

# 医学シソーラス MeSH 生起頻度(1960-2011)による 研究動向の把握

野添篤毅<sup>1)</sup>、西村飛俊<sup>2)</sup>

1) 愛知淑徳大学名誉教授、2) 愛知淑徳大学大学院

**目的：**医学文献データベース MEDLINE には 2011 年で年間 736,191 件の書誌データが蓄積され、総蓄積文献数は 1963 万件に達している。これらの書誌データには統制語彙としてのシソーラス Medical Subject Headings (MeSH) に登録された用語(2012 年度の登録語数、26,582)が付与されている。本研究は米国国立医学図書館(NLM)が提供している MEDLINE ベースライン・データファイル(MEDLINE Baseline Repository: MBR ; <http://mbr.nlm.nih.gov/>)に含まれる索引語付与の統計データ（1965 年—2010 年）から巨視的に医学研究の動向を探ろうとする。

**方法：**MBR ファイルでは、すべての MeSH 用語の文献への付与状況が生起頻度データとして年度別に集計される(Frequency Count Files)。生起頻度データは MeSH サブカテゴリーごとと登録語の使用頻度(索引に使用された回数)が年度別に示される。また各々の MeSH 用語について索引語として付与された文献のすべての PMID(PubMed 書誌 ID)が示された Raw Data File も用意されている。この 2009 年原データ(1964-2008 の書誌データ、約 1688 万件)から MySQL データベースを作成し解析材料とした。

まず生起頻度データから MeSH カテゴリー(解剖用語、疾患名、化学物質、診断・治療技術など)および疾患サブカテゴリーにおける 1965 年から 2010 年間の 45 年にわたる用語付与状況の年次変化を分析する。付与件数の実数だけでなく各年度における 1000 文献あたりの各カテゴリー、サブカテゴリーに含まれる MeSH 用語の使用頻度指標も導いた。なお NLM から提供された 1988 年から 2012 年にわたる MeSH 用語の登録語数(カテゴリー別、サブカテゴリー別)の統計データも参照した。また我々が作成した MySQL ファイルからは特定の MeSH 用語群(例：消化器系腫瘍)についてその用語が各年度でどのような推移で付与されているかを検証した。

**結果・考察：**蓄積文献量の増大に従って MeSH 用語の生起頻度が急速に増加することは自明であるが、1000 文献あたりの使用頻度指標(1965-2010)をみると、A カテゴリー(解剖)で 800 前後、C カテゴリー(疾患)では 1000 から 1500 の間を推移している。最も急激な変化を示している E カテゴリー(診断・治療技術)では当初 1200 前後であった指標が近年 2100 前後を記録している。これは現在の医学研究論文で研究手法を明確に記述するという流れと合致している。また、G カテゴリー(生物現象・プロセス)、N カテゴリー(ヘルスケア)なども急速の伸びを示し、現在の研究の力点を表している。

MBR ファイルからの MeSH 用語の文献への付与状況の時系列的な分析によって 50 年間の医学研究の傾向を把握することが可能となった。今後は MySQL ファイルを活用して複数の特定の疾患グループでの研究動向の比較を検討する。