

令和5年6月21日	資料8
第12回匿名医療・介護情報等の 提供に関する委員会	

# 第2回介護DBオープンデータについて（案）

厚生労働省老健局老人保健課

# 目次

1. 介護総合DBオープンデータ作成の背景と目的
2. 本専門委員会で検討すべき事項
3. 追加の集計表について
4. 第2回介護総合DBオープンデータの内容：匿名LIFE情報
  1. 公表形式
  2. 詳細集計の集計項目の絞り込み方針
5. 最小集計単位の原則について
  - 論点1：第2回以降のオープンデータについて
  - 論点2：第1回オープンデータの差し替えについて
6. 今後のスケジュール・第3回以降の検討課題

# 1. 介護保険総合DB（介護DB）オープンデータ作成の背景と目的

令和5年3月6日

第11回匿名介護情報等の提供に  
関する専門委員会

資料2

## 1. 作成の背景

- ◆ 介護DBには、悉皆性が高い匿名介護レセプト情報及び認定調査項目等の詳細なデータである匿名要介護認定情報等が含まれており、介護等分野の研究開発を行ううえで有用である。平成28年12月に、介護保険部会において、データベースをより有効活用するため、NDBと同様に、公益性が高い利用目的の場合には第三者への提供を可能とすることが適当であるとされ、平成30年11月より、高いレベルのセキュリティ要件を課したうえで、データ提供を行ってきた。
- ◆ NDBにおいては、多くの人々がNDBデータに基づいた知見に接することができるよう、NDBデータから汎用性の高い基礎的な集計表を作成し、「NDBオープンデータ」として、これまでに6回公表している。
- ◆ 介護DBに関しては、介護サービスの提供実態に係るデータについては介護給付費等実態統計としてe-Stat等で公表されているものの、要介護認定の結果等に係るデータは公表されておらず、NDBと同様、オープンデータとして公表する意義は大きい。

## 2. 作成の目的

- ◆ 多くの人々が介護DBデータに基づいた知見に接することが出来るよう、介護DBデータを用いて、「介護給付費等実態統計では公表されていない内容」という観点で基礎的な集計表を作成したうえで、公表する。
- ◆ 介護DBデータに基づき、介護サービスの提供実態や要介護認定情報等のデータをわかりやすく示す。

## 3. 第1回オープンデータの報告と今後の予定

- ◆ 2023年1月に第1回オープンデータを公表した。今後は1年に1回程度を目安に集計表を拡充して公表する。
- ◆ 一般からの意見も受け付けているが、現時点では集計の要望は届いていない。

厚生労働省：介護DBオープンデータ,

[https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/hukushi\\_kaigo/kaigo\\_koureisha/nintei/index\\_00009.html](https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/hukushi_kaigo/kaigo_koureisha/nintei/index_00009.html)

## 2. 本専門委員会で検討すべき事項

### 1. 経緯

- ◆ 本専門委員会では、オープンデータとして公表する集計表の内容・形式を確認いただく。

### 2. 具体的な検討内容

- ◆ 第2回オープンデータの方針（前回までの専門委員会で承認済み）
  - ・ 第1回オープンデータに対し、要介護認定情報の時点の追加
  - ・ 第1回オープンデータに対し、LIFE情報の追加
- ◆ **第1回及び第2回オープンデータの集計仕様の変更**（今回の専門委員会で審議）
  - ・ 要介護認定情報に新しい集計区分（施設類型）の追加
- ◆ 第2回オープンデータのLIFE情報の集計仕様（前回までの専門委員会で方針は承認済み）
- ◆ **秘匿処理方法の変更**（今回の専門委員会で審議）

# 目次

1. 介護総合DBオープンデータ作成の背景と目的
2. 本専門委員会で検討すべき事項
3. 追加の集計表について
4. 第2回介護総合DBオープンデータの内容：匿名LIFE情報
  1. 公表形式
  2. 詳細集計の集計項目の絞り込み方針
5. 最小集計単位の原則について
  - 論点1：第2回オープンデータについて
  - 論点2：第1回オープンデータの差し替えについて
6. 今後のスケジュール・第3回以降の検討課題

# 3. 追加の集計表について

## 1. 経緯

◆ 前回の第11回専門委員会では、政策利用・研究利用のニーズを踏まえて、施設類型別の集計について要望をいただいた。第2回オープンデータ公表のタイミングで追加の表についても公表することとした。

➤ 施設類型別（サービス種類コード別）

介護老人福祉施設、介護老人保健施設、介護療養型医療施設、  
地域密着型介護老人福祉施設、介護医療院

## 2. 具体的な集計内容

◆ 認定情報の集計について、第1回オープンデータの集計表に加えて、施設類型別の集計を加える。施設類型は、当該認定データの認定有効期間内に発生した最初の施設サービスのレセプトにより判別する。

新たに追加する集計事項一覧（匿名要介護認定情報） 第1回・第2回オープンデータ

明細 番号	表頭	表側				集計事項
		表1	表2	表3	表4	
		都道府県	性・年齢階級	要介護度	保険者	
19	施設類型；二次判定結果	●	●	－	－	延べ申請件数（施設サービス利用者）
20	施設類型；二次判定結果	●	●	－	－	実申請者数（新規申請・施設サービス利用者）
21	施設類型；性・年齢階級	●	－	－	－	延べ申請件数（施設サービス利用者）
22	施設類型；性・年齢階級	●	－	－	－	実申請者数（新規申請・施設サービス利用者）

注1) 二次判定結果には、要介護度が格納されている。

注2) 施設サービス利用者は、認定有効期間内かつ、集計対象期間内に、施設サービスのレセプトが発生している人。

注3) 施設サービスは、介護老人福祉施設、介護老人保健施設、介護療養型医療施設、地域密着型介護老人福祉施設、介護医療院。

## 集計対象・項目等

- ◆ 公表データ：匿名要介護認定情報
  - ①主要項目の集計結果
  - ②匿名介護レセプト等情報との突合率
  - ③前回と今回の二次判定結果とのクロス集計結果

### ◆それぞれの対象期間、公表項目、集計事項

	①主要項目の集計結果	②匿名介護レセプト等情報との突合率	③前回と今回の二次判定結果とのクロス集計結果
対象期間	2018年度、2019年度*		
公表項目	<ul style="list-style-type: none"> <li>・基礎的な項目 (申請区分、一次判定結果、二次判定結果、基準時間 等)</li> <li>・認定調査項目(74項目)</li> <li>・障害高齢者自立度</li> <li>・認知症高齢者自立度</li> </ul>	<ul style="list-style-type: none"> <li>・該当年度全体の突合率 (項目別の集計はしない)</li> </ul>	<ul style="list-style-type: none"> <li>・今回の申請区分</li> <li>・今回の二次判定結果</li> <li>・前回の二次判定結果</li> </ul>
集計事項	延べ申請件数(延べ人数) 又は 申請者数(実人数) ※実人数は新規申請のみ実施	介護サービス利用者割合 (匿名介護レセプト等情報との突合率)	延べ申請件数(延べ人数) 及び 申請者数(実人数) ※実人数に区分変更等含む

\*「②匿名介護レセプト等情報との突合率」については、匿名要介護認定情報に対して突合させる匿名介護レセプト等情報の範囲は、認定有効期間に含まれる範囲。匿名要介護認定情報のレコードごとに認定有効期間が異なる。

「③前回と今回の二次判定結果とのクロス集計結果」については、前回の匿名要介護認定情報の範囲は、介護DBに含まれるうち最も古い時点(2009年4月)から最新時点までの範囲。認定有効なレコードの中で、個人別・認定申請日別に並び替えた際に、集計対象の匿名要介護認定情報の1つ前のレコードを個人別に集計対象の個票に突合させる。

## 1. 集計対象・項目等

- ◆ 公表データ：①匿名要介護認定情報 ②匿名LIFE情報
- ◆ 匿名LIFE情報を公表する目的：
  - ・ 匿名LIFE情報の登録情報を明らかにする。
  - ・ 利用者の状態、ケアの内容に関する主なデータの入力状況を把握する。
- ◆ ①、②の対象期間、公表項目、集計事項

	①匿名要介護認定情報	②匿名LIFE情報
対象期間	2020年度、2021年度 (時点の判定項目:「認定申請日」)	2021年度 (時点の判定項目:「初回確定日」)
公表項目	<ul style="list-style-type: none"> <li>・2018年度・2019年度と同様の項目</li> <li>・基礎的な項目、認定調査項目</li> <li>・障害高齢者自立度、認知症高齢者自立度</li> <li>・匿名介護レセプト情報との突合率</li> <li>・前回と今回との二次判定結果のクロス集計</li> </ul>	<ul style="list-style-type: none"> <li>・LIFE関連加算ごとのサービス事業別算定状況</li> <li>・科学的介護推進体制加算に関連する項目のうち、基礎的な項目・主要項目 (フィードバック票の項目より選定予定※)</li> </ul>
集計事項	<ul style="list-style-type: none"> <li>・2018年度・2019年度と同様の集計事項</li> <li>・延べ申請件数 または 実申請者数</li> <li>・介護サービス利用者割合 (匿名介護レセプト情報との突合率)</li> </ul>	<ul style="list-style-type: none"> <li>・延べレコード数</li> <li>・実事業所数(サービス別・様式別)</li> <li>・実利用者数(サービス別・様式別)</li> <li>・平均内服薬数(LIFEのフィードバック票の定義に準じて集計を実施)</li> <li>・LIFE関連加算の算定率 (介護レセプト情報のみを用いて集計)</li> </ul>

(※) LIFE関連加算で必須となっている項目のうち事業所と利用者へのフィードバック項目から集計対象を選定予定。



## 2. 公表形式

- ◆ 匿名要介護認定情報について、第1回目のオープンデータと同様に、年度別に「都道府県別」、「性・年齢階級別」、「要介護度別」及び「保険者別」の集計を行う。ただし、保険者別は最小集計単位の観点から、一部の項目に限定する。

集計事項一覧（匿名要介護認定情報） 第1回オープンデータと同様

明細番号	表頭	表側				集計事項
		表1	表2	表3	表4	
		都道府県	性・年齢階級	要介護度	保険者	
1	申請区分（申請時）コード	●	●	●	●	延べ申請件数
2	現在の状況	●	●	●	-	実申請者数（新規申請）
3	一次判定結果	●	●	●	-	実申請者数（新規申請）
4	一次判定結果（認知症加算）	●	●	●	-	実申請者数（新規申請）
5	二次判定結果	●	●	-	●	実申請者数（新規申請）
6	要介護認定等基準時間（12区分）	●	●	●	-	実申請者数（新規申請），平均要介護認定等基準時間
7	主治医意見書	●	●	●	-	実申請者数（新規申請）
8	第1群 身体機能・起居動作	●	●	●	-	実申請者数（新規申請）
9	第2群 生活機能	●	●	●	-	実申請者数（新規申請）
10	第3群 認知機能	●	●	●	-	実申請者数（新規申請）
11	第4群 精神・行動障害	●	●	●	-	実申請者数（新規申請）
12	第5群 社会生活への適応	●	●	●	-	実申請者数（新規申請）
13	その他 過去14日間にうけた特別な医療	●	●	●	-	実申請者数（新規申請）
14	障害高齢者自立度	●	●	●	-	実申請者数（新規申請）
15	認知症高齢者自立度	●	●	●	-	実申請者数（新規申請）
16	匿名介護レセプト情報との突合率	●	●	●	-	介護サービス利用者割合（突合率）（新規申請）
17	今回の申請区分（申請時）コード；前回の二次判定結果	-	-	●	-	延べ申請件数
18	今回の申請区分（申請時）コード；前回の二次判定結果	-	-	●	-	実申請者数

# 目次

1. 介護総合DBオープンデータ作成の背景と目的
2. 本専門委員会で検討すべき事項
3. 追加の集計表について
4. 第2回介護総合DBオープンデータの内容：匿名LIFE情報
  1. 公表形式
  2. 詳細集計の集計項目の絞り込み方針
5. 最小集計単位の原則について
  - 論点1：第2回オープンデータについて
  - 論点2：第1回オープンデータの差し替えについて
6. 今後のスケジュール・第3回以降の検討課題

## 4. 第2回介護総合DBオープンデータの内容：匿名LIFE情報

### 1. 公表形式

- ▶ 匿名LIFE情報について、年度別に「性・年齢階級別」、「サービス種類別」及び「市町村（事業所所在地）別」の集計を行う。ただし、保険者別は最小集計単位の観点から、一部の項目に限定する。薬品コードから利用者の服薬種類数を集計する。

集計事項一覧（匿名LIFE情報）

明細 番号	集計対象の項目（表頭）	属性情報（表側）			集計事項
		表5	表6	表7	
		性・年 齢階級	サービ ス種類	市町村 （事業 所所在 地）	
1	LIFE関連加算の種類別の登録状況	●	●	－	延べレコード数
2	LIFE関連加算の種類別の登録状況	－	●	－	実事業所数
3	LIFE関連加算の種類別の登録状況	●	●	－	実利用者数
4	LIFE関連加算の算定率	－	●	－	実事業所数，算定率
5	日常生活自立度（科学的介護推進情報）	－	●	－	実利用者数
6	ADL（科学的介護推進情報）	－	●	－	実利用者数，Barthel index平均値
7	栄養（科学的介護推進情報）	－	●	－	実利用者数，BMI平均値
8	認知機能(DBD13)（科学的介護推進情報）	－	●	－	実利用者数，DBD13平均値
9	意欲(Vitality index)（科学的介護推進情報）	－	●	－	実利用者数
10	事業所ごとの利用者1人あたりの内服薬の平均種類数の分布状況（科学的介護推進情報（服薬情報））	－	●	●	事業所別の利用者1人あたり内服薬の平均種類数の最小値，最大値，平均値，標準偏差

注1) 要介護度は必ずしも全ての様式に含まれていないことから第2回オープンデータでは公表しない。

都道府県別や保険者別は件数が少ない懸念があり集計しないこととした。

注2) 市町村（事業所所在地別）の内服薬の平均種類数は、事業所の分布を集計する内容と、

地域全体の平均を掲載する内容であり、最小集計単位の基準には該当しないと考えられるため、公表を予定している。

## 4. 第2回介護総合DBオープンデータの内容：匿名LIFE情報

### 2. 詳細集計の集計項目の絞り込み方針

「科学的介護推進情報」について、現行のLIFEのフィードバック票（事業所（※1）・利用者（※2））に掲載され、入力必須となっている項目のうち、一部の項目に限定して、各コード別の該当者数、平均値等を集計する。

- 1人の利用者が複数の時点で登録されることが想定されるため、最新時点の情報を集計する。
- 数値データについては、平均値を集計する項目、カテゴリ化する項目を整理する。
- オープンデータを公表する段階までに、公表する項目の絞り込みを実施予定である。
- 前回の第11回専門委員会では、より広い内容の項目を提示しているが、反対意見は無かったため、公表項目の絞り込みについて、引き続き検討を進め、9月の専門委員会で公表予定項目につき協議をさせていただく。

(※1) 厚生労働省：科学的介護推進体制加算 事業所フィードバック票, <https://www.mhlw.go.jp/content/12301000/000949390.pdf>

(※2) 厚生労働省：科学的介護推進体制加算 利用者フィードバック票, <https://www.mhlw.go.jp/content/12301000/001015742.pdf>

# 目次

1. 介護総合DBオープンデータ作成の背景と目的
2. 本専門委員会で検討すべき事項
3. 追加の集計表について
4. 第2回介護総合DBオープンデータの内容：匿名LIFE情報
  1. 公表形式
  2. 詳細集計の集計項目の絞り込み方針
5. 最小集計単位の原則について
  - 論点1：第2回オープンデータについて
  - 論点2：第1回オープンデータの差し替えについて
6. 今後のスケジュール・第3回以降の検討課題

## ■ 最小集計単位の原則について

- ◆ 「匿名介護情報等の提供に関するガイドライン（以降ガイドラインと表記）」に記載されている、公表時の基準に準じている。

### （1）最小集計単位の原則

- ① 原則として、公表される研究の成果物において要介護者等の数が10未満になる集計単位が含まれていないこと（ただし要介護者等の数が「0」の場合を除く）。  
また、集計単位が市町村の場合には、以下のとおりとする。
  - i) 人口2,000人未満の市町村では、要介護者等の数を表示しないこと。
  - ii) 人口2,000人以上25,000人未満の市町村では、要介護者等の数が20未満になる集計単位が含まれないこと。
  - iii) 人口25,000人以上の市町村では、要介護者等の数が10未満になる集計単位が含まれないこと。
- ② 原則として、公表される研究の成果物において介護事業所または市町村の属性情報による集計数が、3未満となる集計単位が含まれていないこと（ただし要介護者等の数が「0」の場合を除く）。

出所：厚生労働省HP：匿名介護情報等の提供について

[https://www.mhlw.go.jp/stf/shingi2/0000198094\\_00033.html](https://www.mhlw.go.jp/stf/shingi2/0000198094_00033.html)

## 5. 最小集計単位の原則について

### 1. 経緯

- ◆ 介護DBのオープンデータは、「匿名介護情報等の提供に関するガイドライン（以降ガイドラインと表記）」に記載されている、公表時の基準に準じている。
- ◆ 最小集計単位の基準（通常は10未満）に該当する場合は、集計値をハイフン「-」に置き換えている。
- ◆ 総数から、秘匿されていない数値を差し引くことで、逆算が可能である場合は、必ず2箇所以上を秘匿した。また、総数が最小集計単位の基準に該当する場合があり、総数自体を秘匿した場合には、内数もすべて秘匿していた。これらの処理により、集計値10以上の箇所が秘匿されている場合があり、集計表全体として、秘匿されるセルが多くなっていた。
- ◆ 第2回オープンデータの公表に際して、秘匿されるセルを少なくするための考え方について協議する。その上で、第1回オープンデータの集計表を差し替えを行うかどうかを協議する。

### 2. 対応案

秘匿セルがある場合、行方向、列方向の2か所以上のセルを秘匿する要件を廃止した上で、以下のいずれかの対策を講じることとする。

（案1）平均値や割合の集計を除く、全ての集計値について、1の位を切り上げ

（案2）平均値や割合の集計を除く、全ての集計値について、四捨五入

（案3）総数のみ、1の位を切り上げ

（案4）秘匿されているセルが、行方向または列方向に1つだけの場合に、当該系列の中の最大値について、1の位を切り上げ

なお、上記のいずれであっても最小集計単位の基準（通常は10未満）に該当する場合は、集計値をハイフン「-」に置き換える（ただし、集計値の0は表示する）。

## 5. 最小集計単位の原則について：案1～3の例示

案1から案3の秘匿処理の例示は以下の通り。

秘匿処理前

	総数	A	B	C
総数	1476	39	372	1065
①	59	11	12	36
②	1302	2	305	995
③	115	26	55	34

(案1の処理) 全セル切り上げ

	総数	A	B	C
総数	1480	40	380	1070
①	60	20	20	40
②	1310	—	310	1000
③	120	30	60	40

(案3の処理) 総数のみ切り上げ

	総数	A	B	C
総数	1480	40	380	1070
①	60	11	12	36
②	1310	—	305	995
③	120	26	55	34

(案2の処理) 全セル四捨五入

	総数	A	B	C
総数	1480	40	370	1070
①	60	10	10	40
②	1300	—	310	1000
③	120	30	60	30

案3の場合、総数だけでなく小計も端数処理する必要がある。

例えば保険者別の集計がある場合、都道府県別は保険者別の小計にあたる。

このため、都道府県別集計値をすべて端数処理する必要が生じる。



## 5. 最小集計単位の原則について：案4の例示

秘匿処理前

	総数	A	B	C
総数	1476	39	372	1065
①	59	11	12	36
②	1302	2	305	995
③	115	26	55	34

秘匿処理（ステップ2）

	総数	A	B	C
総数	1476	39	372	1065
①	59	11	12	36
②	1302	—	305	1000
③	115	30	60	34

秘匿処理（ステップ4）

	総数	A	B	C
総数	1476	39	372	1065
①	59	11	20	40
②	1302	—	310	1000
③	115	30	60	34

秘匿処理（ステップ1）

	総数	A	B	C
総数	1476	39	372	1065
①	59	11	12	36
②	1302	—	305	995
③	115	30	55	34

秘匿処理（ステップ3）

	総数	A	B	C
総数	1476	39	372	1065
①	59	11	12	40
②	1302	—	310	1000
③	115	30	60	34

秘匿処理後（最終形）

	総数	A	B	C
総数	1476	39	372	1065
①	59	11	20	40
②	1302	—	310	1000
③	115	30	60	34

### 論点1

- ◆ 第2回以降のオープンデータにおいていずれの案を採用するか。

## 5. 最小集計単位の原則について：逆算のリスク

第1回オープンデータの差し替えを行う場合、  
すでに公表されているオープンデータを使って逆算できるリスクは次の通り。

秘匿処理前

	総数	A	B	C
総数	1476	39	372	1065
①	59	11	12	36
②	1302	2	305	995
③	115	26	55	34



公表済みの集計表 (第1回)

	総数	A	B	C
総数	1476	39	372	1065
①	59	—	—	36
②	1302	—	—	995
③	115	26	55	34



案4

	総数	A	B	C
総数	1476	39	372	1065
①	59	11	20	40
②	1302	—	310	1000
③	115	30	60	34

例示している組み合わせの場合、  
A①の「11」がわかることで、  
秘匿処理前の値をすべて逆算可能。



案2

	総数	A	B	C
総数	1480	40	370	1070
①	60	10	10	40
②	1300	—	310	1000
③	120	30	60	30

全セル端数処理されているため、  
元の値を逆算することは不可能。  
案1も同様。

## 5. 最小集計単位の原則について：論点2

### 論点2

- ◆ すでに公表しているオープンデータを差し替えるか否か。
- ◆ 差し替える場合、
  - ◆ メリット：第1回目から統一したルールで秘匿された状態になる。
  - ◆ デメリット：すでに公表済みのオープンデータを残している場合、差し替え前後の集計表の突合により、秘匿されている数字がわかる可能性がある。
- ◆ 差し替えない場合、
  - ◆ メリット：差し替えないことで安全性は担保される。
  - ◆ デメリット：2018年度と2019年度は他の年度と比べて、秘匿の仕方が異なる。

	公表済みの集計表との突合による逆算可能性(※)	秘匿処理方法
案1	逆算不可能	簡易
案2	逆算不可能	簡易
案3	逆算可能	簡易
案4	状況により逆算可能	複雑

(※) 第1回オープンデータの差し替え前後の数値を比較して逆算可能かどうか

## 6. 今後のスケジュール・第3回以降の検討課題

### 1. 今後のスケジュール

以下のスケジュールで進める。

- 2023年9月 第13回専門委員会にて、オープンデータの公表内容を協議
- 2023年11月頃 第2回介護DBオープンデータ公表

### 2. 第2回オープンデータには考慮せず、第3回以降に考慮する検討課題

- 匿名LIFE情報の集計仕様の検討
  - ・ 科学的介護推進体制加算以外の加算に関する詳細集計
  - ・ 複数テーブルに類似項目が分かれている場合の集計方法の検討、等
- 匿名要介護情報の集計仕様について
  - ・ 集計時点の追加（2017年度以前の追加）
- 公表対象データの追加検討（匿名介護レセプト等情報）
- 集計単位の検討
  - ・ 保険者別に集計するデータ項目を追加。
  - ・ 二次医療圏別の集計追加の検討。