

平成25年6月7日	資料4
第15回レセプト情報等の 提供に関する有識者会議	

DPCデータの提供について

平成25年6月7日
厚生労働省保険局医療課

DPCデータの第三者提供に係るこれまでの経緯の概要

- 平成22年6月22日に決定された「新たな情報通信技術戦略 工程表（高度情報通信ネットワーク社会推進戦略本部決定）」においては、レセプト情報等データベースの第三者提供に加えて、DPCデータの第三者提供についても提供形態の決定、ガイドライン策定に関する検討を行うこととされている。
- DPCデータは、レセプト情報とは記載内容やデータ様式が異なり、患者住所郵便番号や診断情報（傷病名等）等といったDPCに特有の慎重に扱うべき情報が含まれており、統一的な指針に基づいて提供ルールを検討する必要があることから、平成24年2月10日に開催された第8回レセプト情報等の提供に関する有識者会議において、DPCデータの提供については本有識者会議で検討することとされた。
- 平成24年9月5日の第11回有識者会議において、個人情報に係る影響の評価およびデータ提供のシミュレーションについて、厚生労働科学研究班において検討することとなった。

厚生労働省科学研究班における検討内容の概要

研究班の報告書を元に
事務局が作成

DPCデータ

I 集計表データを を提供する場合

現状

- ・すでに毎年DPC分科会においてDPCデータの集計表は公表しており、問題点は比較的少ない。
- ・患者数集計値については、個人識別防止の観点から、10を下回る場合はマスクされている。
- ・個別医療機関に関する情報はすでに公開されている。

II 個票データを を提供する場合

現状

- ・非常に濃密なデータであるが、検討すべき課題は多い。
- ・DPCデータの特徴として、個別医療機関ごとの集計値が既に公表されていることで、提供される個票データと既存公表データの組み合わせによる特異情報の識別の可能性がある。

検討すべき項目

- 1 個人を識別しうる情報の削除・変換をどうするか。
- 2 医学的に稀少な疾患、手術、処置等を含むデータへの対応をどうするか。
- 3 個別医療機関の識別を許容するか。

許容しない場合：

調査・研究としての用途は著しく限定される可能性がある。

許容する場合：

個人等が特定される可能性が非常に高まるため、使用目的の限定、提供する対象の限定、データ管理の厳格な規定等の対応が必要と考えられる。

研究班案

①患者数等の集計を提供する場合

- ・生年月日等個人識別につながりうる集計軸に関しては、集計粒度に下限をもうけること。
- ・1つのセルあたりの集計値が5または10を下回る場合は、その数値をマスクすること。

②医療機関数等の集計を提供する場合

- ・医師個人の特定を防ぐため、1つのセルあたりの集計値が3または5を下回る場合は、その数値をマスクすること。

研究班案

未解決の問題も残っていることから、試行的なデータ提供を試みて、課題・問題点を明らかにすることが望ましいと考えられる。

DPCデータの提供に係る今後の検討課題(案)

- DPCデータの提供については、以下の様に課題を整理して今後検討を進めることとしてはどうか。

① ガイドラインの整備について

- 個人情報保護法との関連
- 疫学研究に関する倫理指針との関係の整理
- データ提供の種類
- 提供先の範囲について
- データ提供にあたってのセキュリティ要件について
- 成果物の公表基準について 等

② 申請・審査・運用方法等について

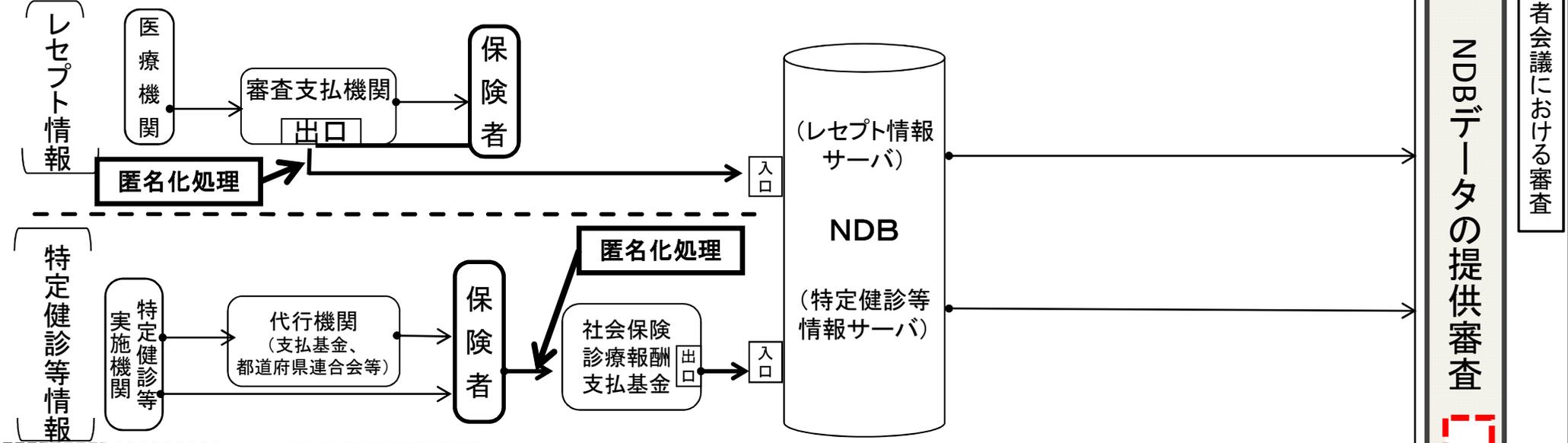
- 申請様式、審査方法、情報提供の手順について 等

③ DPCデータベースの構築等について

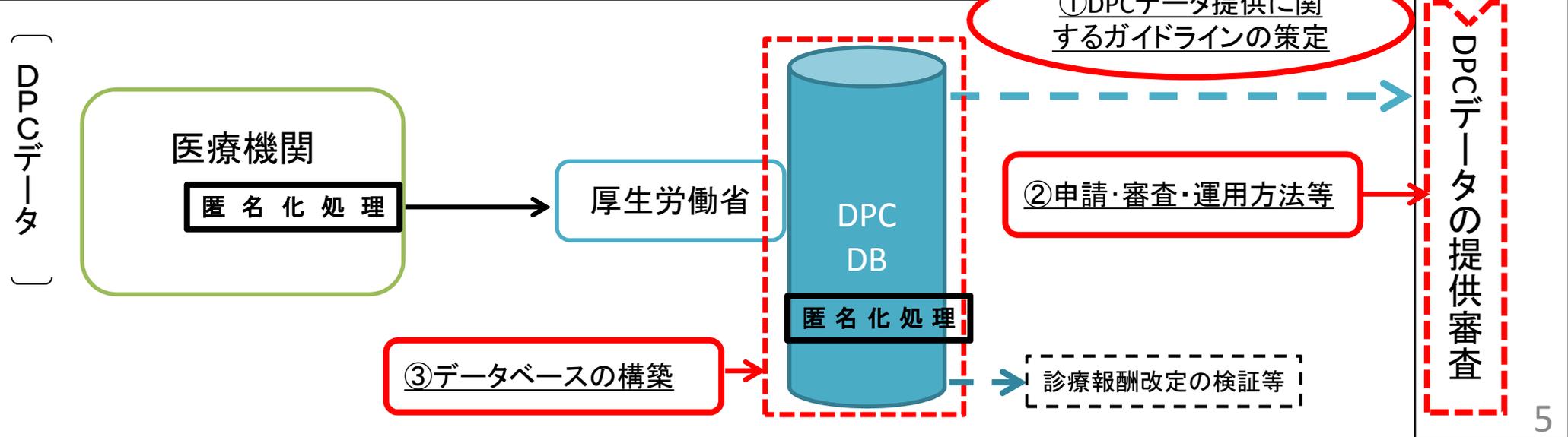
- データベースの様式について
- 構築・管理・運用の体制について 等

DPCデータベースの構築と活用のイメージ(案)

レセプト情報・特定健診等情報の活用方法



DPCデータベースの構築と活用のイメージ(案)



① ガイドラインの整備について（案）

現状

・DPCデータを第三者に提供することを規定するガイドライン等は現時点では存在しない。

【参考：DPCデータについて】（平成24年3月5日 保医発0305第1号 一部抜粋）

提出されたデータについては、厚生労働省に帰属し、個別患者を特定できないように集計した後、医療機関毎に公開されるものである。また、急性期入院医療を担う医療機関の機能や役割を適切に分析・評価するため、中央社会保険医療協議会の要請により適宜活用されるものである。

対応案

○研究班に対し、「レセプト情報・特定健診等情報の提供に関するガイドライン」に基づいてDPCデータの模擬申し出および模擬提供（集計表データおよび個票データ）を行い、以下の項目について検証することとしてはどうか。

- ・「DPCデータの第三者提供に関するガイドライン」の策定
「レセプト情報・特定健診等情報の提供に関するガイドライン」に含めて策定することについて検証を行う。
- ・審査基準の明確化
特定個人又は特定機関の識別情報の削除等について検証を行う。

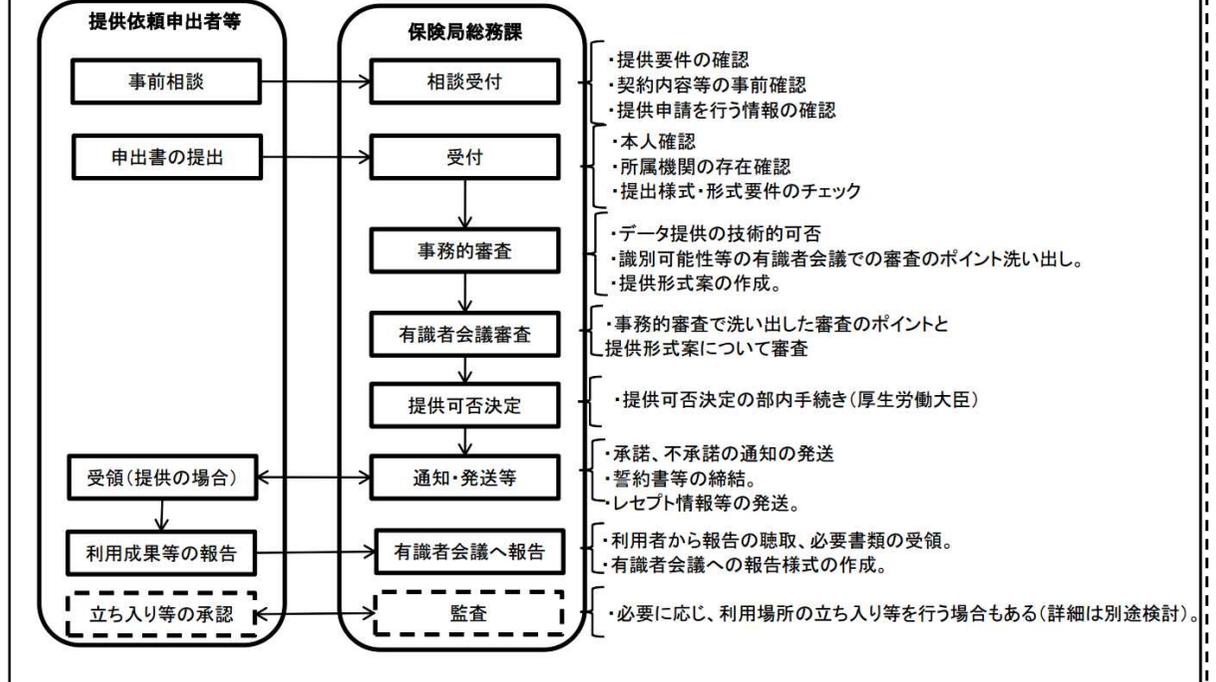
② 申請・審査・運用方法等について（案）

現状

DPCデータの提供方法

現在のところ整備されていない

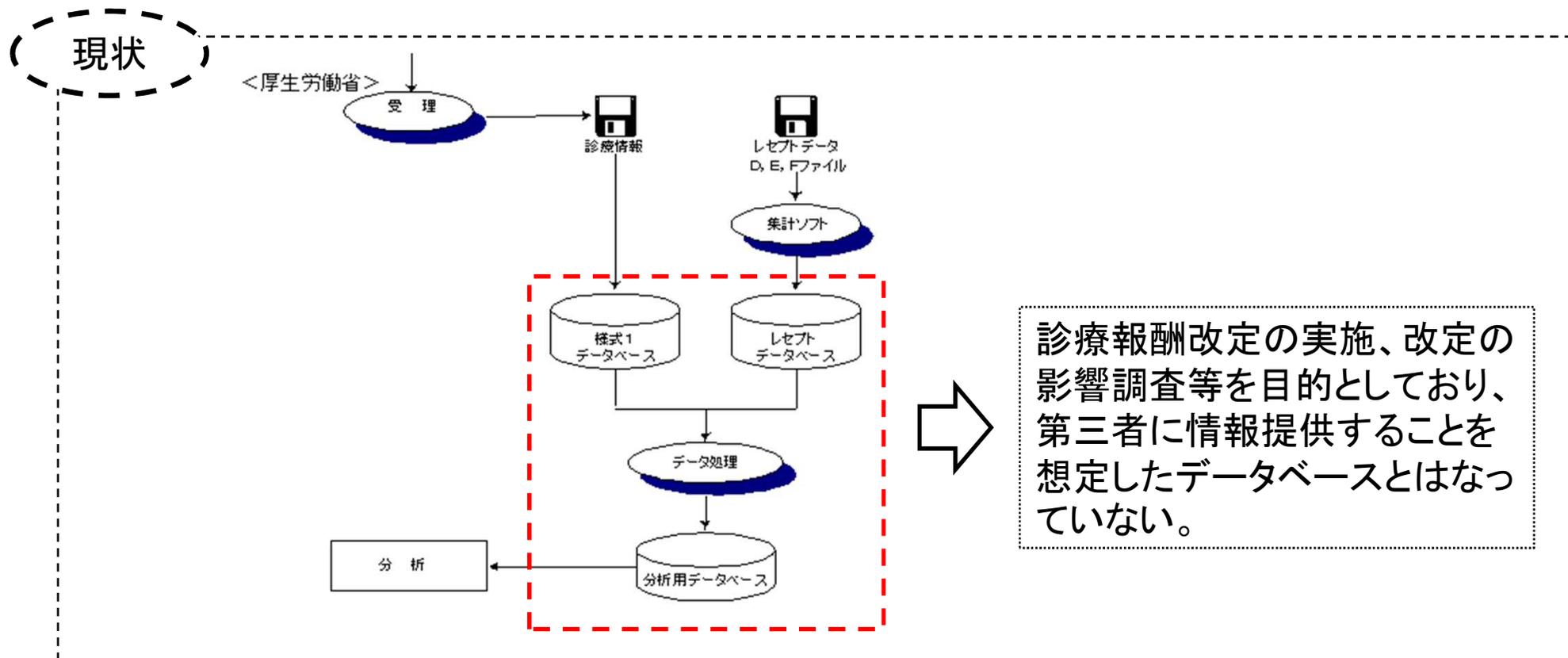
NDBデータの提供方法



対応案

- ・ DPCデータの提供方法については、現在運用されているNDBデータの提供方法の中で行うことを検討することとしてはどうか。

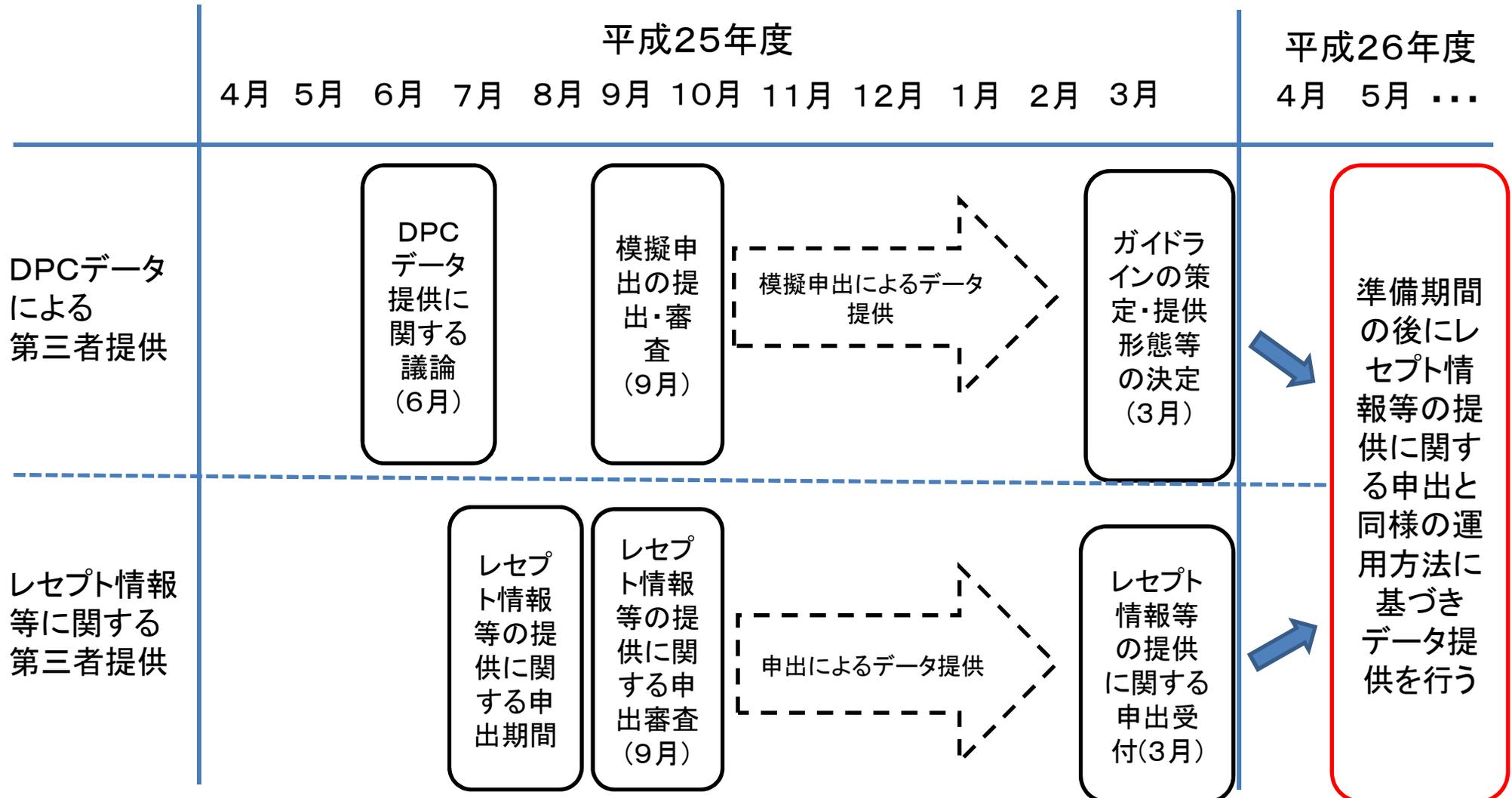
③ DPCデータベースの構築等について（案）



対応案

NDBと同一の方法でDPCデータの第三者提供を可能とするため、NDBと同様の形式でDPCデータベースを構築することとしてはどうか。

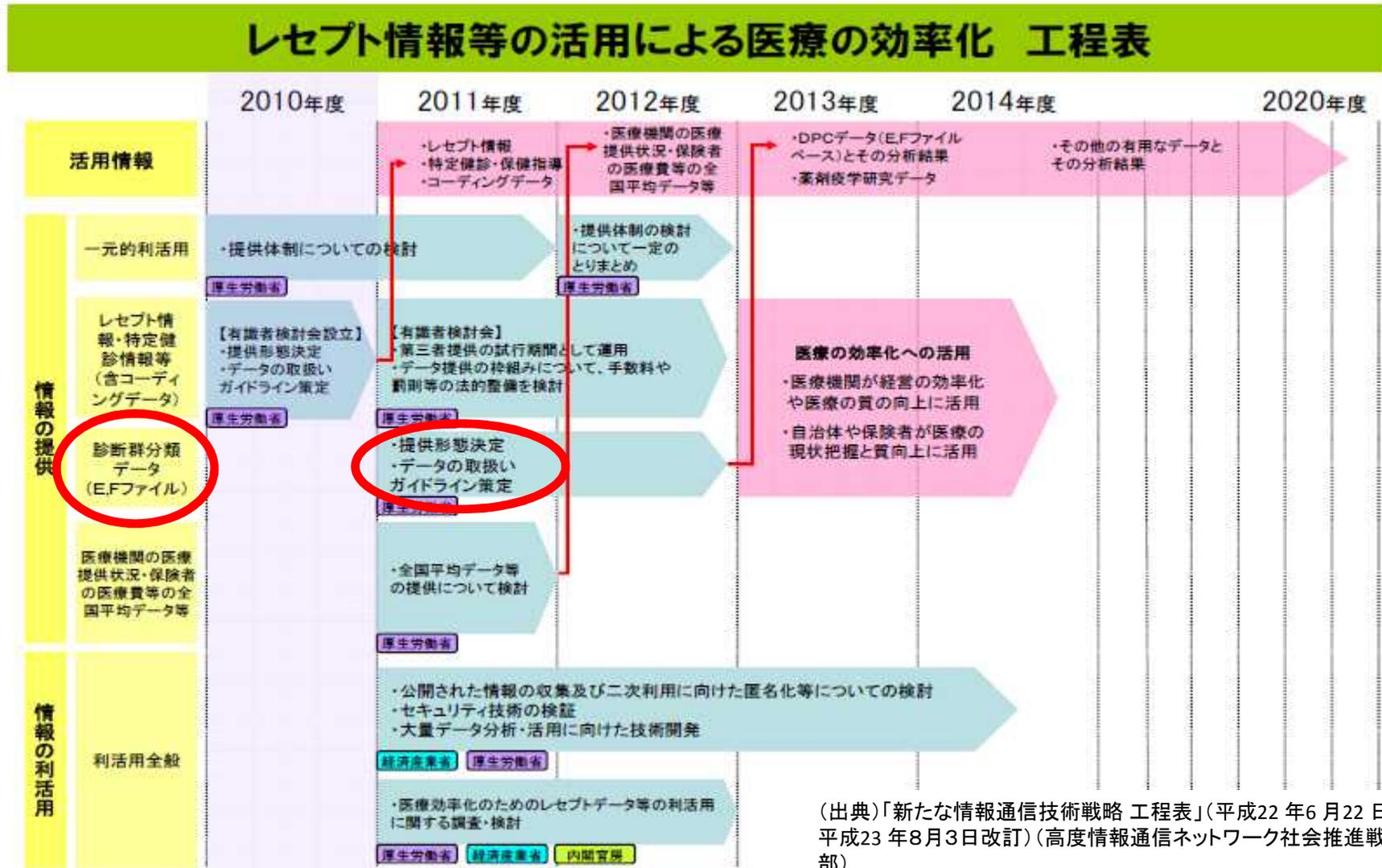
今後のスケジュール(案)



參考資料

DPCデータの提供について

○国が保有するレセプト情報・特定健診等情報データベースとは別に、「新たな情報通信技術戦略 工程表」(平成22年6月22日決定 平成23年8月3日改訂)(高度情報通信ネットワーク社会推進戦略本部)においては、退院患者調査に基づき収集したEファイル・Fファイル等についても提供形態等に関する検討を行うこととされている。



(出典)「新たな情報通信技術戦略 工程表」(平成22年6月22日決定 平成23年8月3日改訂)(高度情報通信ネットワーク社会推進戦略本部)

(参考)「新たな情報通信技術戦略 工程表」(抜粋)

レセプト情報等の活用による医療の効率化

【2010年度の実績】

- レセプト情報等の提供に関する有識者会議を設立し、レセプト情報等の第三者提供のためのガイドラインについて、レセプト情報等の提供に関する基本原則、提供の手続き、提供対象範囲、提供に際しての審査基準等を盛り込んで、2011年3月31日に策定。

厚生労働省：レセプト情報等の提供に関する有識者会議を設立。

レセプト情報等の第三者提供のためのガイドラインを2011年3月31日に策定。同ガイドラインの策定過程において、一元的利活用に関する検討を実施。

【今後の取組】

短期 (2011年度)

- レセプト情報等の第三者提供を試行期間として運用を開始する。また、膨大な関連情報の分析や活用のための技術等の研究開発を実施する。さらに医療効率化のためのデータ利用の在り方についての一次検討を実施し、各種データの一元的利活用に向けた提供体制についても検討を実施する。また、匿名化やセキュリティ技術、大量データ分析・活用に向けた技術開発について検討を開始する。

厚生労働省：早期にデータの提供開始。

引き続き、各種データの一元的利活用に向けた提供体制の検討を実施。有識者による検討会議において、レセプト情報・特定健診情報等について、第三者提供の試行期間として運用を実施。また、データ提供の枠組みについて、手数料や罰則等の法的整備の検討を実施。

医療機関・保険者による医療サービス・保健事業の質の向上及び経営効率化の取り組みを促進するため、医療機関が医療提供状況や経営状況に関する自らの位置付けを把握したり、保険者が自らの被保険者の医療費等の全国的な位置付けを把握できるように、全国平均のデータ等の提供について検討を実施。

医療効率化のためのレセプトデータ等の利活用に関する調査・検討を実施。

経済産業省：各種医療データを時系列で連結できる匿名化についての検討を開始。

厚生労働省と連携して、セキュリティ技術の検証を開始。

大量データ分析・活用に向けた技術開発を開始。

中期 (2012年度、2013年度)

- 各種データを活用し、質が高く、効率的な医療を実現する。また、引き続きレセプト情報等の第三者提供を試行期間として運用を実施する。さらに、一元的利活用に向けた提供体制の検討について一定のとりまとめを実施する。また、匿名化やセキュリティ技術、大量データ分析・活用に向けた技術開発について検討を実施する。

厚生労働省：2012年度中に医療機関が医療提供状況等を把握したり、保険者が自らの医療費等の全国的な位置付けを把握できるような全国平均データ等の提供開始。

2012年度中に医療情報データベース及びその他必要なデータベース間の一元的利活用可能な提供体制の検討について一定のとりまとめを実施。

2012年度についても引き続き、有識者による検討会議において、レセプト情報・特定健診情報等について、第三者提供の試行期間として運用を実施。また、データ提供の枠組みについて、手数料や罰則等の法的整備の検討を実施。

2013年度中にDPCに関するデータ(E、Fファイル等)の医療現場での利活用を推進。

経済産業省：引き続き、各種医療データを時系列で連結できる匿名化についての検討を実施。

引き続き、厚生労働省と連携して、セキュリティ技術の検証を実施。

引き続き、大量データ分析・活用に向けた技術開発を実施。

国が保有するDPCデータの利用の枠組み(案)

	法定目的内利用	法定目的外利用(第三者提供)
利用の根拠や手続を定めた法令等	<ul style="list-style-type: none"> ・診療報酬の算定方法の一部改正に伴う実施上の留意事項について(平成24年3月5日保医発0305第1号) ・DPC制度への参加等の手続きについて(平成24年3月28日保医発0328第1号) 	<div style="border: 2px solid black; border-radius: 25px; padding: 20px; text-align: center;"> <p>レセプト情報等の提供に関する有識者会議において検討を開始 (平成24年2月10日第8回有識者会議)</p> <p>※レセプト情報等の提供方法と同様の利用方法を想定</p> </div>
利用の目的	急性期入院医療を担う医療機関の機能や役割を適切に分析・評価に資するため	
利用できる者	<p style="text-align: center;">厚生労働大臣</p> <p>※公開される集計結果については全国民が利用可能</p>	
利用の手続	<p>中央社会保険医療協議会の要請により、急性期入院医療を担う医療機関の機能や役割を適切に分析・評価を行い、医療機関ごとに公開</p> <p>※公開されるデータは個別患者を特定できないよう配慮</p>	

DPCデータ*の全体像

* DPC/PDPS導入影響評価のための調査(退院患者調査)による調査データを指す。

○ 退院患者調査において、DPC対象病院、DPC準備病院(※1)及び出来高算定病院(※2)が厚労省に提出する情報は、以下のとおり。

内容		ファイル名称	
患者別匿名化情報	簡易診療録情報	様式1	
	診療報酬請求情報	医科点数表に基づく出来高点数情報(入院)	EF統合ファイル
		外来患者の医科点数表に基づく出来高点数情報	外来EF統合ファイル
		診断群分類点数表により算定した患者に係る診療報酬請求情報	Dファイル
		医科保険診療以外の診療情報	様式4
施設情報(病床数、入院基本料等加算、地域医療指数における指定状況等)		様式3	

※1 出来高算定制度で診療報酬請求を行いつつDPC制度に参加するための届出を提出し、その届出が認められた病院(DPC対象病院になる(DPC/PDPSによる支払を受ける)ためには、それ以前に2年間DPCデータを提出しなければならない。)

※2 出来高算定制度で診療報酬請求を行う病院で、DPC準備病院ではない病院(データ提出加算に関する届出を提出し、その届出が認められた病院)

過去の参考資料

退院患者調査の目的

平成24年2月10日
第8回有識者会議

- 診療報酬については、健康保険法第76条第二項の規定に基づき厚生労働大臣が定めることとされており、DPC/PDPSの診療報酬は中央社会保険医療協議会（中医協）の定めるルールに従い計算される「診断群分類点数表」（厚生労働省告示）により支払われている。
- 「診断群分類点数表」は「DPC導入の影響評価に係る調査（退院患者調査）」において提出されたDPCデータに基づいて設定されている。
- 当該調査は、診断群分類点数表の設定（診断群分類（DPC）の妥当性の検証）と、DPC/PDPS導入による診療内容への影響等の評価のための基礎資料を作成することが目的とされている。

DPCデータの提供についての論点等

平成24年3月7日
第9回有識者会議

- レセプトデータではなくDPCデータを活用することのメリット
 - 急性期入院医療の解析に適している(データセットやフォーマットが急性期医療向けに開発/改良されてきた)
 - 医療行為に関する日付情報があり(※)、入院医療における診療プロセスの詳細が分析可能(医療の質的評価、クリティカルパス・ガイドラインと実態診療との比較等)。
※ レセプトデータにも平成24年度から日付情報が記入される予定
 - 支払い(レセプト)にはない患者プロフィール(DPCデータ)による解析が可能
【例】 ・患者住所地の郵便番号(受療動向の把握、地域医療計画の検討)
・がんのTNM分類(がんの病期と診療内容の検討)
 - 主傷病と副傷病を区別した診療内容の分析(DPCデータとレセプトデータの質的相違)
 - ・ DPCは主傷病(唯一)を明確に区別して患者を分類するが、レセプトの記載病名は必ずしも主傷病と副傷病が明確に区別されていない。

具体的な対応について本日検討

- DPCデータの利活用で慎重な検討を要する点
 - 簡易診療録情報(様式1)を含む
 - ・ 診療内容に係る情報は個人情報としてよりセンシティブな取扱いが必要
 - ・ その他の複数の情報(生年月日・住所地郵便番号・入退院日等)を組み合わせることで個人を特定されるリスクが高まる可能性
- DPCデータの限界
 - DPC調査に参加していない施設がある(全医療機関を網羅しているわけではない)

検討課題と今後の対応

1. 様式1のデータを公表するにあたり、慎重な検討が必要

[例]

- ・ 単独である程度個人が特定できる可能性がある項目が含まれている
- ・ 複数の項目を組み合わせることで個人が特定できる可能性がある項目が含まれている
 - i) 様式1に含まれる項目同士を組み合わせる場合
 - ii) 様式1に含まれている項目と、既に公表されている情報(DPC導入の影響評価に関する調査)と組み合わせる場合

○ 様式1の調査項目数は100項目以上に登り、これらの項目ひとつひとつ、又はこれらの項目の組み合わせについて個人情報に係る影響を評価するには、より専門的な検討を行う必要がある。

2. DPCデータ提供形態のあり方について

○ 実際のDPCデータ提供の申し出を想定した技術的なシミュレーションを行う必要がある。

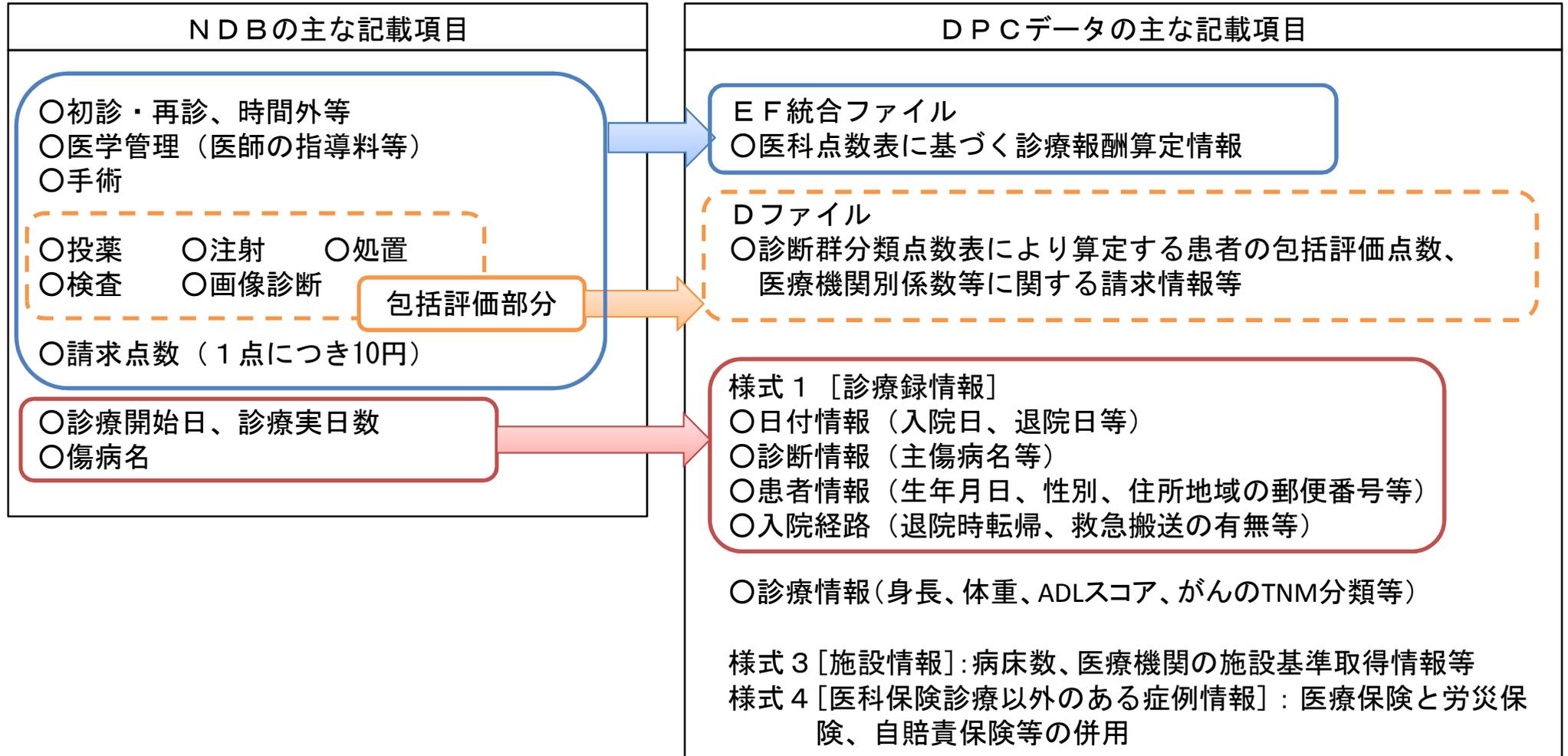


上記のような、個人情報に係る影響の評価とデータ提供のシミュレーションについて、厚生労働科学研究班を活用することとしてはどうか。

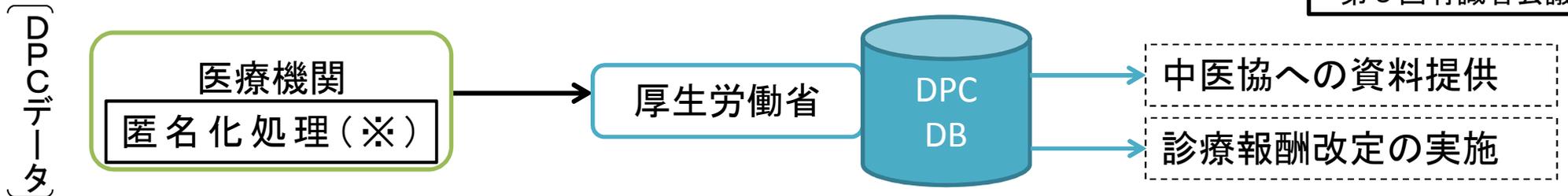
レセプトデータとDPCデータの主な相違点について（1）

平成24年3月7日
第9回有識者会議

1. 記載内容の相違(概要)



2. データ収集時点の相違

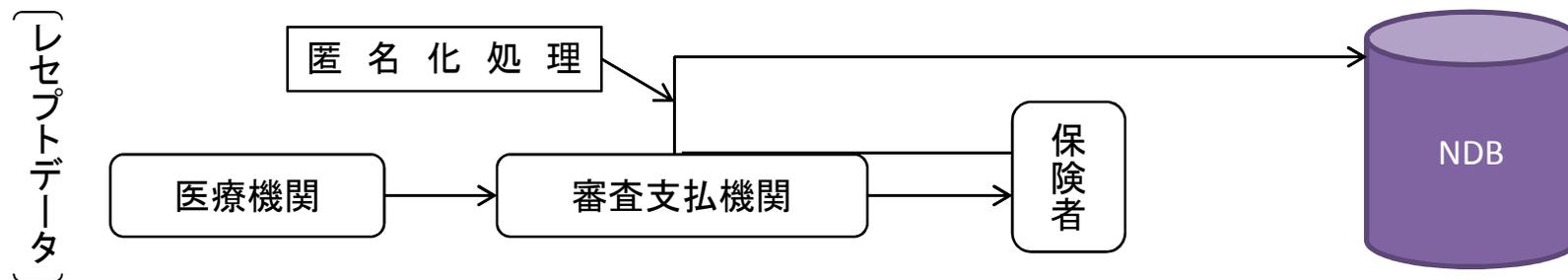


※ DPCデータにおける匿名化処理について

- DPCデータに患者の氏名は含まれず、医療機関毎に同一患者は同じ番号(ID)とすることを要求(当該医療機関においてのみ連結可能な匿名化であり、別の医療機関同士で比較は不可能)。カルテ番号、被保険者証等の記号・番号等、社会的に個人の有する番号は収集対象外。
- 上記以外で患者の属性に係る項目(患者プロフィール)については、特段の加工をしない状態でデータベースに収集。

【例】・ 性別 ・ 生年月日 ・ 患者住所地の郵便番号 ・ 入退院日

【参考】レセプト情報の収集経路と匿名化処理の方法(第1回 当会議 資料2-2より再掲)



※レセプト情報は、審査支払機関の審査後データを匿名化処理した後に収集。

【参考】ナショナルデータベースにおける匿名化処理(第1回資料2-2より再掲)

レセプトデータのうち、以下の項目は、同一人の特定する方策を講じた上で、削除されデータベースに収集される。

- 患者の氏名 ○生年月日の「日」 ○保険医療機関の所在地及び名称 ○カルテ番号等 ○国民健康保険一部負担金減額、免除、徴収猶予証明書の証明書番号 ○被保険者証(手帳)等の記号・番号 ○公費受給者番号

「DPC データ提供の課題について」

調査・研究目的でのレセプト情報等の提供が開始されているが、医療関連情報の有効活用の観点から「DPC 導入の影響評価に係る調査」で収集されているデータ（DPC データ）の調査・研究目的での提供が求められている。そこで、個人情報保護等の観点から DPC データの提供において生じうる問題点や危険性とそれらへの対応方法等を検討した。

まず、データの提供方式において、一定のレベルで集約された集計データを提供する場合と、なんらかの匿名化処置等を施した個票データを提供する場合では、生じうる課題が大きく異なるため、この 2 者を分けて検討を行った。

I. 集計データの提供について

既に中医協 DPC 評価分科会で 1 年分の集計データを毎年公表しているように、集計データの提供における問題点は比較的少ない。集計データの提供にあたっては、集計軸とその粒度、および一つの集計セル内のサンプル数に配慮することで、個人が識別される危険性をとり除くことができると考えられる。一方、DPC 評価分科会での公表では既に医療機関名が公表されているので、個別医療機関が識別されないようにする対応は、基本的に不要と考えられる。

集計軸について、生年月日、年齢、入院日、退院日、手術日、住所地郵便番号等のデータは、粒度を細かくすると個人の識別の危険性が高まるので、集計軸として用いる際には配慮が必要である。例えば、年齢は 5 歳刻み、日付等は曜日、週、月単位、住所地は人口を勘案して市区町村単位等を最小集計粒度とすることなどの配慮が必要である。但し、年齢については、新生児、乳児等の集計ではそれよりも細かい粒度の集計が必要とされるが、その際は、他の軸の粒度等に配慮して、個人が識別される危険性を取り除くことが必要である。

一つの集計セルあたりの患者数集計値の公表に下限を設けることは、個人の識別を防止する手段としては一定の効果は期待できるものの、絶対的に有効な手法とはいえない。仮に一つのセルあたりが一定の数値を下回る時に秘匿する処置を取ったとしても、多数の集計表データを組み合わせることで、その数値を認識しうることもあるからである。また、一部のセルの値をマスクすることで、それより粒度の粗い集計値を得ることができなくなってしまう問題点もある。現在、DPC 評価分科会で公表されている数値は、一つのセルあたり 10 を下回る場合は、マスクされている。この場合公表されている集計表があまり多くないため、個人識別防止の一定の効果は得られているが、一方、合計値等の集計ができなくなっているなど、問題が大きい。

以上の検討から、患者数等の集計値の提供においては、

- | |
|--|
| ①生年月日、年齢、入院日、退院日、手術日、住所地郵便番号等個人識別につながりうる集計軸に関しては、集計粒度の下限を設けること、
②1つのセルあたりの集計値が5または10を下回る場合は、その数値をマスクすること、 |
|--|

の2つの原則を組み合わせることが必要と考えられる。

医療機関数等を対象とした集計では、患者個人の識別につながる可能性はないが、医師等の医療機関の職員個人を識別する情報となる可能性があるため、その点についての配慮が必要である。例えば、特定の手術手技ごとの合併症の発生数などを集計した時に、手術を行った医師個人が識別されないような配慮が必要と考えられる。

そのため、医療機関数等の集計値の提供においては、

- | |
|---|
| ● 1つのセルあたりの集計値が3または5を下回る場合は、その数値をマスクすること、 |
|---|

を原則とすることが適当と考えられる。

II. 個票データの提供について

DPC データは年齢、住所地等個人に関する情報、日々の詳細な診療明細情報および個別医療機関に関する情報を含む非常に濃密なデータであるため、その個票データの提供にあたっては、検討すべき課題が非常に多い。また、DPC データの特徴として、個別医療機関ごとの集計値等が既に公表されていることで、提供される個票データと既存公表データとの組み合わせによる特異情報の識別の可能性についても検討する必要がある。

第一に、個票自体がもつ個人等を識別しうる情報の削除、変換の処置を講ずる必要がある。具体的には、以下のような対応が考えられる。

1. 様式1 データ

- 患者識別番号 →ダミーコードに変換
- 生年月日 →5歳刻み年齢等に変換
- 住所郵便番号 →市区町村等に変換
- 入院日、退院日 →在院日数に変換

2. 様式D、E、F、EF データ

- 実施日 →入院後日数に変換
- マスターID →削除
- 医師ID →削除
- 病棟ID →削除

第二に、個票データを提供する場合においては、医学的に非常に稀な傷病、手術、処置等を含むデータに対する対応を考える必要がある。これらの稀少な病態に関する情報を含む個票データは、稀な傷病名等およびそれらの組み合わせからは個人を識別される可能性があるため、これらを含む個票は提供対象から除外ことが望ましいであろう。しかし、どの程度の水準の稀な傷病等やその組み合わせが個人の識別につながりうるかは不明である。稀少傷病等の絶対的な基準を設けることは困難と考えられるので、今後、検討を続ける必要がある。

第三に、個別医療機関の識別に関する問題を検討する必要がある。この場合、提供される個票データの個別医療機関が識別されることを許容するか否かで分けて検討する。個別医療機関別の傷病別患者数等が公表されていて、特定の地域、傷病毎の患者数、施設基準の取得状況などと公表データを組み合わせることで、個別医療機関が識別されうる可能性が非常に高いため、一定の条件で抽出された個票データから医療機関が識別されることを完全に阻止するためには、以下のような対応が必要と考えられる。

- 医療機関が識別されうる診療報酬の施設基準に係わる加算等の情報をすべて除去
- 患者住所地情報を除去
- 医療機関の異同も含めて全く識別不能となるように医療機関を識別するコードを除去する、あるいは、医療機関の一定の属性情報を付与する場合は、一定の基準でサンプリングされたデータセットとする

このように個別医療機関が識別されないように作成されたデータセットは、当然のことながら医療機関の特性の影響、地域特異性の影響等の分析が困難となるため、調査・研究としての用途は著しく限定されたものとなると考えられる。

一方、個別医療機関の識別を許容する場合は、提供しうる情報の幅を広げることができるが、そのようなデータは自ら「特定の医療機関を受診する個別患者のデータ」となることを鑑み、個人等を識別されうる危険性が非常に高まることを念頭に最大限の注意を払わなくてはならない。

以上 3 つの視点から、個票データの提供に係る課題とその対応案を検討したが、個票データの提供においては、個人識別につながりうる情報の部分削除、ダミー化等いかなる手法を用いても個人を識別しうる情報を完全に除去することは非常に困難と考えられる。したがって、個票データを調査・研究目的等で提供する場合は、データを提供する対象を、個票データ以外のデータでは十分な成果が得られず、かつ学術的、医療政策的に一定の成果が期待される調査・研究に限定するとともに、データ提供を受ける者の情報管理に関する責任を明確化し、提供される個票データの管理等を厳格に規定する必要があると考えられる。

個票データの提供における個人情報の保護や安全性の確保などに関しては、未解決の問題も残っており、実際のデータ提供においては、新たな課題が生じる可能性も残されている。何らかの形で試行的なデータ提供を試みて、それらの課題、問題点を明かした上で、さらなる検討を続けることが望ましいと考えられる。