

第386回 OUG ライフサイエンス分科会公開講座

～第18回情報プロフェッショナルシンポジウムINFOPRO2021 ver.～

The 「医学文献データベース」
その中身と違いを検証する（海外編）

2020年7月2日（金） 13:30～14:00

INFOSTA OUG ライフサイエンス分科会

本日の発表内容

1. はじめに ライフサイエンス分科会の紹介
2. 目的と調査対象のデータベースと調査内容
3. 調査結果
4. 結果のまとめと考察
5. 質疑応答
6. まとめ

日本オンライン情報検索ユーザー会 Online Users Group(OUG)

OUG ライフサイエンス分科会

<https://www.infosta.or.jp/research/oug-life/>



- 医薬分野から広くライフサイエンス全般を対象に検索演習、勉強会、見学会等を開催
- 開催日：毎月第3木曜午後、年間9回程度開催
- 開催場所：現在はWeb（以前は日本図書館協会会議室（茅場町））
- 本会はメンバー全員によるグループ（輪番）制で会の企画・運営を行っております。
- OUG入会手続きは入会申込（OUG/SIG）のページをご覧ください。
OUG会員の方で初めてライフサイエンス分科会に出席される場合は、予め事務局経由にて主査までご連絡ください。
- OUG非会員はビジターとして参加も可能です。（参加費：2000円）

目標： 有料のデータベースと無料のデータベースとを比較・ 検証して特徴を知り有効活用に繋げる

- 文献データベースは、データベースごとに作成機関や収録対象、収録方針、収録方法また検索機能などが異なる。
- こうした点を理解することにより、効果的で正しい活用できる。
- 今は無料で利用できるデータベースの活用が拡大し一般化している。
- そうした無料データベースと有料データベースとを比較検証し、それぞれの特徴を知り有効活用に繋げる。
また、エンドユーザーが接する機会の多い無料データベースについて理解を深めることで、有料データベースの価値や意義を確認し、適切な使い分けを実践する。

検証対象のデータベース

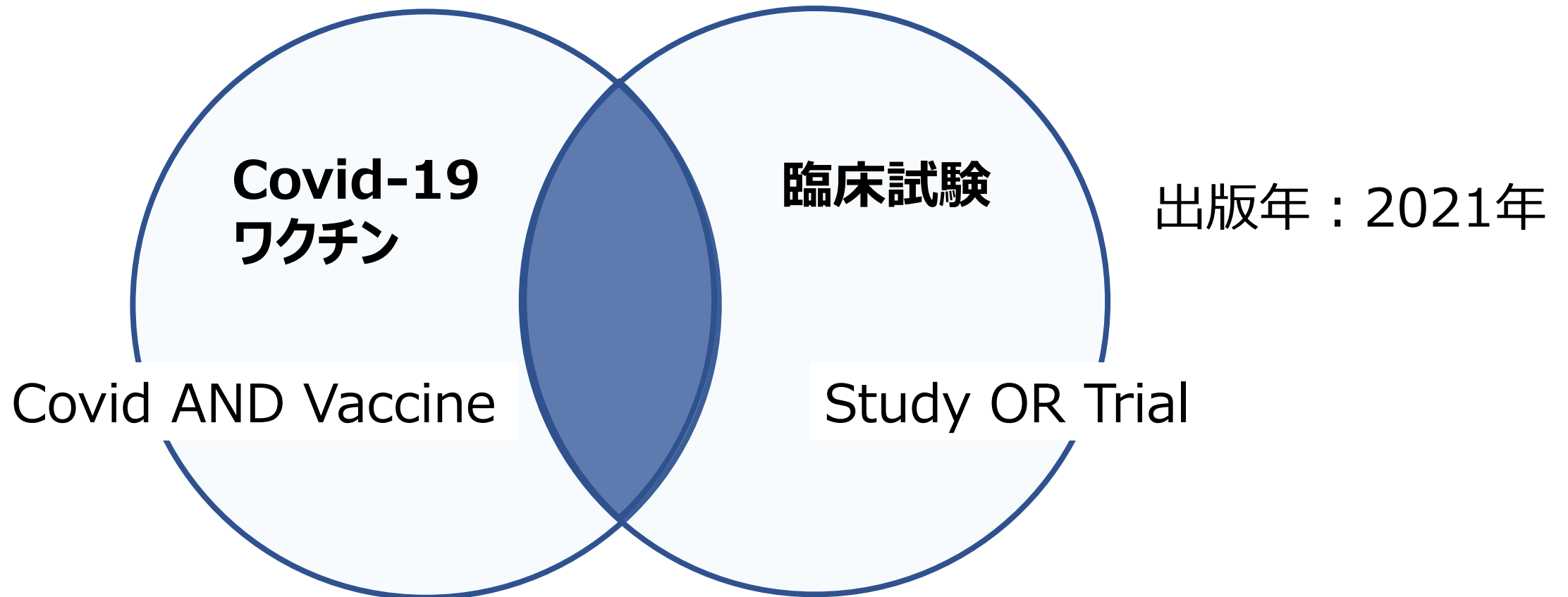
	PubMed	MEDLINE	EMBASE	GoogleScholar	CochraneCENTRAL
作成／提供機関	NLM（米国医学図書館）	NLM（米国医学図書館）	エルゼビア社	Google	国際的非営利団体コクラン／Wiley
収録年	1781～*1	1946～	1947～	不明	不明（サービス開始は1996年）
収録数	約3.210万件	約3,210万件	約3,960万件	不明	約174万件
収録誌*2	約33.966タイトル	約5,600タイトル	約8,100タイトル	不明（雑誌のほかWeb情報も収集している）	不明（PubMed、EMBASE、ClinicalTrials.gov, WHO等から収集）
利用	無料	有料	有料	無料	無料（検索）
一般公開	Web	専用サイト Dialog、EBSCO、 OVID、STNなど	専用サイト Dialog、EBSCO、 Embase.com、 OVID、STNなど	Web	Web

*1 MEDLINEと同じ1946年以降とされているが、出版年で検索してヒットした年を記載した。

*2 PubMedはリストをダウンロード、Medline・Embaseは各データベースの説明ページに掲載の情報より。GoogleScholarは明示されていないが雑誌のほかmedRxivなどのプレプリントや機関リポジトリの情報などもある。

検索式

- 検索式：「タイトル」について以下のキーワードで検索
(covid AND vaccine) AND (study OR trial)
- データベースごとの抄録やキーワードなどの索引情報の影響を受けないよう「タイトル」のみを検索した。また、対象年は2021年とした。



検索式および結果件数

検索は2021年4月6,7,13日に実施

- **PubMed 【87件】** (2021/4/13実施)

("covid*" [Title] AND "vaccine*" [Title] AND ("study" [Title] OR "studies" [Title] OR "trial" [Title] OR "trials" [Title])) AND (2021:2021 [pdat]) 87件

- **MEDLINE 【86件】 / EMBASE 【75件】** (Dialogを用いて2021/4/13実施)

ti(covid* AND vaccine*) AND ti(study OR studies OR trial OR trials) AND (pd(20210101-20211201))

集合	検索式	MEDLINE	EMBASE
S1	ti(covid* AND vaccine*)	1,625	1,541
S2	ti(study OR studies OR trial OR trials)	2,180,172	2,845,187
S3	S1 AND S2 and (pd(20210101-20211201))	86	75

- **GoogleScholar 【96件】** (2021/4/7実施)

allintitle: covid vaccine (study OR studies OR trial OR trials)

"年"は画面上のフィルターで指定

- **CochraneCENTRAL 【22件】** (2021/4/6実施)

(covid* AND vaccine* AND (study OR studies OR trial OR trials)):ti with Publication Year from 2021 to 2021, in Trials

検証内容

- タイトルや雑誌名、ISSN、doi、PMIDなどからデータベースの収録状況（重複収録or単独収録）を確認する。
- 各データベースの結果について、他データベースに収録されなかった（またはヒットしなかった）データがどのようなものかを確認してデータベースごとの収録の傾向を探る。
- 各データベースの機能や収録データの不明点も確認する。
- 4月の検索実施以降の各データベースの収録状況も確認する。

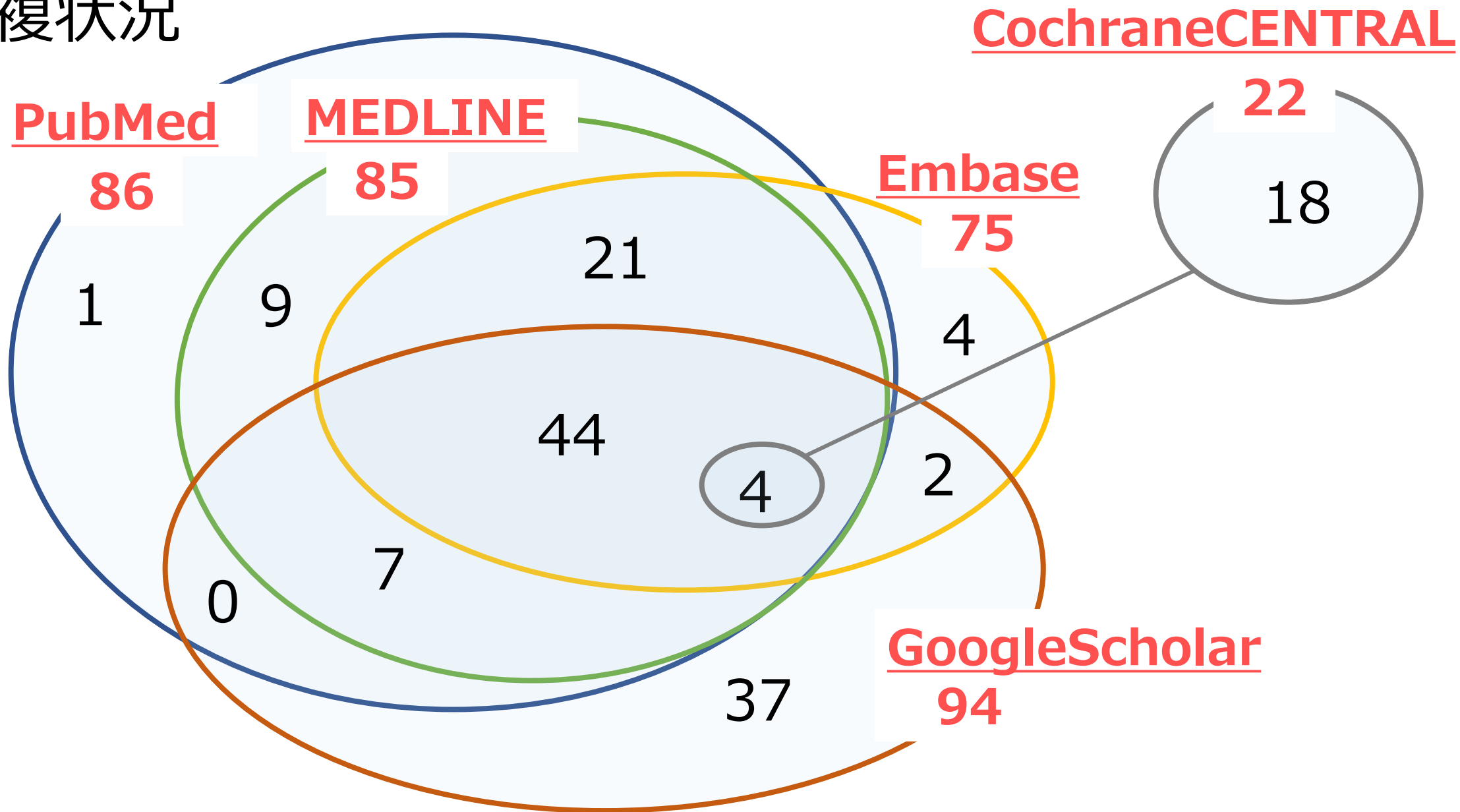
各データベースの検索結果と重複状況

結果件数	PubMed	MEDLINE	EMBASE	GoogleScholar	CochraneCENTRAL
ヒット件数	86(87)	85(86)	75	94 (96)	22
備考	重複2件 (1報)	重複2件 (1報)		重複収録3件 (1報)	

重複収録件数	PubMed	MEDLINE	EMBASE	Google Scholar	CochraneCENTRAL
PubMed	—	85	69	55	4
MEDLINE	85	—	69	55	4
EMBASE	69	69	—	50	4
Google Scholar	55	55	50	—	4
Cochrane CENTRAL	4	4	4	4	—

単独収録件数	10	9	4	37	18
--------	----	---	---	----	----

重複狀況



各データベースの確認結果 PubMed・MEDLINE

- PubMedが86件、MEDLINEが85件で、Embase（75件）より多くGoogleScholar（94件）より少なかった
- PubMedがMEDLINEよりも1件多かったほかは全く同じ結果だった（その1件はその後すみやかにMedlineに収録された）
- プレプリントのmedRxivが4件あった
- そのmedRxiv4件のうち1件は雑誌論文も同時にヒットしていたため、重複として除いて比較・検証することとした
- 4月以降にタイトルが修正されたデータが1件あった

各データベースの確認結果 **EMBASE**

- EMBASEは75件で、PubMed（86件）・MEDLINE（85件）、GoogleScholar（94件）よりも件数が少なかった
- medRxivなどのプレプリントはなく雑誌論文のみが収録されていた
- PubMed・MEDLINEでmedRxivと雑誌論文とで重複収録されていたデータは、雑誌論文のみがヒットした
- 4月以降にPubMed・MEDLINEでタイトルが修正されていたデータはEMBASEでもヒットしており、全く同じ修正がなされていた

各データベースの確認結果 **GoogleScholar**

- 94件と最も件数が多く、雑誌論文のほかプレプリントが多数あった
- GoogleScholar内での重複があった（3件・1報）
- PubMed・MEDLINEでmedRxivと雑誌論文とで重複収録されていたデータは、EMBASEと同様に、雑誌論文のみがヒットした
- PubMed・MEDLINE、EMBASEで4月以降にタイトルが修正されたデータはGoogleScholarでもヒットしていたが、その後もタイトル修正はなかった
- リストは「お気に入り」に保管しないと出力できず、手間がかかった
- 出力したリストは空欄が多く、1件でタイトル欄に抄録があるなど不均質だった
- また、タイトル欄に抄録があった論文のタイトルは今回の検索式でヒットする内容ではなかった（抄録が条件に合致しヒットしていた）

各データベースの確認結果

GoogleScholar

- GoogleScholarのみでヒットした37件の内訳

medRxiv 8、 LANCETのプレプリント誌 5、
Multidisciplinary Digital Publishing Instituteのジャーナル 5、
ResearchSquare 3、 arXiv 2、 F1000Research 2、 ResearchGate 2、
WHOのサイト 1、 大学機関アーカイブ 1、 その他 8

- その他 8件の内訳

- ① Biotechnology to Combat COVID-19 *
- ② Russian Journal of Infection and Immunity= Infektsiya i immunitet
- ③ International Journal of Research Studies in Medical and Health Sciences. *
- ④ Science Midwifery *
- ⑤ Anais Brasileiros de Estudos Turísticos-ABET *
- ⑥ Central African Journal of Public Health *
- ⑦ International Journal of Advanced Research in Economics and Finance *
- ⑧ International Association of Decision Sciences *

* はPubMedCatalogおよびEMBASE収録誌リストになし

各データベースの確認結果 **CochraneCENTRAL**

- 収録対象が疾患の治療や診断の論文に限られるため、検討対象DB中では22件と最も件数が少なく、そのうち臨床試験が22件中18件と大半を占め、また、残り4件の雑誌論文すべてが他データベースでもヒットしていた
- タイトルが同じ臨床試験が2件（1試験）あったが解析目的が異なるものと思われた
- 各臨床試験の文献化に関する情報は把握できないが、有料契約者向けには関連文献の表示機能がある（ただし、該当する臨床試験以外の論文情報も表示されていたため参考情報と思われた）
- またリスト出力機能も有料契約者向けのみであり、Dialogほどの機能はないものの、GoogleScholarよりは使いやすくリストには空欄などもなかった

各DB収録のその後の状況

(5月以降のヒット状況)

PubMed・MEDLINE その後

EMBASEにあるがPubMed・MEDLINEになかった6件について

- その後ヒットしたデータが2件：
1件は追加収録だったが、もう1件はMEDLINE収録日が2021年4月以前であったため、MEDLINE上でタイトル修正があったものと思われた
- MEDLINE上の出版年が2020年のためヒットしなかったものが1件
- その後もヒットしなかったデータ3件の雑誌タイトル：
 - ① Flora (1300-932X)
 - ② Journal of the International AIDS Society (1758-2652)
 - ③ VacciMonitor (1025-028X)

すべてPubMedCatalogにある収録誌
3件ともGoogleScholarでもその後の収録なし

PubMed・MEDLINE その後

GoogleScholarにありPubMed・MEDLINEになかった39件について

- MEDLINE上の出版年が2020年のためヒットしなかったものが1件
- その後ヒットした6件：同じ書誌での収録が3件、書誌が異なるものが3件
- その後もヒットしなかったもの32件の内訳：
medRxiv 8、 arXiv 2、LANCETのプレプリント誌 2、
Multidisciplinary Digital Publishing Instituteのジャーナル 3、ResearchSquare 3、
F1000Research 2、ResearchGate 2、WHOのサイト 1、大学機関アーカイブ 1、
残り8件の雑誌タイトル：
 - ① Biotechnology to Combat COVID-19 *
 - ② Russian Journal of Infection and Immunity= Infektsiya i immunitet
 - ③ International Journal of Research Studies in Medical and Health Sciences *
 - ④ Science Midwifery *
 - ⑤ Anais Brasileiros de Estudos Turísticos-ABET *
 - ⑥ Central African Journal of Public Health *
 - ⑦ International Journal of Advanced Research in Economics and Finance *
 - ⑧ International Association of Decision Sciences *

アンダーラインはEMBASEでもその後の収録なし *はPubMedCatalogになし

EMBASE その後

PubMed・MEDLINEにありEMBASEになかった17件について

- その後ヒットが8件：追加収録と思われるが、うち1件はMEDLINEではmedRxivだったものがEMBASEでは雑誌論文だった
- その後もヒットしなかった9件中medRxivが2件、残り7件の雑誌タイトル：
 - ① JMIR medical informatics(2291-9694) *
 - ② Journal of molecular liquids(0167-7322) *
 - ③ Journal of medical virology(0146-6615)
 - ④ Nature(0028-0836)
 - ⑤ Nature reviews. Clinical oncology(1759-4774)
 - ⑥ Neurology(0028-3878)
 - ⑦ Vaccine(0264-410X)

アンダーラインはGoogleScholarでもその後の収録なし *はEMBASEの収録リストになし

EMBASE その後

GoogleScholarにありEMBASEになかった44件について

- その後ヒットの8件：同じ書誌が5件、書誌が異なるものが3件
 - その後もヒットしなかったものが36件：
medRxiv 10、 arXiv 2、LANCETのプレプリント誌 2、
Multidisciplinary Digital Publishing Instituteのジャーナル 3、ResearchSquare 3、
F1000Research 2、ResearchGate 2、WHOのサイト 1、大学機関アーカイブ 1、
残り10件の雑誌タイトル：
 - ① Vaccine
 - ② Journal of molecular liquids *
 - ③ Nature
 - ④ Biotechnology to Combat COVID-19 *
 - ⑤ International Journal of Research Studies in Medical and Health Science *
 - ⑥ Science Midwifery *
 - ⑦ Anais Brasileiros de Estudos Turísticos - ABET *
 - ⑧ Central African Journal of Public Health *
 - ⑨ International Journal of Advanced Research in Economics and Finance *
 - ⑩ International Association of Decision Sciences *
- アンダーラインはPubMed・MEDLINEでもその後の収録なし *はEMBASEの収録リストになし

GoogleScholar その後

PubMed・MEDLINEにありGoogleScholarになかった31件について

- その後の収録が18件あり、うち1件はResearchGateで、PubMed・MEDLINEと異なっていた
雑誌タイトル Clinical infectious diseases : an official publication of the Infectious Diseases Society of America (1058-4838)
- 出版年が2020年のためにヒットしていなかったものが7件あった
- その後も収録されなかった6件の雑誌タイトル
 - ① Journal of clinical pharmacology (0091-2700)
 - ② International journal of clinical practice (1368-5031)
 - ③ Nature (0028-0836)
 - ④ Nature reviews. Clinical oncology (1759-4774)
 - ⑤ JMIR public health and surveillance (2369-2960)
 - ⑥ Public Health Nurs. (0737-1209)

アンダーラインはEMBASEでもその後も収録されていなかった ただしEMBASE収録リストにあり

GoogleScholar その後

EMBASEにありGoogleScholarになかった25件について

- その後の収録が12件あり、うち1件はResearchGateで、EMBASEとは異なっていた

雑誌タイトル Clinical infectious diseases : an official publication of the Infectious Diseases Society of America (1058-4838)

- 出版年が2020年のためにヒットしていなかったものが7件あった

- その後も収録なしが6件 その雑誌タイトル

- ① Flora (1300-932X)
- ② Journal of the International AIDS Society (1758-2652)
- ③ International Journal of Clinical Practice (1368-5031)
- ④ VacciMonitor (1025-028X)
- ⑤ Nature (1476-4687)
- ⑥ JMIR public health and surveillance (2369-2960)

アンダーラインはMEDLINEでもその後も収録なし ただしPubMedCatalogにある収録誌

収録データの違いについて（詳細）
重複収録／タイトル／出版年

PubMed・MEDLINEの重複データ

同じ文献1報が、**雑誌論文**と**プレプリント**として別々に計2件収録されていたもの

- 雑誌論文
Gray KJ, COVID-19 vaccine response in pregnant and lactating women: a cohort study.
[Am J Obstet Gynecol. 2021 Mar 24:S0002-9378\(21\)00187-3.](#)
doi: 10.1016/j.ajog.2021.03.023. Online ahead of print.
【PMID : 33775692】
- プレプリント
Gray KJ, COVID-19 vaccine response in pregnant and lactating women: a cohort study.
[medRxiv. 2021 Mar 8:2021.03.07.21253094.](#)
doi: 10.1101/2021.03.07.21253094. Preprint.
【PMID : 33758889】

タイトル変更 PubMed・MEDLINE

- EMBASEにありPubMedMEDLINEになかったデータが、PubMed・MEDLINEで5月に新たにヒットしたが、収録日が2020年10月であったため、PubMed・MEDLINE上で4月以降にタイトル修正があったと思われる

Informed consent disclosure to vaccine trial subjects of risk of COVID-19 vaccines worsening clinical disease.

Timothy Cardozo, Int J Clin Pract. 2021 Mar;75(3):e13795.

doi: 10.1111/ijcp.13795. Epub 2020 Dec 4.

【PMID: 33113270】 **最初の収録日: 2020-10-29**

Dialogアップデート: 2020-10-29 2020-11-07 2020-12-04 2021-01-18 **2021-05-08**
2021-05-10

- なおEMBASE上で4月～5月にかけてタイトル変更はなかった
- GoogleScholarで本データは4月の時点でヒットしなかったため比較できない

タイトル変更 PubMed・MEDLINE, EMBASE

- 4月 タイトル変更前

COVID-19 vaccine response in pregnant and lactating women: a cohort study.

Gray KJ, Am J Obstet Gynecol. 2021 Mar 24:S0002-9378(21)00187-3. doi: 10.1016/j.ajog.2021.03.023. Online ahead of print.
【PMID : 33775692】

- 5月 タイトル変更後

Coronavirus disease 2019 vaccine response in pregnant and lactating women: a cohort study.

Gray KJ, Am J Obstet Gynecol. 2021 Mar 24:S0002-9378(21)00187-3. doi: 10.1016/j.ajog.2021.03.023. Online ahead of print.
【PMID : 33775692】

- GoogleScholarはタイトル変更なく4月のままだった

出版年の違い PubMed・Medline, EMBASE

- **MEDLINE上の記述：**

Prioritizing incarcerated populations for COVID-19 vaccination and vaccine trials.

EClinicalMedicine. 2020 Dec 23;31【PMID: 33385124】

最初の収録日: 2021-01-01

Dialogアップデート: 2021-01-01 2021-01-02 2021-01-03 2021-01-29 2021-01-30

- **EMBASE上の記述：**

Prioritizing incarcerated populations for COVID-19 vaccination and vaccine trials.

EClinicalMedicine. Jan 2021; 31

最初の収録日: 2021-01-08

Dialogアップデート: 2021-01-08 2021-05-03

- その後GoogleScholarでもヒットし、GoogleScholarの出版年は2021年だった。

出版年の違い GoogleScholar

- GoogleScholarで2020年だったためにヒットしなかった7件のDialog上のPublication Date

雑誌タイトル	MEDLINE	EMBASE
Emerging microbes & infections	Dec 2021	2021
The Lancet. Infectious diseases	Apr 2021	Apr 2021
Human vaccines & immunotherapeutics	Mar 4, 2021	2021
Annals of internal medicine	Feb 2021	Feb 1, 2021
Human vaccines & immunotherapeutics	Feb 1, 2021	2021
JAMA internal medicine	Feb 1, 2021	Feb 2021
Drug safety	Jan 2021	Jan 2021

結果のまとめと考察

結果のまとめ 1 収録対象について

➤収録について：

- MEDLINEとPubMedではわずかなタイムラグはあるほかは同じであった
- MEDLINE・PubMed・GoogleScholarにはmedRxivが収録されていた
- EMBASEにはmedRxivの収録はなかった
- EMBASEにはMEDLINEも収録されるはずだがタイムラグがあるようだった
- このほかにも、データベースごとに収録のタイミングが違っていた
- GoogleScholarにはプレプリントが多い点をはじめ他データベースと収録対象に違いが認められたが、逆に、PubMed・MEDLINE・EMBASEにはあってもGoogleScholarにはないものがあった
- CochraneCENTRALは文献のほか臨床試験情報があることを確認した

結果のまとめ 2 収録内容について

➤タイトルについて：

- PubMed・MEDLINEで収録当初に本来のタイトルが登録されず修正されたと思われるデータがあった
- 一旦、登録されたデータの修正が、インデキシング部分だけでなく「タイトル」においても行われていたデータがあった

➤出版年について：

- 出版年がデータベースにより異なる場合があった

➤重複について：

- プレプリントが雑誌論文として掲載されることにより、データベース上に別データとして新たに収録されるものがあった

考察

- データベースごとに収録誌が異なることは予め想定されたが、同じ雑誌論文でも出版年やタイトルが異なるケースがあり注意が必要である
- タイトルのような重要な情報でも収録後に変更があり、電子入稿やプレプリントの拡大により、データベース上のデータも流動的な様子が窺えた
- 検索に際しては、検索式と検索実施日、結果の出力といったエビデンスの保管が重要である
- 出版年やタイトルのブレは作成者側での調整・統一が望ましい
- 同じ論文の重複収録についても、防止策または利用者向けに何らかのガイドが望ましい

質疑応答 と まとめ