

令和6年12月18日	資料6
第19回 匿名医療・介護情報等の提供に関する委員会	

# 定型データセット開始後のデータ提供について

厚生労働省老健局老人保健課

# 目次

1. 定型データセットの経緯
2. 定型データセット導入の前後比較
3. 定型データセットの現状評価  
(参考) 定型データセットの概要

## 定型データセットの経緯

提供データの形式として、申出者ごとに異なる抽出条件で作成する特別抽出が最も多く、申出から提供まで平均して1年、最長で2年半を要する事例もあった。

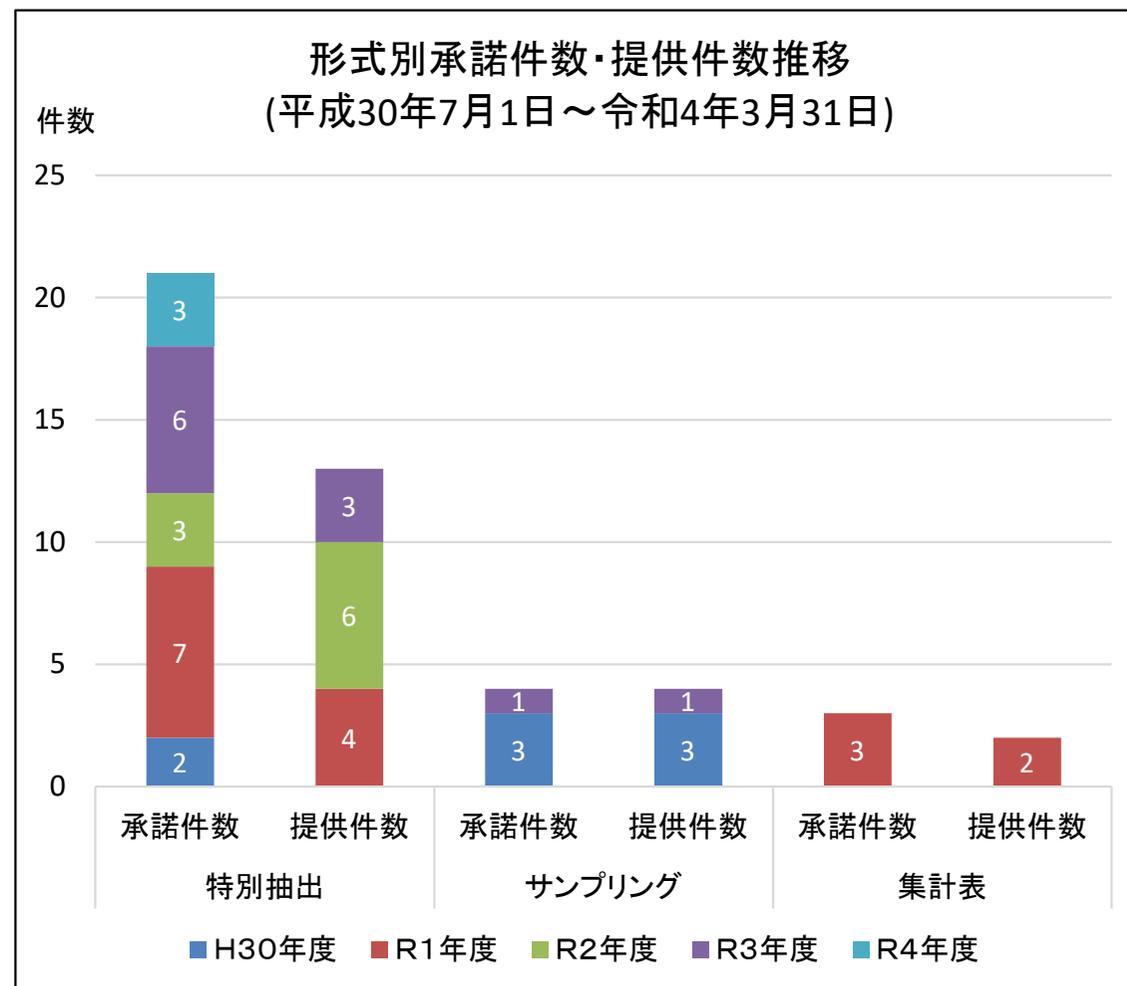
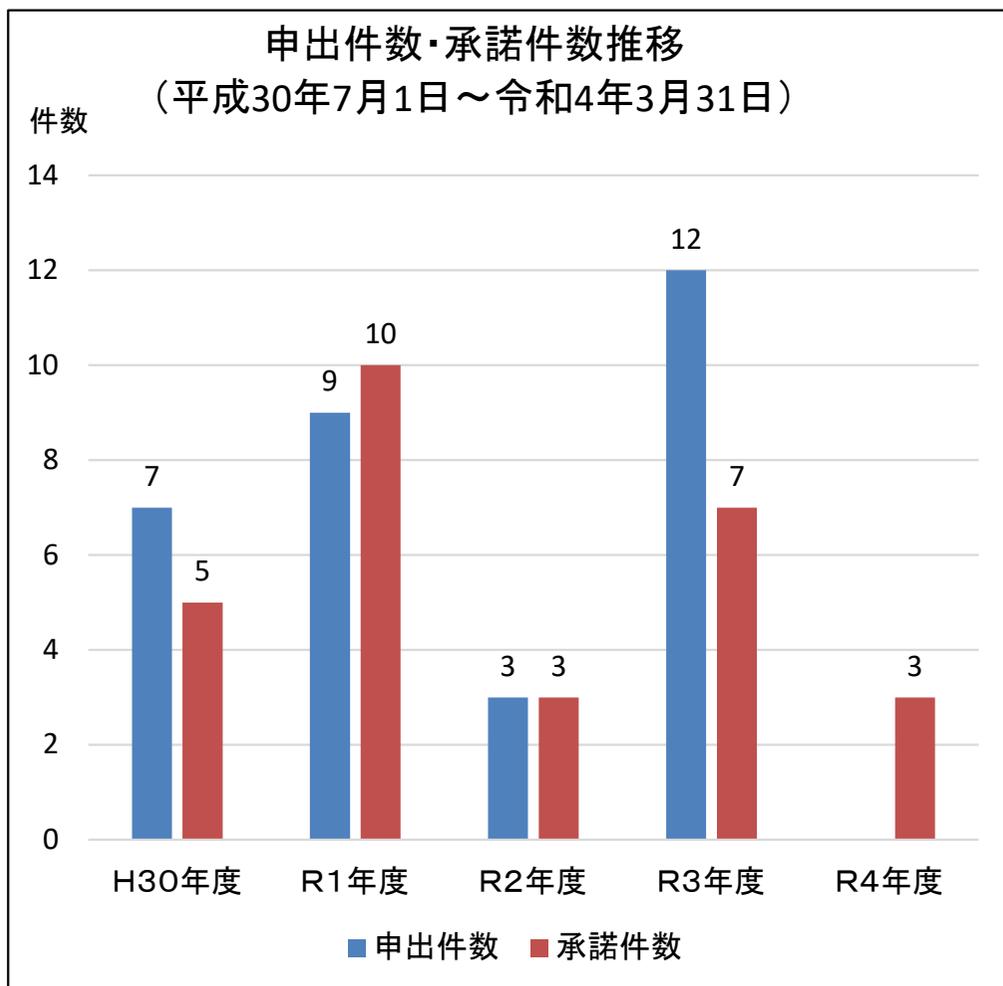
**審査から提供までの日数が全体の約75%を占め、データ抽出・整備・確認作業に半年以上の時間を要していた。**

これらの状況を踏まえて、**あらかじめ抽出された「定型データセット」を整備することとした。**定型データセットについては特別抽出と同じデータ形式ではあるが、申し出ごとにSQL（抽出条件）を作成して抽出するのではなく、事前に全データを抽出して整備し、申出者が使用する変数の入った帳票を、あらかじめ準備したデータとして提供するものである。

また、定型データセットの整備と併せて、提供したデータの目的外利用を防ぐ観点から、ガイドラインの改定を行った。

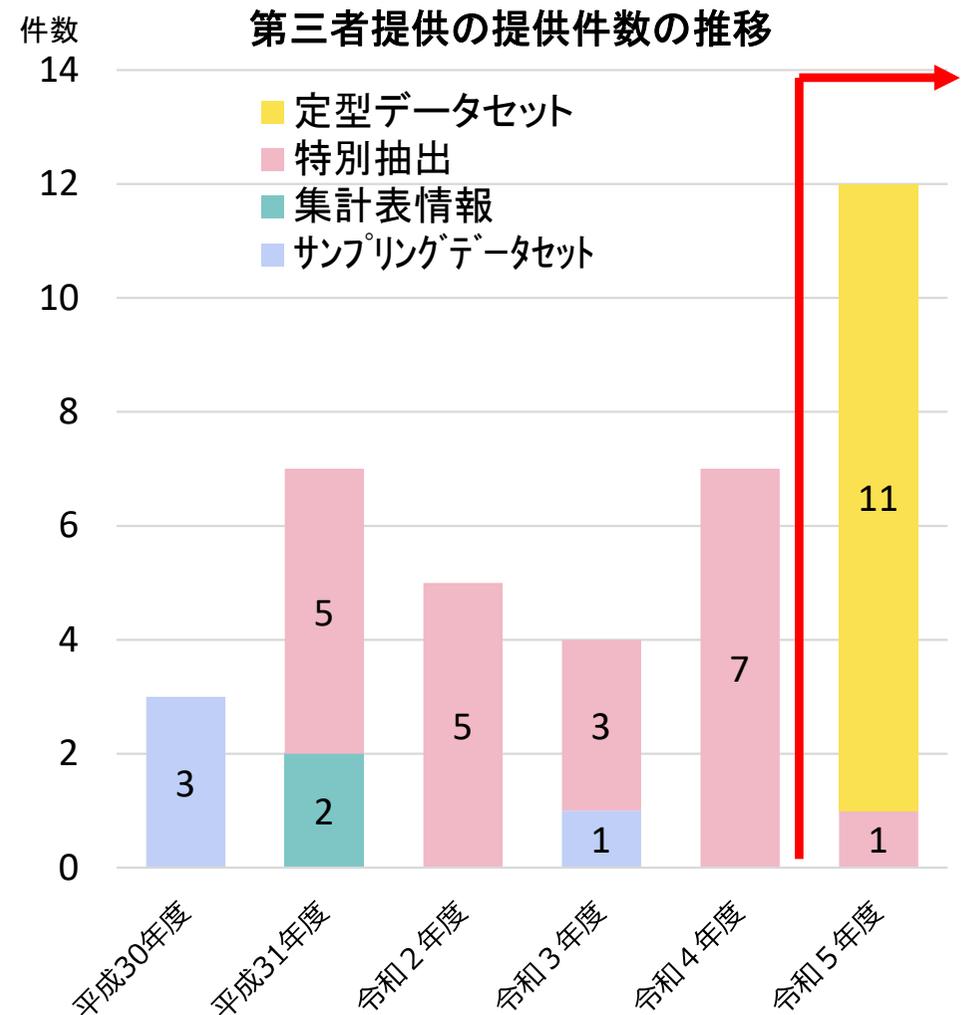
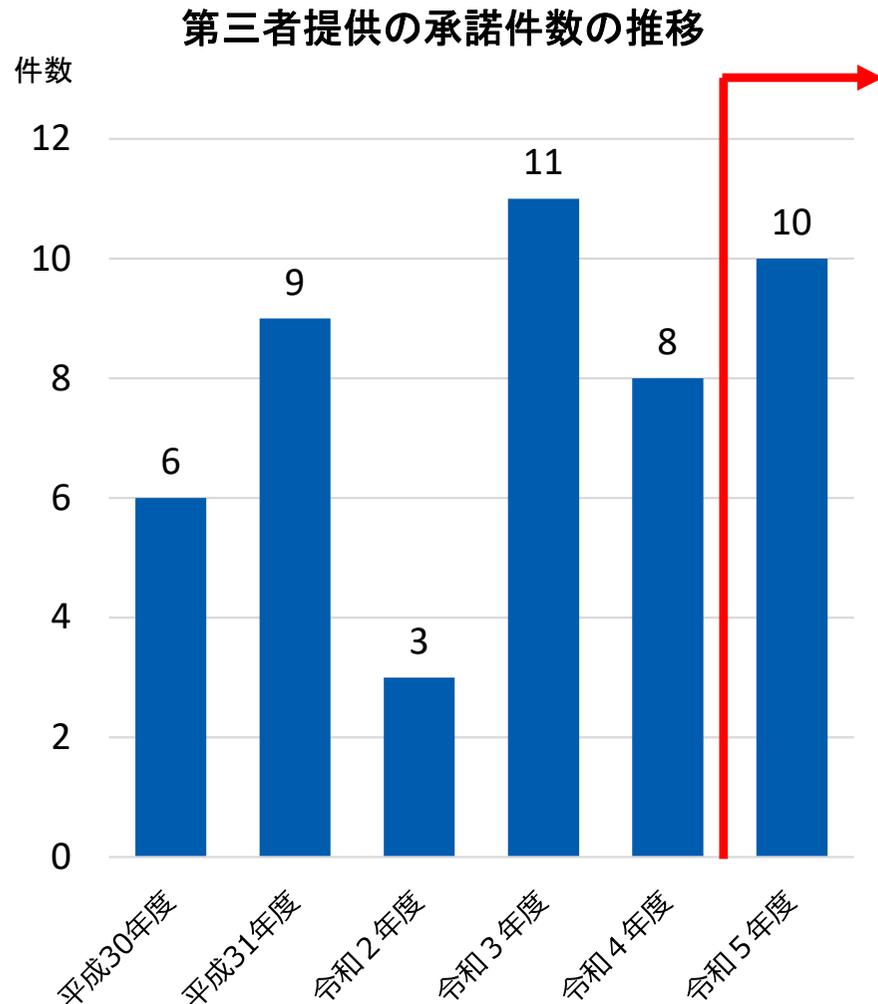
# 定型データセット導入前の申出・承諾件数と提供件数について

- 令和2年の新型コロナウイルス感染症流行初期に一時的に申出件数、承諾件数ともに落ち込んだものの、増加傾向にある（左図）
- 一方、特に特別抽出案件において、申出・承諾件数の増加に提供件数が追いついていない現状がある（右図）



# 定型データセット導入後の申出・承諾件数と提供件数について

- 承諾件数に比べて、提供件数が少ない場合がほとんどであったが、定型データセットを導入した令和5年度は提供件数が承諾件数を上回っている。
- 待ち行列を解消しつつ、承諾後の早期提供ができるようになった。



出典：厚生労働省老健局「匿名介護情報等の提供について（報告）」，社会保障審議会介護保険部会（第112回）資料2，令和6年3月28日より抜粋。矢印を追加。

# 申請からデータ提供までの平均日数内訳

- 第三者提供の中で最も多く申出があり、主に研究者が利用する特別抽出について、平均で申請から提供まで1年程度の時間を要していた(最長2年半)。
- その内訳として、審査日から提供日までの日数が全体の75%の時間を占めており、特にデータ抽出、その後のデータ整備、確認作業に半年以上の時間を要していた。
- 定型データセットでは、特別抽出に比べて、利用依頼日から提供までの時間が短縮された。なお、審査継続の申出があったため、申請日から審査日までの時間が長くなっている。
- 個別事情が発生していない案件(3件)については、審査後約3か月、承諾後に利用依頼を受けてから約2か月で提供している。

	申請日 ～提供日	申請日 ～審査日	審査日～提供日			
			審査日 ～承諾日	承諾日～ 利用依頼日	利用依頼日～ 提供日	
特別抽出 (N=13)	367	80	275	24	55	213
定型データセット (N=9)	335	135	200	62	39	99
定型データセット (N=3) 個別事情がない案件※	219	91	128	43	23	62

※以下等の個別事情があった案件をを除いた場合

- 提供できる範囲を超えて、未来の時点が含まれる場合
- 案件により誓約書・依頼書の提出に時間がかかる場合
- 条件付き承諾の条件解消の対応に時間がかかる場合
- 手数料の支払い完了までに時間がかかる場合

出典：(特別抽出、サンプリングデータセット、集計表情報) 平成30年7月1日から令和4年3月31日の第三者提供申出・提供情報より老人保健課にて作成  
(定型データセット) 令和5年9月1日以降に承諾された申出・提供情報より作成

# 參考資料

# 「定型データセット」のデータ概要（1）

## ○ データの概要

- 「新たな提供形式」は「定型データセット」とする。
- 第三者提供対象の全テーブル・全項目のデータセット。年2回作成。
  - 介護レセプト情報は「サービス提供年月」ベースで格納する。
  - 要介護認定情報と台帳情報は、全テーブルを差し替える。最も古い時点から格納する。
  - LIFE情報は、初回確定日（LIFEシステム側で自動記録）に基づいて格納する。
- 提供データに含まれる集団は、提供データにおける「サービス提供年月」の期間に有効な認定情報に絞り、第1号被保険者と第2号被保険者に限定する。
- データ抽出以降の過誤調整は反映させない。
- 月遅れ請求は、データ作成日（抽出開始時点）までに格納されているデータを含める。データ作成後に新たに月遅れ請求が届いたとしても、作成済みデータの作り直しは実施しない（例えば2024年2月作成時に、すでに作成済みの2022年12月データの作り直しは実施しない）。

	各データのサービス提供年月の範囲		
(データ完成時期)	2012年4月～2022年12月	2023年1月～2023年6月	2023年7月～2023年12月
1回目（2023年8月）			
2回目（2024年3月）	この期間内で、データ抽出開時点の2023年4月までに格納されている、月遅れ請求を含める。		この期間内で、データ抽出開時点の2024年4月までに格納されている、月遅れ請求を含める。
3回目（2024年10月）		この期間内で、データ抽出開時点の2023年10月までに格納されている、月遅れ請求を含める。	

注) サービス提供年月でデータを区切るため、対象範囲に含まれないサービス提供年月の月遅れ請求は入らない。

# 「定型データセット」のデータ概要（2）

## データ範囲と提供スケジュール

### ○ データの範囲

- 対象集団のレコードを事前に抽出した上で、これらの対象集団に紐づくデータを抽出する。

### ○ 提供スケジュール

	1回目	2回目	3回目以降①	3回目以降②
対象集団 (認定有効期間)	認定有効期間に 次の期間を含む 2012年4月 ～2022年12月	認定有効期間に 次の期間を含む 2023年1月 ～2023年6月	認定有効期間に 次の期間を含む yyyy年7月 ～yyyy年12月	認定有効期間に 次の期間を含む yyyy+1年1月 ～yyyy+1年6月
認定情報 (★) (認定申請日)	2009年4月 ～2022年12月	2009年4月 ～2023年6月 <b>以前のデータは廃棄</b>	2009年4月 ～yyyy年12月 <b>以前のデータは廃棄</b>	2009年4月 ～yyyy+1年6月 <b>以前のデータは廃棄</b>
介護レセプト情報 (サービス提供年月)	2012年4月 ～2022年12月	2023年1月 ～2023年6月	yyyy年7月 ～yyyy年12月	yyyy+1年1月 ～yyyy+1年6月
台帳情報 (★) (異動年月日)	2022年12月31日以前	2023年6月30日以前 <b>以前のデータは廃棄</b>	yyyy年12月31日以前 <b>以前のデータは廃棄</b>	yyyy+1年6月30日以前 <b>以前のデータは廃棄</b>
LIFE情報 (データ入力日※)	2021年4月 ～2022年12月	2023年1月 ～2023年6月	yyyy年7月 ～yyyy年12月	yyyy+1年1月 ～yyyy+1年6月
提供予定時期	提供可	提供可	yyyy+1年 の夏頃を予定	yyyy+1年 の冬頃を予定

★データ抽出のたびに更新するデータ。最新のスナップショット（過去からの履歴あり）を保存する。

★印の無い場合には、データ抽出のたびに追加する。

※データ入力日は、LIFE上で入力された日付。LIFEシステム側で自動的に記録されているデータ。

# 「定型データセット」のデータ概要（3）

## 個別データ項目の設定

### ○ 個別データ項目の設定

- 作り置き of データを作成する上で加工する content と、提供時に申出者ごとに異なる加工をする content があり、それぞれ実施する。

データ項目	個別の設定内容
ID4（NDB連結用ID）、 ID5（NDB連結用ID）	提供先番号（個人ID）とID4、ID5の対応表を提供する。 （データ提供時に申出ごとに匿名化）
提供先番号（個人ID）	保険者番号と被保険者番号由来のIDを組み合わせたデータ。 データ提供時に申出ごとに匿名化する。
事業所番号、介護支援専門員番号	データ提供時に申出ごとに匿名化する。
保険者番号、市町村コード	必要に応じて、データ提供時に申出ごとに匿名化する。
職員数、利用者数定員	事業所台帳サービス情報の数値を標準の階級値に変換する。
年齢階級コード、生年月日	共通の年齢階級コードに変換する。 （最下位65歳未満、最上位95歳以上の5歳刻み）
都道府県コード	各テーブルにおいて、保険者番号の上2桁から都道府県コードを作成する。 保険者番号のない事業所台帳関連のテーブルは、事業所番号の上2桁から生成する。

（合議体番号は、個人特定の可能性が低いため匿名化していない。）

# 「定型データセット」のデータ概要（４）データ形式

- 介護DBデータのテーブル別にCSVファイルを作成する。介護レセプト情報はテーブル別に加えてサービス提供年月別にファイルを分ける。
- データ項目順は別添 8 の順番通りとする。提供データに含めない項目は別添 8 で提供不可となっている項目のみで、特別抽出同様に空白の列で提供する。
- 介護レセプト情報を連結するための「給付実績情報連結キー」を各介護レセプト情報の右端（最後の項目）に格納する。

## 介護DBデータの第三者提供対象テーブル一覧

### 要介護認定情報

要介護認定情報
---------

### 介護レセプト情報

給付実績情報（基本情報レコード）
給付実績情報（明細情報レコード）
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）
給付実績情報（特定診療費・特別療養費情報レコード）
給付実績情報（居宅サービス計画費情報レコード）
給付実績情報（福祉用具販売費情報レコード）
給付実績情報（住宅改修費情報レコード）
給付実績情報（高額介護サービス費情報レコード）
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）
給付実績情報（社会福祉法人軽減額情報レコード）
給付実績情報（ケアマネジメント費情報レコード）
給付実績情報（所定疾患施設療養費等情報レコード）
給付実績情報（明細情報（住所地特例）レコード）
国保連合会保有給付実績情報（基本摘要情報レコード）
給付実績情報（集計情報レコード）
給付管理票情報

### 台帳情報

事業所台帳情報（基本情報）
事業所台帳情報（サービス情報）
事業所台帳情報（介護支援専門員情報）
保険者台帳情報
広域連合情報（行政区情報）
市町村固有情報
受給者台帳情報

### LIFE情報

利用者情報
科学的介護推進情報
科学的介護推進情報（既往歴情報）
科学的介護推進情報（服薬情報）
栄養・摂食嚥下情報
栄養ケア計画等情報
口腔衛生管理情報
口腔機能向上サービス管理情報
興味関心チェック情報
生活機能チェック情報
個別機能訓練計画情報
リハビリテーション計画書（医療介護共通部分）
リハビリテーション計画書（介護）
リハビリテーション会議録（様式 3 情報）
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）
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）
褥瘡マネジメント情報
排せつ支援情報
自立支援促進情報
薬剤変更情報
薬剤変更情報（既往歴情報）
ADL維持等情報
その他情報

# 定型データセットの特徴

- 「定型データセット」は、「特別抽出」と同じデータ形式ではあるが、毎回カスタマイズしたSQLを作成して抽出するのではなく、研究利用可能な形で事前に全データを抽出して整備しておくもの。
- 承諾から抽出する形ではなく、申請者が使用する変数の入った帳票をすでに抽出されたデータから提供するため、提供までの日数を大幅に削減することが可能。

	①特別抽出	②定型データセット	③サンプリングデータセット	④集計表
抽出単位	個票	個票	個票	集計 (単位は任意)
研究利用	○	○	△	△
他のDBとの連結解析	○	○	×	×
委員会における審査から提供にかかる期間 (平均)	約1年	約2か月 (想定)	約3か月	約4か月
分析に使用する 集団の抽出	済 (都度SQLを作成して 抽出)	研究者が実施	済	済
データ容量 (申出1件あたり (実績) )	10GB~1TB	10GB~3TB	200MB~2GB	500KB~3GB