

DPCデータの提供について

平成26年3月20日（レセプト情報等の提供に関する有識者会議）

厚生労働省 保険局医療課

本日の資料の構成

1. これまでの概要等
 2. 伏見班への模擬提供の結果報告
 - 「集計表データ」の提供について
 - 「個票データ」の提供について
 3. 模擬提供を通じて
 - 伏見研究班からの報告
 - セキュリティ監査の実施について
 4. 今後の対応(案)
- + 過去の参考資料等

1. これまでの概要等

前回の「レセプト情報等の提供に関する有識者会議（平成25年6月7日開催）」の概要

- レセプト情報等の提供に関する有識者会議（平成25年6月7日開催）において、厚生労働省科学研究班（伏見班）から、第三者提供を行う上での問題点等について研究結果が報告された。
- 研究結果報告書の中で、個票データについては、「未解決の問題も残っているため、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」とされている。
- 研究班からの模擬申出・模擬審査を通じて審査基準の明確化等を検討することとなった。



「レセプト情報等の提供に関する有識者会議審査分科会（平成25年9月20日開催）」の概要

- 研究班（伏見班）より提出された集計表・個票それぞれの模擬申出書について、模擬審査を行い、集計表1件、個票1件、合計2件を承認し、実際にデータの提供を行った。



今回の「レセプト情報等の提供に関する有識者会議（平成26年3月20日）」

- 研究班（伏見班）から試行的データ提供を通じた検証結果について報告を受け、問題・課題を明らかにし、ガイドラインの策定の方針等を含め、今後のDPCデータ提供の在り方について検討する。

2. 伏見班への模擬提供の結果報告

平成25年9月20日の審査分科会における模擬審査に基づき、以下の2本の模擬申し出についてDPCデータの模擬提供を行った。

- ① 集計表データ
産業医科大学 公衆衛生学教室
久保達也・松田晋哉
- ② 個票データ
京都大学 医療経済学分野
猪飼宏・今中雄一

(集計表)

① 久保先生 (産業医大)

DPCデータを用いた若年性胆管がんの記述疫学研究

研究概要

- 印刷事業場での有機溶剤(1,2-ジクロロプロパン)暴露による職業性胆管がんが社会問題となっている。
- ところが今般、問題となっている50歳未満の若年発症胆管がんに関する報告及び情報は未だ限られており、その疫学的特性は不詳である。
- 同特性に関する情報は、臨床医療のみならず労働安全衛生の向上のためにも必要な情報である。
- 本研究では、DPCデータにおける様式1の情報を利用して胆管がんの年齢・地理分布等に関する疫学的記述を行った。

方法

- 様式1のみ
- Inclusion Criteria
 - [主病名]、[入院契機病名]、[医療資源病名]、[入院時併存症病名] [入院後発症疾患名]のいずれかに"C221" (肝及び肝内胆管の悪性新生物, 肝内胆管癌)、"C24\$" (その他及び部位不明の胆道の悪性新生物)の症例
- Exclusion Criteria
 - 同一患者の重複を可及的に除外しつつ様式1の臨床情報を効率的に補足するために、「施設コード」「データ識別番号」を連結した文字列(以下、RecIDと呼ぶ)を使用してRecIDが同一の場合は、「入院年月日」が最も早い入院記録のみを解析対象とした

申請における研究者側の工夫

- 症例数不足によるマスクを回避するために
 - 50歳未満の症例数ではなく、全症例数と50歳以上症例数を開示請求し、引き算により少数例を把握しようとした。
 - 症例数ではなく割合(%)を開示請求し、全体症例数に割合を掛けることで少数例を把握しようとした
 - 都道府県毎の集計以外に地方別(関東地方等)の集計を請求した
- ダブルカウントの問題を回避するために
 - 施設コード」「データ識別番号」を連結した文字列(RecID)を利用して同一施設内の個人を特定した
 - ただし複数施設受診のダブルカウントは回避不能のため研究計画自体の集計単位を個人ではなく入院とすることとした

結果1-1：患者の性・年齢・病期分布

年齢・病期	患者数	うち男性患者数	肝内胆管癌患者数†	胆道癌患者数‡	初発患者数	早期患者割合※
30歳未満	85	43	25	60	17	38.5%
35歳未満	70	34	20	51	23	38.9%
40歳未満	171	105	42	130	79	28.6%
45歳未満	306	185	97	213	132	23.4%
50歳未満	489	308	108	384	226	31.1%
55歳未満	896	590	241	662	485	27.6%
60歳未満	1,696	1,122	441	1,265	951	27.0%
65歳未満	4,260	2,917	1,017	3,267	2,459	31.2%
70歳未満	5,100	3,364	1,138	3,988	2,975	35.2%
75歳未満	6,684	4,373	1,423	5,311	3,896	37.0%
80歳未満	8,015	4,958	1,520	6,541	4,587	37.4%
85歳未満	6,970	3,958	1,061	5,934	3,853	39.6%
85歳以上	6,433	2,730	798	5,656	3,515	40.2%
不明	-	-	-	-	-	-%

※：病期が判明している症例に占める病期Ⅱ以下の患者割合

†：C221が病名コードに含まれた患者（病期毎の集計においては医療資源病名のみで判定）

‡：C24\$が病名コードに含まれた患者（病期毎の集計においては医療資源病名のみで判定）

※：審査基準の検証を行うため、10未満のセルについても提供を行い、公表時にマスクをしている。（以下同じ。）

公表の基準に基づき、10以下のセルはマスクしている
（実際には10以下のセルもふくめ提供）

ダブルカウントを排除した形でも全国の胆管がん症例が豊富に記述された

結果1-2:患者の病期分布

年齢・病期	患者数	うち男性患者数	肝内胆管癌患者数†	胆道癌患者数‡	初発患者数	早期患者割合※ 50歳未満患者割合
0	63	41	-	37		-
I	2,871	1,777	248	2,434		1.8%
INOS	-	-	-	-		-
IA	98	56	-	67		-
IB	76	48	-	64		-
IBI	-	-	-	-		-
II	3,257	2,035	485	2,499		1.5%
II NOS	-	-	-	-		-
II A	76	50	-	63		-
II B	66	47	-	-		-
III	3,271	2,008	571	2,356		1.8%
III NOS	13	-	-	-		-
III A	59	37	-	21		-
III B	60	35	-	40		-
III C	-	-	-	-		-
IV	485	317	82	210		2.3%
IV NOS	778	444	223	474		3.0%
IV A	2,844	1,668	547	2,007		2.0%
IV B	3,950	2,310	1,414	2,049		2.8%
IV C	-	-	-	-		-
不明	5,622	3,292	970	4,143		2.2%
未入力(NULL)	17,568	10,507	967	2,817		3.6%

※: 病期が判明している症例に占める病期Ⅱ以下の患者割合

†: C221が病名コードに含まれた患者(病期毎の集計においては医療資源病名のみで判定)

‡: C24\$が病名コードに含まれた患者(病期毎の集計においては医療資源病名のみで判定)

公表の基準に基づき、10以下のセルはマスクしている
(実際には10以下のセルもふくめ提供)

✓病期入力状況(不明・未入力が56.3%と過半数)であったが、豊富な病期情報が包括的に示された

結果2：入院時併存・合併症

	N	初発肝内胆管癌患者†に占める割合	初発胆道癌患者数‡に占める割合
B： 50歳未満患者集計			
原疾患			
初発肝内胆管癌患者†	149		
初発胆道癌患者数‡	267		
合併症（入院時併存症）			
胆のう胆管及び膵の障害	200	26.85%	59.93%
胆管結石症	11	-	-
胆道のその他の疾患	163	21.48%	49.06%
胆管炎	56	-	-
慢性ウイルス肝炎	13	-	-
慢性C型肝炎ウイルス感染	-	-	-
慢性B型肝炎ウイルス感染	-	-	-
肝硬変	-	-	-
アルコール性肝硬変	-	-	-
肝線維症・肝硬化症	-	-	-
胆汁性肝硬変	-	-	-
胆のう胆管及び肝の先天奇形	14	-	-
胆管の先天奇形	14	-	-
肝吸虫感染	-	-	-
潰瘍性大腸炎	-	-	-
クローン病	-	-	-

※：病期が判明している症例に占める病期Ⅱ以下の患者割合

†：G221が医療資源病名にコードされた初発患者のうち当該合併症を持つ患者の割合

‡：G24\$が医療資源病名にコードされた初発患者のうち当該合併症を持つ患者の割合

公表の基準に基づき、10以下のセルはマスクしている

（実際には10以下のセルもふくめ提供）

- ✓ 初発患者のみを対象に胆管癌のリスク因子として知られる合併症コードがどの程度発生しているかを集計
- ✓ 一般の有病率と比較することでリスク因子の検討が可能となる可能性がある（容易ではないが）
- ✓ 少数症例を対象とした集計でも割合で開示請求することでN数を同定可能である

結果3-1 患者の地理的分布(地方)

	患者数	男性患者数	女性患者	50歳以上患者数	50歳未満患者数
地方別集計					
北海道地方	2,307	1,389	918	2,267	40
東北地方	3,756	2,185	1,571	3,663	93
関東地方	11,507	7,046	4,461	11,122	385
中部地方	6,786	4,071	2,715	6,615	171
近畿地方	7,241	4,321	2,920	7,022	219
中国地方	2,564	1,542	1,022	2,508	56
四国地方	1,440	844	596	1,414	26
九州沖縄地方	5,427	3,195	2,232	5,300	127
不明	147	94	53	-	-

公表の基準に基づき、10以下のセルはマスクしている
(実際には10以下のセルもふくめ提供)

- ✓女性患者数と50歳未満患者数は引き算により算出した
- ✓人口動態調査等と組み合わせることで地域の有病率を検討することが可能と考えられた

結果3-2 患者の地理的分布(都道府県)

	患者数	男性患者数	女性患者	50歳以上患者数	50歳未満患者数		患者数	男性患者数	女性患者	50歳以上患者数	50歳未満患者数
北海道	2,307	1,389	918	2,267	40	滋賀県	416	250	166	-	-
青森県	499	295	204	482	17	京都府	809	489	320	788	21
岩手県	647	393	254	633	14	大阪府	2,891	1,796	1,095	2,782	109
宮城県	707	425	282	688	19	兵庫県	1,736	966	770	1,695	41
秋田県	539	296	243	527	12	奈良県	441	257	184	426	15
山形県	535	299	236	524	11	和歌山県	399	240	159	-	-
福島県	829	477	352	809	20	鳥取県	272	165	107	-	-
茨城県	862	518	344	846	16	島根県	310	182	128	-	-
栃木県	661	389	272	634	27	岡山県	613	366	247	603	10
群馬県	594	345	249	580	14	広島県	928	558	370	902	26
埼玉県	1,769	1,089	680	1,703	66	山口県	441	271	170	-	-
千葉県	1,534	965	569	1,490	44	徳島県	286	173	113	-	-
東京都	3,494	2,080	1,414	3,364	130	香川県	357	218	139	-	-
神奈川県	2,593	1,660	933	2,505	88	愛媛県	425	236	189	414	11
新潟県	802	485	317	786	16	高知県	372	217	155	-	-
富山県	447	278	169	431	16	福岡県	1,827	1,059	768	1,771	56
石川県	425	227	198	410	15	佐賀県	292	164	128	-	-
福井県	302	178	124	-	-	長崎県	550	320	230	529	21
山梨県	240	145	95	-	-	熊本県	649	391	258	635	14
長野県	975	595	380	956	19	大分県	459	269	190	-	-
岐阜県	665	380	285	645	20	宮崎県	360	213	147	-	-
静岡県	1,189	733	456	1,166	23	鹿児島県	833	501	332	-	-
愛知県	1,741	1,050	691	1,688	53	沖縄県	457	278	179	442	15
三重県	549	323	226	531	18	不明	147	94	53	-	-

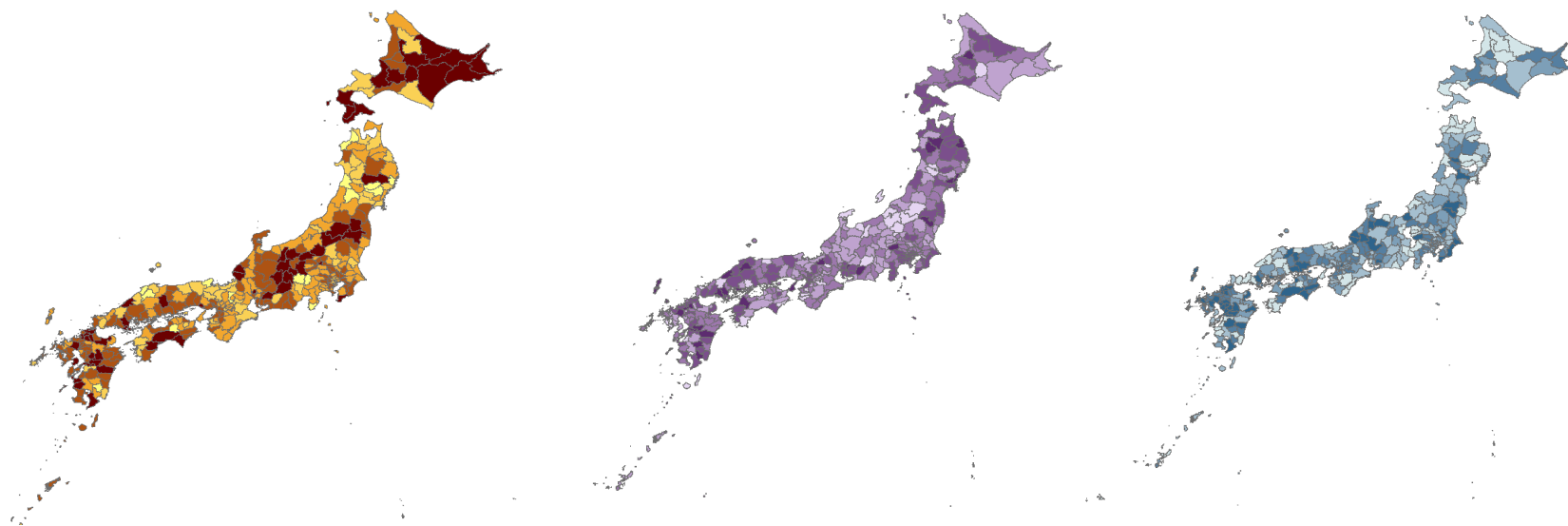
公表の基準に基づき、10以下のセルはマスクしている
(実際には10以下のセルもふくめ提供)

- ✓ 結果3-1と同様に女性患者数と50歳未満患者数は引き算により算出した
- ✓ N数が豊富なDPCデータにより全都道府県における若年発症胆管がん患者数が記述可能であった

(個票データ) ②京都大学医療経済学分野

患者居住地二次医療圏単位でみた地域別医療パフォーマンスの分析

心筋梗塞・心不全・脳梗塞の3疾患につき、主要な医療の質指標を患者居住地二次医療圏別に分析した。



(個票データ) ②京都大学医療経済学分野

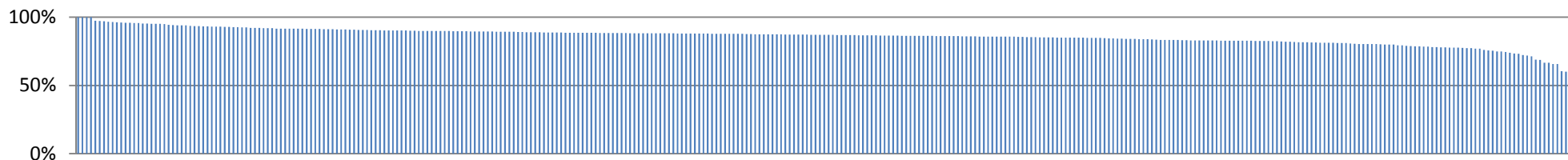
臨床指標の例:

1. 急性心筋梗塞患者における
 - ・入院後早期アスピリン投与割合
 - ・アスピリン投与割合
 - ・βブロッカー投与割合
 - ・ACE阻害剤もしくはアンギオテンシンⅡ受容体阻害剤の投与割合
2. 心不全患者における
 - ・βブロッカー投与の割合
 - ・ACE阻害剤もしくはアンギオテンシンⅡ受容体阻害剤投与の割合

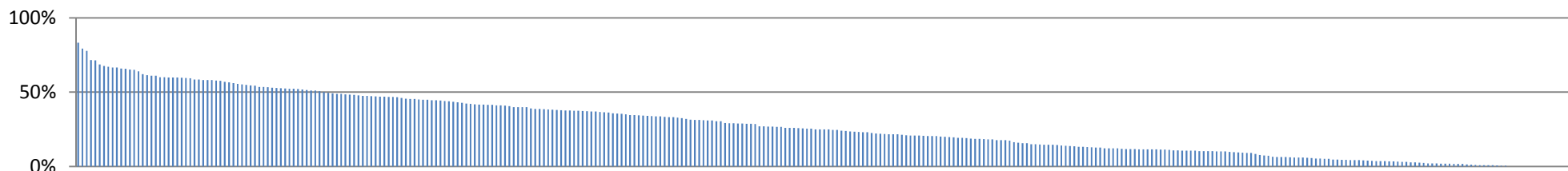
3. 急性心筋梗塞(再発性心筋梗塞含む)患者・慢性心不全患者における
 - ・心臓リハビリ実施割合
4. 脳卒中の診断で入院し、
 - ・リハビリ治療のための評価を受けた症例の割合
 - ・リハビリ治療を受けた症例の割合
 - ・入院後早期にリハビリ治療を受けた症例の割合
5. 脳梗塞(TIA含む)の診断で入院し、
 - ・第2病日までに抗血栓療法を受けた症例の割合
 - ・入院中に未分画ヘパリンを投与されなかった症例の割合

(個票データ) ② 京都大学医療経済学分野

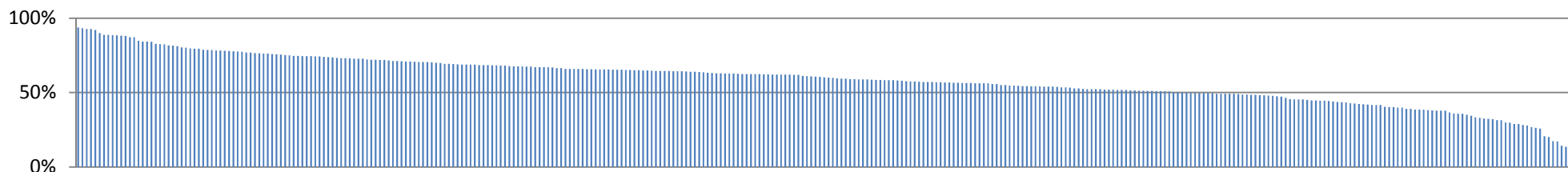
急性心筋梗塞患者における入院後早期アスピリン投与割合



慢性心不全患者・急性心筋梗塞(再発性心筋梗塞含む)患者における心臓リハビリ実施割合



脳梗塞の診断で入院し、入院後早期にリハビリ治療を受けた症例の割合



3. 今回の模擬提供を通して

①伏見研究班からの報告

- DPCデータの個人・医療機関特定可能性についての検証
- 審査基準(案)の検証
- セキュリティについて

②セキュリティ監査の報告

①伏見研究班からの報告

患者、医療機関等の特定可能性について

- DPCデータを審査する際は、「患者」、「医療機関」、「保険者」、「医師個人」等にとって不利益な情報が公表されないよう、以下の視点から審査を行われたが、他に留意すべき個別事項、共通事項がある。

視点	患者情報	医療機関情報	保険者情報	医師個人情報
留意すべき個別事項	・ 郵便番号 ・ 生年月日 等	・ 毎年公開されている集計表との組み合わせ 等	・ 比較的小規模な保険者 等	・ 医師コード 等
留意すべき共通事項	・ 稀少な疾患 ・ 稀少な術式・処置 ・ 稀少な薬剤・医療材料の使用 等 ・ 報道等で先に患者の個人情報 が周知された症例			

患者、医療機関等の特定可能性について

- 患者個人の特定可能性について

実際には行っていないが、新聞等で医療機関や病状が詳細に報道された症例や、著名人で生年月日等個人情報公開されている人物について、現状のマスク化にもかかわらず複数条件の絞り込みでレコードを同定できうるケースは予想された。

今回の集計表請求では個人の特定は不能であった。ただし粒度の高い集計表を組み合わせることで特定が可能になる可能性は当然ある。

- 医療機関の特定可能性について

医療圏別に各医療機関を症例規模順で並べると、大規模な施設に限っては、病床規模やDPC評価分科会の退院患者調査との関連から一定程度照合・類推できる可能性があった。

今回の集計表請求では医療機関の特定は不能であった。ただし粒度の高い集計表を組み合わせることで特定が可能になる可能性は当然ある。

- 医師個人の特定可能性について

手術や手技等で地域に限られた医師のみが行っている状況が学会等で発表された場合、二次医療圏単位での分析では一定程度類推できる可能性があった。

提供データの安全性について

【① 全般的な事項】

①-1 [DPCデータ特有の基準]

- ・「DPC導入の影響評価に関する調査（以下「退院患者調査」という。）」の結果報告において毎年公表されている集計内容（MDC別・救急医療入院、診断群分類毎の集計、疾患別・手術別集計等）と重複する申請については、提供は行わない。もし必要とする場合は、その理由が明確に申出書に記載されているか確認したうえで、慎重な審査を行う。
- ・「DPCデータ以外の情報（申出者が保持する情報等）」と結びつけて集計を行う申出は、慎重な審査を行う。
- ・退院患者調査の公表情報以上に患者・医療機関等が特定される恐れがある申出は、慎重な審査を行う。
- ・提供対象医療機関はDPC対象病院までとし、DPC準備病院、出来高算定病院のDPCデータの提供は行わない。

①-2 [NDBデータと同様の基準]

- ・多数の項目を用いた探索的研究を目的とした申出は、慎重な審査を行う。
- ・「複数の研究」が1申出に盛り込まれている場合は、慎重な審査を行う。
- ・研究に際して抽出項目の指定や研究目的との関連については、慎重な審査を行う。
- ・「必要最小限」の範囲で必要な情報を提供することから、具体的な個々のデータ集計方法、解析方法まで分解した記述がない場合は、慎重な審査を行う。

項目の組合せにより症例を同定することは完全には防ぎ得ない。希少疾患や希少な手技などデータ項目の組合せで容易に小さなセルを発生できるケースは、地理情報を粗くする等の工夫を要する。

集計表に関する提供データの安全性について

【② 提供に関する審査基準】

- ・生年月日については、年月までとする。
- ・入院年月日・退院年月日については、在院日数・術前日数での提供も可とする。
- ・患者住所地域の郵便番号については、原則提供を行わないこととし、経年データ分析等を行うため必要な場合は、2次医療圏までの提供とする。
- ・施設コード、データ識別番号（患者情報）提供については、原則提供を行わないこととし、提供を行う場合は新たな通し番号を付番する。
- ・保険者番号の提供については、原則提供を行わないこととし、経年データ分析等を行うため必要な場合は、提供することとする。

－ 集計表において各データ項目（生年月日等）で粒度を落としても、多様な項目の組み合わせにより症例を絞り込んで同定できることを完全には防ぎえない。
⇒従って、研究の意義を損なうことなく、かつ個人や医療機関の特定ができないようにする普遍的な審査基準を策定することは困難」である。
⇒提供データにおける匿名性とは別に、抽出過程で取り扱っているデータのマスキング条件については、第三者の同席や第三者による操作ログのチェックなどの対策を前提に、ある程度緩和できるのではないか。

集計表に関する提供データの安全性について

【③ 公表に関する審査基準】

- ・生年月日については5歳刻みとし、85歳以上については同一の集計とする。
- ・患者の公表単位は、公表される研究の成果物において最も狭い地域区分の集計単位は2次医療圏とする。
- ・医療機関の所在地の公表単位は研究の成果物において最も狭い地域区分の集計単位を都道府県単位とする。
- ・保険者番号の公表については、4制度（被用者保険、国民健康保険、後期高齢者、共済組合）とする。
- ・**最小集計単位**については、研究の成果物において、1つのセルあたりの集計値が患者の数が10を下回る場合、医療機関の数が5を下回る場合は、その数値をマスクする。

- ・ **最少集計単位での匿名化を通じて、最少セルが一定基準を下回る場合にはマスクさせるほかに、周辺のセルと融合させる等の配慮も同等に有効であろう。**
 - ・ **保険者番号の公表については、4制度（被用者保険、国民健康保険、後期高齢者、共済組合）に限らず、研究に応じて独自のグループ分け(数グループ程度)を行っても安全性には影響を与えないのではないか。**
 - － **むしろ、小規模の保険者や公費医療などを単独でラベルしないルールがより実効的。**
 - ・ **すでに医療機関に関する情報は、厚労省の退院患者調査の結果報告で一定程度公表されているので、医療機関の所在地は特定されてもよいのではないか。**
 - ・ **公表時に匿名化を義務付ける現在の提供方法でよい。提供時に匿名化を行うとすれば解析に大きな制約が加わると懸念される。**
- 一定の症例数制限を設けても、背反的事象や割合の公開請求を行うことでデータは容易に算出されてしまう。10未満をマスクして提供を行ったほうがいいのか、全部提供した後に公表時に利用者がマスクをしたほうがいいのかという議論があるが、上述のことから、公開時に利用者がマスクすることが個人情報保護の観点から適切と思われる。

集計表に関するセキュリティの安全性について

【セキュリティの整備】

DPCデータの提供を行うにあたって、利用者に対し一定のセキュリティ水準を具備することを求めることとしている。

○利用にあたっての基本的な条件（案）

- ・提供したレセプト情報等の情報システム等への複写は、前段階でのデータが消去されない限り原則1回のみ。この原則は、厚生労働省から提供されたレセプト情報等の元データだけでなく当該元データから作成される全ての中間生成物も含め適用される。
- ・利用・保管場所は、あらかじめ申し出られた施設可能で入退室管理を行っているスペースのみとし、原則として持ち出されないこと。
- ・レセプト情報等を複写した情報システムはインターネット等の外部ネットワークには接続しないこと。
- ・レセプト情報等は事前に申し出られた利用者以外の者が利用してはならないため、これを担保するための情報システムの認証等の措置も必要。
- ・学部、研究室などの合理的な範囲内でガイドライン等のルールを定めた運用管理規程も必要。
- ・運用管理規程の運用状況を確認するための内部監査（自己点検）規程も必要。

- ①当施設で複数の研究者が出入りする環境でデータ利用を行った場合、NDBデータ専用のサーバー設置スペースの確保、生体認証による入退室記録に200万円（当施設での一例）、情報セキュリティマネジメントの運用自体に数人月を要し、大きな負担を要している。
- ②受け手がセキュリティを構築・維持するのは大変なので、DPCデータを用いた研究を普及するためには、サンプリングデータでの提供やオンサイトセンターを整備することが望ましいのではないかと。
- ③データ利用期間終了後も論文査読等でデータの再検証を求められる可能性があるため、利用期間終了時に加工データも併せて返納し、期間限定・あるいは理由を限定したうえで、後日オンサイトセンターにおいて参照できる環境が必要ではないかと。
利用者がセキュリティ基準を満たす体制を構築・維持することは労力、コストの観点から非常な負担である。日本全体の研究資源の効率的な活用という視点からも、DPCデータを用いた研究を普及するためにはオンサイトセンターを整備することが望ましい。

個票に関する提供データの安全性について

【② 提供に関する審査基準】

- ・ 生年月日については、年月までとする。
- ・ 入院年月日・退院年月日については、在院日数・術前日数での提供も可とする。
- ・ 患者住所地域の郵便番号については、原則提供を行わないこととし、経年データ分析等を行うため必要な場合は、2次医療圏までの提供とする。
- ・ 施設コード、データ識別番号（患者情報）提供については、原則提供を行わないこととし、提供を行う場合は新たな通し番号を付番する。
- ・ 保険者番号の提供については、原則提供を行わないこととし、経年データ分析等を行うため必要な場合は、提供することとする。

○ 日付情報に下記の加工を事前に加えることで、同定可能性を下げてはどうか。

- 生年月日は提供せず、代わりに入院時年齢・退院時年齢を提供する。
 - 入院年月[日]は原則的に提供せず、代わりにE/Fファイルを含むすべての日付情報を、入院日からの通算日数に変換して提供する。（年月情報は季節変動や経年データ分析等を行うなど特定の研究目的に限って提供する。）
 - 患者住所地郵便番号は提供せず、二次医療圏に変換して提供する。
 - ただし下記のいずれかの情報は併せて提供しても良いのではないかと。
 - 医療機関郵便番号 あるいは
 - 住所地市区町村が医療機関と同一であるか
 - 自治体等が行政上の必要から分析を行う場合には住所地郵便番号を提供しても良いのではないかと。
- - 個票において各データ項目（生年月日等）で粒度を落としても、多様な項目の組み合わせにより症例を絞り込んで同定できることを完全には防ぎえない。

⇒従って、「研究の意義を損なうことなく、かつ個人や医療機関の特定ができないようにする普遍的な審査基準を策定することは困難」である。

⇒提供データにおける匿名性とは別に、抽出過程で取り扱っているデータのマスク条件については、第三者の同席や第三者による操作ログのチェックなどの対策を前提に、ある程度緩和できるのではないかと。

個票に関する提供データの安全性について

【③ 公表に関する審査基準】

- ・生年月日については5歳刻みとし、85歳以上については同一の集計とする。
- ・患者の公表単位は、公表される研究の成果物において最も狭い地域区分の集計単位は2次医療圏とする。
- ・医療機関の所在地の公表単位は研究の成果物において最も狭い地域区分の集計単位を都道府県単位とする。
- ・保険者番号の公表については、4制度（被用者保険、国民健康保険、後期高齢者、共済組合）とする。
- ・**最小集計単位**については、研究の成果物において、1つのセルあたりの集計値が患者の数が10を下回る場合、医療機関の数が5を下回る場合は、その数値をマスクする。

- ・ **最少集計単位でのk匿名化を通じて、最少セルが一定基準を下回る場合にはマスクさせるほか、周辺のセルと融合させる等の配慮も同等に有効であろう。**
- ・ **保険者番号の公表については、4制度（被用者保険、国民健康保険、後期高齢者、共済組合）に限らず、研究に応じて独自のグループ分け(数グループ程度)を行っても安全性には影響を与えないのではないか。**
 - － **むしろ、小規模の保険者や公費医療などを単独でラベルしないルールがより実効的。**
- ・ **すでに医療機関に関する情報は、厚労省の退院患者調査の結果報告で一定程度公表されているので、医療機関の所在地は特定されてもよいのではないか。**
- ・ **公表時に匿名化を義務付ける現在の提供方法でよい。提供時に匿名化を行うとすれば解析に大きな制約が加わると懸念される。**

個票に関するセキュリティの安全性について

【セキュリティの整備】

DPCデータの提供を行うにあたって、利用者に対し一定のセキュリティ水準を具備することを求めることとしている。

○利用にあたっての基本的な条件（案）

- ・提供したレセプト情報等の情報システム等への複製は、前段階でのデータが消去されない限り原則1回のみ。この原則は、厚生労働省から提供されたレセプト情報等の元データだけでなく当該元データから作成される全ての中間生成物も含め適用される。
- ・利用・保管場所は、あらかじめ申し出られた施設可能で入退室管理を行っているスペースのみとし、原則として持ち出されないこと。
- ・レセプト情報等を複製した情報システムはインターネット等の外部ネットワークには接続しないこと。
- ・レセプト情報等は事前に申し出られた利用者以外の者が利用してはならないため、これを担保するための情報システムの認証等の措置も必要。
- ・学部、研究室などの合理的な範囲内でガイドライン等のルールを定めた運用管理規程も必要。
- ・運用管理規程の運用状況を確認するための内部監査（自己点検）規程も必要。

- ①入退室管理を行っているスペースにおいて、関係者以外の入室を制限したり、目隠しを行うなど、窃視防止の対策が必要ではないか。
- ②当施設で複数の研究者が出入りする環境でデータ利用を行った場合、NDBデータ専用のサーバー設置スペースの確保、生体認証による入退室記録に200万円（当施設での一例）、情報セキュリティマネジメントの運用自体に数人月を要し、大きな負担を要している。
- ③受け手がセキュリティを構築・維持するのは大変なので、DPCデータを用いた研究を普及するためには、サンプリングデータでの提供やオンサイトセンターを整備することが望ましいのではないか。
- ④データ利用期間終了後も論文査読等でデータの再検証を求められる可能性があるため、利用期間終了時に加工データも併せて返納し、期間限定・あるいは理由を限定したうえで、後日オンサイトセンターにおいて参照できる環境が必要ではないか。

オンサイトセンターに期待される機能

- 生に近いデータを閲覧できる環境
- データサイエンティスト(専任)により、データ特性を踏まえた解析計画のカウンセリング
- 終了したプロジェクトについて、中間テーブル等のファイルを保管し、申請に応じて参照できる環境

※全国で複数拠点の運用が望ましい。

※人員配置は利用実績に応じて定期的に見直す必要がある。

オンサイトセンターで配慮すべき セキュリティ上の課題

- 不適切なデータ参照を防ぐため、SQL文のログを残し、担当者が事後チェックできる環境づくり
- 研究者がデータを持ち帰る際に、個人特定可能性が十分に低いことを適切に判断できる体制
- 生に近いデータを安全に保管・運用するための堅牢な情報セキュリティマネジメント
- 研究利用を許可されたものを正確に認証する仕組み

②セキュリティ監査の報告

【セキュリティ監査の目的について】

ガイドライン策定にあたり、想定しているセキュリティ項目の他に盛り込む項目があるのか、模擬申出書に記載されたセキュリティが実際に確保されているか、セキュリティ監査を実施する方法等を検証するため、厚生労働省科学研究班（伏見班）のご協力を得て、実際にDPCデータの利用場所、保管場所に職員等を派遣し、模擬的にセキュリティ監査を実施した。

【検査の概要】

①日時：平成25年11月15日（金）13時30分～18時10分

②場所：京都大学大学院 医療経済学分野

③監査実施者：厚生労働省保険局医療課職員2名及び総務課職員1名（※）

※レセプト情報等の提供で既に監査を行っていることから、総務課職員をアドバイザーとして助言を得た。

④検査内容

○書類に記載されたセキュリティ要件の確認

○現場での運用状況の確認

セキュリティ監査の概要

【安全管理がなされていた点】

- 厚生労働省から提供された元データはセキュリティ基準が高いサーバ室に保管されており、生体認証による入退室管理がなされていた。研究室へ入室する場合もID認証管理を行っており、登録された職員等の帯同がなければ入室することはできない
- DPCデータの利用は研究室内のサーバ室内のみであった
- DPCデータの分析等の業務は保管サーバから作業用PCへ教室内LANを介さず専用ネットワークを用いて情報を受け渡しを行っている。また、格納サーバに格納したデータは保管サーバから呼び出すことができるが、情報が書き換えられないよう情報の流れを一方向となるよう工夫がされていた
- ウィルス対策ソフトを更新する際はDPCデータを外付けHDDに退避させ教室内LANに接続するなど、インターネットの接続によるウィルス感染を防止する対策をとっていた

【対応が必要と思われる点】

- サーバ室への入退は生体認証により利用者が制限されているが、監視する位置に防犯カメラ、監視カメラは設置されていなかった
- サーバ室への入退において、携帯電話等の情報機器の持ち込みに伴う情報漏洩への対策については明確に定められていなかった。

データ提供を行うための留意すべき点

【セキュリティに関すること】

- 生体認証による入退室管理などセキュリティが高いサーバ室等が必要がある
- サーバ室を設置したとしても維持・運用にコストや人的資源がかかり、研究者個人が負担することは難しい

【提供に関すること】

- DPCデータはDPC評価分科会において「DPC導入の影響評価に関する調査」として一定程度のデータを公表しているため、研究者は提供申出をする場合は確認する必要があると考える
- 研究者は研究内容、提供項目等を明確にする必要があるが、どのような研究に有効なデータかを知る必要があると考える
- 個票を提供する場合、匿名化をしたうえで提供を行うが複数の条件で絞り込むことで、ニュースや新聞など公表されている情報から個人が特定される恐れがあると考える
- 匿名化をしたうえで提供するが10未満のセルが多数存在した場合、本来行いたかった研究ができなくなる恐れがあるためある程度制限を設けずに提供することも検討すべきだが、その場合セキュリティ審査を厳しくする必要があると考える

4. 今後の対応（案）

現状の整理

- ① 個票データの提供において、研究の意義を阻害せずに安全性が確保できるような普遍的な「審査基準」を作るのは難しいと考えられる

[(NDBデータと比較した)DPCデータ特有の問題点]

- ・ すでに公表されている医療機関別のDPC集計データが存在し、組み合わせが可能
 - ・ DPCデータを提出する病院は少ない(約1,800病院)
 - ・ 様式1(簡易診療録情報)等において診療情報等の個人特定性が高い濃密な情報が含まれる
- ② DPCデータ提供のための厚生労働省側の体制の整備がまだなされていない
 - ・ 第三者提供向けのデータベースの構築（予算要求中）
 - ・ セキュリティ監査の体制の整備 等
 - ③ 個票データの取り扱いは研究者にとっても大きな負担となりうる
⇒ 申請可能な者が限定され、DPCデータを用いた研究の進展において障壁となる可能性がある

模擬提供を踏まえた今後の方針(案)

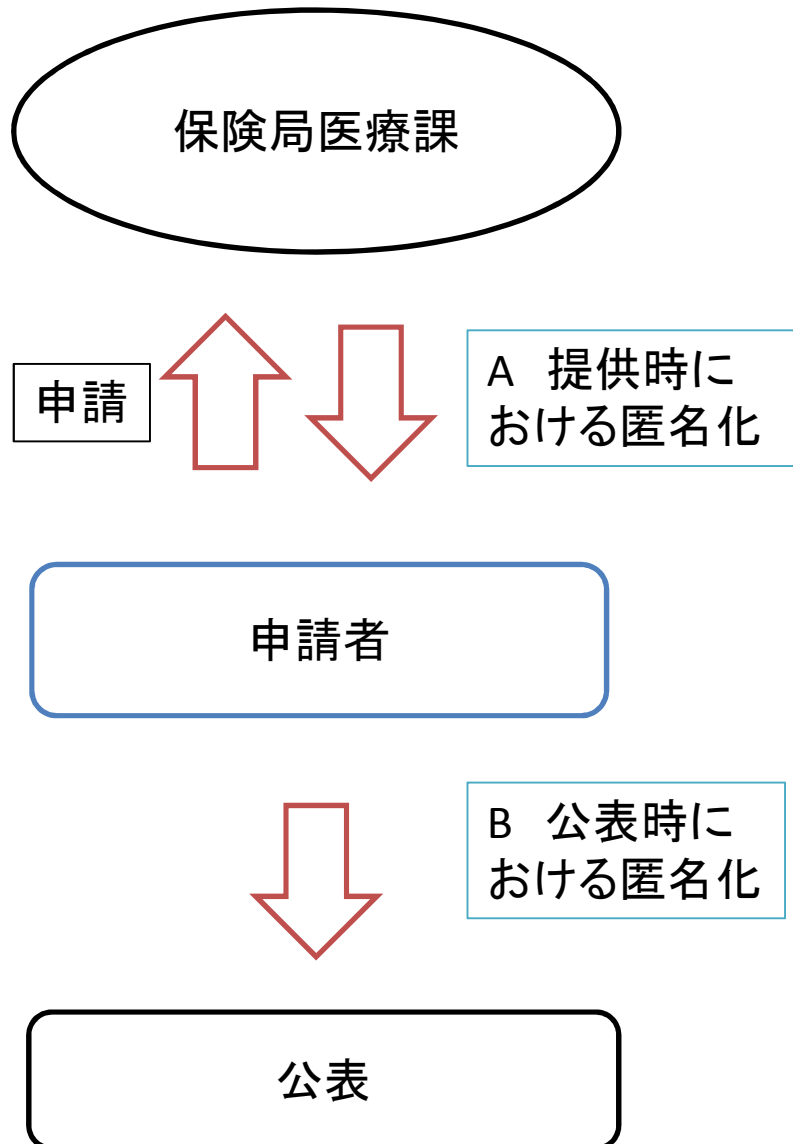
今後の方針(案)

1. 平成26年度より、本格的な運用に向けた準備期間として、比較的安全性が高いと考えられる「集計表」の提供をまず開始することとしてはどうか。
2. 「個票データ」の提供方法については、「オンサイトセンター」における提供、「サンプリングデータ」による提供等を中心に、今後も引き続き検討を進めていくこととしてはどうか。

考え方

- 「個票データ」の提供については、実用的な審査基準の構築において困難な点があり、まずは集計表に関する審査・提供の実績を積み重ねた上で、検討していくのが妥当ではないか。
- 厚生労働省の提供体制が十分に整っていないこともあり、個票データの提供を多数実施することが難しい状況である。
 - ・ 個票データの切り出し作業は時間がかかる
 - ・ セキュリティ監査が必要となる
- 「個票データ」をそのまま第3者に提供するためには、申請者が高額なセキュリティ設備を有することが必要となるため、申請可能な者を限定してしまうことにつながるため、DPCデータを用いた研究を発展させるにあたり、必ずしも優れた方法であるとはいえない。
(検討すべき提供方法 (案))
 - オンサイトセンター (個票データを自由に取り扱えるセキュリティの高い環境の整備)
 - サンプリングデータセット (利用者の教育を含めた様々なデータセット)

集計表の提供の方向性



○平成26年度上半期を目途に、ガイドライン等を明確化した上で、募集を開始。

○データ提供のための事務局の提供体制や申請の個数・申請内容等の状況を踏まえつつ、下記の観点や「NDBデータの提供の際の優先順位の考え方」等に基づき、総合的に判断した上で、優先度が高いものから提供していくこととしてはどうか。

(例)

- ・公共性が高いもの
- ・探索的でないもの

審査基準の明確化①～優先順位について(案)

平成24年3月7日
第9回レセプト情報等の
提供に関する有識者会議

資料2

○今後、データ提供の申出が増加していくことも考えられるが、申出が非常に多くなり、公益性やセキュリティ要件の面で審査を了する申出が増えた場合、全ての申出についてデータ抽出を行うことは困難となることも考えられる。

(参考) 第1回申出における医政局指導課のデータ抽出では、全国の半年間における医科・DPCレセプト4億7,000万件の抽出作業に約200時間を要した(営業日で2週間程度)。

○こうしたことから審査の基準とは別に、各申出にデータ提供の優先順位をつけ、順位の高いものから順番に対応できるものまでデータ提供を行うこととしてはどうか。評価の方法については、有識者会議の委員に協力を仰いで行う部分と事務局において行う部分とを設け、事前に各項目について点数化した上で有識者会議の審査に諮ることとしてはどうか(イメージは次ページ)。

【評価項目】

以下のうち、①と②については、有識者会議の一部の委員があらかじめ評価を行い、それ以外については事務局において、評価を行った上で有識者会議に諮ることとしてはどうか。

①学術的な期待度

・申出された研究により、学術的に有意義な結果が得られる期待度が高いかどうか。

②研究内容の簡潔さ・明解さ(複雑・難解なものとなっていないか)

・複雑な研究内容や仮定を置いている研究であるかどうかについて判断することとして、理解が得やすいか否か。

③具体的な政策への反映を想定しているか(単なる基礎資料か否か)

・具体的な政策への反映を想定しているものかどうか。

(例) 医療計画の策定の基礎資料とするため都道府県への提供を想定 など。

④地域の範囲(全国か、地域限定の研究か)

・全国規模のデータベースという性質を活かす観点から、地域限定をした研究よりは、全国規模の研究を行うものを優先してはどうか。

⑤活用するデータ量・規模(大量のデータを使用するものか否か、抽出に要する見込み時間で判断)

・上記を第1回申出における処理時間約200時間を一応の標準的な処理時間として、データ抽出に要する時間についての評価を行うこととしてはどうか。

(参考)優先順位付けのイメージ

平成24年3月7日	資料2
第9回レセプト情報等の提供に関する有識者会議	

【点数付けの考え方】

- 事務的な作業が膨大であったとしても極めて学術的に有益な研究であれば、データ提供の優先順位上考慮する必要があることから、有識者委員の学術面での評価点数と事務局による実務面での評価点の最高点を同一とする。
- 両者の点数を勘案した総合評価の下にデータの提供の可否を決定。

	項目	評価
通常審査部分	研究内容の公益性	
	セキュリティ要件	
	
有識者委員が評価	①学術的な期待度	<input type="checkbox"/> 低い(1) <input type="checkbox"/> やや低い(2) <input type="checkbox"/> 普通(3) <input type="checkbox"/> ある程度高い(4) <input type="checkbox"/> 高い(5)
	②研究内容の簡潔さ・明解さ	<input type="checkbox"/> 難解(1) <input type="checkbox"/> やや難解(2) <input type="checkbox"/> 普通(3) <input type="checkbox"/> ある程度明解(4) <input type="checkbox"/> 明解(5)
優先順位評価部分	有識者委員評価点合計	(2~10)
	③具体的な政策への反映	<input type="checkbox"/> 想定していない(0) <input type="checkbox"/> 想定している(3)
	④研究地域の範囲	<input type="checkbox"/> 地域限定(0) <input type="checkbox"/> 全国(2)
	⑤活用するデータ量 ※200時間を一応の標準処理時間とする。概ね300時間超がデータ量が多い。100時間未満が少ない。	<input type="checkbox"/> データ量が少ない(~100時間程度)(5) <input type="checkbox"/> 普通(100~300時間)抽出作業(3) <input type="checkbox"/> データ量が多い(300時間~)(0)
	事務局評価部分	(0~10)

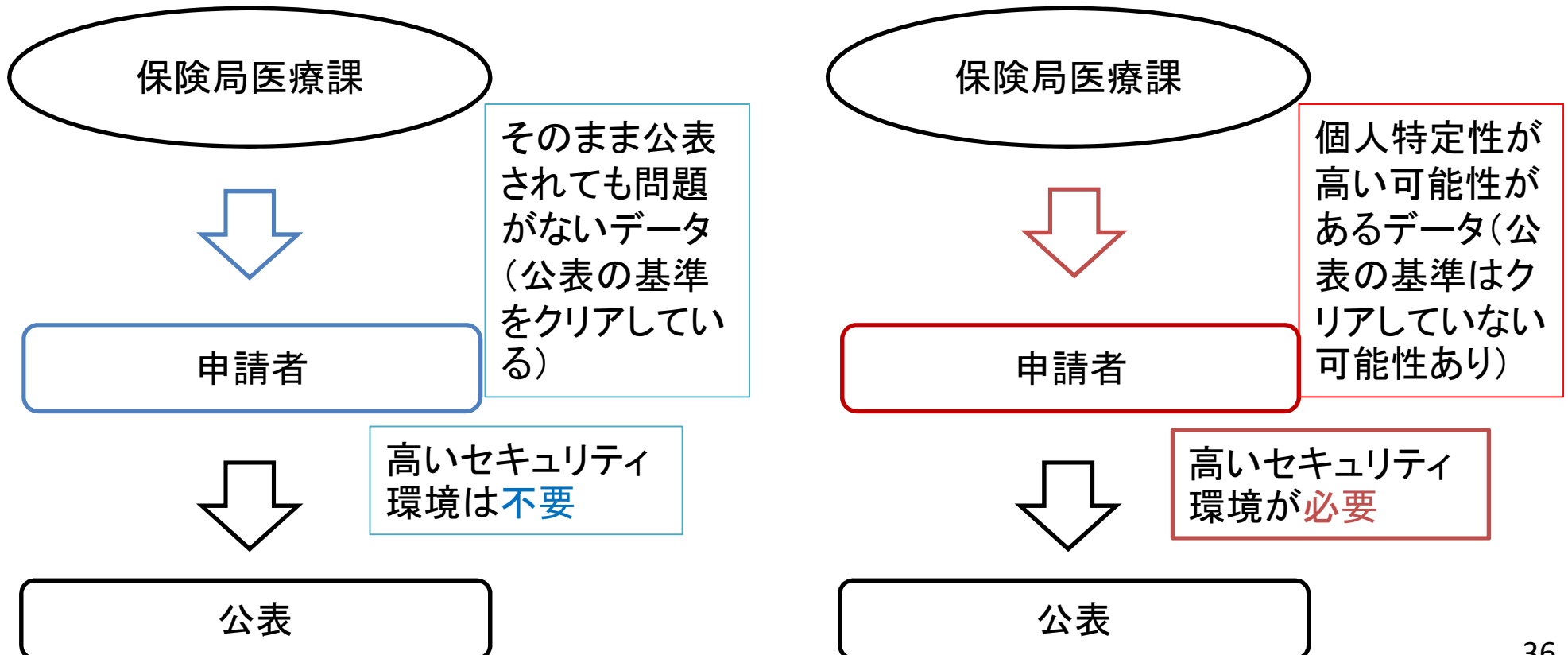
総合評価	可・否
------	-----

集計表データの提供にかかる論点①

【セキュリティの考え方】

「集計表」による提供の場合であっても、提供の際に最小集計単位によってマスクされない場合、個人特定性が高いデータが提供される恐れがあり、その場合は申請者側に高いセキュリティが要求されるのではないか。

平成26年度の準備期間における集計表の提供については、「公表の基準」を満たしたものを提供することについてどのように考えるか。



集計表データの提供にかかる論点②

○医療機関別情報の公表にかかる問題

- ・退院患者調査の結果報告において、一定程度の個別医療機関の情報が公開されていることから、医療機関は特定されても問題がないのではないかという意見がある。
- ・ただし、退院患者調査における公表の範囲を大きく超える場合は、医療機関にとって不利益となる情報が公開されてしまうという問題が発生する可能性がある。
- ・DPCデータ模擬審査基準においては医療機関に関する最小集計単位は「5」であったが、NDBデータの提供においては、「3」となっている。

- ・DPCデータの公表の基準の医療機関に関する最小集計単位は「3」とすることとしてはどうか。

集計表データの提供にかかる論点③

【外部データ(マスタ)を用いた集計】

申請者独自で準備する外部データ(マスタ)と突合して変換して集計する申請について、どのように考えるか。

申請者独自で準備する外部データ(マスタ)と突合して変換をする集計を行うことで、研究意義の高い結果となる可能性がある。

ただし、外部データと突合して集計することによって、個人特定性や医療機関特定性等が高くなる可能性がある。

ただし、元のDPCデータをN対1で集約する場合等、外部データ(マスタ)の性質によっては、安全性において問題がないと考えられる事例もあると考えられる。

集計表データの提供にかかる論点④

「内訳」に係るデータの取り扱い

患者数	うち男性患者数
63	41
2,871	1,777
-	-
98	56
76	48
-	-
3,257	2,035
-	-
76	50
66	47
3,271	2,008
13	-
59	37
60	35
-	-
485	317
778	444
2,844	1,668
3,950	2,310
-	-
5,622	3,292
17,568	10,507

「内訳」に係るデータは、原則としてすべての要素を申請することについてどのように考えるか。
 (例:この場合、男性患者数を申請する場合は、女性患者数も申請する)

「割合」に係るデータの取り扱い

50歳未満患者割合
-
1.8%
-
-
-
1.5%
-
-
1.8%
-
-
-
2.3%
3.0%
2.0%
2.8%
-
2.2%
3.6%

「割合」に係るデータが必要である場合は、原則として「分母」の値と「分子」の値を申請することについてどのように考えるか。

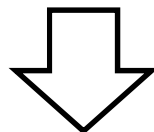
集計表データの提供にかかる論点⑤

○行政利用の場合の審査方法について

NDBデータの提供に関するガイドラインにおいては、下記の通り定められている。

- 厚生労働省の各部局が利用する場合、「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」(以下「ガイドライン」という)では以下においては有識者会議の審査を省略することができる定められている。

- (1)厚生労働省の各部局が、その所掌する事務の範囲内で政府が開催する審議会等に提出する資料の作成のために利用する場合
- (2)厚生労働省の各部局が、その所掌する事務の範囲内で作成し、公表する統計資料の作成のために利用する場合
(「第6 提供依頼申出に対する審査 7 有識者会議の審査を省略することができる利用」より)



・厚生労働省の各部局が今後策定する予定であるガイドラインに規定された目的でDPCデータを利用する場合は、有識者会議の審査を省略することとしてはどうか。

・ただし、利用に際しては事務局で現行の有識者会議における審査と同等の内容の審査を実施するとともに、有識者会議に申出内容を報告することとし、必要に応じて利用の成果についても報告を行うこととしてはどうか。

3 厚生労働省から提出される申出の審査について

平成25年9月20日
第16回レセプト情報等の
提供に関する有識者会議

これまでの対応

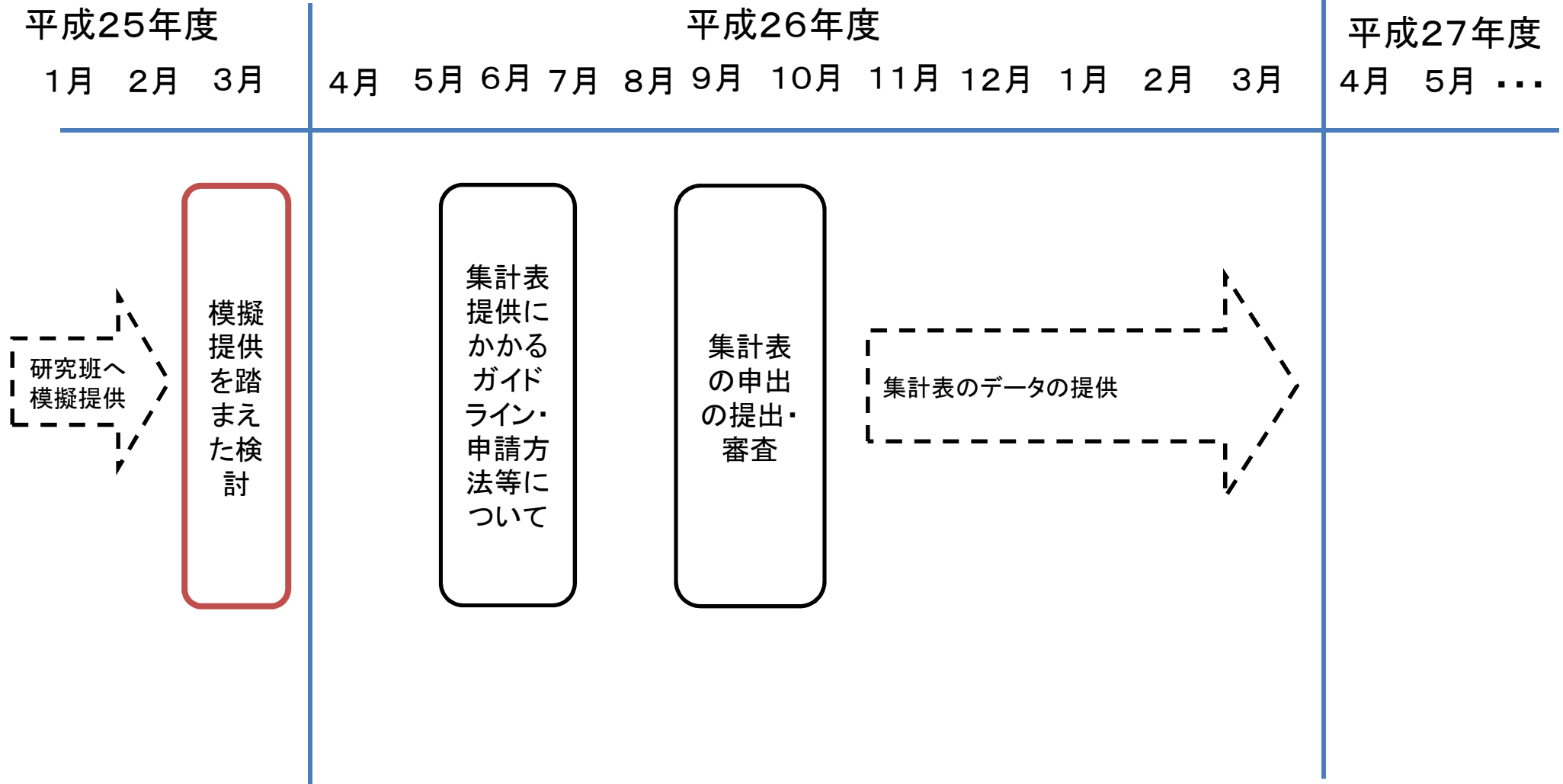
- 厚生労働省の各部局が利用する場合、「レセプト情報・特定健診等情報の提供に関するガイドライン」(以下「ガイドライン」という)では以下においては有識者会議の審査を省略することができる定められている。
 - (1)厚生労働省の各部局が、その所掌する事務の範囲内で政府が開催する審議会等に提出する資料の作成のために利用する場合
 - (2)厚生労働省の各部局が、その所掌する事務の範囲内で作成し、公表する統計資料の作成のために利用する場合
(「第6 提供依頼申出に対する審査 7 有識者会議の審査を省略することができる利用」より)
- 試行期間においては、厚生労働省の各部局からどのような申出がなされるか明確でなかったこと等に鑑み、一般の申出者と同様に有識者会議の審査を経たうえでデータ提供の承諾／不承諾が決定されていた。

現在の運用にみられる実績および課題

- 実績:これまでに厚生労働省の各部局から寄せられ、承諾された申出は、以下のとおりである。
 - ・「医療計画の見直しにあたっての適正な受療医療圏などの検討について」(医政局指導課)
 - ・「難病指定研究及び小児慢性特定疾患指定研究」(健康局疾病対策課・雇用均等・児童家庭局母子保健課)
 - ・「ナショナルデータベースを用いたがん治療の費用対効果評価」(保険局医療課)
- 課題:ガイドラインに規定されているデータの利用申出手続きが、主として研究機関に所属する研究者を申出者と想定しているため、厚生労働省や国の行政機関・都道府県にとっては遂行困難な手続きがある。
 - 例:データの利用は「疫学研究に関する倫理指針」の適用対象とされているが、省内各部局の申出は研究ではなく事業の一環であり、研究として位置づけにくい。また、省内に倫理委員会が設置されておらず、倫理委員会の評価を受けられない。
 - 例:利用者の変更がある場合はその都度有識者会議に申出し、審査を受けることとなっているが、省庁では利用者の変更が人事異動という形式で突然かつ頻繁に発生するため、持ち回り審査の時期を待つ時間的余裕がない。

- 厚生労働省の各部局がガイドラインに規定された目的で利用する場合は、有識者会議の審査を省略することとしてはどうか。
- ただし、利用に際しては事務局で現行の有識者会議における審査と同等の内容の審査を実施するとともに、有識者会議に申出内容を報告することとし、必要に応じて利用の成果についても報告を行うこととしてはどうか。
- 厚生労働省以外の国の行政機関及び都道府県からの申出については、厚生労働省の各部局と同様の取扱いとすべきかどうか。

今後のスケジュール(案)



参考 審査基準(案)

2. 審査基準（案）

- 前回の有識者会議（平成25年6月7日）において、「今後、模擬審査を進める上で、たたき台でもよいので何らかのガイドラインがあった方がよい」という指摘があったことから、事務局で審査基準（案）を作成した。

⇒ 今回の模擬審査では、当該審査基準（案）に照らし合わせ、審査を行うこととしてはどうか。

（※ 審査基準のたたき台の構成）

- ・ 基本的な考え方
- ・ 審査方針
 - ① 全般的な事項
 - ② 提供にかかる基準
 - ③ 公表にかかる基準

(参考) DPCデータ提供にかかる審査方針の基本的な考え方

平成25年9月20日
第1回有識者会議分科会

データ提供にあたり審査が必要となる背景

- DPCデータ提供にあたり、以下の者にとって不利益とならないよう配慮する必要がある。
 - ①患者個人
 - ②医療機関
 - ③その他（保険者、医師個人等）



審査の観点

- 上記のような者にとって不利益とならないよう、データ提供にあたり以下の観点から審査を行う必要がある。
 1. 申出者の適格性
 2. 申出目的の妥当性
 3. 研究デザインの妥当性
 4. 申出者のセキュリティ環境の妥当性
提供されたデータが外部に漏洩しないようセキュリティの高い環境となっている必要がある
 5. 提供データの安全性
セキュリティの問題等により、仮に不特定多数に広まってしまった場合であっても、国民に与える不利益が最小限である必要がある

提供データの安全性の基準について、模擬申出を通じた
検証を行う

DPCデータを審査するうえでの基本的な考え方（案）

平成25年9月20日
第1回有識者会議分科会

審査の視点について

- DPCデータを審査する際は、「患者」、「医療機関」、「保険者」、「医師個人」等にとって不利益な情報が公表されないよう、以下の視点から審査を行うことが必要であると考えられる。

視点	患者情報	医療機関情報	保険者情報	医師個人情報
留意すべき個別事項	・郵便番号 ・生年月日 等	・毎年公開されている集計表との組み合わせ 等	・比較的小規模な保険者 等	・医師コード 等
留意すべき共通事項	・稀少な疾患 ・稀少な術式・処置 ・稀少な薬剤・医療材料の使用 等			

特定性について留意すべき点

- 「患者」「医療機関」「保険者」「医師個人」等の特定は、それぞれがその他の者の特定につながる可能性があることに留意する必要がある。
（例：医療機関が特定されると、その医療機関に係る公知な情報によって、その医療機関で働く医師個人が特定される）

DPCデータ提供・公表にかかる審査方針（案）

平成25年9月20日
第1回有識者会議分科会

【① 全般的な事項】

[DPCデータ特有の基準]

- ・「DPC導入の影響評価に関する調査（以下「退院患者調査」という。）」の結果報告において毎年公表されている集計内容（MDC別・救急医療入院、診断群分類毎の集計、疾患別・手術別集計等）と重複する申請については、提供は行わない。もし必要とする場合は、その理由が明確に申出書に記載されているか確認したうえで、慎重な審査を行う。
- ・「DPCデータ以外の情報（申出者が保持する情報等）」と結びつけて集計を行う申出は、慎重な審査を行う。
- ・退院患者調査の公表情報以上に患者・医療機関等が特定される恐れがある申出は、慎重な審査を行う。
- ・提供対象医療機関はDPC対象病院までとし、DPC準備病院、出来高算定病院のDPCデータの提供は行わない。

[NDBデータと同様の基準]

- ・多数の項目を用いた探索的研究を目的とした申出は、慎重な審査を行う。
- ・「複数の研究」が1申出に盛り込まれている場合は、慎重な審査を行う。
- ・研究に際して抽出項目の指定や研究目的との関連については、慎重な審査を行う。
- ・「必要最小限」の範囲で必要な情報を提供することから、具体的な個々のデータ集計方法、解析方法まで分解した記述がない場合は、慎重な審査を行う。

DPCデータ提供・公表にかかる審査方針（案）

平成25年9月20日
第1回有識者会議分科会

【② 提供に関する審査】

- ・ 生年月日については、年月までとする。
- ・ 入院年月日・退院年月日については、在院日数・術前日数での提供も可とする。
- ・ 患者住所地域の郵便番号については、原則提供を行わないこととし、経年データ分析等を行うため必要な場合は、上3桁までの提供とする。
- ・ 施設コード、データ識別番号（患者情報）提供については、原則提供を行わないこととし、提供を行う場合は新たな通し番号を付番する。
- ・ 保険者番号の提供については、原則提供を行わないこととし、経年データ分析等を行うため必要な場合は、提供することとする。

【③ 公表に関する審査】

- ・ 生年月日については5歳刻みとし、85歳以上については同一の集計とする。
- ・ 入院年月日、退院年月日等の「年月日」については、「年月」までとする。
- ・ 患者の公表単位は、公表される研究の成果物において最も狭い地域区分の集計単位は2次医療圏とする。
- ・ 医療機関の所在地の公表単位は研究の成果物において最も狭い地域区分の集計単位を都道府県単位とする。
- ・ 保険者番号の公表については、4制度（被用者保険、国民健康保険、後期高齢者、共済組合）とする。
- ・ **最小集計単位**については、研究の成果物において、1つのセルあたりの集計値が患者の数が10を下回る場合、医療機関の数が5を下回る場合は、その数値をマスクする。

D P C データ提供・公表にかかる審査方針（案）

平成25年9月20日
第1回有識者会議分科会

○テキスト入力される項目について

D P C データには主傷病名などコードではなく、テキスト入力している項目がある。これらの項目については、「〇〇との合併症」、「〇〇に転移」、「山田 太郎」、「生年月日」など個人につながる情報が入力されている恐れがあることから、提供にあたっては慎重な審査を必要とする。

○様式 1

- ・主傷病名
- ・入院の契機となった傷病名
- ・医療資源を最も投入した傷病名、医療資源を2番目に投入した傷病名
- ・入院時併存症名1・2・3・4、入院後発症疾患名1・2・3・4
- ・点数表コード
- ・手術日
- ・自由記載欄

○Dファイル

- ・診療行為名称
- ・診療行為名称
- ・医療機関係数

○E F 統合ファイル

- ・診療明細名称

○外来E F 統合ファイル

- ・診療明細名称

○様式 4

- ・医療保険外との組合せ 等

過去の参考資料

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

DPCデータ

研究班の報告書を元に
事務局が作成

平成25年6月7日
第15回有識者会議

現状

I 集計表データを を提供する場合

- ・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。

現状

II 個票データを を提供する場合

- ・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。

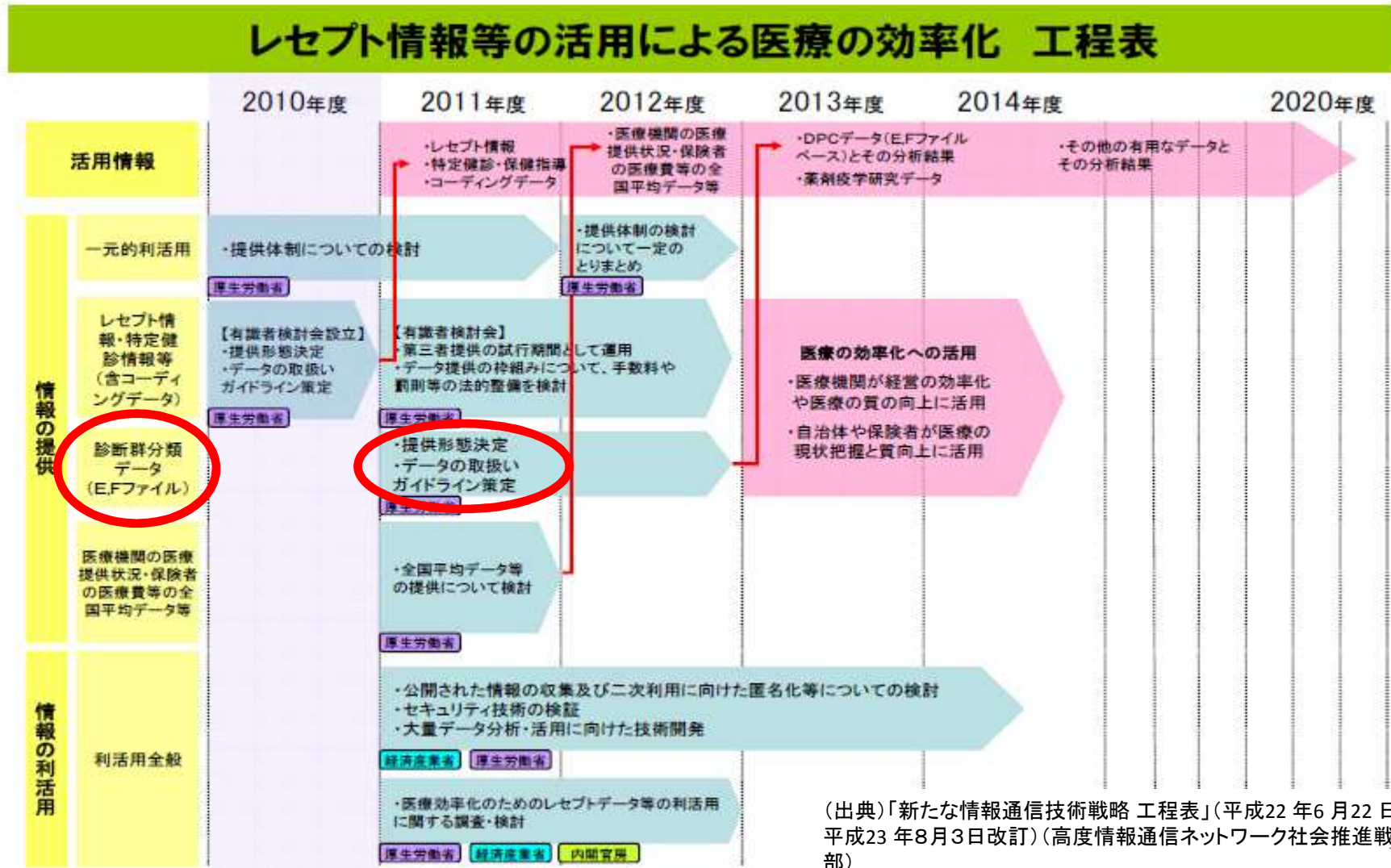
研究班案

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データの提供について

平成25年6月7日
第15回有識者会議

○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)

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。

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

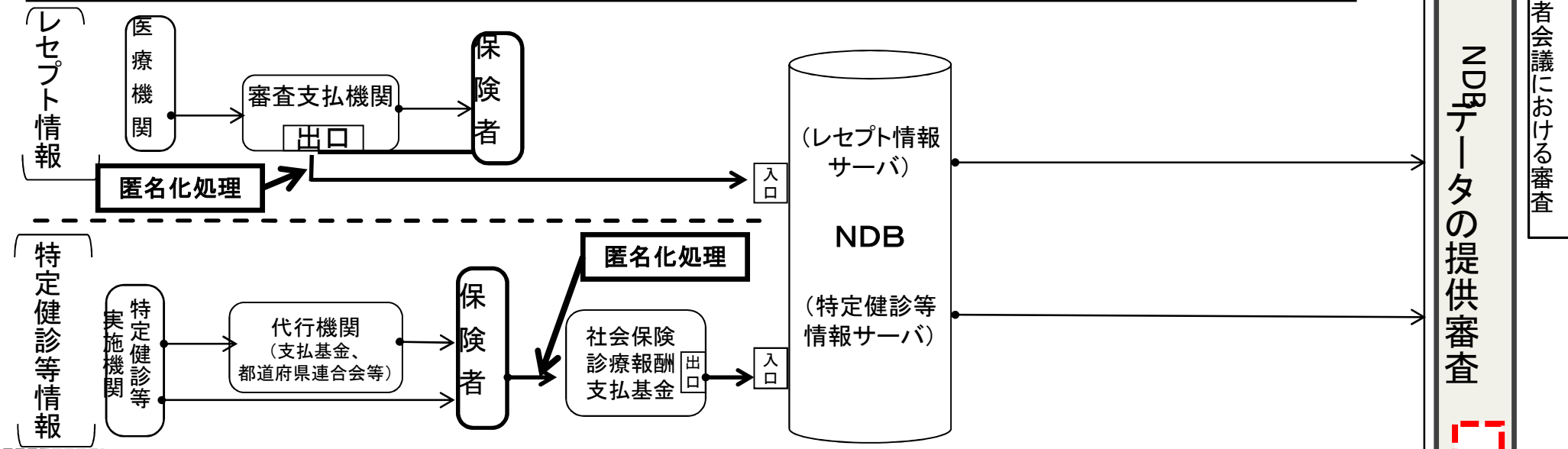
③ DPCデータベースの構築等について

- 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

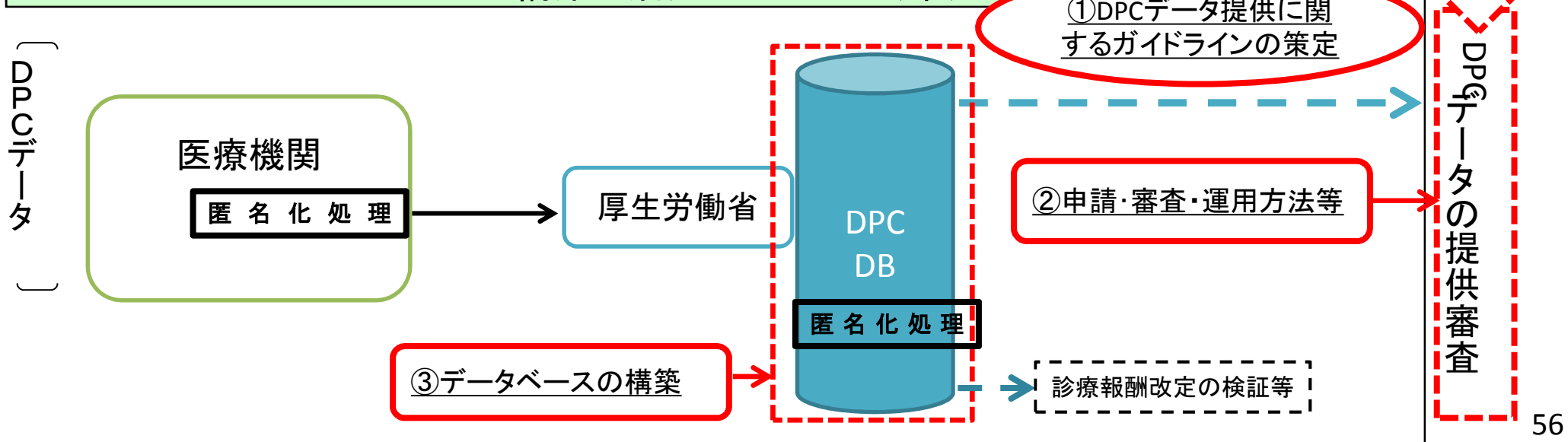
DPCデータベースの構築と活用のイメージ(案)

平成25年6月7日
第15回有識者会議

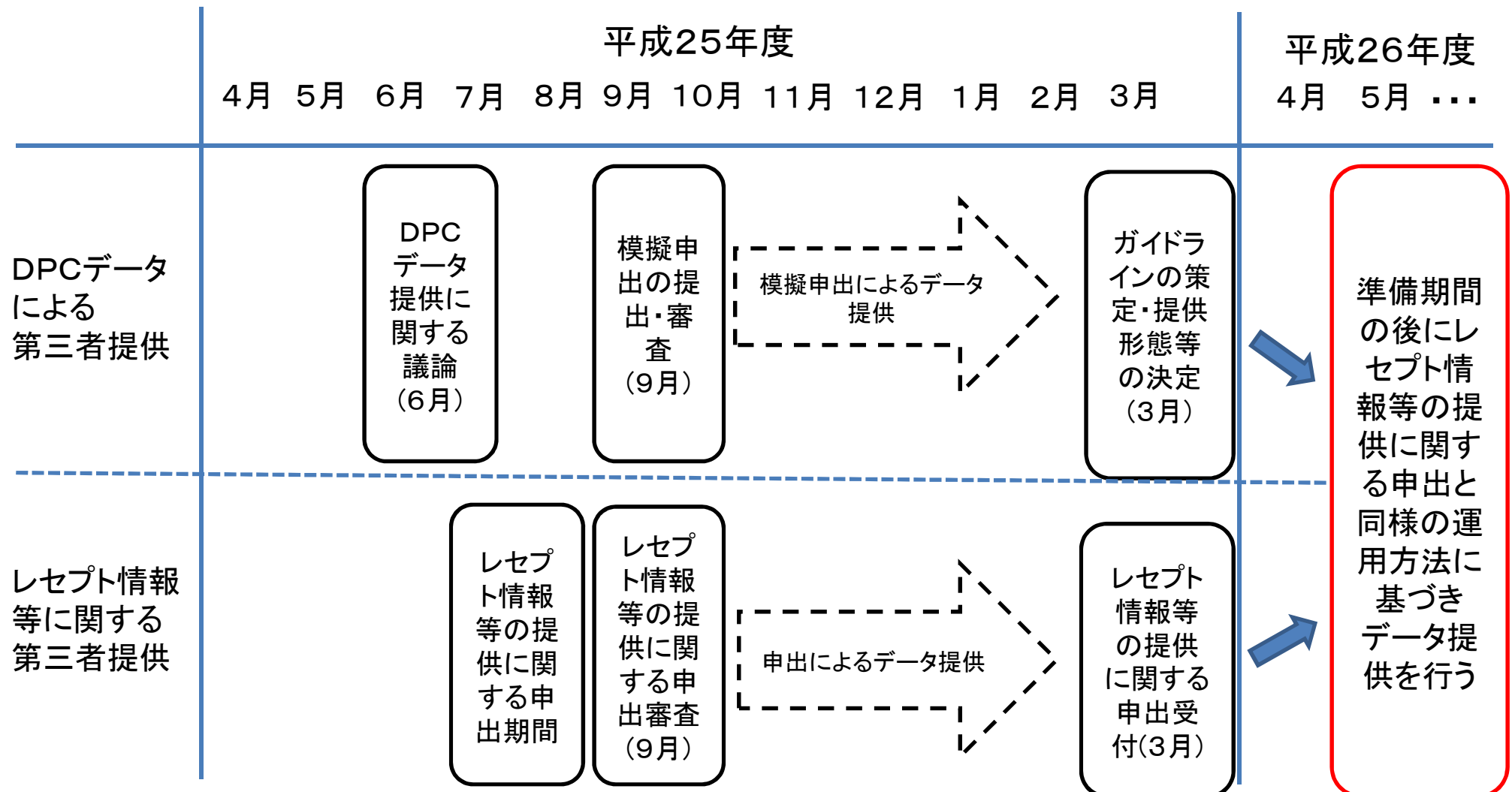
レセプト情報・特定健診等情報の活用方法



DPCデータベースの構築と活用のイメージ(案)



今後のスケジュール(案)



退院患者調査の目的

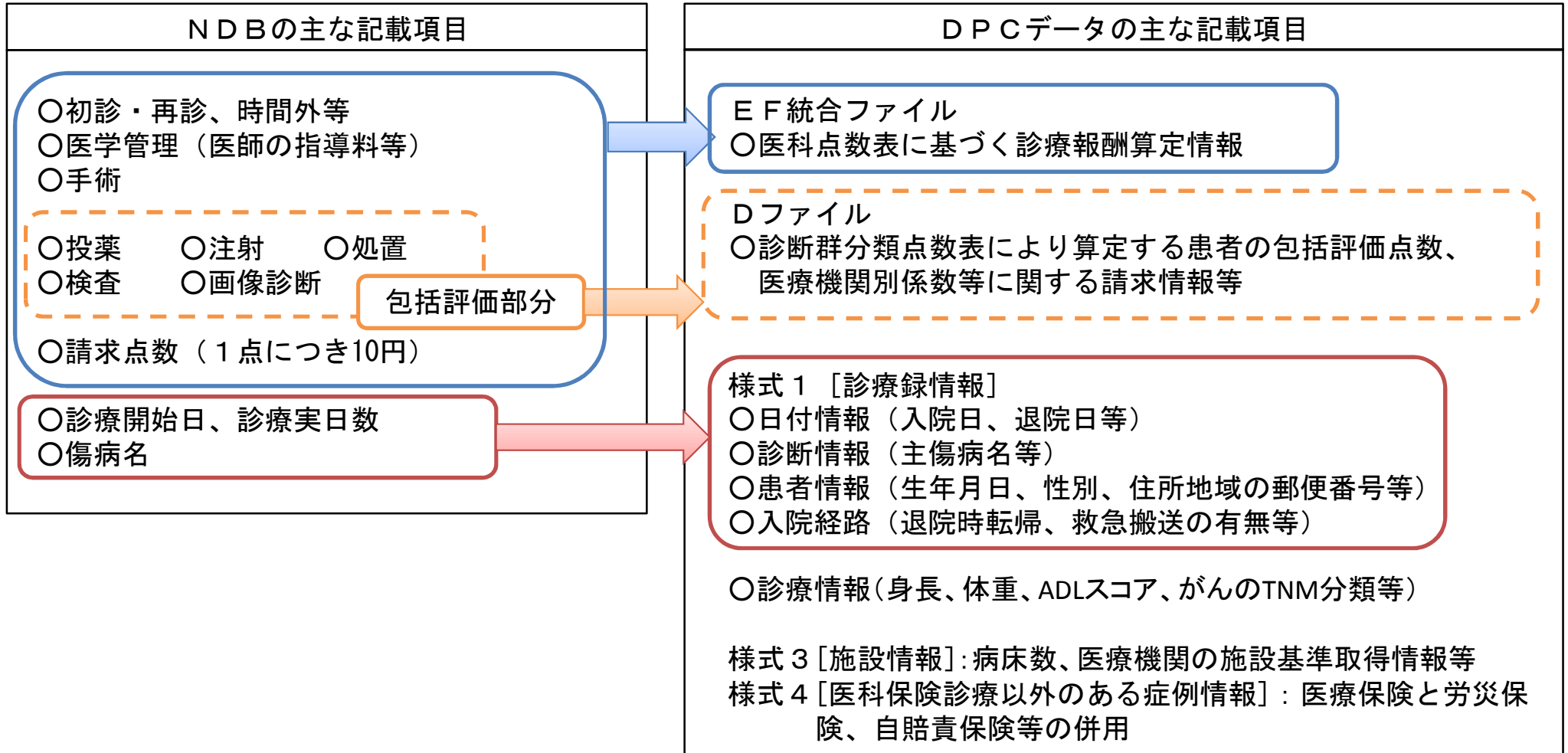
平成24年2月10日
第8回有識者会議

- 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。

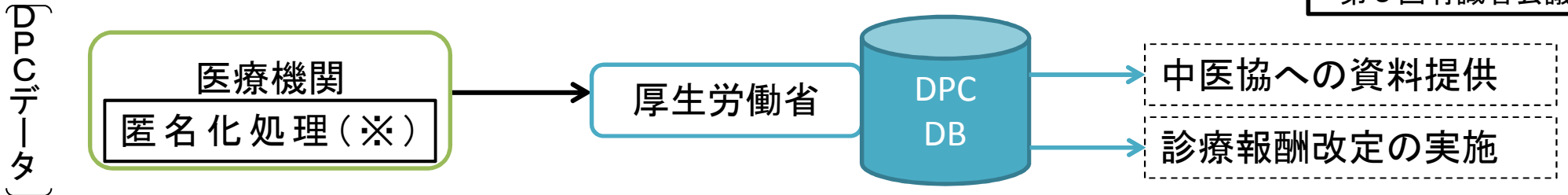
レセプトデータとDPCデータの主な相違点について（1）

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

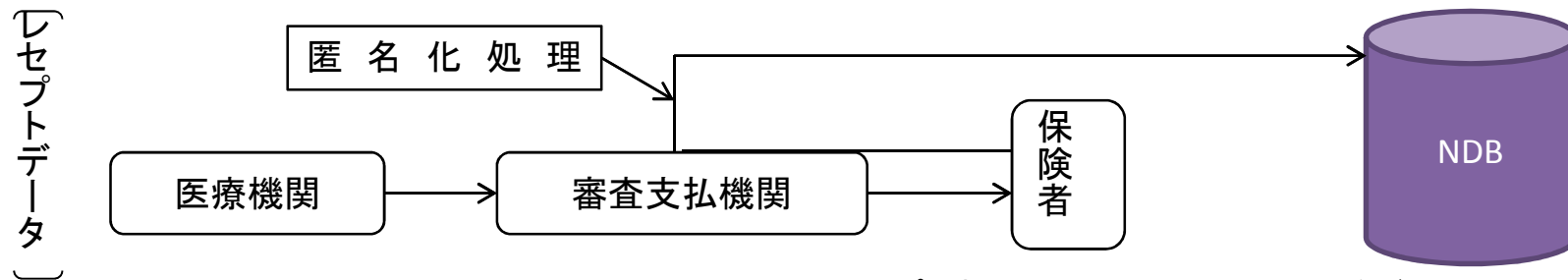


※ DPCデータにおける匿名化処理について

- 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

(参考)DPC導入の影響評価にかかる調査(退院患者調査)の結果報告で毎年公表されている内容一覧

(※概ね、調査年度の次年度の秋頃に結果が公表している)

参考資料①(経年変化に関する集計)

- 医療機関別集計
 - 平均在院日数
 - 救急車による搬送の有無
 - 救急医療入院
 - 他院よりの紹介、退院先、退院時転帰の状況
 - 再入院の状況
 - 手術件数 等
- 精神病棟の集計
- 医療圏別MDC患者数 等

参考資料②(当該年度データの集計)

- 診断群分類別集計
 - 年齢別件数
 - 退院時転帰
 - 在院日数
 - ICD10内訳
 - 実施された手術 等
- 医療機関別集計
 - 疾患別・手術別集計
 - MDC別・手術有無別・処置1有無別集計 等

(参考)DPCデータとレセプトデータの違いのまとめ

	DPCデータ	レセプトデータ
個人特定可能性	<ul style="list-style-type: none"> 患者住所地域の郵便番号等、直接的に特定に結びつく情報が含まれている 匿名化したデータ識別番号を格納しているがカルテ番号など個人が特定される情報を入力している可能性がある 診療録情報（がんのTNM分類等）が含まれている 	<ul style="list-style-type: none"> 氏名、生年月日の「日」、医療機関の所在地及び名称、カルテ番号、被保険者証（手帳）等の記号・番号等、個人が特定される恐れがある情報は削除している 同一人でのレセプトを連結できるように、氏名、被保険者（手帳）等の記号・番号等を基に匿名化を2重に行ったうえで格納している
医療機関特定可能性	<ul style="list-style-type: none"> 施設コードなど医療機関の特定に結びつく情報が含まれている 調査対象となる医療機関数が少ないため、特定可能性が高い 各医療機関に関する情報は一定程度公開されている 	<ul style="list-style-type: none"> 医療機関コード、電話番号など医療機関の特定に結びつく情報が含まれている 調査対象となる医療機関が多く、特定可能性は低い
情報の公開	毎年、退院患者調査の参考資料として一定程度結果が公表されている	公開されていない
年間レコード数	1,000万件 ※平成24年度	医科（約96億件） DPC（約1,167万件） ※平成24年4月～平成25年3月審査分
レコード単位	入院別	月別
対象となる医療機関	DPC対象病院（約1,500） DPC準備病院（約300） 出来高病院（約50）	病院（約8,500） 診療所（約88,000） 歯科（約71,000） 薬局（約55,000） ※平成25年6月診療分
1レコード単位あたりの項目数	様式1（130項目） EFファイル（31項目）	医科レセプト（約53項目） DPCレセプト（約40項目）

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）