

ると考えられる。また、米国と韓国のオンサイトセンタの機能は、来年度開設予定の日本のオンサイトセンタのサービスを考える上で有益であると考えられた。

#### **F. 研究発表**

満武巨裕（参考人）、平成 24 年度～平成 25 年度厚生労働科学研究「汎用性の高いレセプト基本データセット作成に関する

研究」の進捗状況について（資料 3-1）、第 17 回レセプト情報等の提供に関する有識者会議，平成 25 年 12 月 19 日

#### **G. 知的所有権の取得状況**

該当なし

参考資料1. 各国別レセプト情報の二次利用（第三者提供）状況

	日本	韓国	台湾	
格納データ	データアーカイブ機関等 (主な提供データの名称)	厚生労働省 提供データの正式名称はない。(レセプト情報・特定健診等情報データベース)	HIRA(The Health Insurance Review and Assessment Service) 提供データの名称 ・全体(外来)患者標本資料(100万人):約3%抽出 ・入院患者標本資料(70万人):約13%抽出 ・高齢者65歳以上標本資料:約10%抽出 ・小児・青少年(19歳以下)患者標本資料:約20%抽出	National Health Insurance Research Database (全民健康保険研究資料庫)
	機関又はデータの概要	厚生労働省保険局が、高齢者の医療の確保に関する法律第16条2項に基づき、保険者よりレセプトデータおよび特定健診・特定保健指導等データを集め、データベースとしている。 レセプトデータについては2013年4月現在、全国の全レセプトの92.2%を格納している。	DVDで提供している。 一年単位のデータ。 一年ごとにIDは完結している。複数年のDVDを購入しても、連結できない。	台湾では、1995年に国民健康保険プログラムが導入され、2007年時点で総人口2,296万人中、2,260万人がプログラムに加入。このプログラムに基づく、診療報酬支払いのためのレセプトデータを、国民健康保険局(Bureau of National Health Insurance: BNHI)が収集しデータベースを構築しており、その管理を国家衛生研究院(National Health Research Institutes: NHRI)が行っている。 データの利用を希望する研究者は、「一般申請」と「特別申請」の2つの方法で申請ができる。
	所有データ数(量)	2013年4月現在: レセプト情報等…約60億件 特定健診等情報…約9000万件	サンプルデータセット:4600万人からのサンプル(全受療者):韓国人口5100万人 明細書一般内訳:約3000万件 診療内訳:約2800万件 傷病内訳:約3000万件 処方箋詳細内訳:約1500万件	1996年からの全レセプトデータ
	運営経費の負担 (スポンサー)	国庫	国庫	データベースの運用費用は、国家衛生研究院が負担。ユーザーから手数料を徴収しているが、全体経費の2割以下である。
	データ提供開始時期又は アーカイブ設立時期	データ提供開始: 2011年 データ格納開始時期: 特定健診等情報…2008年度~ レセプト情報…2009年度~	2012年	1996年からの全レセプトデータ 1995年からの医療機関の属性データ(医療機関の特定はできない)
データ提供の根拠	本来目的での利用: 高齢者の医療の確保に関する法律 目的外利用: 高齢者の医療の確保に関する法律第16条第2項の規定に基づき保険者及び後期高齢者医療広域連合が厚生労働大臣に提出する情報の利用及び提供に関する指針(厚生労働省告示第424号)	公共データの提供および利用活性化に関する法律第三条(基本原則)の規定に基づき、公共データを誰もが便利に利用できるよう提供	学術研究を活性化させ、国民健康保険の付加価値を高めるためにデータを提供している。(根拠法は確認できず)	
民間提供の有無 提供している場合は、提供先について(例:「シンクタンクに限定」、「無制限」等)	提供されるデータを用いた研究の実施に要する費用の全部又は一部を国の行政機関から補助されている者	公共データの提供および利用活性化に関する法律第三条(基本原則)の規定に基づき、HIRAの統計分析チームで公共・民間部門に提供	民間への提供は、禁止している。 過去に製薬会社に、治験業務サポート等の研究目的として試行的に提供した時期があった。しかし、現在は研究者にしか提供していない。	
個人情報保護に関する取扱の根拠	レセプト情報等は匿名化がなされているので個人情報にはあたらない情報として位置づけられているが、個人を特定しうる情報が含まれているため、実運用で「個人情報に準ずる」情報として、「行政機関の保有する個人情報保護の保護に関する法律」に則った取扱を行っている。	個人情報には当てはまらない情報として認識されているが、個人を特定できる情報が含まれているため、実際の運用においては(個人情報に準ずる)情報として「行政機関が有する個人情報保護の保護に関する法律」に基づき取り扱っている。	個人情報コンピュータ処理保護法(電腦處理個人資料保護法(1995年))に則った取扱いをしている。(The Computer-Processed Personal Data Protection Law)	
漏洩・目的外利用等の不適切利用に対する対応及びその根拠規程	厚生労働省と提供依頼申出者・利用者との契約違反として取り扱う。違反内容については、ガイドライン(局長通知)により規定を行う。	申込者が利用誓約書の事項を違反した場合、あらゆる不利益および民事刑事上の処罰を甘受し、今後HIRAの資料を利用することに制限をおく	個人情報コンピュータ処理保護法(電腦處理個人資料保護法(1995年))に罰則・罰金規定を定めている。 同時に、誓約書などの規定に違反した研究者に対しては、その研究者の研究機関の長、研究者本人、共同研究者の全員に利用停止およびデータの返却を通知する。	
提供されるデータ	特別抽出: 依頼に応じて個票を抽出、提供 サンプリングデータセット: 1ヶ月分の医科入院、DPC、医科入院外、調剤レセプトを、抽出及び匿名化処理等を行ったうえで提供 集計表情報: 抽出対象項目に対し、三次元程度の区分を施した集計表を申出に応じ作成し、提供	匿名化抽出:原資料からそれぞれを無作為に抽出したのちに提供する。 サンプリング:原資料から提供が許容される範囲内で患者および療養機関に代替Keyを与えて個人情報を保護したのち、一年単位で抽出して提供。	(I) 基本資料データ(基本資料檔):集計データ (II) 系統抽出データ(系統抽樣檔):個票データ 入院レセプトファイル(DD)の5%抽出 外来レセプトファイル(CD)の0.2%抽出 (III) 特定主題データ(特定主題分檔):個票データ 特定疾患や調査目的に沿って抽出したデータファイル。悪性新生物、糖尿病、精神疾患、交通事故、リハビリ、漢方薬(中醫藥)ファイル等 (IV) ランダム抽出データ(抽樣歸人檔):個票データ 被保険者ファイル(ID)から、ランダムサンプリングしたもの。2000年は20万人を抽出し、(1996年~2007年)入院、外来、調剤レセプトデータを提供している。2005年の100万人、2010年の100万人の3種類が存在している。 (V) 教育用データ(教學用資料檔):個票データ 2000年のレセプトデータから1,000人をランダムサンプリングしたもの。教育用に無償提供。	
データ提供に関する審議機関(開示方針、匿名化方法等の検討・決定機関)	レセプト情報等の提供に関する有識者会議での助言を経て、厚生労働大臣が決定する。		一般申請は、行政審査。特別データ等については、行政審査を経た上で、専門家による審査を行う。	

参考資料1. 各国別レセプト情報の二次利用（第三者提供）状況

	日本	韓国	台湾
匿名化方法（一般）	<p>保険者からデータセンターに送付する際にハッシュ化（SHA-256）し、データセンターでも再度変換処理のうえハッシュ化し、格納している。</p> <p>なお、第三者提供時にはさらにハッシュ化を行い、第三者に提供されたデータ間での紐付けができないようにしている。</p>		<p>個人情報保護及びデータ安全保障観点から、中央健康保険局が2007年よりIDの匿名化を実施している。研究者に提供する際は、プロジェクトごとに再度匿名化をしている。</p>
匿名化方法（加工したデータセット）	<p>サンプリングデータセット：1ヶ月分のデータを外来、入院ごとに比率を設定して抽出し、高額レセプト等の削除、保険者種別等の空欄化、希少傷病名等の匿名化等を行っている。</p>	<p>サンプリングデータセット：それぞれのデータセットの抽出率で一年単位、患者別に抽出し、新たな代替Keyを与え、希少疾患および法定伝染病は削除などを実施して提供（約4500万人から140万人を抽出）。</p>	<p>個人情報として氏名、生年月日、身分証明書番号、婚姻、教育、職業、健康、病歴、財務状況等について匿名化を行っている。</p>
利用条件・手続き	<p>本来目的での利用： 医療費適正化対策推進室および都道府県が、医療費適正化計画に資する資料を作成する際に、適宜利用する。</p> <p>目的外利用： ガイドラインに規定された要件を満たす者（国、都道府県、研究者等）が、厚生労働省保険局に申出を行い、有識者会議において申出内容等を審査した上で、最終的に厚生労働大臣が承諾／不承諾の決定を行う。</p>	<p>申込者は、HIRAホームページを通して、申込書および利用誓約書を作成後、担当者に送付。担当者は、誓約書確認後、申込者に提供されるサンプルデータセットにシリアルナンバーを与えたのち提供</p> <p>申込者は、学術研究用以外の目的で使用することを禁止</p> <p>販売および貸出を行わない誓約書を提出</p>	<p>国家衛生研究院全民健康保険研究資料庫（National Health Insurance Research Database）使用申請概要に記載</p>
提供形態	<p>本来目的での利用： 医療費適正化対策推進室および都道府県において、抽出依頼をデータセンターに行い、データセンターからはエクセル形式の集計表が適正化室宛てに電子的に送付される。</p> <p>目的外利用： 申出者から、データセンターに対しに厚労省を介して個票の抽出もしくは集計表の作成が依頼され、データセンターからは、申出者において準備した電子媒体（DVDもしくは外付けHDD）に複写されたデータが書留もしくは手渡しにて送付する。</p>	<p>申込者がHIRAホームページを通して、申込書および利用誓約書を提出し、手数料を支払うと、担当者はDVDにコピーをし、シリアルナンバーを与えたのち、郵送するか直接提供する。</p>	<p>CDおよびDVD</p> <p>国家衛生研究院全民健康保険研究資料庫（National Health Insurance Research Database）使用申請概要に記載</p>
提供フォーマット	CSV	text	csv
データの利用料金	なし	一年分一枚（DVD）が30万ウォン。	あり C-ROM申請：一枚につき2,500元ニュー台湾ドル 特別データ処理：読み取り回数及び読み取り量を基準とし、1GB分のデータの読み取りごとに800元ニュー台湾ドル
データの提供実績	2011年度：1件 2012年度：13件	2012年度：69件 2013年度：115件	2013年は合計270件。ただし、一般申請と特別申請の合計。
データの寄託元、寄託手続き	<p>保険者、後期高齢者医療広域連合 レセプト情報： 審査支払機関から厚労省保険局に提供される。</p> <p>特定健診等情報： 各保険者から最終的に支払基金を通して厚労省保険局に提供される。</p>		全データが衛生院（日本の厚生労働省）から提供される。
当該サイトトップページ、Mission Statement、年次報告書	<a href="http://www.mhlw.go.jp/seisakunitsuite/bunya/kenkou_iry/iryuhoken/reseputo/">http://www.mhlw.go.jp/seisakunitsuite/bunya/kenkou_iry/iryuhoken/reseputo/</a>		<a href="http://nhird.nhri.org.tw/">http://nhird.nhri.org.tw/</a>
アーカイブスタッフ	厚生労働省 6名（他業務と兼務） ペンダー（富士電機） SE3名等	記録保管管理スタッフ：HIRAから2名（兼務）	兼任（13名）：NHIR 外部専門家（12名） 外部委員（12名）

## 参考資料 2. 韓国 HIRA におけるレセプト情報提供の概要

### 1 章 背景および目的

#### 1. 背景

- 医療分野での IT 技術の発達および医療の利用率の増加に伴い、保健医療関連のデータが急増
  - － 我が国（韓国）では、生活の質の向上や高齢化社会の進展に伴い高品質の医療サービスの需要が増加し、医療分野での IT 技術力の発達などの要因により保健医療関連のデータが急増している。
- 公共機関が保有するビッグデータは未来の中心的価値として重要性が急浮上している
  - － データ活用に対する社会的関心の向上とともに、医療情報の提供や支援活動などサービス拡充のニーズが持続的に増加。
  - － 公共機関が保有する多様なデータは、国家政策の策定および国民の健康増進に関する研究の基礎資料として活用することが可能。
- 公共ビッグデータを求める社会的ニーズの増大
  - － 資料活用の活性化を図るための提供チャネルの多様化および研究者支援システムの拡大など。
  - － 情報提供のインフラを構築するため、政府全体で「インテリジェンス」を確保できる保健医療関連の「ビッグデータ」のプラットフォーム構築方案を検討する必要がある。
- 近年の政府政策は、国民パーソナライズ型サービスを行うため公共ビッグデータを積極的に共有、公開する傾向
  - － 政府および政府の傘下機関は、公共ビッグデータを活用した国民パーソナライズ型サービスおよび科学的行政の具現などに対する公共機関間の協力について共通認識の形成を図っている。
- HIRA は、多様かつ膨大なデータを構築、保有しており、保健医療の発展のため積極的な公開に取り組んでいる
  - － HIRA は、全国民の 98%が登録されている我が国の保健医療を代表する健康保険請求データであるだけでなく、保健医療に関する様々なデータを運用している。

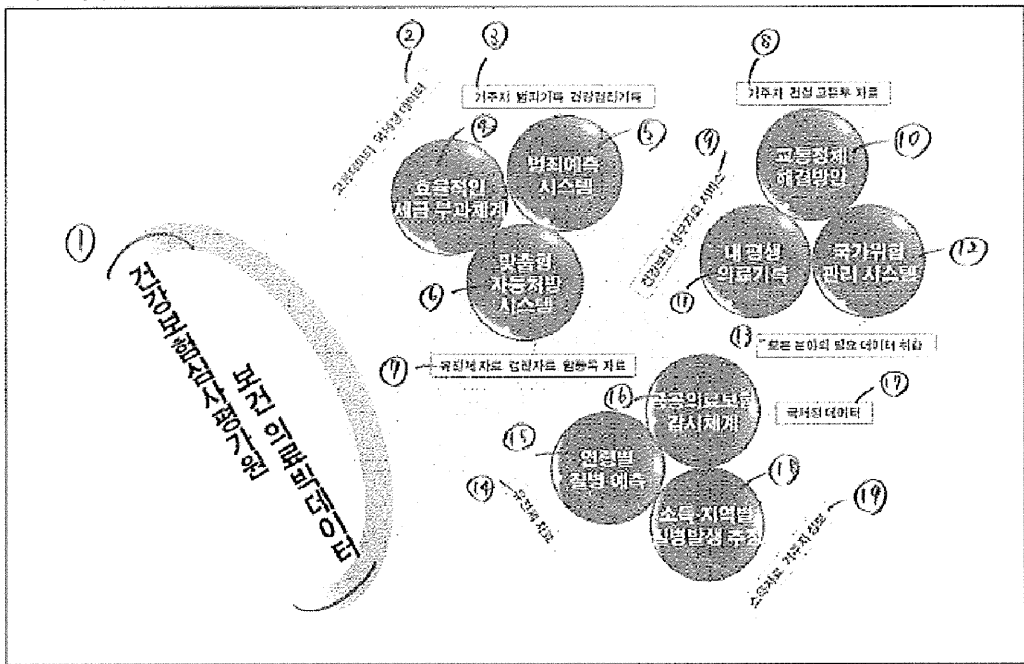
#### 2. 目的

- HIRA における「保健医療ビッグデータプラットフォーム」構築方案の検討
  - － HIRA の未来の重要価値創出のための情報化方案として、R&D 強化に向けた IT インフラの構築、連携サービスによるデータの多様性の拡大、需要者中心のパーソナライズ型コンテンツの提供による顧客サービスの最大化の実現が挙げられる。
  - － 保健医療ビッグデータプラットフォームの定義を、IT インフラの構築だけでなくデータの多様性の拡大、サービス改善にまで拡大し、実行する方案を検討。

- ビッグデータプラットフォーム実現のための IT インフラの構築
  - － 保健医療ビッグデータプラットフォームの実現に向け、ビッグデータデータベース、ビッグデータ分析システム、HIRA-dream ポータルを構築。
- ビッグデータプラットフォームの実現を通じた資料連携サービスの構築により、提供可能な情報の多様性を拡大
  - － 公共データの共有、公開を行うため、機関間における資料連携システムを構築
  - － 需要者中心の情報提供だけでなく、政府における保健医療政策の発展および保健医療分野やその他学術研究の活性化を図るため、様々な情報を提供
- ワンストップサービスによる顧客サービスの最大化を実現
  - － データの取得や分析、利用手続きなど手軽に資料の利用ができるよう、利便性に優れたワンストップサービスを導入することで研究環境を整備し、需要者が常時利用できるような情報提供に必要な制度的基盤を構築する。
- パーソナライズ型健康情報コンテンツの開発
  - － 審査評価院が保有する健康情報データを体系的に分析し、需要者のニーズに合わせたパーソナライズ型健康情報コンテンツを開発し、提供する方案を検討。
- 医療ビッグデータ分析専門家の養成
  - － 保健医療分野のビッグデータの持続的かつ体系的な活用に向け、医療ビッグデータ分析専門家養成プログラムを開発。
- 保健医療分野の多様な情報提供による R&D の活性化
  - － 多様な情報を需要者中心のパーソナライズ型サービスとして提供し、公共機関が保有する情報をもとに新事業の需要創出および保健医療産業の生産性を増大する。
  - － 民間経済の活性化を図るため公共部門の役割を強化する。
  - － 全国民を対象とした健康保険請求資料が保健医療分野の国家政策の策定および学術研究の基礎資料として活用されることによって、保健医療情報の根拠および分析に基づいて価値を創出し、保健医療分野の国家政策の策定および意思決定に寄与することが期待される。

<図>ビッグデータプラットフォーム構築による多様な価値の創出

<그림> 빅데이터 플랫폼 구축을 통한 다양한 가치 창출



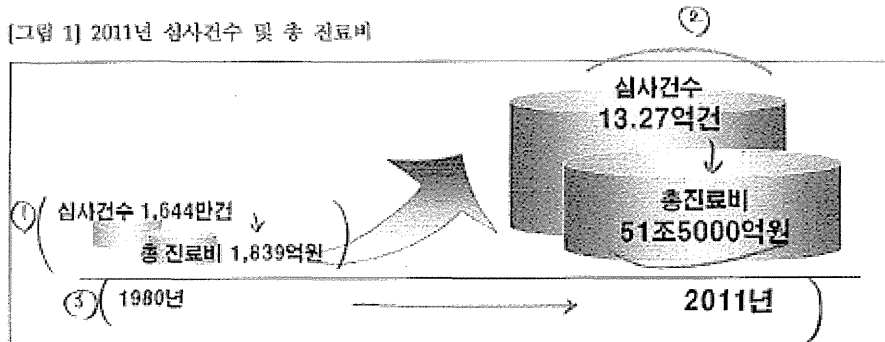
- ① HIRA 保健医療ビッグデータ
- ② 雇用データ、国税庁データ
- ③ 居住地、犯罪記録、健康検診記録
- ④ 効率的な課税システム
- ⑤ 犯罪予測システム
- ⑥ パーソナライズ型自動処方システム
- ⑦ ゲノム資料、検診資料、がん登録資料
- ⑧ 居住地、建設交通部資料
- ⑨ 健康保険請求資料サービス
- ⑩ 交通渋滞解決方
- ⑪ 生涯カルテ
- ⑫ 国家危険管理システム
- ⑬ 全分野における必要データの統合
- ⑭ ゲノム資料
- ⑮ 年齢別疾病予測
- ⑯ 公共医療保険監視システム
- ⑰ 国税庁データ
- ⑱ 所得・地域別疾病発生の推定
- ⑲ 所得資料、居住地情報

## 2章 保健医療資料の現況および概要

### 1. HIRA における保有資料の概要

- 「健康保険請求資料」とは、医療機関において患者の診療費用のうち「国民健康保険」が負担する部分について支給を受けるため HIRA に保険給付請求を行う際に発生する情報をいう。
  - － HIRA は、医療機関が請求する「請求明細書」に対する診療費の適正性を審査し、その結果を「国民健康保険公団」に通知し、支給(医療費支払)を要請する。
- 我が国の年間「健康保険請求患者数」は、2011 年は 45,804,866 人で住民登録人口 50,734,284 人のうち 90.3%に該当する
  - － 健康保険請求に対する審査件数および請求診療費総額は少しずつ増加を続け、2011 年の審査件数は約 13 億件、総診療費は約 51 兆 5 千億ウォンに達する。

[図 1] 2011 年審査件数および診療費



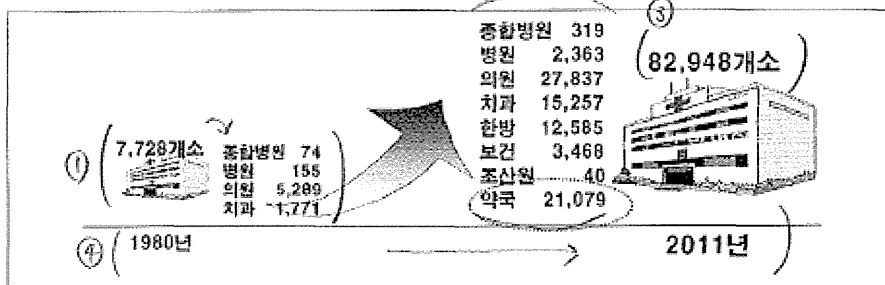
○ 건강보험에 등록된 요양기관 현황은 1980년대 7,728개소에서 2011년 기준으로 82,948개소로 증가

- ① 審査件数 1,644 万件、総診療費 1,839 億ウォン
- ② 審査件数 13.27 億件、総診療費 51 兆 5000 億ウォン
- ③ 1980 年、2011 年

- 健康保険に登録されている療養機関の現況は、1980年代の7,728か所から2011年には82,948か所まで増加

[図2] 2011年健康保険登録療養機関の現況

[그림 2] 2011년 건강보험 등록 요양기관 현황



① 7,728 か所

総合病院	74
病院	155
医院	5,289
歯科	1,771

②

総合病院	319
病院	2,363
医院	27,837
歯科	15,257
漢方	12,585
保健	3,468
助産院	40
薬局	21,079

④ 82,948 か所

⑤ 1980年、2011年

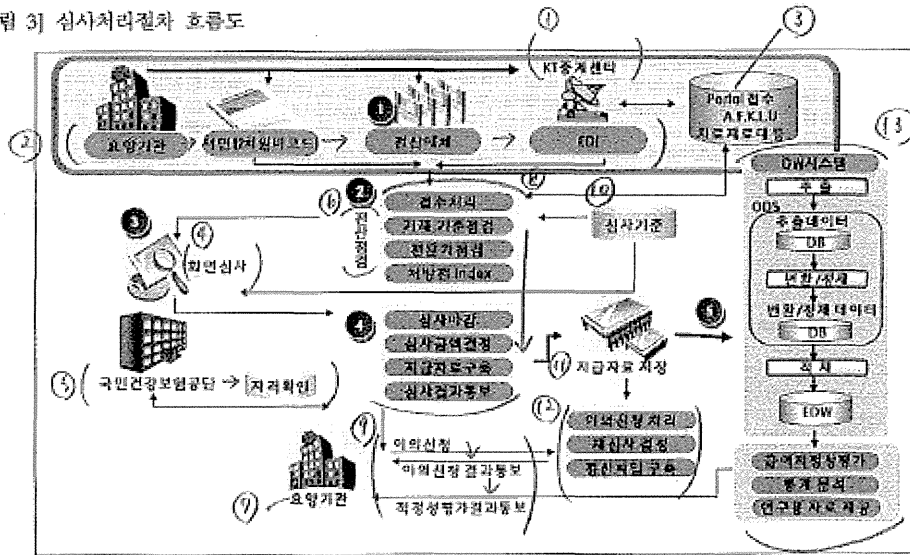
- 病院は患者を診療した場合、給付償還のため HIRA に [図3] の手続きに従って療養給付費用を請求

- 一 療養機関から診療費請求明細書を受理したときは、データ点検および審査段階を経てデータを累積



[図 3] 審査処理の手続きの流れ

[그림 3] 심사처리절차 흐름도



- ① KT 中継センター
- ② 療養機関、書面 (2次元バーコード)、データ媒体、EDI
- ③ ポータル受理、A.F.K.L.U.、治療材料代等
- ④ 画面審査
- ⑤ 国民健康保険公団、資格確認
- ⑥ データ点検
- ⑦ 療養機関 (医療機関)
- ⑧ 受理処理、記載、基準の点検、専門家による点検、処方せんインデックス  
審査締切、審査金額の決定、支給資料の構築、審査結果の通知
- ⑨ 異議申立て、異議申立て結果の通知、適正性評価結果の通知
- ⑩ 審査基準
- ⑪ 支給資料の保存
- ⑫ 異議申立ての処理、再審査の決定、精算ファイルの構築
- ⑬ DW システム、抽出、抽出データ、変換/生成、変換/生成データ、ロード  
給付適正性評価、統計分析、研究用資料の提供

○ [図 3] のように蓄積されたデータは DW (Data Warehouse) システムに保存され、給付適正性評価や統計分析、研究資料の提供用として活用

ー 収集された資料には、全ての診療、処置、処方内訳が含まれており、対象となる資料は大きく分けて療養給付費用請求明細書資料と療養機関現況資料とで構成される。

○ 療養給付費用請求明細書資料とは、医療機関および薬局などで患者に診療または調剤し

た後、療養給付費用請求方法に従って作成した診療内訳が記載された資料をいう。

- － 請求明細書は、EDI、データ媒体（フロッピーディスク、CD）または書面で請求することができ、健康保険患者、医療給付患者、報奨国費患者の診療費請求内容を全て含む。

<表 3> 健康保険請求資料の主要パラメータ情報

一般事項	分類コード	受診者氏名、受診者代理キー、加入者番号、事業所番号、療養機関分類コード
	細部事項	傷病、手術の有無、診療科目、療養開始日、入院日数、外来訪問日数、処方せん数、処方日数、初診および再診回数、診療結果、療養給付費用総額、本人一部負担金、保険者負担金など
診療細部内訳		診療費項目別内訳（行為、医薬品、治療材料別に細分化して記載） 処方調剤詳細内訳（個別薬品別処方および調剤内訳）

- 提出された療養機関現況資料には、療養機関の一般現況、病床現況、医療人材現況、医療装備現況を含む
  - － 療養機関現況資料は、請求明細書資料とともに HIRA データベースに保存管理され、統計資料として活用。
  - － 療養機関現況資料は、HIRA に請求した療養給付費用を審査、評価するのに必要な基礎資料として活用するため、少なくとも月 1 回、療養機関から法定書式「療養機関現況通知書」および「療養機関変更事項通知書」をデータで受け取る。

<表> 療養機関現況の主要パラメータ情報

一般現況	開設者の人的事項、住所、設立形態、救急医療機関、開放病院などの運営現況および開設診療科目など
病床現況	入院室、特殊診療室（集中治療室、手術室、救急室、人工透析室、無菌治療室、隔離病棟など 12 種類）、昼病棟など
医療人材現況	医師、歯科医師、漢方医師、薬剤師、看護師、臨床病理技術士、放射線技師、栄養士など 24 分類
医療装備現況	CT、MRI、PET など放射線診断および治療、検査、理学療法、手術および処置、漢方装備など 167 種

- HIRA は、健康保険請求資料および療養機関現況資料の他にも様々な資料を<表>のとおり保有している。

<表>HIRA 保有資料現況

資料名	資料内容	資料件数 (容量)	活用内容
合計：15種			
療養給付費用請求明細書 DB	療養機関（病・医院、薬局）の診療費請求明細書（一般内訳、傷病内訳、診療内訳、院外処方内訳）	7,581 百万件 (33TB)	療養給付費用審査、評価指標算出および各種統計生産
療養機関現況 DB	療養機関現況情報（療養機関記号、療養機関名、所在地、電話番号、郵便番号、設立日時、種別区分、医師数などの人材現況、CT 装備数などの装備現況、病床数など）	8 万件 (0.1GB)	療養機関別一般現況および請求資料連携分析活用
医薬品処方調剤情報 DB	療養機関において、受診者に処方・調剤する併用禁止、年令禁止などの処方せん情報	116 億件 (6.5TB)	医薬品の安全性関連情報を医師や薬剤師にリアルタイムで提供および DUR 運営現況のモニタリング
医薬品生産実績 DB	医薬品生産内訳（製造者、製造年月、生産量、生産金額など）	50 万件 (3GB)	国内医薬品生産現況分析
医薬品輸入実績 DB	医薬品輸入内訳（輸入者、輸入年月、輸入量、輸入金額など）	8 万件 (1GB)	国内医薬品生産現況分析
医薬品供給実績 DB	医薬品供給内訳（供給業者、購入業者、供給日、供給量、供給金額など）	12 億件 (6TB)	医薬品流通現況および医薬品供給・請求比較分析
安全常備医薬品コンビニエンスストア情報 DB	安全常備医薬品販売先情報	17 千件 (0.2GB)	安全常備医薬品販売先の管理
医薬品 RFID Tag 情報 DB	RFID tag 付医薬品情報	10 百万件 (24GB)	RFID 付医薬品情報の提供および管理
療養機関総合情報 DB	一般（医療機関名、記号、代表者）人材（氏名、住民番号、資格、勤務）装備（種類、コード）	22 百万件 (4TB)	療養給付費用支給のスライド制などに活用
統計資料 DB	健康保険統計年譜（健康保険請求件数、診療費現況など） 診療費統計指標（健康保険／医療給付審査実績など）	20 万件 (84MB)	統計の発刊 保健医療研究 情報公開
疾病統計 DB	傷病小分類統計（3 段傷病の患者数および診療費など） 傷病細分類統計（4 段傷病の患者数および診療費など）	186 万件 (478MB)	統計の発刊 保健医療研究 情報公開
行為統計 DB	行為別統計（報酬、保険登載薬、治療材料などの診療コードに基づく実施回数および金額）	50 万件 (262MB)	統計の発刊 保健医療研究 情報公開
病院評価情報 DB	急性心筋梗塞症／急性期脳卒中／高血圧／糖尿病／血液透析／医療給付精神科／手術の予防的抗生剤／手術別診療量／診療結果／帝王切開／冠動脈バイパス術／抗生剤処方率／幼・小児中耳炎抗生剤／抗生剤処方率／注射剤処方率／薬品目数／処方薬品費／療養病院／大腸がん評価結果	4 億件 (50GB)	情報公開 保健医療研究 療養機関医療のクオリティ向上の管理
療養病院患者評価表 DB	療養病院患者評価表（療養機関、患者情報事項、入院日、血圧、認知機能、身体機能、排せつ機能、疾病診断、健康状態、皮膚状態、特殊処置）	12 百万件 (6GB)	療養病院定額報酬の算定および療養病院審査時に活用
死亡推定者資料	死亡資料（人的事項、死亡日、市道コード、福祉対象有無）	1 百万件 (100MB)	行政情報共同利用センター 死亡情報の連携

- 健康保険請求資料を活用する際の注意事項として、給付が認められた医療利用の内訳のみ含まれているため、非給付内訳または処方せんがなくても購入が可能なアスピリンなどの一般医薬品に対する情報については、請求資料から確認することはできない。
- 診断名の正確性に対する研究者の考慮が必要
  - － 診断名の正確性は外来よりは入院患者、多頻度軽症患者よりは重症患者に高く現れ、医院よりは総合病院の療養機関においてさらに高い傾向がみられる。
  - － 診断名および施術において医師の個人差、慣習的要因を完全に排除することは困難なことから、資料の特性、患者の医療利用の形態や疾病の固有特性、医師の診療過程および臨床環境、病院の電算システムおよび請求過程、健康保険給付制度などを十分に把握してこそ正しい解釈が可能である。
  - － 診断名の妥当度などを周期的に評価し、請求資料の持続的な品質管理を通して信頼性を高めることも重要な課題。

## 2. (韓国) 国内保健医療資料の保有現況

- 機関別保有資料源の形態および特徴に応じて、標本方式ではなく全国を対象とする母集団の資料源として、健康保険請求資料 (HIRA)、中央がん登録資料 (国立がんセンター)、資格資料 (国民健康保険公団)、全国多文化家族実態調査 (韓国保健社会研究院)、乳児母性死亡調査 (韓国保健社会研究院)、死亡原因資料 (統計庁)、新生児資料 (統計庁)、居住地および死亡資料 (安全行政部)、所得水準資料 (国税庁) などで区分される。
  - － 標本単位資料は、資料連携時に必要な事前同意を獲得しやすく、その容量が小さいため機関同士の同意が得られれば USB メモリのみでも移動が可能。

<表>機関別保有保健医療資料

資料名	収集方式	収集周期	生産機関
健康保険請求資料	全国民対象	毎日随時データ発生	健康保険審査 評価院
中央がん登録資料	全国 (がん患者対象)	初回発生以降、 医療機関の移動 あるいは完治時に更新	国立がんセンター
国民健康保険公団 健康診断資料	該当検診対象者	2年に1回あるいは 1年に1回検診	国民健康保険公団
資格資料	全国民対象	変動事項発生時	
国民口腔健康実態 調査	学校標本 (未就学児および 小中高生対象)	3年に1回	保健福祉部
国民健康栄養調査	世帯標本(全国民対象)	毎年1回	疾病管理本部
地域社会健康調査	世帯標本(全国民対象)	毎年1回	
青少年健康状態 オンライン調査	学校標本 (中1～高3年生対象)	毎年1回	
退院損傷深層調査	病院標本(退院患者対象)	毎年1回	
高齢者実態調査	世帯標本(高齢者対象)	3年に1回	韓国保健社会研究院
ベビーブームの生活実 態 および福祉欲求	世帯標本 (1951～1964年生まれ 対象)	2010年1回	
障害者実態調査	世帯標本 (障害者および 非障害者対象)	2000年、2005年、 2008年、2011年 年度別1回	
全国多文化家族 実態調査	全国(多文化家庭対象)	2009年1回	
全国出産力および 家族保健福祉 実態調査	世帯標本(既婚および未婚 男女対象)	3年に1回	
乳児母性死亡調査	全国(乳児母性死亡者対象)	約3～4年に1回	
次上位階層実態調査	世帯標本(次上位階層対象)	2007年1回	韓国雇用情報院
患者調査	医療機関標本 (全国患者対象)	毎年1回	
韓国福祉パネル	世帯標本(全国民対象)	2006年基準1年毎に 追跡調査	
韓国医療パネル	世帯標本(全国民対象)	2008年および2009年 半期別追跡調査	
高齢化研究パネル	人口標本 (45才以上の国民対象)	2005年基準2年毎に 追跡調査	
韓国労働パネル	世帯標本(全国民対象)	1998年基準1年毎に 追跡調査	
社会調査	世帯標本(全国民対象)	毎年1回	
統計庁死亡原因資料	全国(死亡者対象)	死亡時発生	統計庁
統計庁新生児資料	全国(新生児対象)	出生時発生	
居住地および死亡資料	全国 (移住者および死亡者対象)	居住地移動および 死亡時発生	安全行政部
所得水準資料	全国民対象	変動事項発生	国税庁

- 全国を対象とする母集団の資料源のうち、特定階層（乳幼児、がん患者など）以外の全国民を対象とする資料源は、「健康保険請求資料」、「所得水準資料」程度として再度制限
  - － 全国単位、全国民対象の資料源において、資格資料および所得水準資料の場合、随時ではなく変動事項がある場合のみにアップデートが行われるため、データ量自体はそれほど多くはないものの、健康保険請求資料の場合は全国民を対象とし随時生成され、膨大な容量を有するため相当な初期投資費用<sup>1)</sup>（約 700 億ウォン）および維持、補修費用が発生（年 60 億ウォン以上、データシステム職 160 人以上）

<表>母集団の資料源

資料名	収集対象	収集時期
健康保険請求資料	全国民対象	毎日随時データ発生
資格資料	全国民対象	変動事項発生時
中央がん登録資料	がん患者対象	(初回発生以降) 医療機関の移動あるいは 完治時に更新
統計庁死亡原因資料	死亡者対象	死亡時発生
統計庁新生児資料	新生児対象	出生時発生
居住地および死亡資料	移住者および死亡者対象	居住地移動および死亡時発生
所得水準資料	全国民対象	変動事項発生時
全国多文化家族実態調査	多文化家庭対象	2009 年 1 回
乳児母性死亡調査	乳児母性死亡者対象	約 3~4 年毎に 1 回

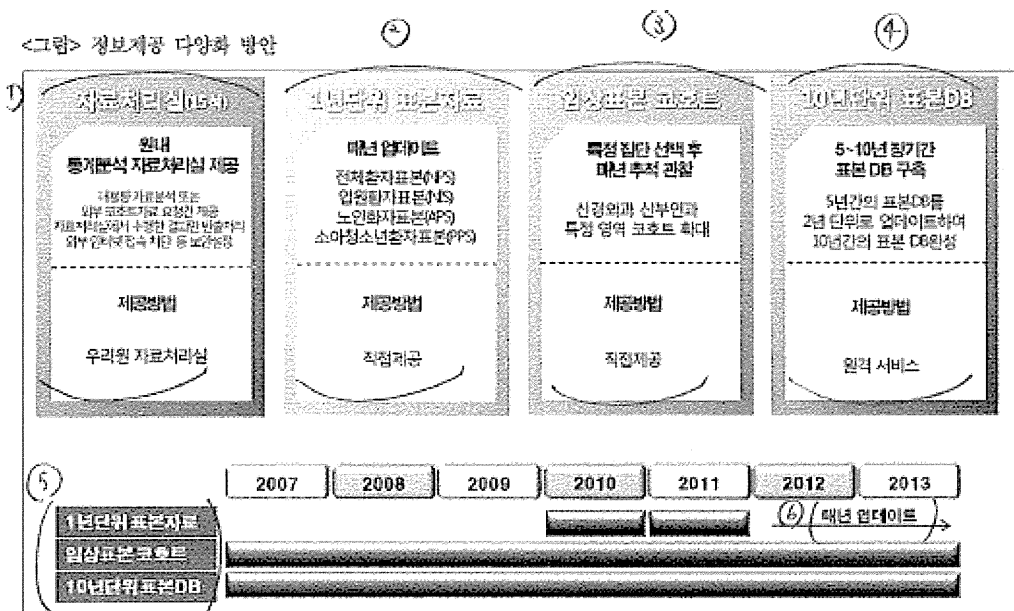
- 健康保険請求資料はビッグデータの構築に重要な要素
  - － 健康保険請求資料は、個人情報だけでなく医療の利用情報を時間の流れに沿って確認することができ、ビッグデータの構築において重要な要素として作用
- 健康保険請求資料の分析および管理には、莫大な維持補修費用がかかる
  - － 毎日新規に生成されるためその容量は膨大で、莫大な維持補修費用が発生し、また資料が流出した際に個人情報を侵害する危険性が高い
  - － 健康保険請求資料は、機関間の移動が非効率的で、重複投資のおそれがある

<sup>1)</sup> 健康保険審査評価院における請求資料管理のための投資および維持、補修費用基準

## 2. HIRA における保有情報の提供（開放）に関する多様な方法

- 社会的な公共情報活用の要求増加に伴う HIRA 保有の公共情報の積極的な開放および公開
  - － 情報公開を通じた国民の知る権利の充足および公共データ活用の活性化を支援
  - － 政府 3.0 の趣旨に足並みを揃え、多様なチャンネルを通して HIRA 保有の資料を段階的に開放
- HIRA 資料提供サービスの多様化の要求（多様なチャンネル、データの連携）
  - － 領域別、段階別データ拡大の構築
  - － 資料提供領域の多様化を通して、研究者の資料の活用を最大化
  - － 資料処理室：院内に研究者専用の統計分析資料処理室を提供
  - － 遠隔接続サービス：当院に仮想 PC を接続し、統計分析が行える環境を提供
  - － 標本資料の拡大開放：標本資料の多様化を通して、研究者の資料へのアクセスおよび活用を最大化
- 需要者中心の情報開放（アクセス性、便宜性の拡大）
  - － 需要者の便宜性およびアクセス性を考慮し、様々な情報を効率的に提供する具体的な方案が必要
  - － 費用対効果の高い量的、質的データの提供により、新たな価値の創出が可能
  - － 保健医療情報の積極的な開放および公開により、国民の知る権利を高める
  - － 公共データの民間活用の活性化
- 保健医療分野の研究の活性化に向けた保健資料の提供インフラに対する必要性を強調
  - － 審査、評価の革新および保健医療政策の決定（または政策研究）時に開放された公共データの分析を通して科学的根拠を提供
  - － 臨床学会とのワーキンググループ会議を通して需要者中心の研究データセットとして拡張
  - － 費用対効果の高い保健医療データの提供により、保健医療分野の研究の活性化を支援
- 現在の HIRA の情報提供の方法により、患者標本資料の提供および資料処理室の利用を通して資料を提供
  - － HIRA 内において、研究者専用に統計分析のための資料処理室 15 席を使用空間として保有しており、需要の増加時には資料処理室を拡大する計画
  - － 一般の研究者を対象に提供される患者標本資料の多様化を通して、研究者の資料へのアクセスおよび活用を最大化
- 研究者の便宜性を図るため、HIRA に仮想 PC を接続して統計分析が行える遠隔接続サービスを 2013 年 11 月から提供する計画

< 図 > 情報提供の多様化方案



① 資料処理室 (15 席)

院内

統計分析資料処理室の提供

大容量の資料分析または外部コホート資料の要請件の提供

資料処理室で行われた結果のみ搬出処理

外部インターネット接続の遮断などセキュリティの設定

提供方法

当院資料処理室

② 1 年単位の標本資料

毎年アップデート

全体患者標本

入院患者標本

高齢者標本

小兒青少年患者標本

提供方法

直接提供



### ③臨床標本コホート

特定集団を選択した後、毎年追跡観察

神経外科、産婦人科

特定領域のコホートの拡大

提供方法

直接提供

### ④10年単位の標本データベース

5～10年長期間標本データベースの構築

5年間の標本データベースを2年単位でアップデートし、10年間の標本データベースを完成

提供方法

遠隔サービス

### ⑤1年単位の標本資料

臨床標本コホート

10年単位の標本データベース

### ⑥毎年アップデート

## 1) 資料処理室を利用したパーソナライズ資料の提供

#### ○ 詳細かつ具体的な資料の提供が可能（大規模資料）

- － 研究者の研究目的に合った資料をパーソナライズして製作し、直接提供
- － 個人情報保護が弱い膨大な資料の提供および個人情報の連携が必要なコホート資料の提供
- － 全国民を対象とした情報であるため、発症の少ない希少疾患の分析が可能
- － 資料算出可能期間：最新審査決定分（5か年度基準）

#### ○ 国策事業および公共研究の支援（外部機関の調査資料および健康保険請求資料の連携）

- － 非営利学術、研究または公益機関における学術、研究目的などの場合に提供（国家機関、公共研究機関、大学、病院、研究センター、学会など）
- － 国家、行政機関において、業務遂行に関連して要請する場合

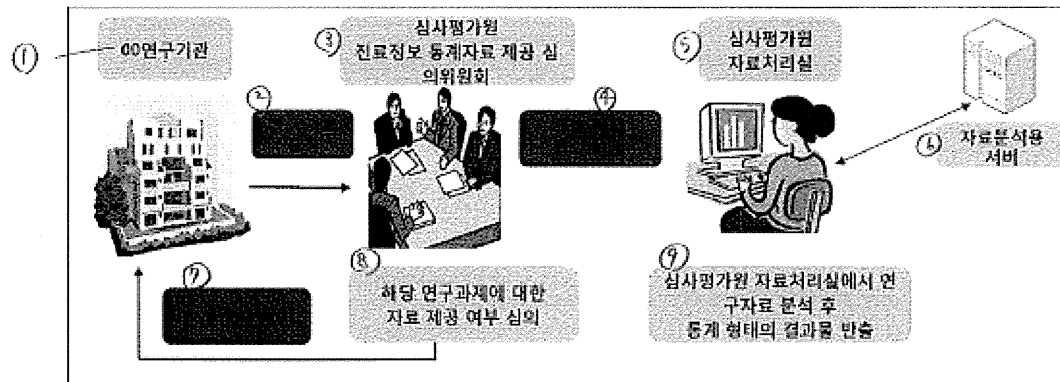
#### ○ 情報保護措置後に提供

- － 氏名、療養機関の名称および療養機関記号など、個人情報および個別法人、団体の情報は、識別不可能な形態に変形して提供 ⇒ 住民登録番号および療養機関記号は、別途の代替番号を付与し重複性を排除

- 審査評価院内に設置された資料処理室で分析できるよう提供
- 健康保険請求明細書の記載事項の範囲内において資料の連携が可能
  - 患者の包括的な事前同意の獲得や、健康保険請求資料および連携に関する同意を獲得した場合に連携が可能
  - 非給付、全額本人負担の内訳は算出不可

[図] 資料処理室の運営手順

[그림] 자료처리실 운영 절차



- ① 〇〇研究機関
- ② 研究資料の要請
- ③ 審査評価院、診療情報統計資料の提供、審議委員会
- ④ 資料の提供 審議議決
- ⑤ 審査評価院 資料処理室
- ⑥ 資料分析用サーバー
- ⑦ 審議否決
- ⑧ 該当研究課題に対する資料提供の可否を審議
- ⑨ 審査評価院資料処理室で研究資料を分析した後、統計形式の結果物を搬出

## 1) 1年単位標本資料の開発の提供

- 事前情報公開の一環として、一般の研究者に重複した統計資料の提供業務を削減し、資料を効率的に使用することができるよう事前に加工、公表された資料を提供（標本資料）
- 「HIRA」で「資料処理室」を運営しているが、直接来訪して利用しなければならないという煩わしさがあり、アクセス性および便宜性の側面において限界が生じていたことから、直接提供することが可能な患者標本資料を開発
  - － 年間約 10 億件以上の膨大な容量の資料は、使用者の保存容量、処理速度など受け入れ能力の限界により研究者への時期適切な資料確保が不可能である
  - － 多様な需要層に対するアクセス性および便宜性、即時性を確保するための手段の一つとして、我が国の健康保険請求資料に対する入院患者標本資料（HIRA-NIS）を 2010 年 12 月に開発し、2012 年 6 月から提供サービスを開始
- 2013 年 6 月現在、3 種類の標本資料を追加で開発し、提供中
  - － 全体患者標本資料を通して重症疾患のような入院診療を研究するには代表性が不足
  - － 標本資料の種類を、全体患者標本（NPS）、入院患者標本（NIS）、高齢者患者標本（APS）、小児・青少年患者標本（PPS）に分類して提供。
  - － 患者の特性および代表性が反映されるよう特定階層に対して別途標本を抽出することにより、その特定階層だけが持つ患者の特性および代表性を高め、研究に対する活用度を向上させる。

<表>標本資料の種類および算出基準

標本資料の種類		算出基準
4 月提供開始	HIRA-NIS	1 年単位入院患者約 70 万人（13%）、外来患者約 40 万人（1%）
	HIRA-NPS	1 年単位全体患者約 140 万人（3%）
6 月提供開始	HIRA-APS	1 年単位 65 才以上患者約 100 万人（20%）
	HIRA-PPS	1 年単位 20 才未満患者約 110 万人（10%）

※各患者の標本資料の標本限界値は、患者数 150 万人または領域別 20%以内を基準

- 患者標本資料の制限点は全ての標本資料における共通の限界点であり、標本資料内の観測値は確率に応じて抽出される資料であるため、適正水準以上の標本数を確保してこそ代表性、有意性が保障される。
  - － 患者標本資料において、特定年齢層の希少疾患発生頻度の場合、標本抽出頻度が非常に少ないため代表性および説明力に欠けるおそれがある。
  - － 標本資料の説明力は、傷病の発生頻度が上がるほど高く、発生頻度が下がるほど低くなる。
- 患者標本資料の制限点を解決するための方案として、臨床標本コホート資料の開発および遠隔サービスの提供方案を開発
  - － 臨床標本コホートは、1 年単位の患者標本資料を拡張して患者の医療使用履歴を時系

列に沿って確認することができ、活用可能な研究の範囲が広がる。

- － 遠隔サービスの場合、個人情報保護に関する処理を行った後は、大きな制約なく健康保険請求資料の大部分について使用することが可能。

## 2) 臨床標本コホートの構築計画

- 臨床学会とのワーキンググループを通して、個人情報保護される範囲内で多年間の標本追跡コホート資料を構築
  - － 1年単位の患者標本資料では、分析が不可能な希少疾患など長期間のフォローアップが必要な患者で、より専門的かつ臨床的な領域を対象
  - － 個人情報を保護するため、研究者ごとに異なる患者識別代理キーを付与し、資料の移動経路を追跡することができ、研究者間のデータの連携が行えないよう個人情報の保護に焦点を当てながら開発を進めている。
- 2013年度12月までに、脊椎手術患者を5年間追跡した脊椎手術コホート資料および産婦の出産日を基準に過去1年前から出産後の5年間を追跡した産婦人科コホート資料を構築

<表>臨床標本コホートの種類

コホート名称	内容
脊椎手術コホート	○ 2007年度整形外科および神経外科脊椎手術（診断）患者を対象に5年間の追跡コホート資料を構築
産婦人科コホート	○ 2008年度に出産した産婦49万人に対する標本産婦を抽出 ○ 2008年度に出産した産婦の出産日を基準に過去1年間の診療内訳を追加（後向きコホート） ○ 産婦の出産日を基準に産婦と新生児の5年間の診療内訳を追加（前向きコホート）

- 産婦人科コホートの構築により、産婦が持つ疾患（高齢産婦、妊娠性糖尿など）により新生児の死亡率など新生児に及ぼす影響を把握することができ、新生児の難治性疾患による家計負担率を把握することが可能
  - － 長期間のデータ構築の場合は、小児から成人になるまでの疾患の持続的把握が可能
- 産婦人科コホートを構築する場合、産婦人科・小児科で多くの研究が活発に行われることが予想される
- 今後、医薬品分野および内科分野のコホート資料などにより多様化を計画

## 3) 大規模患者の長期間標本データベース（遠隔サービス）の構築計画

- あらゆる研究が可能となり、標本資料に対する様々な需要を満たす広範囲かつ包括的な領域の研究用データベースの構築を計画
- 最小500万人の大規模患者の長期間標本データベースはあらゆる研究が可能で、標本資