

令和4年度厚生労働省 老人保健健康増進等事業

「介護DBの解析・利活用を実践しうる人材を育成するためのプログラム作成等に関する調査研究事業」

「介護DBを用いた研究を開始する際のガイド」

令和5年3月

つくば医療介護サービス研究機構株式会社

内容

はじめに.....	5
I. 介護 DB の利用申請について.....	6
1. 介護 DB 情報サイトとは.....	6
2. 介護 DB 申出の手順.....	8
1) 申出の概要.....	8
2) 事前相談とスケジュール.....	12
3) 申出における留意点.....	13
4) 申請に必要な書類.....	14
3. まとめ.....	18
参考資料.....	18
II. 介護 DB のデータ構造.....	20
1. 介護 DB のデータ構造を理解する.....	20
1) 介護 DB 情報はデータ種類ごとに異なるデータテーブルに分割保存されている.....	20
2) 各データテーブルは各データテーブルに固有の情報（変数）を含んでいる....	22
3) 各データテーブルは利用者 ID などの値をキーとして結合が可能である.....	23
4) 必要なデータテーブルと変数を事前を選択する.....	24
2. 介護 DB に含まれているデータ項目について.....	24
1) 「匿名要介護認定情報」に含まれるデータ項目.....	25
2) 「匿名介護レセプト等情報」に含まれるデータ項目.....	27
3) 「匿名 LIFE 情報」に含まれるデータ項目.....	28
3. まとめ.....	30

4. 引用.....	30
III. 介護 DB を使用した研究例の紹介.....	31
1. 研究事例①：介護保険施設におけるケアの質を評価した研究（利用者の要介護重 度化に関連する介護福祉施設の特徴）.....	32
書誌情報.....	32
1) 研究の動機.....	32
2) 必要な作業.....	32
3) 手順.....	33
4) 介護 DB の長所・短所.....	36
5) 工夫した点.....	36
2. 研究事例②：介護報酬給付実績データウェアハウスの開発.....	38
書誌情報.....	38
1) 概要.....	38
2) 得られた知見.....	38
3) 初学者へのフィードバック.....	39
3. 研究事例③：自己負担率と介護サービスの利用量との関連を検討した研究.....	40
書誌情報.....	40
1) 概要.....	40
2) データ処理.....	40
3) 初学者へのフィードバック.....	40
4. 研究事例④：在宅介護のフォーマルケア時間の推計.....	41
書誌情報.....	41
1) 概要.....	41
2) データ処理.....	41

3) 初学者へのフィードバック	42
5. 研究事例⑤: 病院看取り率の経年変化 (在宅介護利用者と施設介護利用者の比較)	43
書誌情報.....	43
1) 概要	43
2) データ処理	43
3) 初学者へのフィードバック	44
6. 研究事例⑥: 介護保険受給者台帳の資格喪失記録は死亡の代理変数となるか...	45
書誌情報.....	45
1) 概要	45
2) データ処理	45
3) 初学者へのフィードバック	45

はじめに

- 厚生労働省は介護保険総合データベース（以下、介護 DB）で収集した匿名要介護認定情報、匿名介護レセプト等情報、匿名 LIFE 情報の第三者提供を平成 30 年度に開始している。
- 介護 DB には、高齢者が利用した介護サービスの詳細情報や高齢者の状態像について貴重な情報を多く含んでおり、幅広い主体による利活用により、社会への還元に資する成果が求められている。
- 本ガイドは、介護 DB の活用を検討している若手研究者や介護現場スタッフ等を対象に、介護 DB の活用を開始する前に理解しておくべき内容を解説した入門書である。
- 介護 DB を適切に分析するためには、まず必要なデータテーブルやデータ項目を適切に選択し、厚生労働省へ利用申請を行う必要がある。そのためには、まず介護 DB のデータ構造や申請方法の理解が不可欠である。また、実際に介護 DB を活用した事例を通して、注意すべき点等を理解することも重要である。
- 本ガイドが、これらの理解の一助となれば幸いである。

本ガイドは令和 4 年度厚生労働省老人保健健康増進等事業「介護 DB の解析・利活用を
実践しうる人材を育成するためのプログラム作成等に関する調査研究事業」の一環で作
成しました。

1. 介護 DB の利用申請について

1. 介護 DB 情報サイトとは

介護 DB 情報サイトとは、厚生労働省によって介護 DB の提供に関する情報がまとめられている web サイト¹⁾ (以下、介護 DB 情報サイト) であり、正式名称は「匿名介護情報の提供について」となっている。1～7 までの介護 DB 情報に関する情報や必要書類がダウンロードできるようになっており (図 1)、このサイトに収載されている資料を見れば介護 DB 申請に関する内容が網羅できるようになっている。

匿名介護情報等の提供について

※本ページは令和2年10月1日以降の手続きに関するページです。令和2年9月30日以前の要介護認定情報・介護レセプト等情報の提供に係る手続きについては、[こちら](#)をご覧ください。

1 匿名介護情報等の提供等について 匿名介護情報等の提供等について(PDF形式:2.109KB) 匿名介護情報等の第三者提供に係る手続きについて(PDF形式2.805KB)	介護DBやNDBとの連携について、提供されるデータ形式、介護DB等の第三者提供に係る手続きについて等、介護DB帝位教に関する概要のpdfが掲載されている。
2 ガイドライン ガイドライン [PDF形式:1.177KB]	データの提供に係る事務処理の明確化及び標準化並びに審査の基準を定めたガイドラインのpdfが掲載されている。
3 提供申出・事前相談様式 事前相談様式(XLS:18KB)	提供申出に関する事前相談の様式が掲載されている。この様式をダウンロードし、必要事項を記載した上で、事前相談として事務局へ提出する。
4 申出に必要な書類 様式1 匿名介護認定情報等の提供に関する申出書 [XLS形式:97KB] 様式1-2 匿名介護認定情報等の提供に係る手数料免除申出書 [XLS形式:22KB] 別添8 申出依頼テンプレート (抽出) 申出依頼テンプレート (抽出) V2.0 [XLS形式:249KB] 別添8 申出依頼テンプレート (集計) 申出依頼テンプレート (集計) V2.0 [XLS形式:255KB] 第三者提供用データベース コード定義表V2.0 [XLS形式:288KB] サービス項目コードにつきましては、下記のリンク (WAM NET) をご参照ください。 https://www.wam.go.jp/gyosei/Shiryou/detail-list?bun=020050010 承諾通知が出た後に必要な書類 様式3 匿名介護認定情報等の利用に関する依頼書 [DOC形式:46KB] 様式5 匿名介護認定情報等の利用に関する誓約書 [DOC形式:86KB] 様式6 匿名介護認定情報等の受領書 [DOC形式:50KB] 研究終了時に必要な書類 様式10 匿名介護認定情報等のデータ積層報告書 [DOC形式:58KB] 様式12 匿名介護認定情報等の利用実績報告書 [DOC形式:48KB] 申請変更時等に必要な書類 様式7 所屬等変更届出書 [DOC形式:47KB] 様式8 匿名介護認定情報等の提供に関する申出書の記載事項変更依頼申出書 [DOC形式:50KB] 様式13 匿名介護認定情報の提供に関する申出書の経過届出書 [DOC形式:47KB]	申出等に必要な書類のうち、指定の書式である様式1、様式1-2、別添8が掲載されている。また、承諾通知が出た後に必要な書類、研究終了時に必要な書類、申請変更時に提出が必要な書類も掲載されている。作成に際してはここからダウンロードする。
5 審査スケジュール 審査スケジュール [PDF形式:146KB]	審査スケジュールが記載されたpdfが掲載されている。
6 関連する審議会 社会保障審議会 (介護保険部会匿名介護情報等の提供に関する専門委員会) 社会保障審議会 (医療保険部会・介護保険部会) (匿名医療・介護情報等の提供に関する委員会)	介護DB関連の審議会のリンクが貼られている。
7 介護DBオープンデータについて 介護DBオープンデータ	介護DBオープンデータのリンクが貼られている。

図1 介護DB情報サイトの各項目とその内容

2. 介護 DB 申出の手順

以下では申出の概要、事前相談とスケジュール、申請に必要な書類について、介護 DB 情報サイトと対応させながら説明する。

1) 申出の概要

申出～承諾までの流れ、申出に関連する用語の定義やデータの提供等、基本的な事項について説明していく。介護 DB 情報サイトでは「1 匿名介護情報等の提供について」「2 ガイドライン」に記載されている。

<p>1 匿名介護情報等の提供等について</p> <p>匿名介護情報等の提供等について[PDF形式：2.109KB] 匿名介護情報等の第三者提供に係る手続きについて[PDF形式2.805KB]</p>	<p>2 ガイドライン</p> <p>ガイドライン [PDF形式：1.177KB]</p>
---	--

・申出～承諾までの流れ

まず申出に先行して、介護DB第三者提供事務局（以下、事務局）に事前相談をする必要がある。この事前相談によって研究計画の詳細決定や申出書類の作成相談を行い、提供申出書及び添付書類の最終確認を行う。完成した書類を締め切り前に提出し、提出された書類は事務局及び専門委員会にて研究目的の公益性・分析実施可能性・個人の識別可能性等について審査される。専門委員会での議論を踏まえ、申出に対する承諾または不承諾の決定がなされる。「承諾」と判断された場合は、事務局より「承諾通知書」および「誓約書」等の提出依頼が届くので各種手続きを進める。なお、事務局が示す修正事項に対応した時点で正式な承諾となる「条件付き承諾」と判断される場合もある。

「審査継続」と判断された場合は、専門委員会で指摘された情報が記載された事務連絡が送付され、「不承諾」と判断された場合は、そのままの研究デザインでは提供できないということなので、改めて詳細を考え直す必要がある。

・用語の定義

ガイドラインでは「取扱者」「担当者」「提供申出者」がそれぞれ定義されている。ここでいう「取扱者」とは介護 DB を扱う者のことであり、「担当者」とは筆頭の「取扱者」、「申出者」とは「取扱者」と「担当者」が所属する組織のことである。つまり「申出者」は、人物ではなく組織であることに注意が必要である。なお、申出者の範囲は①公的機関（国の行政機関、都道府県及び市区町村）、②大学その他の研究機関（大学および研究開発独立行政法人等）、③民間事業者等となっている。

その他の定義についてはガイドライン p1~3 に記載されているので参照されたい。

・提供形式

2023 年 4 月現在、提供形式は「特別抽出」「サンプリングデータセット」「集計表情報」の 3 種類がある（表 1）。

図1 提供するデータの種類²⁾

	特別抽出	サンプリングデータセット	集計表情報
基本的なイメージ	申出者の指定した抽出条件に従って匿名要介護認定情報等をデータベースから抽出する。	予め一定程度の割合で抽出した匿名介護レセプト等情報に対して、さらに安全性に配慮した工夫を施した上で提供される。	匿名要介護情報認定情報について、提供申出者の指定した抽出条件及び集計条件に従って抽出したデータに対して一定の集計処理を加えた集計表を提供する
提供データ	個票	一部匿名化等を行った個票	集計表
含まれているデータ項目例	匿名要介護認定情報・匿名介護レセプト等情報に含まれている、ほぼすべての項目	希少な情報があらかじめ匿名化・削除された介護レセプト情報等	集計表
時系列での分析	可能	不可能	集計軸に時系列を含めれば可能
地域単位の分析	可能（市町村単位まで）	可能（市町村単位まで）	可能（市町村単位まで）
想定される取扱者像	匿名要介護認定情報等を用いた研究に一定の知見があり、申出内容や抽出条件を吟味し、大量のデータを高速に処理することを想定している取扱者	匿名要介護認定情報等を用いた研究に関心はあるが経験がまだ十分でなく、データの特徴や各項目の概要を把握したいと考えている取扱者	集計された結果を必要とし、データ処理を行うことを想定していない取扱者
利用にあたり具備すべきセキュリティ	データ利用時に、情報セキュリティマネジメントシステムを確実に運用できる利用環境を整える	特別抽出で求められるセキュリティ水準と比較してある程度具備しやすいセキュリティ水準での利用が可能	

「特別抽出」は利用者が事前に申告した条件で提供されるデータ形式である。個票で提供され、また、複数年のデータの提供が可能であり、介護DBの研究利用として多く想定される形式である。しかしながら、提出書類が多く、提供に際しては申請毎に研究内容、抽出期間及びセキュリティ要件について、専門委員会で審査を行う必要がある。特に、対象者が極めて限定される可能性がある場合、多数の項目を用いた探索的研究の場合、「全ての変数を求める」という要望の場合、複数の研究が1申出に盛り込まれている場合などは審査会で慎重に審査される可能性がある。また、患者数やレセプト数などが10未満の値は公表不可であり、データ利用時のセキュリティ要件が定められている（「医療情報システムの安全管理に関するガイドライン」³⁾に準じた措置が必要）。また、データ提供までの期間は、審査会の承諾後1年程度を要している。

「サンプリングデータセット」は特別抽出とは異なり一部のデータ（「基本情報レコード」「明細情報レコード」「居宅サービス計画費情報レコード」）について、1ヶ月分のデータから無作為に抽出し、さらに安全性に配慮した工夫（10件未満のサービスコード

を匿名化、事業所番号・担当介護支援専門員番号の空欄化、年月日を含む項目は一律で「〇月 1 日」として提供、などの匿名化処理)を施した上で提供されるデータ形式である。データは 1 ヶ月分の提供となっているため時系列での分析は不可能である。特別抽出と比較しこのようなデメリットがあるものの、専門委員会での審査では探索的な研究を認める等、特別抽出における審査ほどの研究目的の限定は求められない。データ利用時のセキュリティ要件も特別抽出で求められる水準よりは少なく、また審査会の承諾後は比較的早くデータが提供されるというメリットもある。

「集計表情報」は提供申出者の申出に応じて、一定の集計を加えた上で、集計表情報として提供されるデータ形式である。集計表情報は提供申出者の申出に従い、厚生労働省が最も狭い地域性の集計単位を市町村として一定の集計を加え、原則として、内容が簡易なものであって表数も少数であるものに対して行うこととしている。集計表情報では、その情報から個人を特定することはほぼ不可能であるものの、他データとの照合により照合したデータの個人が特定される可能性を否定できないことから、専門委員会による審査を経ることとなっている。

・手数料

提供申出にかかる手数料は、人件費等を踏まえた時間単位の金額（1 時間まで毎に 5,900 円）に、作業に要した時間を乗じて得た額となっている。作業に要した時間とは、申出処理業務（申出書類確認・専門委員会への諮問手続・データの抽出条件の精査等）とデータ抽出業務（SQL 作成・テスト実施・結果の検証等）に要した時間となっている。提供申出に係る全ての提供申出者が、① 公的機関、② 大学そのほか研究機関または民事事業者等のうち補助金等を充てて匿名要介護認定情報を用いて研究又は業務を行う者、③ ①又は②から当該申出に係る業務の委託を受けた者（再委託を含む）、のいずれかに該当する場合には、手数料は免除される。

ただし、手数料の免除・非免除については、確認すべき詳細な規定がある（国からの補助金でない場合には免除されない等）ので事務局に必ず確認する必要がある。

・NDB との連結データの申出を行う場合

NDB との連結データの申請をする場合は事務局にその旨を伝え、NDB 申請の事務局に対しても同様の申出を別途行う必要がある。それにより、共通の ID が付与された NDB、介護 DB が提供される。

2) 事前相談とスケジュール

次に事前相談とスケジュールについて説明していく。介護 DB 情報サイトでは「3 提供申出・事前相談様式」「5 審査スケジュール」に記載されている。



事務局への事前相談は、提供申出書の提出に先立ち必ず行う必要がある。事前相談の書式（図 2）は「3 提供申出・事前相談様式」からダウンロード後、必要事項を記載する。事前相談では、申出書の書き方や集計項目の設計等について疑問点を相談し、研究計画の詳細を決定する。その後、公表されている締切日に間に合うように提供申出書を提出する。その後、事務局によって提供申出書の確認がなされ、専門委員会で審査される。2022 年度における提供申出書や審査会のスケジュールを表 2 に記載した。事前相談開始～研究計画書の詳細決定までで約 1 ヶ月～2 ヶ月かかり、また、各審査日程に応じた事前相談の締切日も設定されているため、なるべく早めに事前相談を開始すると良いだろう。

匿名要介護認定情報等の提供に関する事前相談	
相談者 所属機関名:	
相談者 氏名:	
相談者 メールアドレス:	
事前相談番号(再相談・追加相談で番号が発行済みの場合):	
NDBとの連結の有無について (当てはまるものに「■」)	<input type="checkbox"/> 連結利用を予定 <input type="checkbox"/> 連結利用の予定なし <input type="checkbox"/> 未定
相談内容	
※質問を簡潔かつ具体的にご記入ください	

図2 事前相談の書式

表2 2022年度における各種スケジュール

審査月	提出締切日※	事前相談の締切日
2022年6月	2022年4月28日(木)	2022年4月15日(金)
2022年9月	2022年7月29日(金)	2022年7月8日(金)
2022年12月	2022年10月28日(金)	2022年10月7日(金)
2023年3月	2023年1月27日(金)	2023年1月6日(金)

※事前相談を経た上での、書類の最終提出締切日

3) 申出における留意点

申出に際し、重要な点・注意すべき点を以下に述べていく。

1点目は、とにかく早めに事務局に連絡するということである。単純な質問であったとしても、不明点があれば早めに事務局に連絡した上で申出の対応を進めていくことをおすすめする。

2点目は、「どのデータを」「何のために」使うかを明確にするということである。研究計画を構造化し、それを根拠としてデータを指定することが必要である。とはいえ、最初は難しいので、経験者と協働するとよいだろう。

3点目は、セキュリティ要件や学内プロセスを要確認するということである。セキュ

リティ要件は審査会で審査されるポイントでもあるため、詳細に記載する必要がある。
また、別添 1 申出者の証明書は、共同研究の場合は組織の数だけ必要であるため、発行に時間を要すことも考慮して、締め切りまでに間に合うように対応する必要がある。

4 点目は、承諾後の話になるが、承諾後の書類提出は早急に対応するというのである。特に「様式 3. 匿名要介護認定情報等の利用に関する依頼書」と「様式 5 匿名要介護認定情報等の利用に関する誓約書」の両方が受理されてからデータの抽出作業に入るため、承諾後は速やかに提出する必要がある。

最後は、承諾から提供までにタイムラグがあるということである。現時点で 1 年程度の待ち時間があり、すぐに研究を始められるわけではないという点に留意が必要である。

4) 申請に必要な書類

最後に申請に必要な書類について説明していく。介護 DB 情報サイトでは「4 申出に必要な書類」に記載されている。

4 申出に必要な書類

[様式1.匿名要介護認定情報等の提供に関する申出書 \[XLS形式：97KB\]](#)

[様式1-2.匿名要介護認定情報等の提供に係る手数料免除申出書 \[XLS形式：22KB\]](#)

別添8.申出依頼テンプレート（抽出）

[申出依頼テンプレート（抽出） V2_0 \[XLS形式：249KB\]](#)

別添8.申出依頼テンプレート（集計）

[申出依頼テンプレート（集計） V2_0 \[XLS形式：255KB\]](#)

[第三者提供用データベース_コード定義表V2_0 \[XLS形式：288KB\]](#)

サービス項目コードにつきましては、下記のリンク（WAM NET）をご参照ください。

<https://www.wam.go.jp/gyoseiShiryou/detail-list?bun=020050010>

希望するデータの提供形式によって必要書類が異なり、表 3 に一覧が記載されている。
書式が指定されているものは「4 申出に必要な書類」からダウンロードして必要事項を

記載する。

様式1には研究計画(表4)や必要な抽出データ(表5)を記載するシートなどが含まれている。必要な抽出データについては、個人や介護事業、保険者の特定につながる項目の利用の有無と利用する場合の理由を記載する必要がある。別添1の提供申出者の証明書は、申出者が公的機関等あるいは個人の場合は担当者の身分証明書等の写しが、法人等の場合は登記事項証明書等が必要となっている。別添2は運用に関する各種規定である。記入例がHP上の資料3)に示されているのでそれを参考に作成するとよいだろう。別添3～6は必須ではなく該当時に必要な書類となっており、別添7は倫理委員会の写しである(サンプリングデータセットの申請には不要)。別添8(表6)は介護DBテーブルのカラムのリストとなっており、特別抽出と集計表情報の場合に、出力してほしいデータ項目は「出力」欄にチェックをして提出する。別添9は詳細な公表形式であり、現時点で想定している内容をイメージ案として提出する。なお、最終的に公表する図表等を別添9と完全に一致させる必要はないが、公表の前に厚生労働省に確認し、許可を受ける必要がある(前述した「10未満の値」など、個人の特定に繋がり得るとして公表不可とされている値がなければ、結果の内容に関わらず許可される)。

表3 提出書類一覧⁴⁾

様式/別添	書類題目	備考	特別抽出	集計表情報	データアップロード
指定	様式1	匿名要介護認定情報等の提供に関する申出書	● 必須	● 必須	● 必須
	様式1-1	匿名要介護認定情報等を利用した研究に関する承認書	● 必須	● 必須	● 必須
	様式1-2	匿名要介護認定情報等の提供に係る手数料免除申出書	○ 該当時必須	○ 該当時必須	○ 該当時必須
書式自由	別添1	・提供申出者の証明書（提供申出者が公的機関等あるいは個人の場合は担当者の身分証明書等の写し、提供申出者が法人等の場合は登記事項証明書等） ・担当者及び代理人の本人確認、提供申出者に所属していることを確認できる書類の写し	● 必須	● 必須	● 必須
	別添2-1	運用フロー図	● 必須	● 必須	● 必須
	別添2-2	リスク分析・対応表	● 必須	● 必須	● 必須
	別添2-3	運用管理規程	● 必須	● 必須	● 必須
	別添2-4	自己点検規程	● 必須	● 必須	● 必須
書式自由	別添3	所属組織の個人情報保護に関する規程（プライバシーポリシー、情報セキュリティポリシー等）	○ 任意	○ 任意	○ 任意
	別添4	厚労科研費交付決定通知の写しなど、公共性の高い研究であることを示唆する書類	○ 該当時必須	○ 該当時必須	○ 該当時必須
	別添5	提供申出者における過去の研究実績を証明するもの（論文の写し等）	○ 任意	○ 任意	○ 任意
	別添6	外部委託先との守秘義務契約の写し	○ 該当時必須	○ 該当時必須	○ 該当時必須
	別添7	倫理委員会承諾書の写し	● 必須	○	○
	別添8	申出依頼テンプレート（抽出）←※特別抽出用 申出依頼テンプレート（集計）←※集計表情報用	● 必須	● 必須	○
	別添9	詳細な公表形式	● 必須	● 必須	○
その他	その他適宜必要な書類	○ 任意	○ 任意	○ 任意	

表4 研究計画書（様式1より）

●研究計画	
(4)-1 研究の名称 研究概要が分かるように具体的に記述	
(4)-2 研究の内容	
(4)-3 研究の概要 下記の項目を参考に具体的に記述する。 ①研究の背景となる基本情報 これまでの先行研究を含めた当該分野の知見や研究のビジョン、研究を行う必要性を記述する ②研究の目的 研究により明らかにしたい内容を具体的に記述する ③研究計画 ・研究対象集団（選択・除外基準等） ・研究デザイン（統計解析法等。実態把握のみの場合は省略可） ・データ抽出条件（必要な項目と分析での利用方法等） ・アウトカム（介護費等） ④研究によって期待される効果 本研究で期待される結果やその意義について記述する	1 ページ
(4)-4 研究の計画及び実施期間 匿名要介護認定情報等の受領から削除までのスケジュールについて、データ加工・分析・レポート作成・事務局による公表審査・公表時期（査読期間を含む）・削除時期等を具体的に記述する。	
(4)-5 他の情報との連結の有無	
(4)-6 外部委託等の有無等	なし 「あり」を選択した場合

表5 抽出データ（様式1より）

「提供データの種類」欄で、「特別抽出情報」又は「集計表情報」を選択した場合
個人や介護事業所、保険者の特定につながる以下の項目の利用有無と利用する場合の理由について

利用する項目	当該項目が必要な理由等
<input type="checkbox"/> 事業所番号 必ず通番化します。 必要な理由を記載してください。	
<input type="checkbox"/> 介護支援専門員番号 必ず通番化します。 必要な理由を記載してください。	
<input type="checkbox"/> 合議体番号 必ず通番化します。 必要な理由を記載してください。	
<input type="checkbox"/> 保険者番号 主に市町村を意味します。 必要な理由を記載してください。	保険者番号が必要な理由(匿名化しない場合はその理由も記入してください。)
<input type="checkbox"/> サービス種類コード 必要な理由に加えて、利用する介護サービスの名称(通所介護等)を記載してください。書ききれない場合はリストを添付してください。	
<input type="checkbox"/> サービス項目コード 必要な理由に加えて、利用する加算・減算や提供サービスを記載してください。書ききれない場合はリストを添付してください。	3 ページ

表6 介護DBのカラムリストの一部（別添8より）

レコード識別名 : 給付実績情報（明細情報レコード）(DT1111_D1)

項番	データ項目名(日本語)	型	桁数	備考	出力
1	交換情報識別番号	文字	4		
2	入力識別番号	文字	4		
3	レコード種別コード	文字	2		
4	証記載保険者番号	文字	6		
5	提供先番号	文字	64	匿名化後の個人ID	
6	サービス提供年月	文字	6		
7	事業所番号	文字	10	通番（匿名化）	
8	整理番号	文字	10		
9	サービス種類コード	文字	2		
10	サービス項目コード	文字	4		
11	単位数	文字	4		
12	日数・回数	文字	2		
13	公費1対象日数・回数	文字	2		
14	公費2対象日数・回数	文字	2		
15	公費3対象日数・回数	文字	2		
16	サービス単位数	文字	6		
17	公費1対象サービス単位数	文字	6		
18	公費2対象サービス単位数	文字	6		
19	公費3対象サービス単位数	文字	6		
20	摘要	文字	20		×
21	決定後単位数	文字	4		

3. まとめ

本章では介護DB情報サイトに書かれている内容を説明し、介護DBの申出方法・申出における留意点を説明してきた。

提出書類が多く作成に慣れない部分もあると思うが、事務局への事前相談を早めに余裕を持って作成できるようにスケジュールを立てるとよいだろう。

参考資料

- 1) 厚生労働省「匿名介護情報等の提供について」

https://www.mhlw.go.jp/stf/shingi2/0000198094_00033.html

(2023年3月23日アクセス)

- 2) 厚生労働省「匿名介護情報等の提供について」 1 匿名介護情報等の提供等について
匿名介護情報等の提供等について

https://www.mhlw.go.jp/stf/shingi2/0000198094_00033.html

(2023年3月23日アクセス)

- 3) 厚生労働省「医療情報システムの安全管理に関するガイドライン 第5.2版（令和4

年3月)」

https://www.mhlw.go.jp/stf/shingi/0000516275_00002.html

(2023年3月23日アクセス)

- 4) 厚生労働省「匿名介護情報等の提供について」1 匿名介護情報等の提供等について
匿名介護情報等の第三者提供に係る手続きについて

https://www.mhlw.go.jp/stf/shingi2/0000198094_00033.html

(2023年3月23日アクセス)

II. 介護 DB のデータ構造

1. 介護 DB のデータ構造を理解する

介護 DB のデータ構造を理解するためには、以下で解説するようなデータテーブルに関するイメージをもつことが重要である。すなわち、介護 DB の情報はデータの種類ごとに異なるデータテーブルに分割して保存されていること、各データテーブルは各データテーブルに固有の情報（変数）を含んでいること、そして各データテーブルは利用者 ID などの値をキーとして結合することが可能となっていること、である。

1) 介護 DB 情報はデータ種類ごとに異なるデータテーブルに分割保存されている

匿名要介護認定情報、匿名介護レセプト等情報、匿名 LIFE 情報に関するデータは、データの種類ごとにデータテーブルという単位に分割され、表形式で保存されている。匿名要介護認定情報および匿名介護レセプト等情報のデータテーブル一覧を表 1 に示す。

表 1. データテーブルの一覧（匿名要介護認定情報、匿名介護レセプト等情報）

No	テーブル名	テーブル ID
1	要介護認定情報	D_NINTEI
2	給付実績情報（基本情報レコード）	DT1111_H1
3	給付実績情報（明細情報レコード）	DT1111_D1
4	給付実績情報（緊急時施設療養・緊急時施設診療情報レコード）	DT1111_D2
5	給付実績情報（特定診療費・特別療養費情報レコード）	DT1111_D3
6	給付実績情報（食事費用情報レコード）	DT1111_D4
7	給付実績情報（居宅サービス計画費情報レコード）	DT1111_D5
8	給付実績情報（福祉用具販売費情報レコード）	DT1111_D6
9	給付実績情報（住宅改修費情報レコード）	DT1111_D7

10	給付実績情報（高額介護サービス費情報レコード）	DT1111_D8
11	給付実績情報（特定入所者介護サービス費用情報レコード）	DT1111_D9
12	給付実績情報（社会福祉法人軽減額情報レコード）	DT1111_DA
13	給付実績情報（ケアマネジメント費情報レコード）	DT1111_DB
14	給付実績情報（所定疾患施設療養費等情報レコード）	DT1111_DC
15	給付実績情報（明細情報（住所地特例）レコード）	DT1111_DD
16	国保連合会保有給付実績情報（基本摘要情報レコード）	DT1111_DF
17	給付実績情報（集計情報レコード）	DT1111_T1
18	給付管理票情報	DT1121
19	事業所台帳情報（基本情報）	DT5142
20	事業所台帳情報（サービス情報）	DT5243
21	事業所台帳情報（介護支援専門員情報）	DT5245
22	保険者台帳情報	DT5441
23	広域連合情報（行政区情報）	DT5641
24	市町村固有情報	DT5541
25	受給者台帳情報	DT5341

例えば、表 1 の No3 給付実績情報（明細情報レコード）という名前のデータテーブルには、介護給付の実績情報のうち明細情報に関するデータ（どのような介護サービスが何回・何日提供されたのか等）が「DT1111_D1」というファイル名で保存されている（例：2023年3月の給付実績_明細情報であれば”DT1111_D1_2023-03.csv”というファイル名）。

同様に、匿名 LIFE 情報で保存されているデータテーブルの一覧を表 2 に示す。

表 2. データテーブルの一覧（匿名 LIFE 情報）

No	テーブル名	テーブル ID
1	利用者情報	SERVICE_USER_INFO
2	科学的介護推進情報	FORM_0000_2021
3	科学的介護推進情報（既往歴情報）	FORM_0001_2021
4	科学的介護推進情報（服薬情報）	FORM_0002_2021

5	栄養・摂食嚥下情報	FORM_0100_2021
6	栄養ケア計画等情報	FORM_0110_2021
7	口腔衛生管理情報	FORM_0210_2021
8	口腔機能向上サービス管理情報	FORM_0220_2021
9	興味関心チェック情報	FORM_0310_2021
10	生活機能チェック情報	FORM_0320_2021
11	個別機能訓練計画情報	FORM_0330_2021
12	リハビリテーション計画書（医療介護共通部分）	FORM_0410_2021
13	リハビリテーション計画書（介護）	FORM_0420_2021
14	リハビリテーション会議録(様式3 情報)	FORM_0430_2021
15	リハビリテーションマネジメントにおけるプロセス管理票 (様式4 情報)	FORM_0440_2021
16	生活行為向上リハビリテーション実施計画書(様式5 情報)	FORM_0450_2021
17	褥瘡マネジメント情報	FORM_0500_2021
18	排せつ支援情報	FORM_0600_2021
19	自立支援促進情報	FORM_0700_2021
20	薬剤変更情報	FORM_0800_2021
21	薬剤変更情報(既往歴情報)	FORM_0801_2021
22	ADL 維持等情報	FORM_0900_2021
23	その他情報	FORM_8000_2021

2) 各データテーブルは各データテーブルに固有の情報（変数）を含んでいる

実際に上述したデータファイル（DT1111_D1_2023-03.csv）を、統計解析ソフトで読み込んだときのイメージを図1に示す。各列は変数名を示しており、データテーブルの種類ごとに固定されている。図1では、給付実績明細情報のデータテーブルなので、利用者が受けた介護サービスの種類や項目を示すコード値やサービスを提供した日数・回数に変数名となっている。一方、行は同じ利用者（匿名化ID）でも、サービスを提供した事業所（事業所番号）、サービス提供年月、サービス種類、サービス項目が異なれば、

異なる行としてデータが保存されていることに注意が必要である。つまり、1行がどのような粒度で構成されているかは、データテーブルの種類やデータテーブルの抽出の仕方によって異なるので、データ解析時には留意が必要となる。

利用者ごと、事業者ごころ、年月ごとにサービスの利用状況が把握可能。

匿名化ID	事業所番号	サービス提供年月	サービス種類コード	サービス項目コード	日数・回数
fajsgkru57	0123456	201905	11	7311	1
fajsgkru57	0123456	201905	11	1111	20
fajsgkru57	9874565	201905	13	1020	56
fajsgkru57	9874565	201906	13	4001	1
lltmlkgfd44	9874565	201906	13	1020	59
lltmlkgfd44	5688221	201904	52	6110	30
lltmlkgfd44	5688221	201904	52	6122	1
...

サービス種類コードと項目コードの組み合わせで、介護報酬で設定されている詳細なサービス提供情報が把握可能。

図1 給付実績情報（明細情報レコード）のデータテーブルのイメージ

3) 各データテーブルは利用者IDなどの値をキーとして結合が可能である

データテーブルは、利用者ID（匿名化ID）や事業所IDの変数の値を結合キーとして、突合することが可能である。データテーブル間には関係性があることのイメージを図2に示す。例えば、このイメージでは、ある利用者がいつ要介護認定を受けて（D_NINTEI）、その後どのような介護サービス（DT1111_D1）あるいはどのようなリハビリテーション計画が立案されたのか（FORM_0410_2021）等について情報を収集することが可能となる。

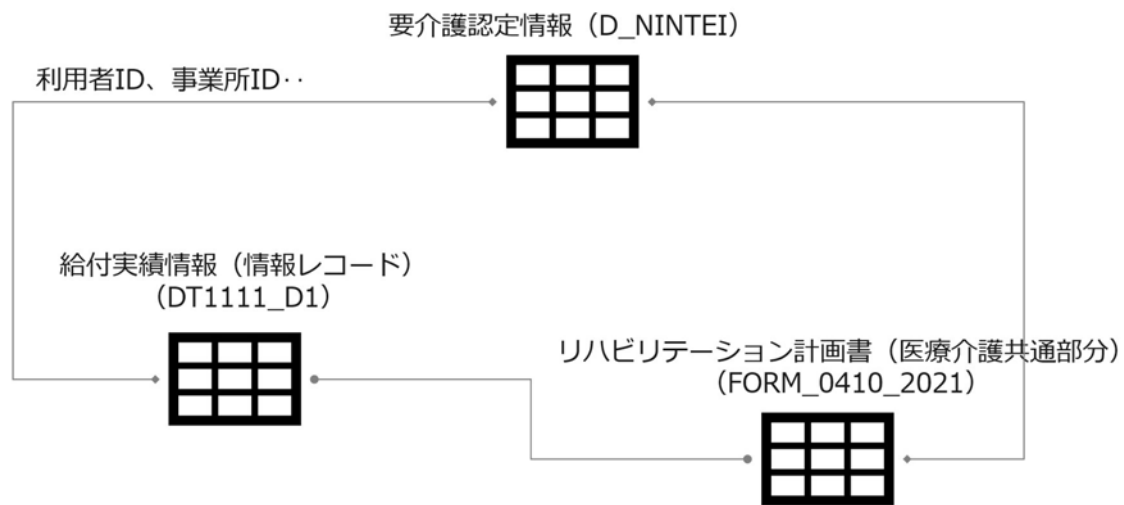


図2 データテーブル間の関係性についてのイメージ

4) 必要なデータテーブルと変数を事前を選択する

介護DBのデータ構造について、データ解析を行う上で初学者が理解しておくべき内容を概説した。介護DBの情報はデータの種類ごとに異なるデータテーブルに分割して保存されていること、各データテーブルは各データテーブルに固有の情報（変数）を含んでいること、そして各データテーブルは利用者IDなどの値をキーとして結合することが可能となっていることを理解することが重要である。

実際に介護DBを扱う際は、これらの介護DBのデータ構造に関するイメージをもった上で、研究や調査のテーマに応じて、必要なデータテーブルと必要な変数を取捨選択する必要がある。取捨選択を行わずに、全データテーブルと全変数を一括して利用することは、セキュリティおよびデータ量の膨大さの観点から不可能である。

2. 介護DBに含まれているデータ項目について

介護DBに含まれる変数は、高齢者が介護保険制度下で介護サービスを利用するプロセスで発生する情報である。よって、そのプロセスを理解することが、介護DBに含まれる変数を利用する上で重要である。以下では、介護DBに含まれる主要な変数（デー

タ項目)の一部について、データ発生のプロセスと関連づけて、紹介する。

1) 「匿名要介護認定情報」に含まれるデータ項目

匿名要介護認定情報は、主に市区町村が要介護認定に用いた調査の結果からなる情報である。図3に匿名要介護認定情報が発生するプロセス[1]を示す。

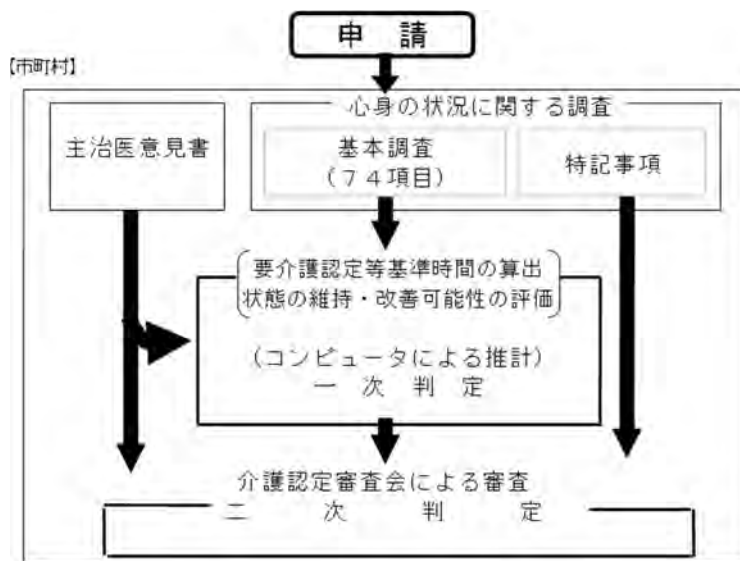


図3. 匿名要介護認定情報が発生するプロセス[1]

特筆すべき点は、申請時点の年齢区分、性別、申請区分（新規、更新など）、最終的に判定された要介護認定の区分に加えて、「心身の状況に関する調査」における74項目の基本調査（身体機能・起居動作、生活機能、認知機能、精神・行動障害、社会生活への適応および過去14日間に受けた医療）や認知症高齢者の日常生活自立度などについての情報が登録される点である。これらの項目の一覧と入力値の例を、それぞれ図4と図5に示す。これらの74項目の基本調査や認知症高齢者の日常生活自立度などの情報は、大規模に評価し収集することが困難な情報であり、利用者の生活機能がある程度把握することが可能となる。

身体機能・起居動作	生活機能	認知機能	精神・行動障害	社会生活への適応	過去14日間に つけた特別な医療
1 麻痺(左-上肢)	21 移乗	33 意思の伝達	42 被害的	57 薬の内服	63 点滴の管理
2 麻痺(右-上肢)	22 移動	34 毎日の日課を理解	43 作話	58 金銭の管理	64 中心静脈栄養
3 麻痺(左-下肢)	23 えん下	35 生年月日をいう	44 感情が不安定	59 日常の意思決定	65 透析
4 麻痺(右-下肢)	24 食事摂取	36 短期記憶	45 昼夜逆転	60 集団への不適応	66 ストーマの処置
5 麻痺(その他)	25 排尿	37 自分の名前をいう	46 同じ話をする	61 買い物	67 酸素療法
6 拘縮(肩関節)	26 排便	38 今の季節を理解	47 大声を出す	62 簡単な調理	68 レスビレーター
7 拘縮(股関節)	27 口腔清潔	39 場所の理解	48 介護に抵抗		69 気管切開の処置
8 拘縮(膝関節)	28 洗顔	40 徘徊	49 落ち着きなし		70 疼痛の看護
9 拘縮(その他)	29 整髪	41 外出して戻れない	50 一人で出たがる		71 経管栄養
10 寝返り	30 上衣の着脱		51 収集癖		72 モニター測定
11 起き上がり	31 スポン等の着脱		52 物や衣類を壊す		73 じょくそうの処置
12 座位保持	32 外出頻度		53 ひどい物忘れ		74 カテーテル
13 両足での立位			54 独り言・独り笑い		
14 歩行			55 自分勝手に行動する		
15 立ち上がり			56 話がまとまらない		
16 片足での立位					
17 洗身					
18 つめ切り					
19 視力					
20 聴力					

図 4. 心身の状況に関する調査における 74 項目の基本調査の一覧

認定調査票（基本調査）

1-1 麻痺等の有無について、あてはまる番号すべてに○印をつけてください。（複数回答可）

1. ない	2. 左上肢	3. 右上肢	4. 左下肢	5. 右下肢	6. その他（四肢の欠損）
-------	--------	--------	--------	--------	---------------

1-2 拘縮の有無について、あてはまる番号すべてに○印をつけてください。（複数回答可）

1. ない	2. 肩関節	3. 股関節	4. 膝関節	5. その他（四肢の欠損）
-------	--------	--------	--------	---------------

1-3 寝返りについて、あてはまる番号に一つだけ○印をつけてください。

1. つかまらないでできる	2. 何かにつかまればできる	3. できない
---------------	----------------	---------

1-4 起き上がりについて、あてはまる番号に一つだけ○印をつけてください。

1. つかまらないでできる	2. 何かにつかまればできる	3. できない
---------------	----------------	---------

1-5 座位保持について、あてはまる番号に一つだけ○印をつけてください。

1. できる	2. 自分の手で支えればできる	3. 支えてもらえればできる	4. できない
--------	-----------------	----------------	---------

図 5. 基本調査の入力値の例[2]

また、要介護認定情報のデータは要介護認定の新規申請や更新申請が行われる度に発生するので、要介護認定情報の推移を評価することも可能である。二次判定結果などの情報についてのデータテーブルのイメージを図 6 に示す。留意点は、次の申請までのインターバルが数年間に及ぶことがあり、要介護認定情報の変化に鋭敏でない可能性がある

る点である。一方、介護度が重度化した場合には、給付を受ける必要性から変更申請することが多いと予想され、重度化については鋭敏に反映できるという議論もある[3]。

保険者	ID	申請日	申請区分	年齢区分	性別	二次判定日	二次判定結果	認定有効期間 (開始)	認定有効期間 (終了)	障害高齢者 自立度	認知症高齢者 自立度
A市	12345	2011-12-05	1	8	1	2012-01-08	22	2012-01-01	2013-12-31	3	6
A市	12345	2012-05-28	3	8	1	2012-06-29	23	2012-05-28	2013-04-30	4	6
A市	12345	2013-05-24	2	8	1	2013-06-28	23	2013-06-01	2015-05-30	3	6
A市	12345	2015-02-28	2	9	1	2015-04-02	24	2015-04-02	2017-03-30	3	7
...

図 6. 要介護認定情報の解析データテーブルのイメージ (基本調査 74 項目などは省略)。

2) 「匿名介護レセプト等情報」に含まれるデータ項目

匿名介護レセプト等情報は要支援・要介護認定者が介護保険制度下で利用した介護サービスに関して、サービス事業者が保険者へ介護報酬を請求する際の明細情報である。事業所単位、月単位および利用者単位で作成され、サービス利用者の基本情報（性別、生年月日、要介護状態区分など）や利用した介護サービス内容（サービスの種類や加算項目）および提供回数に関するデータ項目を含んでいる。表 1 に示したように介護 DB には匿名介護レセプト等情報に関する多くのデータテーブルが存在する (No.2—25)。

主要な情報として、①給付実績情報（基本情報レコード）DT1111_H1 と②給付実績情報（明細情報レコード）DT1111_D1 に着目する。これらのデータテーブルのデータ項目は、図 7 に示す通り介護給付費明細書に記載された情報である。特に、②給付実績情報（明細情報レコード）DT1111_D1 の介護種類コード（2 桁）と介護項目コード（4 桁）を組み合わせることにより、通所介護サービスでは利用者の状態に応じた個別機能訓練計画書を作成した上で個別機能訓練を実施した場合に算定可能な「個別機能訓練加算」、あるいは口腔機能が低下又はそのリスクのある利用者に対し専門職スタッフが口腔機能改善管理指導計画を作成した上で口腔機能向上サービスの提供した場合に算定可能な「口腔機能向上加算」など、より詳細な介護サービスの内容を分析に用いることがで

きる。おおざっぱな介護サービス種類に加えて、加算項目の情報を組み合わせることにより、より詳細な介護やケアの内容について分析可能となることが重要である。

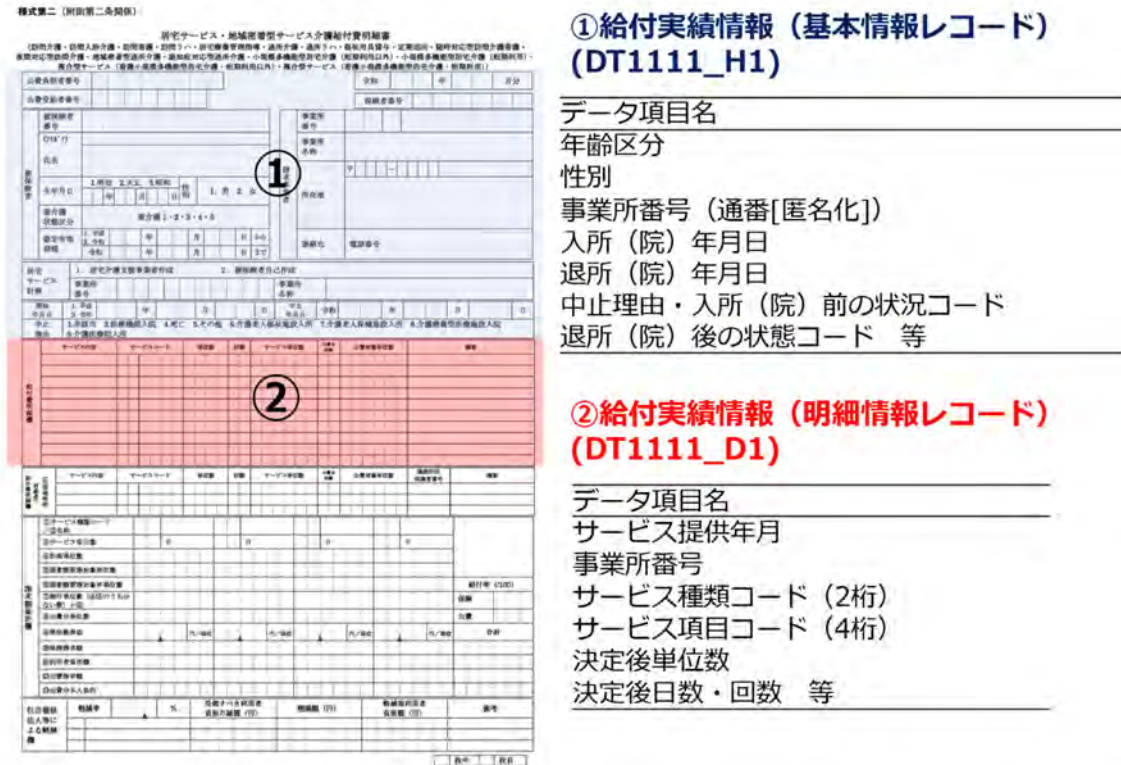


図 7. 介護給付費明細書と匿名介護レセプト等情報に含まれる情報の対応 [4]

3) 「匿名 LIFE 情報」に含まれるデータ項目

匿名 LIFE 情報で保存されているデータテーブルの一覧を表 2 に示した。LIFE はサービス提供事業所が利用者の状態や提供している介護やケアの計画・内容を一定の様式で入力すると、インターネットを通じてその内容が厚生労働省へ送信され、当該施設等へ分析結果 (フィードバック) が送られるシステムである。

例として、科学的介護推進情報 (FORM_0000_2021) のデータテーブルに着目する。このデータテーブルに含まれるデータ項目は、科学的介護推進加算を算定するための要件である科学的介護推進に関する評価 (図 8) に記載された情報である。

科学的介護推進に関する評価（通所・居住サービス）

評価日 令和 年 月 日
 前回の評価日 令和 年 月 日
 記入者名

氏名 職 評定者番号 令和 年 月 日
 利用者番号 令和 年 月 日
 事業所番号 性別 男 女

居住歴（前回の評価時より変化のあった場合は記載）住居変更

事業所名（施設）
 1. 業種名（ /日）（地方期間 年 月 日～ 年 月 日）
 2. 業種名（ /日）（地方期間 年 月 日～ 年 月 日）

居住環境
 同居家族等 なし あり（配偶者 子 その他）（複数選択可）
 家族等が介護できる時間 ほとんど毎日 平日程度 2～3時間程度 必要な時に手をさす程度 その他

生活機能
 ・食事 10 15 20 25 30
 ・椅子とベッド間の移動 10 15 20 25 30
 ・歩行 5 10 15 20 25 30
 ・入浴 5 10 15 20 25 30
 ・平地歩行 5 10 15 20 25 30
 ・階段昇降 10 15 20 25 30
 ・更衣 10 15 20 25 30
 ・履物コントロール 10 15 20 25 30
 ・排泄コントロール 10 15 20 25 30

在宅環境の有無等（複数選択可）
入所/サービス継続中
中止（中止日： / /）
居宅（専） 介護老人福祉施設入所 介護老人保健施設入所 介護施設入所 介護療養型医療施設入院
医療機関入院 死亡 その他

※前記サービスを利用する場合（介護サービスを回用しなくなった場合は、その他にチェック）

身長（ / ） cm 体重（ / ） kg 褥瘡の有無（複数選択可） なし あり
1段階の褥瘡状態
2段階の褥瘡状態
3段階の褥瘡状態
4段階の褥瘡状態
5段階の褥瘡状態

評価者（専） なし あり（専任日： 年 月 日）（免状日： 年 月 日）
 ※前回の入力時には訓練性判断の責任、二回目以降の入力時は前回の評価後の訓練性判断の責任について記載

認知症の診断 なし あり（診断日 年 月 日）軽度 中程度 重度 その他

認知症（認知症の診断または疑いのある場合に記載）
 まったくない ほとんどない ときどきある よくある 常にある

・日常の生活に支障をきたさない
 ・特別な事情がないのに夜中起き出す
 ・特別な理由もないのに人に話しかけようとする
 ・やたらと歩きまわる
 ・同じ動作をいつまでも繰り返す

認知症（訓練性判断）
 ・同じ事を何度も何度も繰り返す
 ・よく物をなくしたり、置き場所を間違えたり、隠したりする
 ・記憶、物忘れがひどい
 ・おぼろけのしるし
 ・間違いのあるいは季節に合わない不適切な服を着る
 ・言葉をさかすかしたる
 ・物を盗め込む
 ・引き出しや部屋の扉をみすみみ開けてしまう

Vitality Index
 ・新記録 自分から挨拶する、話しかける 挨拶、呼びかけに対して返答や笑顔が見られる 反応がない

認知症（訓練性判断）
 ・記憶 いつも記憶が曖昧している 最近ささいなことを忘れている 自分から記憶することは少ない
 ・食事 自分の選んで食べようとする 見られると食べようとする
 ・排泄 トイレに誘われてから行く 自分でトイレに行く 時々、尿意を感じて行く
 ・歩行 いつも自ら歩行を促される、あるいは自分で歩行、排便を行う 時々、尿意を感じて行く
 ・入浴 自分で入浴を促される、あるいは自分で入浴、排便を行う 時々、尿意を感じて行く
 ・平地歩行 自分で歩行を促される、あるいは自分で歩行、排便を行う 時々、尿意を感じて行く

※注）任意項目の記載のない項目は必須項目とする

図 8. 科学的介護推進に関する評価の記載様式 [5]

LIFE 情報を分析に用いることにより、介護レセプト等情報や要介護認定情報では評価できない、さらに詳細な介護やケアの内容や利用者の状況（栄養状態、褥瘡有無、リハビリテーション計画など）を把握することが出来るようになる。なお、2021年度の介護報酬改定で、この取組に参加すると加算が算定できるようになったばかりであり、今後の発展に注目したい。

3. まとめ

介護 DB には、高齢者が利用した介護サービスの詳細情報や高齢者の状態像について貴重な情報を多く含んでいる。介護 DB を分析・解釈し、実社会で活用するためには、必要なデータテーブルやデータ項目を適切に選択する必要がある。そのためには、介護 DB のデータ構造に関する理解が不可欠である。

4. 引用

1. 厚生労働省 「要介護認定に係る制度の概要」

<https://www.mhlw.go.jp/topics/kaigo/nintei/gaiyo1.html>

(2023 年 4 月 7 日アクセス)

2. 厚生労働省 「認定調査員テキスト 2009 改定版 (令和 3 年 4 月改定)」

<https://www.mhlw.go.jp/content/000819416.pdf>

(2023 年 4 月 7 日アクセス)

3. Jin X, Tamiya N, Jeon B, Kawamura A, Takahashi H, Noguchi H. Resident and facility characteristics associated with care-need level deterioration in long-term care welfare facilities in Japan. *Geriatr Gerontol Int.* 2018;18(5):758-766.

4. 厚生労働省 「認定調査員テキスト 2009 改定版 (令和 3 年 4 月改定)」

https://www.wam.go.jp/content/wamnet/pcpub/top/appContents/20190611_01.html

(2023 年 4 月 7 日アクセス)

5. 独立行政法人福祉医療機構 WAM NET 「介護給付費請求書等の様式 (令和元年 5 月 7 日改正省令施行対応)」

https://www.wam.go.jp/content/wamnet/pcpub/top/appContents/20190611_01.html

(2023 年 4 月 7 日アクセス)

III. 介護 DB を使用した研究例の紹介

介護 DB の研究には、研究仮説の設定、研究のデザイン、データの入手、データの処理や統計解析など多くのステップがあり、膨大な介護 DB の中から研究成果として評価される結果を出すためには様々な工夫が必要です。この章では、これまで介護 DB の研究実績を複数発表してきた研究者が、実際の事例を通して、介護 DB をどのように加工分析して研究を実施したのか、またその研究の中で気づいた工夫すべき点をご紹介します。

1. 研究事例①：介護保険施設におけるケアの質を評価した研究（利用者の要介護重度化に関連する介護福祉施設の特徴）

書誌情報

Jin X, Tamiya N, Jeon B, Kawamura A, Takahashi H, Noguchi H. Resident and facility characteristics associated with care-need level deterioration in long-term care welfare facilities in Japan. *Geriatr Gerontol Int.* 2018

1) 研究の動機

高齢化に伴って重度介護者、認知症の高齢者が増える中で、介護問題は、高齢者の生活に係る大きな不安要因となっています。介護保険法の施行により、介護サービスに係る民間業者の参入が急増し、介護サービスの供給は充実してきたものの、それに伴い、サービスの質の保障が重要な課題となってきました。このような状況の中、問題解決の一助となるのが、介護DBを活用してケアの質を評価することです。介護DBには、介護サービス利用者が受けたサービスに情報と利用者状態をリアルタイムで蓄積しており、利用者がサービスを受けたことで利用者状態がどう変化していたかを追跡できるからです。質の高いケアを提供している施設であるほど、利用者の状態の重度化が少ない傾向があります。以下に、介護DBの情報を活用する際の手順や注意すべき点に焦点を当てて、介護施設のケアの質を評価した研究実例を紹介します。

2) 必要な作業

作業順番	作業内容
1	介護DB データテーブルの内容を確認しよう
2	介護DB に含まれる施設サービスの情報を把握しよう
3	研究課題、研究デザインを特定しよう

4	必要に応じて介護施設の特徴が含まれている介護 DB 以外のデータとの照合しよう
5	データの準備・クリーニングをしよう
6	データを分析しよう

3) 手順

(1) 介護 DB データテーブルの内容を確認しよう

介護 DB は、研究者が研究目的で収取したデータではないため、二次データ分析の特徴を持っています。データ分析者はデータ収集の時間とコストが節約できる分、研究仮設の構築に労力をかける必要があります。まずは、介護 DB の収集頻度、データ構造、含まれる変数についての把握から始まります。これに関しては、第 II 章ですでに説明していますので、ここでは省略しますが、適切な研究デザインを選択する上で最も重要な第 1 ステップであることは心がけましょう

(2) 介護 DB に含まれる施設サービスの情報を把握しよう

本章では、介護 DB に含まれる施設サービスの情報に焦点を当てて、説明します。ドナベディアンは、医療福祉サービスの質の評価する際には、「構造(Structure)」、「過程(Process)」、「結果(Outcome)」の三つの側面から行うことを提唱しています。

構造とは介護サービスが提供される条件を構成するものであり、主に事業所台帳情報の基本情報 DT5142, サービス情報 DT5242 に含まれています。プロセスは、介護がどのように提供されたか、或いは提供した介護サービスの内容であり、給付実績情報基本情報レコード,DT1111_H1 と明細情報レコード、DT1111_D1 に多く含まれています。アウトカムとは提供された介護サービスの個人または集団の変化のことであり、利用者の身体状況の変化に関する情報は、給付実績情報(基本情報レコード,DT1111_H1)と要介護認定情報(D_NINTEI)に含まれています。表 1 は介護 DB に含まれるサービス情報の詳細

ですので、ご参考ください。

(3) 研究課題、研究デザインを特定しよう

介護 DB に含まれる情報を読むと、利用者の要介護度が経時的に把握できることが分かります。そのため、利用者の要介護度の重度化に関連する施設特徴を検証できることも分かります。介護保険制度では、利用者の状態が悪化したら、要介護度認定を随時に行いますが、通常有効期間は2年です。利用者の状態変化を把握するには2年の追跡が必要であり、データも最低期間2年で申請しましょう。筆者は、2012年10月に介護福祉施設に入所した利用者を対象とし、2年後である2014年9月までのデータを用いました。アウトカムは、要介護重度化の有り無しの2値であり、利用者の要介護が2年後に重度化した場合、要介護度重度化としました。

(4) 必要に応じて介護施設の特徴が含まれている介護 DB 以外のデータも利用しよう

利用者一人当たりの看護職員数、介護職員数は先行研究でもケアの質に影響する重要な要因と報告されているが、介護 DB にはその情報が含まれておりません。職種別の人員配置が含まれているデータは、介護サービス施設・事業所調査や介護サービス情報公表システムのデータが挙げられます。介護 DB は個人レベルでその他のデータと照合することはできないが、施設が特定できない状態であれば、それらの情報を利用できる可能性があります。その場合、データ申請者が施設番号と施設特徴をカテゴリ化した変数からなる対応表を作り、介護 DB 申請窓口に、介護 DB にそのカテゴリ変数をつけてもらうように依頼する方法があります。例えば、利用者一人当たり介護職員が多い施設と少ない施設に分けて、施設番号と各カテゴリのリストを作り、介護 DB 窓口に相談するなどです。

(5) データの準備・クリーニングをしよう

介護 DB を特別抽出でデータを受領した場合、データの容量は数百 GB や 1 TB 以上になります。介護給付費の生データは月ごとに 1 つの CSV ファイルで受領されますので、3 年の期間を申請した場合は、ファイル数は百個以上になります。この場合、直接に統計分析ソフト R もしくは STATA や SPSS を利用するのは、必要なデータを抽出するまでかなりの時間を要します。現時点で筆者が考える効率的な方法は、Microsoft SQL Server によりデータベース化した上で、前処理を行い、統計分析は、R、STATA や SPSS を用いることである。Microsoft SQL Server は、大量なデータの保存が可能であり、必要なデータを迅速に検索・抽出することが可能となり、データ活用を加速できます。新しいソフトを学ぶことに難しさを感じる人もいますが、介護 DB 分析に必要な Microsoft SQL server のコマンドは限られており、概ねデータベースの作成、データの検索、データの操作（データを並べる、グループ化する、データの追加・削除）、データの結合ができれば良く、すべてを学ぶ必要はありません。

(6) データ分析をしよう

本研究のアウトカムである、利用者の要介護度重度化の分布や施設特徴の分布を記述することから始まります。次に、単変量分析を行い、各施設特徴と利用者要介護度重度化との関連を見ます。例えば、利用者要介護度重度化を被説明変数とし、施設規模（大きい施設と小さい施設）を説明変数としたロジスティクス回帰分析を行います。最後に、利用者要介護度重度化を被説明変数とし、利用者特徴を共変量、施設特徴を説明変数とした多変量ロジスティクス回帰分析を行い、関連を検証します。筆者は、同じ施設に入所した利用者に相関があることを考慮した、マルチレベルロジスティクス回帰分析を行いました。

4) 介護 DB の長所・短所

介護 DB は、介護保険制度運營業務において発生する記録であり、すべての利用者が対象となり、悉皆性を持っています。また、それぞれの事業所にとってもデータ収取への負担がないメリットがあります。一方で、介護報酬に係る情報のみ含まれており、介護施設のケアの質の評価も限られています。例えば、衛生管理、感染症対策、外出支援、嗜好を尊重したケアの実施に関する記録はないです。また、利用者の医療情報が含まれてなく、介護サービスの利用の効果検証が難しいことが短所の一つです。例えば、介護サービスの効果を検証する際には、その介護サービスを利用した群と利用していない群を追跡し、利用者の身体状況への変化を比較する場合があります。ただし、サービス利用群と利用しない群はサービス利用の有無以外の背景因子（身体状況、経済状況など）が近いほど真の効果の測定がしやすいです。背景因子が異なる場合、サービス利用効果はバイアスが含んでしまい、2群の比較は難しいです。厚生労働省では、介護 DB と NDB の連結データの提供を始めていますので、連結データを活用することをお勧めします。

5) 工夫した点

介護 DB を活用するには、ビッグデータ利活用の基盤・環境、研究実践能力だけではなく、介護保険制度及び介護の現場も十分理解する必要性があります。また、要介護者は医療と介護のニーズを併せて持つ特徴があり、医療・看護・介護の多分野の専門家の意見も重要となり、研究者一人がすべてを行うのは難しいです。「つくば医療介護レセプト研究会」は、医療・介護レセプトを用いて研究を行う多分野の専門家が集まり、レセプトの活用方法及び課題を整理しています。筆者は、この研究会で収集した情報が、研究に最も役に立ちました。

表 1. 介護 DB に含まれる施設サービスの情報

構造		過程		結果	
施設特徴	データに含まれる変数名	施設特徴	データに含まれる変数名	利用者状態	データに含まれる変数名
事業の終了	事業廃止年月日	提供したサービスの内容 (加算により一部把握可能)	・療養食加算 ・短期集中リハ加算 ・個別機能訓練加算 など	要介護度重度化	被保険者情報要介護状態区分コード
ユニット型・従来型	施設等の区分コード	サービスの利用量 (介護費で算出)	保険請求額＋保険利用者負担額	在宅復帰	退所（院）後の状態コード
法人種別	法人等種別コード	福祉用具の利用	サービス項目コード	入院	退所（院）後の状態コード
介護老人保健施設種類（基本型・在宅強化型・療養型）	人員配置区分コード			死亡	退所（院）後の状態コード
人員配置（加算・減算により一部把握可能）	・常勤専従医師配置加算 ・医師の欠員による減算 など			認知症発生	認知症高齢者自立度
運営体制 (加算により一部把握可)	・医療連携体制加算 ・介護職員処遇改善加算 ・サービス提供体制強化加算等				

2. 研究事例②： 介護報酬給付実績データウェアハウスの開発

書誌情報

佐藤幹也, et al. 介護報酬給付実績データウェアハウスの開発. 医療情報学, 2003, 23.6: 483-490.

1) 概要

この研究は、自治体が介護保険の保険者として所有している介護保険事務処理システムのデータ（介護 DB の元データ）の構造を分析して 2003 年に報告したものです。介護保険制度が導入された直後に利用可能だった介護データの給付実績情報（介護 DB における DT111）は、それぞれ独自のカラム構造をとる基本情報（H1）、集計情報（T1）、明細情報（D1）などデータ種別ごとに一つのテーブルとして整然とは記録されておらず、様々なカラム構造をとるレコードにデータ種別コードが付与されたものが 1 枚の CSV ファイルに格納されたものでした。

この給付実績情報を研究用データセットとして加工するため、データ種別を示すコードを用いて複数のテーブルに切り分けて別々のテーブルとして抽出し、これを更にリレーショナルデータベースとして SQL Server 上で再構築しました。

2) 得られた知見

給付実績情報（DT111）は、介護報酬の給付請求の支払い実績を、要介護者 1 人 1 事業所 1 か月を単位として記録した静態統計です。給付実績情報は、レコード種別ごとに異なるカラム構造をとるレコードから構成されています。給付実績のレコード種別には基本情報（H1）、集計情報（T1）、各サービスの明細情報（D1、D2、D3...）などがあり、それぞれの情報がサービス種類コードやサービス項目コードを連結キーとする階層構造をとっています。介護報酬の請求に過誤や再審査が生じた場合は、単一のサービス利用に対して複数のレコードが給付実績情報に発生します。受給者台帳（DT5341）は、要

介護認定や更新、異なる保険者をまたいだ転居、死亡などのイベントが生じたときにレコードが発生する動態統計です。

3) 初学者へのフィードバック

給付実績情報（DT1111）が月次の静態統計である一方、受給者台帳（DT5341）はイベント毎にレコードが発生する動態統計なので、そのままでは両者を突合することはできません。両者を使って研究を実施するためには、受給者台帳を月次データとして加工し、受給者台帳の月次断面をとってから両者を突合する必要があります。

給付実績情報には過誤修正や再審査が行われる前のレセプトも含まれるので、給付実績情報を集計する場合は過誤回数、再審査回数、申請月などのカラムを用いて正しいレコードだけを抽出する（過誤修正や再審査によって上書きされる前のレコードを捨てる）必要があります。

給付実績情報の過誤修正を行ったり、受給者台帳を月次データに加工したりするデータクリーニングや、統計解析に用いるデータセットの切り出しには、リレーショナルデータベースを使用するのが便利です。

3. 研究事例③：自己負担率と介護サービスの利用量との関連を検討した研究

書誌情報

Sato M, Hashimoto H, Tamiya N, Yano E. The effect of a subsidy policy on the utilization of community care services under a public long-term care insurance program in rural Japan. *Health Policy*. 2006;77(1):43-50.

1) 概要

この研究は、在宅要介護者の訪問介護と通所介護の利用量を、訪問介護の自己負担率の違いで比較したものです。この研究の結果、訪問介護の自己負担が3%に減額される経過措置の対象者は、非対象者に比べて訪問介護の利用額が多く通所介護（経過措置の対象ではない）の利用額が少ないことが明らかになりました。

2) データ処理

自治体が所有する介護給付実績(DT1111)等を解析しました。介護給付実績(DT1111)の集計情報(T1)からは、在宅介護サービス利用額を抽出してサービスの種類毎に3か月平均利用額を算出しました。介護給付実績(介護DBのDT1111と同等)の基本情報(H1)からは年齢、性別、要介護度、公費負担者番号を使用しました。公費負担者番号から公費の法定番号(公費負担者番号の左側2桁)抽出しこの経過措置の対象か否かをフラグ化しました。データの前処理にはSQL Serverを、解析にはSASを使用しました。

3) 初学者へのフィードバック

暴露因子(群分け)とアウトカム(比較指標)を明確にして分析研究を実施する方が、記述研究を実施するよりも海外の学術誌には掲載されやすいようです。公費負担者番号の左側2桁が公費負担の法別番号になっています。これを用いて生活保護、自立支援医療や経過措置等のフラグを作成できます。

4. 研究事例④：在宅介護のフォーマルケア時間の推計

書誌情報

佐藤幹也, et al. 全国の介護保険レセプトを用いた在宅介護のフォーマルケア時間推計.
日本公衆衛生雑誌, 2019, 66.6: 287-294.

1) 概要

この研究は、在宅介護の利用実績を介護時間に換算してこれをフォーマルケアによる介護時間とし、これと主介護者のインフォーマルケアによる介護時間とを要介護度別に対比した研究です。この結果、居宅系介護サービスおよび短期入所サービスのフォーマルケア時間及びインフォーマルケア時間は要介護度とともに増加すること、通所系介護サービスのフォーマルケア時間および介護保険サービスによる総フォーマル時間は要介護3で頭打ちとなり要介護4-5ではむしろ減少することが明らかになりました。

2) データ処理

統計法33条に基づいて入手した全国の介護給付費実態調査の調査票（介護給付実績を同調査の調査票として利用することを承諾しなかった自治体のものを除いたもの）と、国民生活基礎調査介護票の調査票をこの研究に使用しました。

在宅介護サービスのサービス項目（サービス種類の下位コード）ごとに標準介護時間を設定して、サービス項目マスタを作成しました。このサービス項目マスタと介護給付実績（介護DBのDT1111に相当）の明細情報（D1）とをサービス種類コード及びサービス項目コードを使って突合してから介護時間を利用者ごとに集計して、在宅介護サービスによるケア時間（フォーマルケア時間）に換算しました。

介護給付実績（DT1111）の基本情報（H1）から性別と要介護度を抽出して利用者の層別化に利用しました。男女別要介護度別に1日当たりの平均フォーマルケア時間を集計し、これと国民生活基礎調査の介護票から集計した主介護者の1日当たりの介護時間

(インフォーマルケア時間)とを性別要介護度別に比較しました。データの前処理には SQL コマンドで SAS をリレーショナルデータベース化したものを用いました。データ解析にも SAS を使用しました。

3) 初学者へのフィードバック

研究の群分け(説明変数)に用いる変数が介護 DB のデータにはあまり含まれていません。しかしサービス種類、サービス項目、都道府県番号等の介護給付実績に含まれるカラム(もしくはその一部)に記録されたコードをキー項目とするマスタを作成し、これと介護給付実績(DT1111)や受給者台帳(介護 DB における DT4341)等のデータと連結してカラムを追加すると研究に使用できる説明変数を増やすことができます。

介護 DB のデータを大規模な経年データとして一括処理するためには、SQL 言語を使用してリレーショナルデータベースでデータを加工する必要があります。しかし統計ソフトの中には SQL 言語を使用してデータをリレーショナルデータベースとして取り扱うことができるものもあり、単月データ、小規模データ等の前処理に有用です。

5. 研究事例⑤：病院看取り率の経年変化（在宅介護利用者と施設介護利用者の比較）

書誌情報

Sato M, Tamiya N, Jin X, Watanabe T, Takahashi H, Noguchi H. Impact of a Japanese Incentivization Program to Increase End-of-Life Care Outside of Hospitals. J Am Med Dir Assoc. 2021;22(2):329-333.

1) 概要

これは、在宅介護を受けて死亡した要介護者と施設介護を受けて死亡した要介護者との間で、病院看取り率の経年変化を比較した研究です。この結果、2006年以降に導入された様々な看取り関連施策によって施設で介護された要介護者の病院看取り率が著しく減少し、それによって要介護者全体全体の病院看取り率も減少したことが明らかになりました。

2) データ処理

統計法 33 条に基づいて入手した全国の介護給付費実態調査の調査票および人口動態統計死亡票を使用しました。介護給付費実態調査の受給者台帳（DT5341）と人口動態統計死亡個票とを前者の性別、生年月日、資格喪失年月日、介護保険者番号の一部（左側から 3～5 桁目が自治体コードとなっている）と後者の性別、生年月日、死亡年月日、住所地の自治体コードとで連結して研究データセットに用いました。

対象者が死亡した年の 4 月の受給者台帳（DT5341）の月次断面から年齢、性別、要介護度を抽出してデータセットに追加しました。対象者が死亡した年の 4 月の介護給付費実績（DT1111）の集計情報（/T1）からサービス種類コードを抽出してサービスの利用の有無をデータセットに追加し、在宅介護サービスだけが追加できた者を在宅群、施設介

護サービスだけが追加できた者を施設群としました(在宅介護サービスと施設介護サービスの両者が追加された者は研究対象から除外しました)。人口動態統計死亡票から死亡の場所(病院とそれ以外の施設または自宅)を抽出してデータセットに追加しました。SAS を前処理とデータ解析に使用しました。

3) 初学者へのフィードバック

どのような集団の間で何を比較するのか(研究の Participants、Exposure、Confounders/comparison、Outcome、いわゆる PECO は何なのか)というリサーチクエスチョンが明確である方が学術研究として評価されやすいです。

効率よく研究用データセット作成にも PECO を意識することが大切です。まずは Participants (対象集団) のデータセットを確定させて、このデータセットに Exposure/ Confounder/ Outcome の変数を追加すると良いと思います。

6. 研究事例⑥：介護保険受給者台帳の資格喪失記録は死亡の代理変数となるか

書誌情報

佐藤幹也, et al. 介護保険受給者台帳の資格喪失記録を死亡代理変数として使用することの妥当性の検討. 日本公衆衛生雑誌, 2022, 69.8: 617-624.

1) 概要

これは、人口動態統計死亡票の発生を死亡のゴールドスタンダードとして、介護保険受給者台帳の資格喪失情報が死亡の代理変数として使用することの検査特性を評価した研究です。この研究の結果、受給資格喪失の死亡に対する感度は 100%、特異度は 99.9%、陽性反応的中率は 88.5%、陰性反応的中率は 100% となり、受給資格喪失は死亡の代理変数として許容しうることが明らかになりました。

2) データ処理

前述の研究と同様に、介護給付費実態調査の受給者台帳 (DT5341) の全ての月次断面と人口動態統計死亡票とを性別、生年月日、死亡年月日/資格喪失年月日、自治体コードを用いて連結できたものを死亡例、連結できなかったものを生存例としました。受給者台帳 (DT5341) の月次断面から年齢、性別、要介護度および需給資格喪失の有無を抽出。ODBC を用いて R Studio から SQL Server に格納されたデータを操作して前処理を行い、処理済みのデータセットを R Studio から読み込んで分析しました。

3) 初学者へのフィードバック

R studio 環境で R と SQL Server とを ODBC で直結すると、データの管理と前処理から分析、レポート出力までを一元化できます。効率よくデータベースを研究に用いるためには、作業の再現性を確保することが大切です。R Studio 環境からスクリプトを用いてデータベースを操作するようにすると、作業の再現性が高まり前処理プロセスの修正

も容易になります。

令和4年度厚生労働省 老人保健健康増進等事業

介護DBの解析・利活用を実践しうる人材を育成するためのプログラム作成等に関する調査研究事業

「介護DBを用いた研究を開始する際のガイド」

令和5年3月