

データ提供にあたってのポイント

平成22年11月25日

厚生労働省保険局総務課

(目次)

○ 前回の主な御議論	2
○ データの提供類型について	3
(1) 個票情報について	4
(2) 集計表情報について	13
(3) 全体スケジュール等	16

前回の主な御議論

- 医療機関コードの提供をどのように考えるか。
- 今回策定するガイドラインは、どの程度の期間を対象としたものとするか。
- 申請の受付件数(月10件程度)や方法についてどのように考えるか。
- データの提供類型(個票か集計表か)によって審査基準や提供対象範囲を変えるべきではないか。
- 成果物の2次利用に関する監視や評価も必要ではないか。
- 提供対象先として、一般社団・財団法人を含めると対象先が広くなりすぎるのではないか。
- データ提供の事務の一部について大学等の研究機関へ委託することも考えられるのではないか。
- データ処理を行う者は事前に登録された者に限定し、委託を認めないこととしてはどうか。

データの提供類型について

- レセプト・特定健診等情報データベースのデータ提供にあたっては、
 - ①データベースに収載されている個別のデータ毎に提供する個票情報と、
 - ②上記の個票情報を事務局において一定の集計を加えた集計表情報、の2つの提供類型を想定。
- 提供類型によってデータに含まれる情報の内容・性格が異なることから、それぞれに応じた審査基準や事務手続きを設定する必要。

(参考)統計法のデータ提供類型

提供の類型	情報の種類	提供先
調査票情報の提供 (統計法33条)	統計調査によって集められた情報のうち、文書、 図画又は電磁的記録に記録されているもの。	公的機関と補助金を受けた 研究者等。
委託による統計の作成 (オーダーメイド集計) (統計法34条)	行政機関等が一般からの委託に応じ、その 行った統計調査に係る調査票情報を利用して、 統計の作成等を行ったもの。	一般 ※ 学術研究の発展に資す ると認める場合など。
匿名データの作成・提供 (統計法35・36条)	一般の利用に供することを目的として調査票情 報を特定の個人又は法人その他の団体の識別 ができないように加工したもの。	一般 ※ 学術研究の発展に資す ると認める場合など。

(1) 個票情報について

匿名データの提供(統計法)における識別情報の取扱い

○統計法第2条

12 この法律において「匿名データ」とは、一般の利用に供することを目的として調査票情報を特定の個人又は法人その他の団体の識別(他の情報との照合による識別を含む。)ができないように加工したものをいう。

○全国消費実態調査、社会生活基本調査、就業構造基本調査及び住宅・土地統計調査に係る匿名データの作成について(諮問) 平成20年12月22日(内閣府統計委員会への総務大臣の諮問)

2. 識別情報

(1) 地域区分

一般に、統計分析では、調査客体の属性に着目して分析を行う構造分析と、調査客体の地域に着目して分析を行う地域分析の、二つのアプローチがある。

公的統計の統計表編成においては、全国レベルの集計では属性に関する詳細な分類を用いるのに対して、地域別の集計ではそれよりも粗い分類を用いることがある。

これは、集計する統計の精度とともに、地域と属性の両方で詳細に集計し過ぎることにより調査客体が識別される可能性を回避することにも配慮したものである。

○住宅・土地統計調査

個人の職業・産業等の情報が含まれていないこと、10%リサンプリングであること等を踏まえ、都道府県単位でのデータを提供

○そのほか3統計(全国消費実態、社会生活基本、就業構造基本)

全国を2ブロック(3大都市圏とそれ以外)に分けて、ブロック単位でのデータを提供

匿名データの提供とレセプト情報等データベースの比較

匿名データの提供(統計法)

- 匿名データ(統計法)は他の情報との照合によって特定の個人や法人等の識別可能性をなくすために、特に地域性に関わる情報について厳格な処理(統計局4調査では、他の秘匿措置等とも併せて総体的に検討し、全国2ブロック又は都道府県単位での提供)を行っている。
- 一方で、「一般の求め」により、学術研究及び高等教育目的等については、学術研究者をはじめ学生や国際機関にも一定の条件の下で提供を広く認めている。

レセプト・特定健診等情報データベース

- 研究目的によっては、地域性に関わる情報が必要となる場合もあるとの意見。(例:地域での医療連携や医薬分業の状況など。)
- データベースの中の地域性に関わる情報は、都道府県コード、医療機関・薬局コード、保険者番号(※)、特定健診機関コード、特定健診・保健指導対象者の居住地の郵便番号がある。
 - ※ 市町村国保の保険者番号は地域性に関わる情報となると考えられる。また、健診受診者の郵便番号はそもそも提供しない。
- 個票情報のガイドラインのポイントでは、提供対象先として学生、営利企業や国際機関を含めないこととして、匿名データの提供より狭く設定。

医療機関・薬局コード及び保険者番号の取扱い

行政機関保有個人情報^①を学術研究のために提供できる理由

専ら統計の作成や学術研究のために保有個人情報を利用する場合には、特定個人が識別できない形で用いられるものが通常であり、個人の権利利益を侵害するおそれが少なく、かつ、公共性も高いと考えられることから、利用目的以外の利用・提供の原則禁止の例外としたもの。

(「行政機関個人情報保護法の解説(増補版)」総務省行政管理局)

目的が統計の作成や学術研究であっても特定個人を識別しうる分析・研究方法については審査に当たって抑制的に考える必要があるのではないか。

個別の医療機関コードの情報を提供することが、患者個人の方の特定につながるような場合があるならば、そのような事態は回避する必要。

また、比較的小規模な保険者の保険者番号にも留意する必要がある。

○ 医療機関・薬局コード及び保険者番号の提供は原則行わないこととし、経年データを分析する場合には医療機関等に新たな通し番号等を付番することを検討してはどうか。

○ ただし、例外的な場合として、地域性の分析・調査にのみ用いる場合に、その目的に照らして最小限の範囲内で有識者会議における審査を経て提供できる場合を認めてはどうか。

なお、その場合においても、医療機関等の個別の同意がある場合等、有識者会議が特に認める場合を除いて、公表される成果物の中に特定の医療機関・薬局及び保険者を識別できる資料・データ等は盛り込まないこととし、違反した場合には、他の不適切利用への措置(データ提供禁止)よりも重いペナルティー(氏名・機関の公表)を設けることとしてはどうか。

(参考) 米国でのデータ提供における識別情報の取扱い

米国のメディケア関連情報について、Limited Data Set(患者の情報は匿名化:前回の資料)における提供の類型は、大きく分けて、診療報酬請求書ベースのStandard Analytical Files (SAFs)と入院患者についてのデータのみ在院期間ベースでまとめたMedicare Provider Analysis and Review (MEDPAR)がある。

個別の医療機関の情報については、NPI(National Provider Identifier)が提供されている。

(注)NPIとは、Health Insurance Portability and Accountability Act (HIPAA)に基づき定められた10桁の数字からなる医療機関毎のID。医療提供者やヘルスケアのクリアリングハウスは、HIPAAに基づく事務・財務上の処理には、このIDを使用しなければならないこととなっている。

Limited Data Set の提供類型

提供類型	識別情報の取扱い
Standard Analytical Files (SAFs)	<ul style="list-style-type: none">○ 特定の診療日は記載されず、日付は全て4半期又は1年間でまとめられている。○ 患者の方の年齢は5年刻みで表記。○ 医師の特定番号(UPIN)は暗号化されている。○ 最も狭いレベルの地域の識別情報は郡(county)。 ※ 患者の方の住所として郡単位まで記載。
Medicare Provider Analysis and Review (MEDPAR) file	<ul style="list-style-type: none">○ 患者の年齢や医師の番号については上記と同様。○ 患者の方の入退院日は、4半期毎でまとめられている。○ 最も狭いレベルの地域の識別情報は州(State)。 ※ 医療機関の情報として、所在する州、40程度の医療機関のタイプ (Short Stayか、アルコール病院など)等が記載。

(出典) ResDACのホームページより。(http://www.resdac.org/Medicare/Data_File_Descriptions_LDS.asp) など。

(参考1) SAFsのコード表(NPIの記述)

*Research Data Distribution Center
LDS Inpatient SNF Claim Record
Data Dictionary*

<i>Variable Name</i>	<i>Label</i>
<i>ORGNPINM</i>	<i>Organization NPI Number</i> On an institutional claim, the National Provider Identifier (NPI) number assigned to uniquely identify the institutional provider certified by Medicare to provide services to the beneficiary. NOTE: Effective May 2007, the NPI will become the national standard identifier for covered health care providers. NPIs will replace the current legacy provider numbers (UPINs, NPIs, OSCAR provider numbers, etc.) on the standard HIPPA claim transactions. (During the NPI transition phase (4/3/06 – 5/23/07) the capability was there for the NCH to receive NPIs along with an existing legacy number. NOTE1: CMS has determined that dual provider identifiers (legacy numbers and NPIs) must be available on the NCH. After the 5/07 NPI implementation, the standard system maintainers will add the legacy number to the claim when it is adjudicated. Effective May 2007, no NEW UPINs (legacy number) will be generated for NEW physicians (Part B and Outpatient claims) so there will only be NPIs sent in to the NCH for those physicians. SAS ALIAS: ORGNPINM STANDARD ALIAS: ORG_NPI_NUM

EXPANDED MODIFIED MEDPAR FILE

UPDATED MARCH 2010

88 National Provider ID (NPI)

Description:

Effective May 23, 2007, the National Provider Identifier (NPI) number is assigned to uniquely identify the institutional provider certified by Medicare to provide services.

Coding Scheme:

Ten-position character field.

Original Source:

Uniform Bill

Limitations:

The MEDPAR File contains only inpatient hospital records. Provider numbers are validated against a file of Medicare-certified providers by the intermediary. However, this process is not repeated when the MEDPAR file is constructed.

(出典):

<https://www.cms.gov/LimitedDataSets/Downloads/SAFidsSNFNov2009.pdf>

(参考2) MEDPARのコード表(NPIの記述)

(出典):

<https://www.cms.gov/LimitedDataSets/Downloads/FY09MedPARDec09Layout.pdf>

個別医療機関・保険者の属性情報について

データベース上の個別医療機関・保険者等の属性情報

医療機関	<ul style="list-style-type: none">○都道府県コード(01:北海道、02:青森)○医療機関コード(個別の保険医療機関毎に定められたコード)○診療科コード(旧総合病院は平成22年3月診療分まで記載)○病棟区分(01:精神病棟、02:結核病棟、07:療養病棟のいずれかに該当する場合のみ記載。)○病床数○薬局コード(調剤レセプトのみ)
保険者	<ul style="list-style-type: none">○保険者番号(各保険者固有の番号)

直接的な識別情報である医療機関コードを削除したとしても、都道府県と病棟区分や病床数との組み合わせで個別の医療機関が推定できる可能性が完全になくなる、とは言いきれないと考えられる。

(例:〇〇県に結核病棟が極めて少ない場合、病床数が具体的な数(111など)で示される場合など。)

ガイドライン上、そもそも他の情報とのリンケージ(照合)を行わない、とする取扱いを明記することにより対応することが考えられる(研究は、あくまでも提供を受けたデータセットのみを用いて行うものとする。)

また、属性情報に応じて一定の処理(数字を丸める、極端なデータは提供しない。)をした上で提供してはどうか。

(参考)他の情報提供での枠組みの取扱い

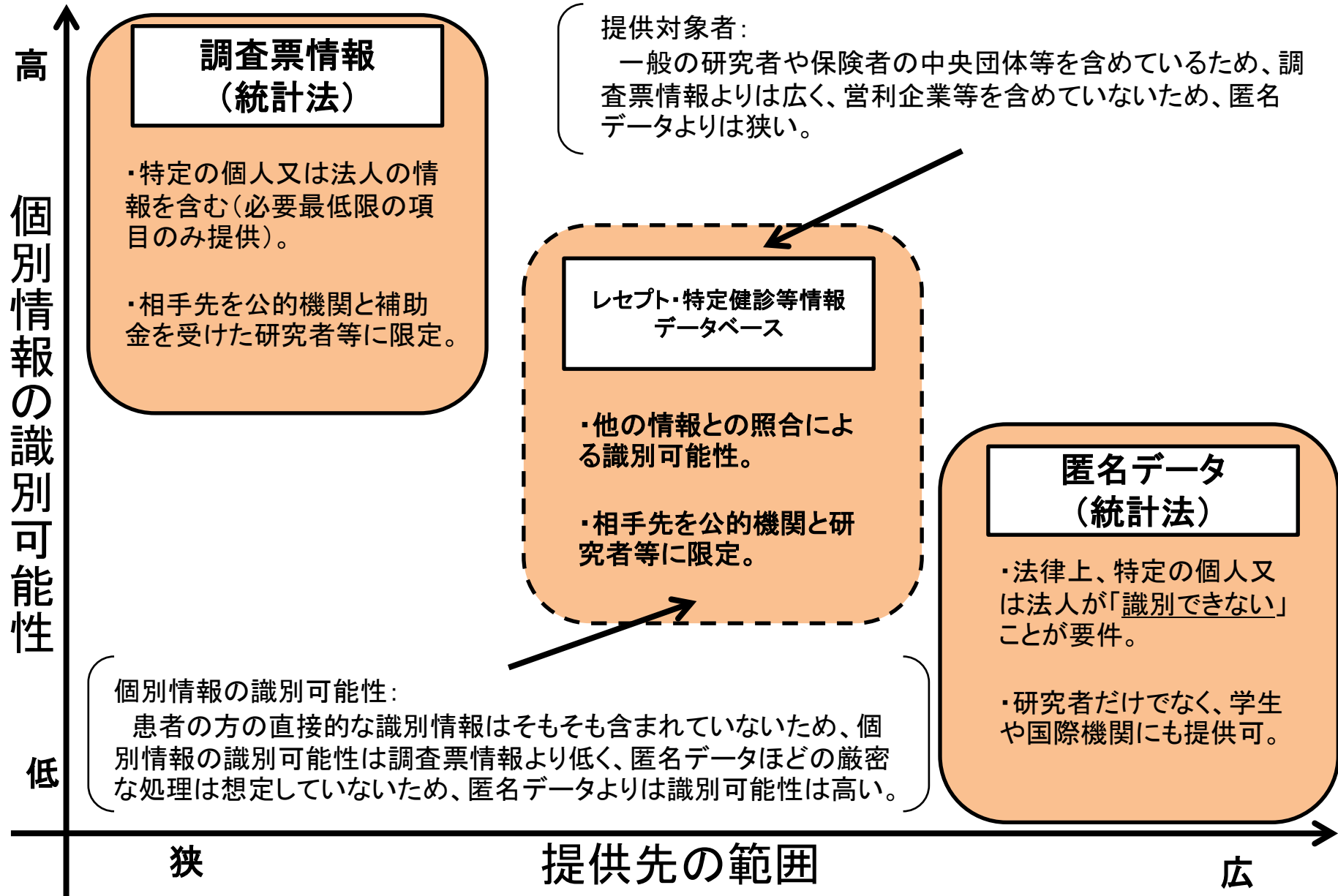
統計法における匿名データやメディケア・メディケイドのLimited Data Setのデータ提供においても他の情報とのリンケージを行わない、とするルールがある。

情報提供類型	取扱いの概要
統計法における匿名データの提供	<p>(匿名データの作成・提供に係るガイドライン) 第7 匿名データ提供依頼申出手続 1 あらかじめ明示しておく事項</p> <p>提供依頼申出手続を行う場合に提供依頼申出者があらかじめ了解しておくべき次の事項を提供機関等はホームページ等において提示し、広く周知する(関連:第4、第6の3、第7の6)。</p> <ul style="list-style-type: none">・ <u>法第36条に基づいて提供されたその他の匿名データ及びその他の個体識別が可能となる可能性があるデータとのリンケージ(照合)を行わないこと</u>
メディケア・メディケイドのデータ提供	<p>(Limited Data Set のData Use Agreement)</p> <p>5. (中略)利用者は、CMSからの許可なく、(この契約書の)セクション4で特定されたファイルに含まれる記録を、<u>その他の受益者固有の情報と照合してはならない</u>。</p> <p>5. (中略) Absent written authorization from CMS, the User shall not attempt to link records included in the file(s) specified in section 4 to any other beneficiary-specific source of information.</p> <p>(出典) ResDACのHPより(http://www.cms.gov/cmsforms/downloads/cms-r-0235l.pdf)。事務局において仮訳したものであり、公定訳を示すものではない。</p>

ガイドラインのポイント(前回提示資料)における情報提供類型

情報提供の類型	提供対象	データの匿名性
調査票情報	<ul style="list-style-type: none"> ○ 行政機関、地方公共団体、独立行政法人等、その他(会計検査院、地方独立行政法人、地方住宅供給公社、地方道路公社及び土地開発公社) ○ 下記のいずれかの統計の作成等を行う者(研究者等) <ol style="list-style-type: none"> 1 行政機関等の公的機関と委託又は共同で行う調査研究に係る統計の作成等 2 その実施に要する費用の全部又は一部を公的機関が公募の方法により補助する調査研究に係る統計の作成等 3 行政機関の長又は地方公共団体の長その他の執行機関が、その政策の企画、立案、実施又は評価に有用であると認める統計の作成等 <p>(統計法第33条、統計法施行規則第8～9条)</p>	<p>個人情報を含む個票ベースの情報を提供。 ※個人情報保護法は適用除外であり、統計法の枠組みで保護。</p>
匿名データ	<ul style="list-style-type: none"> ○ 一般からの求め(統計法第36条) <p>※利用目的が、学術研究・高等教育の発展や国際社会における我が国の利益の増進に資する場合。</p>	<p>調査票情報を特定の個人又は法人その他の団体の識別(他の情報との照合による識別を含む。)ができないように加工したものを提供</p>
レセプト・特定健診等情報データベース	<ul style="list-style-type: none"> ○ 行政機関、研究開発独法、大学、保険者の中央団体、医療サービスの質の向上等を設立目的に含む公益法人に所属する者 ○ その実施に要する費用の全部又は一部を公的機関から公募の方法により補助されている者 	<p>患者の方個人の直接的な識別情報はないが一方で、他の情報との照合による識別は完全には排除されない。</p>

ガイドラインのポイントにおける情報提供類型（イメージ）



(2) 集計表情報について

集計情報の提供イメージ

個票ベースの情報とは別に集計情報として、事務局において一定の集計を行った上で、より広い提供対象者に簡易な手続きでデータベースの情報を提供する仕組みを設けてはどうか。

集計情報(例)

以下のようにデータベース内の個票情報を、都道府県ベースを最小単位として単純集計したものを提供することとしてはどうか。※ 都道府県未満の地域性情報は医療機関・薬局コード、保険者番号、健診受診者の郵便番号であり、これらは提供はしない。

医療機関類型 (〇〇県)	平成〇年〇月 (診療行為Xの回数)	平成〇年〇月 (診療行為Xの回数)	平成〇年〇月 (診療行為Xの回数)
病院(400床以上)	〇〇回	〇〇回	〇〇回
病院(400床未満)	××回	××回	××回
診療所	▲▲回	▲▲回	▲▲回

ただし、都道府県ベースで見た場合にも医療機関等の数が非常に少なくなる集計単位による提供は行わないこととしてはどうか。

※ 最小の集計単位をどのようにするかは、別途検討が必要(例えば、病床数については、病院報告における単位より小さくなってはならない、などが考えられる。)

医療機関類型	平成〇年〇月 (診療行為Xの回数)	平成〇年〇月 (診療行為Xの回数)	平成〇年〇月 (診療行為Xの回数)
病院(1000床以上)	〇〇回	〇〇回	〇〇回
病院(400~999床)	××回	××回	××回
病院(400床未満)	■ ■回	■ ■回	■ ■回
診療所	▲▲回	▲▲回	▲▲回

提供しない

(参考) 病床規模別の施設数

	施設数		対前年		構成割合(%)	
	平成20年 (2008)	平成19年 (2007)	増減数	増減率 (%)	平成20年 (2008)	平成19年 (2007)
病院	8 794	8 862	△ 68	△ 0.8	100.0	100.0
20～49床	1 051	1 093	△ 42	△ 3.8	12.0	12.3
50～99	2 288	2 298	△ 10	△ 0.4	26.0	25.9
100～149	1 433	1 430	3	0.2	16.3	16.1
150～199	1 313	1 295	18	1.4	14.9	14.6
200～299	1 130	1 150	△ 20	△ 1.7	12.8	13.0
300～399	745	763	△ 18	△ 2.4	8.5	8.6
400～499	366	360	6	1.7	4.2	4.1
500～599	200	199	1	0.5	2.3	2.2
600～699	115	120	△ 5	△ 4.2	1.3	1.4
700～799	57	56	1	1.8	0.6	0.6
800～899	33	33	-	-	0.4	0.4
900床以上	63	65	△ 2	△ 3.1	0.7	0.7
一般診療所(有床)	11 500	12 399	△ 899	△ 7.3	100.0	100.0
1～9床	4 026	4 538	△ 512	△ 11.3	35.0	36.6
10～19	7 474	7 861	△ 387	△ 4.9	65.0	63.4

(出典) 平成20年度 病院報告

(3) 全体スケジュール等

試行期間の設定と期間中の審査方法

統計法における匿名データの提供の試行期間(平成16~20年度の4年間)の例を踏まえ、平成23~24年度を試行期間と位置づけた上で、以下のようにしてはどうか。

試行期間における検討

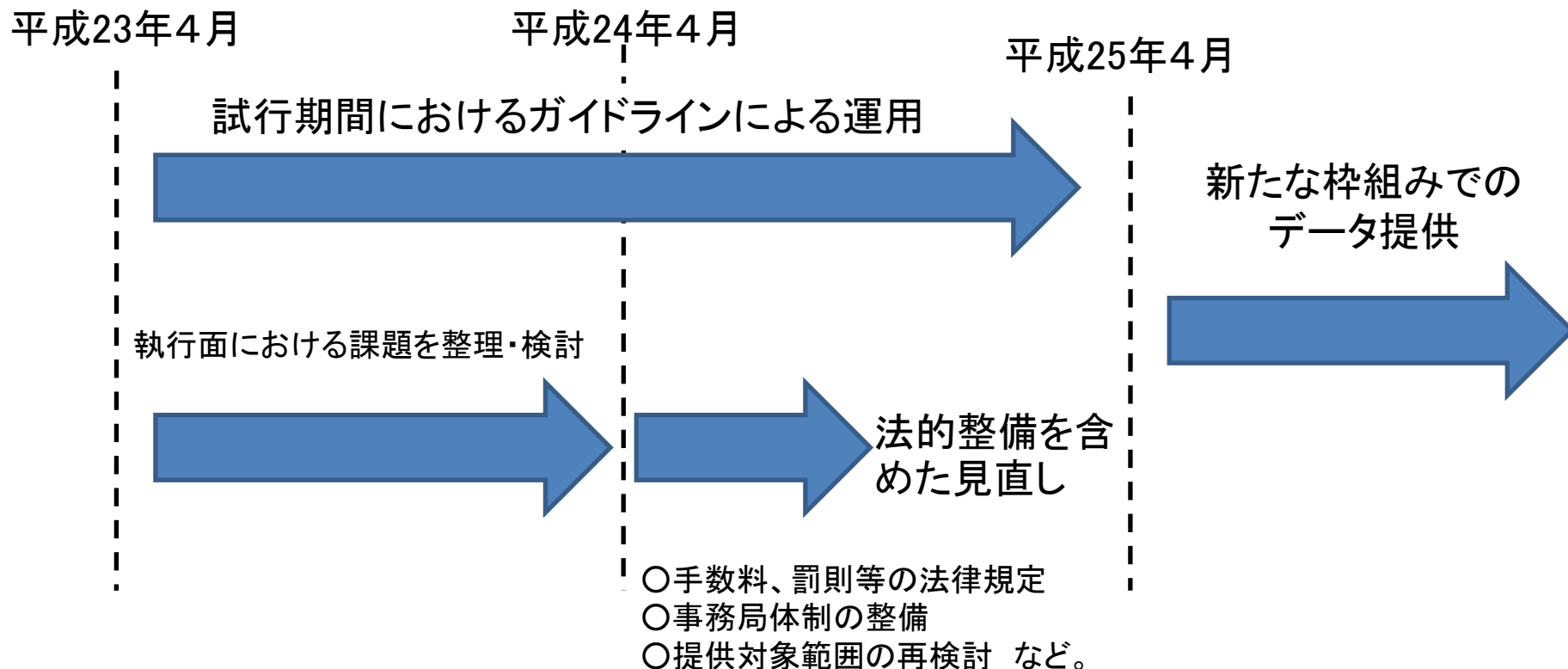
- 申請状況や具体的な審査の実施状況を踏まえて、対象者、審査基準、手続き等を随時見直すことにより法的な枠組みの整備(手数料・罰則等)につなげる。
- 利用者の具体的なニーズの把握・分析(特に集計情報における集計単位など)。
- 提供にあたっての技術的・事務的なコストを把握し、事務局の体制整備を検討。

試行期間の審査方法

- 匿名データの試行的提供の例を勘案し、一定の募集期間(年2~4回各10日程度)を設けた上で、その募集期間内に応募のあったものについて、当有識者会議で審査することとしてはどうか(前回資料では各月10件程度の受付と記載)。
- 集計情報についても、集計単位によっては、特定個人・機関の識別性の問題が生じる可能性もあることから、試行期間においては、提供先は個票情報と同様とし、実際の提供にあたっては、有識者会議の確認を得ることとしてはどうか。

提供にあたってのスケジュール(案)

- 統計法の匿名データ提供における試行的提供期間(平成16年～20年の4年間)も勘案し、以下のようなスケジュールとして、**23～24年度の実施状況を踏まえ**、将来的に法的整備を含めた制度作りを検討することとしてはどうか。
- したがって、現在御議論いただいているガイドラインのポイントについては、あくまで試行期間におけるもの、と位置づけてはどうか。



(参考)事務局の体制

匿名データの提供

仕組みの検討

: 総務省統計局調査企画課 2名

提供の実務

: (独)統計センター 統計データ
高度利用推進室 9名

技術的研究

: 総務省統計局、(独)統計セン
ター研究主幹、等

連携協力機関

: 一橋大、神戸大、法政大、情報・
システム研究機構

このうち専任で担当しているのは
職員2名程度

申請件数: 20件
(22年度上半期)

※この他問い合わせが月20件程度

レセプト・特定健診等情報 データベース

※ 専任の職員は1名。

その他データベース関連業務に携
わっている職員

: 厚生労働省保険局総務課11名

技術的な補助

: 厚生労働省保険局医療課、調
査課

※ 平成23年度定員において専
担の職員を1名増員要求中。

専任の職員1名

御議論を踏まえた修正等

その他、前回での御議論等を踏まえ、ガイドラインのポイント(案)を以下のように修正。
概ねこのポイントに沿って詳細なガイドラインや様式等の作成を行うこととしてはどうか。

- 集計情報の提供に適用されるガイドラインのポイントを別途、作成。
- ガイドラインのポイント(案)に成果物の評価に関する項目を追加。
- 審査基準について以下の要件を追加。
 - ・ レセプト情報の性格を理解した上でレセプト情報を活用することに合理性が認められること。
 - ・ レセプト情報を活用すること以外に研究目的を達成する方法がないこと。
- 中間成果物や提供データの破棄についての規定を追加。
- 一般社団・財団法人については、提供先として広すぎるとの御意見を踏まえ、提供対象先から除外。
- 不適切利用の際の措置として、公表の禁止を追加。
- 申請にあたっての事前相談を受け付ける窓口として提携協力機関の記述を追加。
- 事前に登録されていない者へのデータ分析等の委託禁止を明記。