

業務を効率化させることである。そのために、HIRA は事前に作成した HIRA-NPS を研究者に提供することで、上記の目的を達成しようとしている。日本のレセプト情報等データベース（以下、NDB）から提供されるデータセットは、厚生労働省側で複雑な構造の電子レセプトを研究者の要望に応じて、ある程度分析し易い形式に加工して提供する特別抽出と一月分のサンプリングデータの提供サービスが存在している。加えて提供件数の増加につながると共に負担軽減にもつながると考えられる、汎用性の高い基本データセットの設計と作成が検討され、今年度から試行提供が始まった。したがって、韓国の HIRA-NPS は、今後の日本の汎用性の高い基本データセットの提供サービスを拡充する上で有益であると考えられる。

A. 研究目的

レセプト情報・特定健診等情報データベース（以下、NDB）は、高齢者の医療の確保に関する法律に基づき、医療費適正化計画の作成、検討のための資料を作成することを目的に国（担当部局：厚生労働省・保険局・保険システム高度化推進室）が構築・管理している。NDB は、2009 年度から全ての電子レセプト、2008 年度から導入された特定健康診査（以下、特定健診）・特定保健指導のデータを収集しており、悉皆性の高いデータを保持している。

医療サービスの質の向上を目的とする公益性の高い研究等にも活用するために、本データの研究者等への第三者提供が 2011 年度から試行的に開始された。しかし、第一回のデータ提供に関しては、43 件の申出に対して承諾件数は 6 件であり、申出件数および承諾割合は必ずしも高いものではない。本データの承諾審査を行う「レセプト情報等の提供に関する有識者会議（以下、有識者会議）」からは、次の点が指摘されている。(1)申出者が求めるデータ項目が実際に格納されているデータでは実現困難であった申請が存在した、

(2)データ提供にあたっての各種要件や必要な事項を申出者が十分に把握していない申請が存在する、(3)提供側の情報提供が不十分であった等である。そのために、厚生労働省は、本データの提供依頼申出者からの研究の実現可能性やその具体策等の助言などを行う窓口の設置、データを利用できるオンサイトセンタの設置を検討している。

そこで初年度は、申出者へのレセプトデータ提供サービス機能をすでに有している諸外国（韓国、台湾、米国）の仕組みを調査することで、今後の日本における適切な体制作りについて検討した。

次年度は、2012 年から新たなデータセットの提供を開始した韓国の HIRA-NPS (HIRA National Patient Sample (健保審査評価院・患者標本資料)：以下、HIRA-NPS とする)について検討した。

B. 研究方法

初年度の調査対象国は、レセプトデータを既に研究者へ提供している韓国および台湾である。調査項目は、以下の 6 項目を共通項目とした。レセプトデータベース

名称および担当機関、提供データの概要、データ提供開始時期、民間提供の有無、個人情報保護に関する取扱い根拠、漏洩等の不適切利用に対する対応およびその根拠である。

オンサイトセンタとしては、米国のDOH(U.S. Department of Health & Human Services)傘下のAHRQ(Agency for Healthcare Research and Quality)が保持する医療費関連データMEPS(The Medical Expenditure Panel Survey)を、研究者に対してAHRQ施設内での利用環境を提供しているデータセンタを調査した。

次年度は、韓国のHIRA-NPSについて調査した。韓国は、日本における審査支払機関に該当するHIRA(The Health Insurance Review and Assessment Service (健康保険審査評価院)：以下、HIRAとする)が、HIRA-NPSを作成した。HIRA-NPSは、性別・年齢(5歳単位)区間による患者単位の層化系統抽出したデータセットである。HIRA-NPSの開発目的は、1)健康保険に関する診療情報の利活用の増大、2)情報公開を促進させると同時にHIRAのデータ提供に要する業務を効率化させることである。そのために、HIRAは事前に作成したHIRA-NPSを提供することで、目的を達成しようとしている。

2009年版(該当年度1月から翌年3月までの審決資料を基準に該当年度1年分の診療内訳を構築)のHIRA-NPSには入院患者の約70万人、外来患者約40万人が含まれる。データ項目には、標本抽出患者1年間のすべての診療内訳と処方内訳が含まれる。ただし、算出(抽出)基準は

今後変更されることがあることが、明記されている。

韓国では、HIRA-NPSの提供開始まで、以下の経過をたどっている。2009年度、“診療情報の標本資料提供案に関する研究”を実施し、保健医療関連学会との諮問会議開催を3回行った。その後、HIRAと5つの学会とのMOU締結(2010.12.21)している。さらに、HIRA-NPSの妥当度の検証のためのMOU学会会員対象の資料説明会開催(2011.02.09)、第23回審評フォーラム“健康保険審査評価院患者標本資料を利用した学術シンポジウム”を開催して17の研究結果発表(2011.10.06)している。さらに、個人情報及び手数料算定に対する法規部検討を行った後、HIRA-NPSの提供手続きが開始(2012.04.30)された。

C. 研究結果

(初年度)韓国では、レセプトデータベースの名称をDW(データウェアハウス)と表現しており、特に決まった名称はない。管理・運営は、日本における審査支払機関に該当するHIRA(The Health Insurance Review and Assessment Service)が担当している。提供データは、外来・入院のレセプトをランダムサンプリングしたデータセットである。また、小児と高齢者のレセプトのみのランダム抽出データも提供している。この他に申し出者の要望に応じてDWから独自にデータ抽出を行い、データ提供を行うサービスも有する。データIDは、韓国の国民背番号を利用しており、抽出は1年単位である。データの提供時

期は、2012年からである。データ提供件数は、2012年は69件、2013年は115件であり、DVD媒体で提供される（DVD一枚当たり30万ウォン）。民間に対しては、提供していない。ただし、HIRAの統計分析チームが公共部門とみなしてマスクミ・患者団体等の限られた組織に提供している。個人情報保護に関する取り扱いに関しては、レセプトは個人情報には該当しないデータとして扱われているが、個人を特定できる可能性のある情報を含んでいるために、実際の運用においては（個人情報に準ずる）「行政機関が有する個人情報の保護に関する法律」に準拠している。漏洩等の不適切利用に対する対応およびその根拠としては、申請者が利用誓約書の事項に違反した場合はあらゆる不利益および民事刑事上の処罰を受け、その後の利用制限が課せられる。

加えて、HIRAではレセプトデータを施設内で、研究者・行政関係者が利用できる（民間は利用できない）オンサイトセンタが開設されている。入退室管理など、ISMSの理念に則ったセキュリティが課され、1部屋に約15台のPCがセットされている。オンサイトセンタのPCは、外部にデータウェアハウスと回線がつながっており、利用者は分析できるものの、結果を持ち帰ることはできない。得られた結果は、HIRAが検閲を行い、許可された情報のみが後日、郵送あるいはeメール等で利用者に送られる仕組みになっている。

また、韓国の保険公団（保険者）は、HIRAとは別に研究者に対するレセプトデータ提供機能の検討を始めている。提供するデータは、2002年からのレセプトデ

ータを基に100万人をサンプリングしたコホートデータの試行版を作成した。2014年から、特定の研究者に試行提供が行われる予定になっている。2013年12月には、本データセットの利活用の第1回ワークショップが開催された。

台湾では、1995年に国民皆保健制度が導入され、診療報酬データ（レセプトデータ）等を、国民医療保険局（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。データベース名称は、National Health Insurance Research Database（全民健康保険研究資料庫）である。研究者は、1996年からの全レセプトデータを利用できる。提供データは、5種類である。（Ⅰ）基本資料データ（基本資料檔）であり、これは集計データである。（Ⅱ）系統抽出データ（系統抽樣檔）であり、個票（個人単位）データである。入院レセプトファイル（DD）は5%抽出、外来レセプトファイル（CD）は0.2%抽出となっている。（Ⅲ）特定主題データ（特定主題分檔）は、個票（個人単位）データであり、特定疾患や調査目的に沿って抽出・作成したデータセットであり、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）等の種類がある。（Ⅳ）ランダム抽出データ（抽樣歸人檔）は、個票（個人単位）データであり、被保険者ファイル（ID）から、ランダムサンプリングして、2000年データとして、20万人を抽出し、抽出された被保険者の（1996年～2007年の）入院、外来、調剤レセプトデータを提供している。加えて2005年

単位で 100 万人、2010 年単位で 100 万人のランダム抽出データが作成されたことから、現時点では 3 種類のランダム抽出データが存在する。(V) 教育用データ(教學用資料檔)は、個票(個人単位)データであり、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、教育用に無償提供している。2013 年は、上記のデータセットが研究者へ提供された件数は 270 件である。一方、台湾は、民間への提供は、禁止している(過去に製薬会社へ、治験業務サポート等の研究目的として試行的に提供していた)。個人情報保護に関する取り扱いに関しては、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。漏洩等の不適切利用に対する対応およびその根拠としては、個人情報コンピュータ処理保護法に罰則・罰金規定が定められている。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している。

オンサイトセンタとしては、米国 AHRQ のオンサイトセンタを調査した。米国では医療費に関する調査として、AHRQ が実施している MEPS(Medical Expenditure Panel Survey)がある。MEPS は家計を対象として医療や保健サービスへのアクセス、医療関連の支払い、健康状態、医療保険の加入状況等を調査しているもので、同一対象者に対して 2 年ずつ調査をし、毎年新しい対象者を加える形式をとっている。1996 年までは National Medical Expenditure Survey として 10

年おきに断面調査が実施されていたが、パネルデータの重要性から 1996 年以降は新しい対象者を加え、毎年調査する形式となっている。

AHRQ では、MEPS データセットを研究者に対して提供している。加えて、研究計画が AHRQ によって承認されれば、一般公開されていない機密データ項目がオンサイトセンタを利用することで可能となる。また、AHRQ では、研究者の所属・組織によりデータの提供の可否を決めていない。研究プロポーザルの内容が公益性を満たすかにより提供の可否が決まる。例えば、ランド研究所などの民間の研究組織や非営利団体にもデータを提供しており、博士論文(Dissertation)を執筆中の学生も利用できる。ただし、オンサイトセンタでは、氏名・ID 等の個人の特定コードは提供されない(他の数値に置き換えた形となる)。地域を特定する情報も提供されない。

オンサイトセンタでは、ユーザーの申請書に基づいて AHRQ があらかじめデータセットを作成し、オンサイトセンタ内の PC にコピーされる。ユーザーはオンサイトセンタにて、そのデータセットを利用して分析を行う。データセットは、オンサイトセンタでしか利用できない。利用のためには、利用者は、初回は必ず AHRQ に来なければならない。その後、更新プログラム(SAS 等)を AHRQ に送付すれば、AHRQ 内のデータコーディネーター(常勤 1 名)がプログラムを代わりに実行する(run)ことも可能である。分析結果は AHRQ 内で審査を受けた後、ユーザーにメールで送られる。本方式によって、ユーザーは継続的に

分析を続けることができる。オンサイトセンタの運営コストは、フルタイムスタッフ2名、場所代、光熱費などで約300,000\$ / 年ほどだと推計される。運営コストの正確な公表資料はなく、収入は、約100万ドル/年の政府助成金、ユーザーから300ドル/人を徴収している。オンサイトセンタは、2002年に設立・サービス提供が開始したが、提供件数などの統計は取られていない。オンライセンタの申請件数は、約30件/年であり、約90%が認可される。研究によっては、論文の査読者からのコメントによる再度の分析なども許可しているので、単年度では終わらない。常時、50件の研究PJが進行している状態である。

オンサイトセンタは、セキュリティ入室管理をしている。また、利用者はデータセンタには、メモやマニュアルのみ持ち込が可能で（誓約書で規定）、USBメモリー、携帯電話、スマートフォン、タブレット端末等あらゆるIT機器の持ち込みは禁止されている。ただし、ビデオでの監視はしていない。

（次年度）HIRAが行ったHIRA-NPSの検証結果は、入院病名を高頻度順の100疾患においてHIRA-NPSと母集団の割合が同一であり、HIRA-NPSと母集団の月別患者一人当たりの平均医療費および分散がほぼ一致していたことを報告している。分散分析（一元分散分析）において、95%信頼区間でHIRA-NPSと母集団の平均と分散の集団間の差がなかった。病名では、脳卒中および妊娠中毒症の患者の診療行為について観察している。両方の疾患共に、母集団とHIRA-NPSでは、同様の傾

向が観察されたことを報告している。

D. 考察

（初年度）日本と同様の皆保険制度であり診療報酬点数制度も導入されている韓国と台湾では、研究者への年間のデータ提供件数が100件を超えている。日本は、特別抽出の形態で、依頼に応じてデータセットを抽出・提供しているが、近年はサンプリングデータセットとして、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している。今後、さらに提供件数を増やすためにも、韓国の小児や高齢者のみのデータセットの提供、さらに進んでいる台湾のような5種類のデータセットを準備しておくことも有効な手段であると考えられる。また、韓国HIRAでは基本的に1年単位のデータセットを提供しているが、1年単位の資料では、分析が不可能な希少疾患など長期間のフォローアップの必要性が求められていた。そこで、個人情報保護される範囲内で複数年の標本追跡コホートデータセットの作成も始まっている。

韓国では（現時点において、研究班が得ている情報が限定的なために今度も情報収集を継続するが）、研究者専用統計分析のためのオンサイトセンタ（資料処理室15席）が用意されていることに加え、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提供する計画もある。

米国AHRQでは、医療や保健サービスへのアクセス・医療関連の支払いなどのデータを含むMEPSデータをオンサイトセ

ンタで提供している。オンサイトでは、一般公開されていないデータ項目も利用が可能（全てのデータ項目ではなく、審査承認も必要である）というメリットもあると同時に、オンサイトセンタ内のデータコーディネーターに更新プログラムを送付することで、データコーディネーターが研究者の代わりに実行するサービスもあることが特徴的である。韓国 HIRA においては、研究者が直接来訪して利用しなければならないというアクセス性および便宜性に関する指摘もされている。米国と韓国の知見は、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。

（次年度）これまで、韓国 HIRA のレセプトデータに関する提供サービスは、申し出者の要望に応じて独自にデータセットの提供を行うサービス、外来・入院のレセプトをランダムサンプリングしたデータセット、小児と高齢者のレセプトのみのランダム抽出データを提供していた。これらのデータ提供機能に加えて、HIRA-NPS が加わったことになる。一方で、HIRA-NPS は抽出データなので、抽出頻度が少ない希少疾患の分析が困難であること、期間が 1 年間のため、有病期間が長い慢性疾患などの分析にも適していないことが指摘されている。

現在、日本のレセプト情報等データベース（以下、NDB）から提供されるデータセットは、厚生労働省側で複雑な構造の電子レセプトを研究者の要望に応じて、ある程度分析し易い形式に加工して提供している。そのために、汎用性のあるデータセットを事前作成しておくことで、提供件数

の増加につながると共に負担軽減にもつながると考えられ、レセプト情報等データベースの利用促進のために、汎用性の高い基本データセットの設計と作成が検討され、今年度から試行提供が始まった。したがって、韓国の HIRA-NPS は、日本の基本データセットを考える上で有益な事例であると考えられた。

最後に、HIRA ではレセプトデータをソウル本院で、研究者・行政関係者が利用できるオンサイトセンタが開設されている。加えて、民間企業等に対しても、限定された項目のみであるがオンサイトセンタが利用可能であることが判明した。民間企業に対するデータは、公共データとよばれている。ただし、製薬会社（企業）が利用する場合は、自社製品のみを参照できる。他社製品の情報を閲覧したい場合には、該当する会社（他社）から医薬品情報に関する情報提供同意書が必要となる。

E. 結論

（初年度）米国、韓国、台湾のデータ提供の仕組みを調査し、日本への適用可能性を検討した。韓国に関しては、HIRA に設置されているオンサイトセンタの運用、台湾の 100 万人規模のコホートデータの作成参考になると考えられる。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。（次年度）韓国の HIRA-NPS は、日本の基本データセットを考える上で有益な事例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

・Naohiro Mitsutake, The use of the national healthcare database for

research in Japan, 25th European Medical Informatics Conference, Aug 31st - Sep 3rd, 2014

G. 知的所有権の取得状況

該当なし

参考資料

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CD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な提供	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	層化抽出: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(系統抽標檔):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(中藥)ファイル等 (IV) ランダム抽出データ(抽標歸人權):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。	

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
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。
匿名化方法（一般）	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。 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。		個人情報保護及びデータ安全保護観点から、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。
匿名化方法（加工したデータセット）	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。	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450万5千人から140万人を抽出）。	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。
利用条件・手続き	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。 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。	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 申込者は、学術研究用以外の目的で使用することを禁止 販売および貸出を行わない誓約書を提出	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
提供形態	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。 目的外利用： 申出者から、データセンターに対し厚生労働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を書留もしくは手渡しにて送付する。	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。	CDおよびDVD 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
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	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生労働省に提供される。 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生労働省に提供される。		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	http://www.mhlw.go.jp/seisakunitsuite/bunya/kenkou_iryoku/iryuhoken/reseputo/		http://nhird.nhri.org.tw/
アーカイブスタッフ	厚生労働省 6名（他業務と兼務） ベンダー（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ン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。

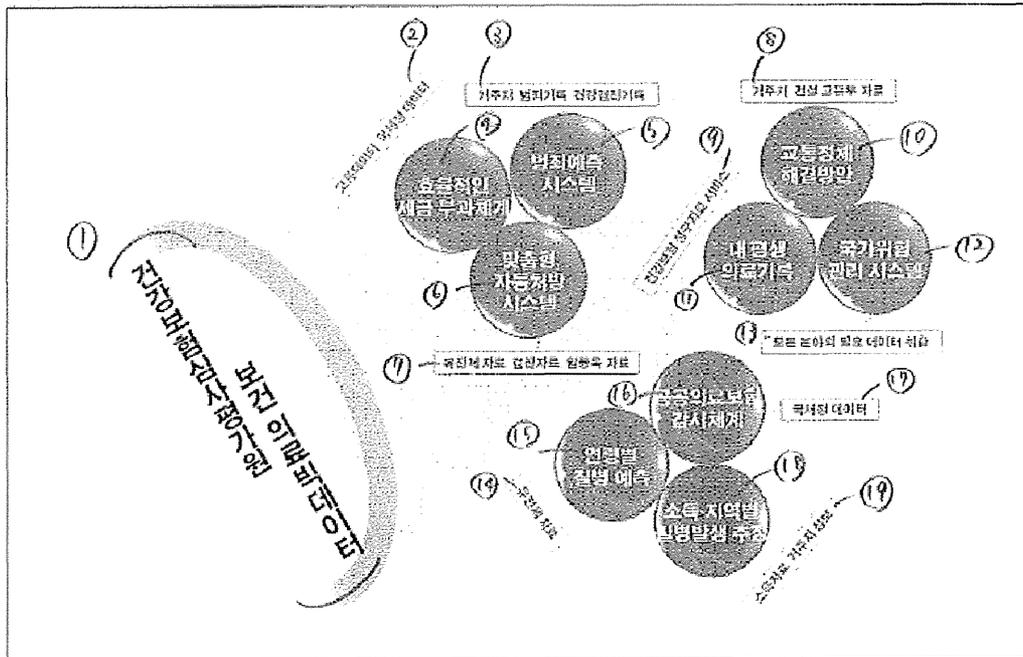
2. 目的

- 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 > 빅데이터 플랫폼構築による多様な価値の創出

< 그림 > 빅데이터 플랫폼 구축을 통한 다양한 가치 창출



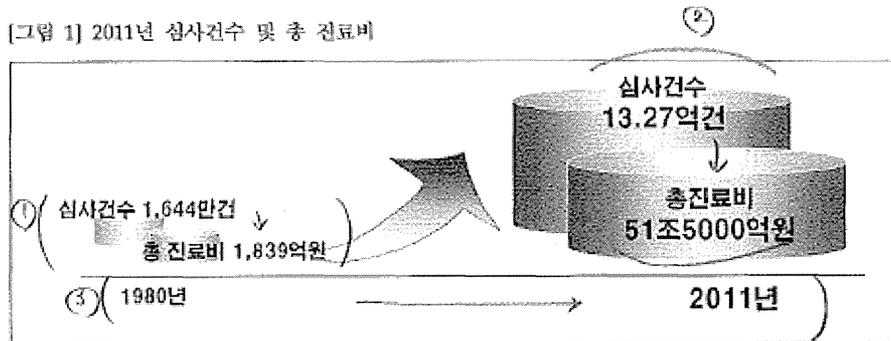
- ① HIRA 保健医療ビッグデータ
- ② 雇用データ、国税庁データ
- ③ 居住地、犯罪記録、健康検診記録
- ④ 効率的な課税システム
- ⑤ 犯罪予測システム
- ⑥ パーソナライズ型自動処方システム
- ⑦ ゲノム資料、検診資料、がん登録資料
- ⑧ 居住地、建設交通部資料
- ⑨ 健康保険請求資料サービス
- ⑩ 交通渋滞解決方
- ⑪ 生涯カルテ
- ⑫ 国家危険管理システム
- ⑬ 全分野における必要データの統合
- ⑭ ゲノム資料
- ⑮ 年齢別疾病予測
- ⑯ 公共医療保険監視システム
- ⑰ 国税庁データ
- ⑱ 所得・地域別疾病発生の推定
- ⑲ 所得資料、居住地情報

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



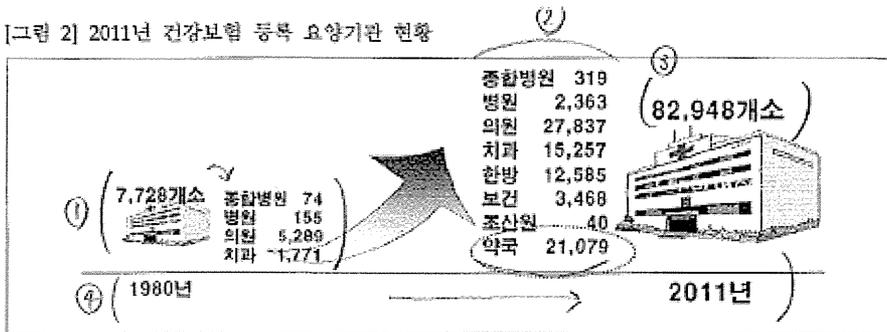
○ 건강보험에 등록된 요양기관 현황은 1980년대 7,728개소에서 2011년 기준으로 82,948개소로 증가

- 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

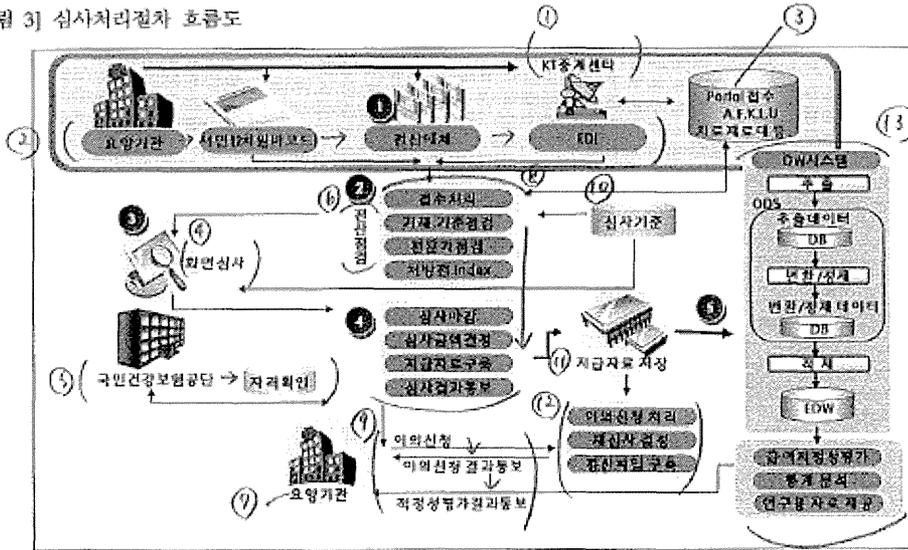
⑤ 1980 年、2011 年

○ 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

- － 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 中繼センター
- 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。

- 一 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
 - 一 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。
 - 一 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、年齢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、排泄機能、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	
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	韓国雇用情報院
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
統計庁死亡原因資料	全国(死亡者対象)	死亡時発生	
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