLA会員限定の公開である。これは冊子版およびその電子版の販売などへの配慮からであるが、一定期間が経過し、その配慮が不要となった時期に、利用制限を設けず一般にも利用できる状態にしたいと考えている。

2) APIの実装

現状のデータベースは、スタンドアロンのシステムで拡張性がない。具体的な計画はないが、APIを実装して、外部サービスとも連携できるようになれば、サービスとしての意義と利便性が向上するのではないかと考えている。

例えば、OpenSearchに対応し、外部からの分類検索に対してその分類に対応する分類表の見出し情報を返すことができるようになる、などである。エクセルのwebservice関数などから参照できるようになれば、プログラミングの心得がなくても、エクセル上に分類を用意すればすぐにそれらの分類に対応する見出し語が分かる。目録の担当者をはじめとして便利に感じる利用者はあるものと思う。

ただし、開発には大きな労力がかかる。それに見合ったニーズがあるのか、あるとしてもどのような機能をどのような形で実装すべきなのかは慎重に考える必要があり、今後、NLMC-DBの利用情報を蓄積・分析したうえで検討されるべきと考える。

VI. おわりに

以上、駆け足ではあるが、NLMC-DBについて紹介した。データベースとして未熟で、修正すべき箇所も多く、まだまだ発展の余地を多く残していると認識しているが、本WGとしてはひとまずその役割を果たしたと考えている。ただ、今年度も本WGは活動を継続し、V.で述べたようなデータやインターフェースの修正を行う予定である。NLMC-DBのおかしな点、ご意見などがあれば、本WGのメンバー宛てにご連絡をいただきたい。

NLMC-DBの作成はJMLAとしてこれまでに類を見ない事業であり、多くのことが試行錯誤の中で行われた。当初の予想に比べると、思いの外苦労は多かった。それだけに、本サービスが今後、今の形を維持するにしろ形を変えるにしろ、末永く多くの人に利用されることを願ってやまない。

最後に、この場をお借りして、NLMC-DBの作成にご協力いただいた方々への謝意を述べたい。

NLMC-DBの実際の構築作業では、松坂重徳様、岸裕紀様を始めとした株式会社インフォレストの皆様にご尽力いただいた。データベースを設置するサーバの契約に関しては、JMLA事務局の宇佐見様、事務局長の谷澤様にお世話になった。また、サービス全般について、JMLA企画・広報委員会委員長の黛様に適宜ご助言をいただいた。みなさまのお力添えでNLMC-DBが実現しました。ありがとうございます。

参考文献・注記

- 1) National Library of Medicine (U.S.), 日本医学図書館協会米国国立医学図書館分類法 (NLMC) 日本語版作成ワーキンググループ, 日本医学図書館協会. 米国国立医学図書館分類法. 2016年版 日本語版. 東京:日本医学図書館協会, 丸善雄松堂学術情報ソリューション事業部 (発売);2017.
- 2) 小嶋智美. 『米国国立医学図書館分類法 2016年版 日本語版』. 医学図書館. 2017;64(2):65.
- 3) 小嶋智美. テキストブック (13) 米国国立医学図書館分類法 2016年版 日本語版:原著のあらましと日本語版の私的活用法. ほすびたるらいぶらりあん. 2018;43(1):35-8.
- 4) 医学中央雑誌刊行会. 医学用語シソーラス[internet]. <https://www.jamas.or.jp/database/thesaurus.html> [accessed 2019-08-05]
- 5) 日本医学会医学用語辞典WEB版[internet]. <http://jams.med.or.jp/dic/mdic.html> [accessed 2019-08-05]
- 6) 検索用のSQLの基本的記述構造は以下ようになる。

```
select (A) from (B) where (C)(D);
```

 (A)で呼び出す対象のカラムとその呼び出し方を指定し、(B)で検索対象のテーブルを指定する。(C)で検索の条件を記述し、(D)で検索結果の表示規則などを指定する。(A)(B)には複雑な式を記述することもできるが、この文の場合はシンプルにカラムやテーブルを指定している。また、この文の場合、(C)はindex_j like '%キーワード%'である。index_jが検索対象カラム、likeは不定要素を含む検索方法の宣言、シングルクォーテーション内が検索キーワードとなり、%部分で任意の文字列を許容する。この文の場合、(D)はorder by dbnoで、検索結果をdbno順に表示する(order)指示となっている。