

大学・大学院での研究活動と文献検索の現場

—MIS38(2023)「継続教育」—

2023.7.2.
諏訪敏幸
mktb@tcct.zaq.ne.jp

略語と注意

注意: 本日の内容はすべて講演者(諏訪)の経験と考えに基づくものです。この講演は「継続教育」という枠組みの下でのものですが、ご紹介する現場事情や意見が国内外の検索実務において一般的あるいは模範的であることを保証しません。また本日の主催者や講演者の所属機関・団体の公式見解でもありません。下に示した用語の定義・解説も講演者流のものです。

略語	スペルアウト	慣用的な日本語訳、および意味
SR	systematic review	システマティック・レビュー ある問題について現時点までの研究結果を系統的に集約したもの。
ScR	scoping review	スコーピング・レビュー ある問題について現時点までの主要な研究状況を系統的に概観するもの
CPG	clinical practice guidelines	臨床ガイドライン 臨床上の問題について現時点で最も適切と考えられる指針を示すもの
RQ	research question	研究疑問 研究対象となる問題の核心を簡潔にまとめたもの
CQ	clinical question	臨床疑問 指針が必要な臨床上の問題点を簡潔にまとめたもの
PICO(s)	Population, Intervention, Comparison, Outcome (Study design)	(日本語なし) 介入研究のSR, ScR, CPGなどで研究対象となるRQ, CQなどの内容を意味的に分節化するためのフレーム(分析枠組)の一つ。他にも問題の性格に応じて種々の枠組がある。

自己紹介

- 1973~2013 大阪大学附属図書館
 - 1998~2013 生命科学図書館参考カウンターで主に文献検索相談を担当(15年間で6000件)
 - いくつかのSR研究、CPGプロジェクトに参加
- 2013~2021 大阪大学大学院人間科学研究科
 - 専門は知識社会学、科学計量学、科学社会学会員、博士(人間科学)
- 2013~ 各地の看護系大学院で文献検索・看護研究のコンサルテーションと講義
 - 併行してWHOや国内の看護系学会のプロジェクトに文献検索で協力
- 現職 大阪大学医学部保健学科招へい教員
京都大学非常勤講師、兵庫医科大学非常勤講師
国立成育医療研究センター共同研究員
- 主著: 『看護研究者・医療研究者のための系統的文献検索概説』(近畿病院図書室協議会, 2013)



今日取り上げる文献検索の「現場」

講演者(諏訪)の視点→	医学系図書館職員として(1998.4~2013.3)	元図書館職員として(2013.4~現在)	研究者として(2013.4~現在)
(主な研究分野)	看護学・医学	看護学	社会学・科学計量学
◆相談活動	セッション数 900セッション/年(2012) 標準的な相談時間 1.5時間/セッション 主な研究者層 大学院生、学部生、看護師、教員、一般市民、医師	50セッション/年 5時間/セッション 教員、大学院生、その他研究者、稀に学部生・医療職	1~5セッション/年 2時間/セッション 大学院生、卒業生
検索関連の主な内容	・先行研究調査 ・SR対象文献の検索 ・投稿誌選択・業績評価 ・査読対応・尺度調査 など	・先行研究調査 ・ScR対象文献の検索 ・概念分析のための文献調査 ・尺度調査 など	・関連研究の調査 ・研究テーマの評価 (周辺の大学院生からの相談は頻繁にあるが、大半は研究テーマ・方法・理論・概念の検討、博論・修論の議論展開や構成に関する精査などで、検索が関係するものは少ない)
◆SR/CPGでの検索協力(相談レベルのものは除く)	1~数件/年	5~10件	
◆その他の協力・支援	博論の引用文献チェック	博論の引用文献チェック 検索関連の質問への回答	
◆自分の研究	(図書館情報学研究)	(プロジェクト参加はあり)	文献分析

※「看護学」は看護学・保健学・助産学。「医学」は臨床医学の他に社会医学・遺伝学・生理学等も含む。以下同じ。

検索は準備と試行錯誤が8割

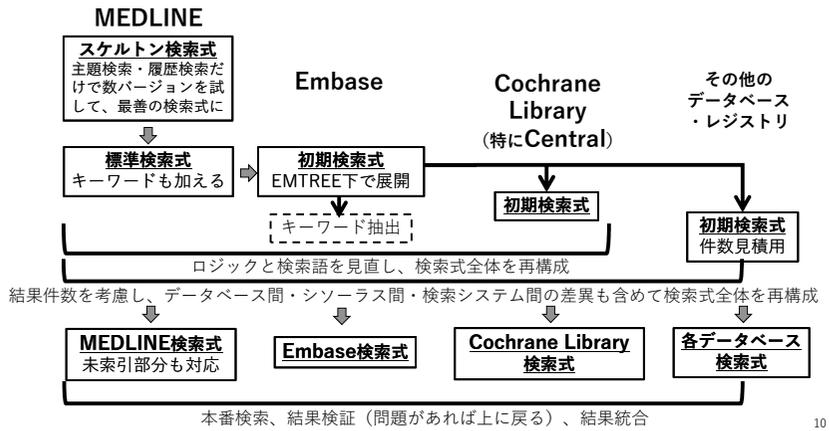
SR (プロトコル段階、プロジェクト遅延による再検索、および後年の改訂については省略)



先行研究調査での検索相談 (MEDLINEとCINAHLを検索する場合)



SRでは複数のデータベースにまたがる試行錯誤



#	Query	Results	24	(microcephal* or microlissencephal* or anencephal* or ((congenital or brain or cerebral or "white matter" or nerv* or neur*) adj3 (malformation* or abnormalit* or defect* or calcification or development* or growth))) .mp.	266351
1	exp Flavivirus/	16541	25	(cephalometr* or (head adj3 (circumference* or size))) .mp.	31460
2	exp Flavivirus Infections/	20393	26	(prenatal or antenatal or fetus or foetus or foetal or gestation* or intrauter* or pregnan* or "expectant mother*") .mp.	1E+06
3	(zika or flavivi* or "flavi vi*" or ...) .mp.	34755	27	embryo* .mp.	376522
4	or/1-3	35485	28	stillbirth* .mp.	10082
5	Chikungunya Virus/	1534	29	neurotroph* .mp.	34460
6	Chikungunya Fever/	584	30	or/24-29	2E+06
7	chikungunya .mp.	2689	31	11 and 30	1811
8	Arboviruses/	3169	32	31 not medline.st.	124
9	arbovirus* .mp.	8293	33	exp Nervous System Malformations/cn, em, ep, et	7727
10	or/5-9	10400	34	TORCH .mp.	930
11	or/4,10	42307	35	Toxoplasma/	11353
12	exp "Congenital, Hereditary, and Neonatal Diseases and Abnormalities"/	1E+06	36	exp Toxoplasmosis/	17866
13	exp "Embryonic and Fetal Development"/	229589	37	exp Viruses/	673651
14	exp Embryonic Structures/	391915	38	exp Virus Diseases/	783448
15	exp Nerve Growth Factors/	41137	39	Aedes/	11697
16	Cephalometry/	24107	40	or/34-39	1E+06
17	Maternal Exposure/	6172	41	33 and 40	243
18	exp Pregnancy/	784084	42	or/23,32,41	1959
19	Pregnant Women/		43	remove duplicates from 42	1954
20	Prenatal Care/	21837	44	43 not (animals.sh. not (humans.sh. or 39))	1216
21	exp Prenatal Diagnosis/	63551	45	limit 44 to (comment or editorial or news or ...)	45
22	or/12-21	2E+06	46	44 not 45	1171
23	11 and 22	1607			

Chibueze, E. C., Tirado, V., Lopes, K. D., Balogun, O. O., Takemoto, Y., Swa, T., Dagvadorj, A., Nagata, C., Morisaki, N., Menendez, C., Ota, E., Mori, R., & Oladapo, O. T. (2017). Zika virus infection in pregnancy: a systematic review of disease course and complications. *Reproductive Health*, 14, Article 28. <https://doi.org/10.1186/s12978-017-0285-6> ※MEDLINE (Ovid) の検索履歴。作図の都合上#3・#45のqueryを一部省略。

(前のスライドの検索式の説明)

- 背景
 - Zika熱は蚊によって媒介されるウィルス性の疾患で、1950年頃にウガンダで発見された。(12995440, 12005441 ... 文献はここでは省略のためPMIDのみ示す)
 - その後インドネシアなど熱帯各地に拡散した。(27708473)
 - この疾患に妊婦が感染すると胎児が小頭症になる危険がある。そのため、2016年のリオ・オリンピックに先立ってこの疾患に国際的に関心が集まった。
 - これに対応してWHOで急速対策会議が持たれ、会議参加各国は次回会議までにこれに関するSRを行って次回の討議材料とすることになった。
 - このSRは異例の短期間(約1か月)で完了させなければならず、そのため文献検索も1週間以下というこれも異例の短期間で行うことが求められた。
 - 日本チームはこのSRを期日までに完了し、その後論文化した。(Chibueze論文=28241773)
- 研究課題の特徴と変則的要素 ※Chibueze論文ではZika感染の小頭症リスクに絞って論文化している
 - Zika感染による小頭症リスクのメタアナリシスを主目的とする観察的SR
 - 小頭症を含む先天性異常を引き起こすTORCHなどの疾患をも併せて視野に入れる探索的性格も併せ持つ
- 検索課題と検索方法における注意事項 (#は検索履歴のステップ番号)
 - この検索の時点でZika VirusおよびZika Virus InfectionのカテゴリーはMeSHにはなかった。
 - 上位分類のFlavivirus属は、日本脳炎・黄熱病・デング熱など錚々たる疾患のウィルスを抱えるグループ
 - 短期間の内に結論を出さなければならぬため、結果件数を抑制したい。
 - Zikaは上位とキーワードで押さえ(#1-#11)、アウトカムで制約する(#12-#23)という変則的手法
 - 社会的注目が集まっている案件なので、新しい論文が急増している可能性がある
 - MeSH未索引(および索引対象外)データに対するキーワード検索を特に重視(#24-#32)

掛け合わせ (AND) の期待値計算で主題間の「距離」を見ながら

※ 1・2・・・をステップまたはブロック、|1|・|2|・・・をその件数、|W| をそのデータベースの全データ件数とする

- 方法1 (絶対的期待値) : 掛け合わせされる2つの検索要素間の親和性を見る
 - |1 AND 2| の絶対的期待値 = $|1| \times |2| / |W|$
 - データベース内で1と2が独立に分布しているという仮定の下での計算だが、そもそも1と2が関連あるという前提なので、これよりは多い結果が出るはず
 - 標準的な結果件数はこの期待値 $\times 10^{0.5 \sim 2.5}$ 程度 (通常は $\times 10^{1.5}$ 程度)
 - 実際の結果件数が標準範囲に比べて過小 → 何かミスマッチ?、またはあまり関心もたれていない?
 - 実際の結果件数が標準範囲に比べて過大 → 重大なノイズ源が存在?
- 方法2 (相対的期待値) : 組み合わせによる親和性の相対差を見る
 - $|3| = |2| \times n$ のとき、|1 AND 3|の相対的期待値 = $|1 \text{ AND } 2| \times n$ が期待値、これより上なら2より3の方が親和的、下なら2の方が親和的

17

先行研究調査で検索するのは (研究者にとって) 他人の研究

- 「自分の (研究者の) 関心から文献を探す」というより、
 - そもそもこの問題についてどんな研究がなされてきたのか
 - その中で自分の研究にできること、すべきことは何か
 - 自分の関心は他の研究との関係でどんな意義を持つかを探るのが先行研究調査、そのための検索
- 同じテーマでも、他人の研究は他人の問題意識、他人の視点
 - 他人を知り、自分の研究を相対化する、新たな気づきを得る

= 検索は探索

※SR/CPGでは、事前のスコーピングとPICO作成段階で探索的部分は解決しておく

18

「正しく検索できているのかわからない」??
 … 検索は探索、検索に「正解」はない

• ただし「まちがった」検索や、より良いーより悪い検索はある

19

ありがちな 神話・認知バイアス・無用な悩みにも答えつつ

- 曰く
 - 先行研究調査は自分と同じテーマの研究を見つけ出すのが目的
 - 検索結果が0件 → 価値がない★
 - 研究題目・RQ・PICOのキーワードはすべて検索に含める
 - ここ数年以内の原著論文以外は不要☆
 - 英語 (・日本語) 以外の読めない言語の文献は不要
 - 学会誌掲載論文は読む価値があり、商業誌や紀要掲載論文は不要★☆
 - SRでは会議録を除外する
 - SRではn数の少ない文献、p値の低い文献を除外する
 - SRでは灰色文献の検索が必須
 - ★は看護系に特有、
 - ☆は学部生・卒後3年研究などの研究初心者への指導としてなら必ずしも誤りではない
- 相談者 (研究者) にこれらの考え方が見えたら、それが誤りであること、またなぜ誤りか、どう考えるべきかを、そのつど丁寧に説明する

20

「誰もしていない研究」は看護研究者の恐怖だが…

- 単に検索方法がまちがってそう見えることも多い
 - 研究初心者（学部生、卒業3年研究）の場合は、本当に意義の無い研究や実行不可能な研究もある
→ これらはそれなりに対処
 - 修士生以上の場合、「誰もしていない」はそもそも見当違いの悩み
 - 「誰もしていない」のはそれ自身が悪いことではない（むしろ望ましい）
 - 本当に「誰もしていない」と思われる検索課題は、検索前に疑問点を詰め、仮説を立てて検索の方向性を考え、検索結果を見ながらさらに検討
 - 視点設定が狭すぎる → より広い視点で検索し直し、周囲との関係を検討
 - 状況設定が狭すぎる → 中心的な問題関心との関係での状況の意味・役割を考える
 - あまり関心を持たれていない問題 → 出発点となる問題関心を軸に、他の研究者・実務者との関心の接点を探る検索に切り替える
 - 日本語文献はあるが英語文献はない → 背景事情の違いを検討し、この問題を他国の社会的文脈にどう位置付けるかを考えるための検索をおこなう
(国際誌への投稿などを考えている場合は特に重要)
- etc.

探索はいつも手探り…何か頼れるものは？

- 国際的「標準」や「メソッド」…必ずしも適切なガイドとならない
 - (記述法の標準だが) PRISMA-S (<http://www.prisma-statement.org/Extensions/Searching>)
 - 私見では、16項目中、概ね適切な指示6項目、何らかの問題あり6項目、不適切4項目
 - 論文での記述はこれに従うべきだが、検索作業の指針としては一部ミスリーディング
 - Flow diagramでは理論的・実践的な混乱も見られる
- JBI 3 Step Search Strategy (<https://jbi-global-wiki.refined.site/space/MANUAL/4687278>)
 - 第3ステップはそれ自体としては妥当だが不十分
(書誌データベース検索がそれだけで完結するものでないことはKnipschild以来のセオリー)
Knipschild, P. (1994). Systematic reviews: Some examples. *BMJ*, 309(6956), 719-721. doi: 10.1136/bmj.309.6956.719
 - 第1ステップは常用の手法だが1つの手法に過ぎず、ここが出発点でもない
 - 第2ステップはこれだけでは意味をなさず、結局このメソッド自体が全体として無意味
- MINDS マニュアル (https://minds.jcqhcc.or.jp/s/manual_2020_3_0) 4.1
 - 昔日に比べて根本的に改善され、検索例も良いが、解説は初心者向けに留まっている。
- 他の指針 (Cochraneなど) や教科書もどれも不十分
- 国際誌の査読コメントもほとんどは力量不足、誤ったコメントも多い
- AIは?
 - 「検索式を作る」のは方向違い? 発想を広げるにはプラスかもしれない

検索システムは試しながら、データベースは観察しながら、シソーラスは常に確認して使う

- 結局、自力でその問題に即して適切な方針を見つけるしかない だが、
 - データベースも検索システムも無条件には信用できないし頼れない
 - 文献分布は生き物、データベースやシソーラスもそれによって変化
… システムの仕様や実際の動作・データベースの項目構成と実際の入力内容・シソーラスの構造はいつ変わるかわからない
 - システムの動作はHelpや説明文書の通りとも限らない
 - 例: 履歴演算の大文字問題、近接演算、MEDLINEの項目種類…
 - 系統的検索の特徴は、
 - ①主題検索を骨格とする、
 - ②検索課題の批判的分析とブロック化、
 - ③構成的検索、
 - ④コントロールと透明化、
 - ⑤評価・検証しながらの試行錯誤、
 - ⑥スクリーニングとの分担、
 - ⑦事後検証と補完・検索を孤立的に運用しない、
- 検索システムの動作やデータベース/シソーラスは常に注意を向けながら使用する; 事前・事中のテスト運用・テスト検索もしばしば行う

信頼できるデータベース/検索システムがほしい…

- 求めるのは
データベースと検索システムの信頼性・シソーラスとコーディングの質>検索と履歴管理の性能
» 採録件数>系統的検索に役立つ付加機能・データベース項目»» その他の付加機能
- データベースの信頼性 ※()内は講演者が常用する書誌データベース9種についての主観的評価の分布
 - エンドユーザに対して情報は十分に公開されているか? (○2△4×3)
 - データの採録方針は合理的か? 採録状況は概ね信頼が置けるか? (○5△1×3)
 - 主たる索引のコーディング基準は検索画面上で明確にされているか? (○1△2×5?1)
 - 主たる索引のコーディングには信頼が置けるか? (○1△5×1?2)
 - 書誌情報には概ね信頼が置けるか? (○7△1×1)
 - エンドユーザからの疑問や指摘を受け止める体制があるか? (○1△1×7)
- 検索システムの信頼性 ※()内は常用する検索システム7種についての主観的な評価の分布
 - エンドユーザに対して情報は十分に公開されているか? (○1△6)
 - 動作は仕様の範囲内で正常か? (○5×1?1)
 - 深刻な動作異常はエンドユーザに即時公開されるか? (○1×3?3)
 - 異常や不具合は迅速に修復しているか? (○2△1×1?3)
 - エンドユーザからの疑問や指摘を受け止める体制があるか? (○1△2×4)
- 系統的検索に適した検索システムの仕様・検索性能 (付加機能も含めて)
 - 多くの検索システムは発見的検索を「基本」と考えており、仕様上の問題も多い