

実験材料の引用に基づく機関影響度の計測

天野 晃¹⁾, 角田裕之²⁾, 石川大介³⁾, 水野澄子¹⁾, 柴田大輔⁴⁾
¹⁾理化学研究所, ²⁾鶴見大学, ³⁾神奈川県立病院機構, ⁴⁾東邦大学

背景: 従来, 引用分析における引用とは文献の引用すなわち, 参照文献を指し, 引用される側から引用する側への影響を示す。一方, 影響という観点では文献の引用以外に施設の供与(有償無償にかかわらない), 材料や技術の提供など多様な形態があるが, これらは影響度として評価されていない。我々は研究評価指標により多角的な観点を導入するための基盤の確立を目指す。

目的: 前述した基盤確立のための第一歩として, 実験材料の提供による影響を計量することを試みる。具体的にはナショナルバイオリソースプロジェクトの各参画機関の影響度(実験材料の“被引用”)計量を行う。

評価対象: 2015年時点でのナショナルバイオリソースプロジェクト参画機関(最上位機関)17機関。

データ: (a)2014年時点でのPubMed CentralのXMLファイルすべての「謝辞」及び「材料」セクション。(b)2015年時点でのNISTEPによる国内研究機関の英組織名バリエーション。(c)2015年時点でのナショナルバイオリソースプロジェクトによる提供実績のある生物実験材料名リスト。

方法: (1)PubMed Centralの全文ファイルより<ack>セクション及びsec-typeの値に“material”を含む<sec>セクションを抽出する。(2)抽出したセクションをセンテンスに分割する。(3)各センテンスに対して,(b)及び(c)を検索し,(b),(c)が共出現するセンテンスを選択する。(4)選択されたセンテンスをマニュアルでチェックし,提供関係が明確な組織名と生物実験材料名のペアを得る。(5)4のセンテンスを含む論文の著者所属機関を得る。(6)4,5より[提供機関]-[生物実験材料]-[被提供機関]のトリプルを得る。

結果: 提供機関名, 生物実験材料名, 被提供機関名の出現数は,それぞれ延べ974(異なり:47),延べ1250(異なり:446),延べ3395(異なり:3082)であった。提供数上位の機関は,理化学研究所,東京大学,京都大学であった。提供数上位の生物実験材料名(名寄せなし)は,C57BL/6(マウス系統),Mouse,HeLa(ヒト細胞系統)であった。被提供数が上位である機関名(名寄せなし)は,Baylor College of Medicine,University of Washington,University of Cambridgeであった。なお,当日の発表においてはより詳細な数値を報告する予定である。

議論: 提供機関においてはNISTEPのデータを活用することにより名寄せが可能となり,各機関につき高い精度で提供数を計上できたと考えられる。一方,生物実験材料名については,Mouseなどの一般名称が,ナショナルバイオリソースプロジェクトの提供リストにも,論文中の生物実験材料名にも見られ,正確な計上を妨げている。被提供機関名においても,世界レベルのNISTEPのような名称バリエーションのリストが存在せず,この場合も正確な計上が困難であると思われる。