

令和 2 年度厚生労働科学研究費補助金
(政策科学総合研究事業 (統計情報総合研究事業))
分担研究報告書

韓国の公衆衛生分野における公的データリンケージの状況

研究協力者 Boyoung Jeon 筑波大学医学医療系 客員研究員
韓国明知専門大学保健医療情報科 助教
研究代表者 伊藤 智子 筑波大学医学医療系 助教
研究分担者 田宮 菜奈子 筑波大学医学医療系 教授

研究要旨

本研究では、韓国の公的機関において行われている各種データベース事業について、公開されている資料をレビューした。その結果、韓国では国民健康保険公団の「国民健康情報 DB」基盤のビッグデータ、4つの公共機関（疾病管理庁、国民健康保険公団、健康保険審査評価院、国立がんセンター）が協力して作った「保健医療ビッグデータプラットフォーム」を通じて、研究者たちが公益目的で活用できるデータが提供されていることがわかった。一方で、データの更新やアクセシビリティにおいて、解決すべき課題があり、今後の動向を注視していく必要があると考えられた。

*本務先 : Department of Health and Medical Information, Myongji College, Korea

A. 研究目的

韓国の公衆衛生分野で、代表的な公的データリンケージの事例は、国民健康保険公団の「国民健康情報 DB」である。国民健康保険公団では、2011 年「国民健康情報 DB」を構築して以来、現在まで様々な分野（医療、健康診断、高齢者、働く女性及び子ども、カスタマイズ型 DB(Customized Research DB)など）に至る、さまざまなビッグデータを研究者に提供している。さらに 2019 年からは、国民健康保険公団だけでなく、様々な機関との連携を通じた公共目的の研究資料を構築した。これを「保健医療ビッグデータプラットフォーム」としており、一部の研究者にデータを開放している。そこで、本報告書では、国民健康保険公団の「国民健康情報 DB」基盤のビッグデータ運営の事例と「保健医療ビッグデータプラットフォー

ム」モデル事業について紹介する。

B. 研究方法

本研究では、韓国の公的機関において行われている各種データベース事業について、公開されている資料をレビューした。

(倫理面への配慮)

本研究で用いるデータは、個人情報を含まない。

C. 研究結果

1. 韓国の健康保険制度の特性

韓国の国民健康保険制度は、全国民が義務的に加入する強制力を持っており、すべての医療提供者は、当然指定制度を通じて健康保険指定医療機関となっている。このすべての医療機関は、患者を診療した後、診療費（自己負担額は除く）を国民健康保険公団に請求している。また、主な診療費

報酬支払制度として行為別の個別支払制度（Fee-for-Service）を運営している。

国民健康保険公団は、単一の保険者であって、加入者（国民全体）の資格と保険料を管理する。また、医療機関、検診機関、長期療養機関等の提供者に診療、検診、療養サービスに対する診療費を支給する。これにより、国民健康保険公団は、全国民5000万人以上の出生から死亡までの資格と保険料資料、病院・医院の利用内訳、健康診断の結果、加入者のがんの登録情報、医療給与（低所得層対象の公共扶助制度）資料、老人長期療養保険資料などの膨大なデータを保有している。

上記のように「単一保険者基盤」の「全国民の義務加入」という特性と「行為別の個別支払制度（Fee-for-Service）」を通じた診療費支払制度は、健康保健ビッグデータの側面において、かなりの強みとなっている。「全国民」の医療利用内訳が、国民健康保険公団の「単一DB」に「行為別個別支払制度」を通じて、医療行為、治療材料、薬剤などの様々な形で蓄積されているからである（パク・ジョンホン、2016）。

2. 国民健康保険公団の「国民健康情報DB」

1) 国民健康情報DBの紹介

人口の高齢化につれて健康保険分野の支出が増加することで、持続可能な健康保険制度のためのエビデンスに基づいた政策形成の必要性が増大し始めた。よって、国民全体の健康保険データを活用し、時系列的レベルで解析することができる健康保険DBの構築と活用が求められた。公益目的の保健研究に対する社会的な必要性を反映して、2011年に初めて、国民健康情報DBが構築された。2012年の標本コホートDBが構築されて以来、さまざまな種類のデータが構築及び公開されて研究者に活用されている。「国民健康情報DB」の構

築プロセスと沿革を図1、表1に示す。

国民健康保険公団のデータの中に「標本研究DB(Sample Cohort DB)」は、国民健康保険ビッグデータに基づいて、需要が高いデータをサンプリングして、情報主体を見つけることができないように、非識別措置（加工）した後、政策及び学術研究用に提供するために規格化したデータセットを意味する。現在、標本コホートDB、健康診断コホートDB、高齢者コホートDB、子ども健診コホートDBは、働く女性コホートDBなど5種のデータセットがあり、2002年から2015年までの資料を提供している。「標本研究DB」は、リモート研究分析システムにアクセスして、時間と空間の制約なしに、データ分析が可能である。「カスタマイズ型DB(Customized Research DB)」は、申請者の研究目的に合わせて、国民健康保険公団からデータを抽出・要約・加工して、情報主体を見つけることができないように措置したデータセットを提供することである。「カスタマイズ型DB」はデータ活用が限られている。外部搬出が不可能であり、国民健康保険公団本部または地域本部内に設置されたビッグデータ分析センター内でのみ資料閲覧、分析を実行した後、結果物である統計表だけを搬出できる（表2）。

2) 健康保険ビッグデータ研究支援のための資料提供状況

国民健康保険公団は、健康保険、長期療養保険、保険料徴収業務などを行っている。これにより、健康保険の資格と保険料徴収資料、患者の医療利用情報などの診療内訳、健康行動、健康診断情報など、約3兆9,000億件の全国民に対する大容量の情報を蓄積し管理している（青年医師、2020）。保健医療分野の研究活性化のために、国民健康保険公団の国民健康情報DBを活用した公益目的の研究支援資料提供件数は、

2014年の77件から835件が増えた2019年の912件で、12倍増加した。これと関連し、国民健康保険公団は「健康保険ビッグデータの構築及び提供システムの運営方法」の特許を取得した（青年医師、2020）（表3）。

韓国の国民健康保険公団が提供した国民健康情報DBのうち、標本研究のDBを活用した研究リストを「Pubmed」で確認してみた。「Pubmed」でTitle/Abstractに「Nationwide Population-Based Cohort Study Korea」が入った論文の件数は図2の通りである。

3) 国民健康情報DB資料連携根拠

データ間の連携時、マッチングキーである住民登録番号（韓国人固有のID、日本のマイナンバーに該当する）は、個人情報に該当するため、資料連携のために、データ間の連携キー（key）に該当する「個人シリアル番号ID」に「転換」して提供することになっている。データの公開レベルに応じて、国民健康情報DBは、次のように分類される（表4）。

3. 保健医療ビッグデータプラットフォーム（4つの機関のデータリンケージ事業）

1) 保健医療ビッグデータモデル事業の紹介

「保健医療ビッグデータモデル事業」とは、保健医療分野の4つの機関のデータを個人単位で連携、公共的な目的の研究に活用できるように研究者に開放する事業である。この4つのデータの提供機関は、疾病管理庁、国民健康保険公団、健康保険審査評価院、国立がんセンターに該当する。この事業の目的は、主要な保健医療公共機関に分散しているビッグデータを共通の継ぎ目（Key）を基盤に連携した後、研究者に閲覧できるようにするものである。ここでKey（例えば、氏名+生年月日+性別）

は、二つのデータを相互に繋げるときに使用する情報を意味する。この事業を通じて、4つの公共機関に分散された情報を個人単位で連携できるにつれ、分散されたデータではできなかった研究を可能にすることが目的である。「保健医療ビッグデータモデル事業」の運営プロセスと沿革は、図3、表5の通りである。

2) 保健医療ビッグデータプラットフォームの資料提供状況

保健医療ビッグデータプラットフォームでは、データカタログを提供する。データカタログは、各機関が提供するデータのデータセット名、データセットの韓国語名、変数の説明及びコード値に関する内容である。研究者は、カタログを通じて研究に必要なデータを確認し、利用申請をする。このデータを利用できる研究者は、中央行政機関及び地方自治団体、公共機関及び地方公共機関、国内の医療機関、学界、研究機関所属の研究者である。データ提供機関別データを紹介したカタログとデータの利用手順は表6、図4の通りである。

3) 保健医療ビッグデータプラットフォーム資料の連携方式

保健医療ビッグデータプラットフォームでは、個人単位の情報連携を行う際に、個人情報の非識別化のために様々な措置を取っている。まず、個人単位の連携のための情報（氏名、生年月日、性別など）と第3の信頼機関（TTP）が発給したランダムなキー値を結合して使用する。ここで、第3の信頼機関（Trusted Third Party）とは、各データ提供機関と分離された第3の機関として、安全な連携キー発給を目的とする機関を意味する。保健医療ビッグデータプラットフォームでは、保健福祉部がTTP役割を遂行する。安全な個人識別キーの管理のために一方向暗号化関数を使用

する。

保健医療ビッグデータプラットフォームで提供するデータと外部機関との連携は不可能である。例えば、民間病院資料などの個人情報が含まれたデータの連携、統計庁の死亡資料などと連携が不可能である。ただし、個人情報ではないデータ（例えば、気象情報、環境資料など）は、分析閉域網に搬入することが可能である。次に、連携された資料を研究者が閲覧したときに、外部にデータを搬出せず、インターネットから断絶された閉鎖環境でのみ閲覧する。閉鎖環境は、国民健康保険公団、健康保険審査評価院が運営する施設を使用する。研究者は、希望する閉域網解析センターを指定して訪問することができ、分析の完了後、個人情報が含まれていない研究結果だけを搬出することができる。

D. 考察

韓国の公衆衛生分野で公的データのリンケージの状況を調べた。国民健康保険公団の「国民健康情報 DB」基盤のビッグデータ、4 つの公共機関（疾病管理庁、国民健康保険公団、健康保険審査評価院、国立がんセンター）が協力して作った「保健医療ビッグデータプラットフォーム」を通じて、研究者たちが公益目的で活用できるデータを提供するために努力していることを確認した。

国民健康保険公団の国民健康情報 DB ビッグデータの場合、研究支援資料の提供がかなり速い速度で増加し、研究者のアクセシビリティが向上している。しかし、国民健康保険公団の「標本研究 DB」の場合、2002 年から 2015 年までの資料を提供しており、最新のデータに対する更新が必要とされているのが実情である。「カスタマイズ型 DB」は、国民健康保険公団で申請者の研究目的に合わせてデータを抽出できるという利点があるが、国民健康保険公団

によって定められたビッグデータ分析センター内でのみ分析が可能であり、データの抽出と分析にかなりの時間がかかる。

保健医療ビッグデータプラットフォームの場合、4 つの機関（疾病管理庁、国民健康保険公団、健康保険審査評価院、国立がんセンター）が機関間の連携を通じてデータの質を向上させようとしたモデル事業である。多機関データ間の連携の際に、個人情報非識別化措置のためにかかりの力を注いだという特性がある。しかし、4 つの機関で提供することにした変数およびデータのほかに、統計庁の死亡データおよびその他の個人情報が要求される資料は、連携されていない。また、データ分析のための閉域網分析センターに訪問しなければならず、研究結果の搬出条件が厳しくて、多くの研究者が利用しやすい環境ではない。まだモデル事業の段階として、今後資料活用を通じた研究の活性化の程度は、時間を置いて見守るべきである。

E. 結論

韓国では国民健康保険公団の「国民健康情報 DB」基盤のビッグデータ、4 つの公共機関（疾病管理庁、国民健康保険公団、健康保険審査評価院、国立がんセンター）が協力して作った「保健医療ビッグデータプラットフォーム」を通じて、研究者たちが公益目的で活用できるデータが提供されていることがわかった。一方で、データの更新やアクセシビリティにおいて、解決すべき課題があり、今後の動向を注視していく必要がある。

F. 研究発表

1. 論文発表 なし
2. 学会発表 なし

G. 知的財産権の出願・登録状況（予定を含む）

1. 特許取得 なし
2. 実用新案登録 なし
3. その他 なし

参考文献

国民健康保険公団。健康保険共有サービス
(<https://nhiss.nhis.or.kr>)

国民健康保険公団のビッグデータ運営室。
「健康保険ビッグデータ開放現状と展望」
韓国保健産業振興院。保健医療ビッグデータ
プラットフォーム

(<https://hcdl.mohw.go.kr>)

パク・ジョンホン。健康保険公団ビッグデータ
活用方法。大韓肝学会臨床研究方法論
ワークショップ。2016

青年医師。

2020.07.21([https://www.docdocdoc.co.kr/
news/articleView.html?idxno=2001038](https://www.docdocdoc.co.kr/news/articleView.html?idxno=2001038))

Pubmed

(<https://pubmed.ncbi.nlm.nih.gov>)

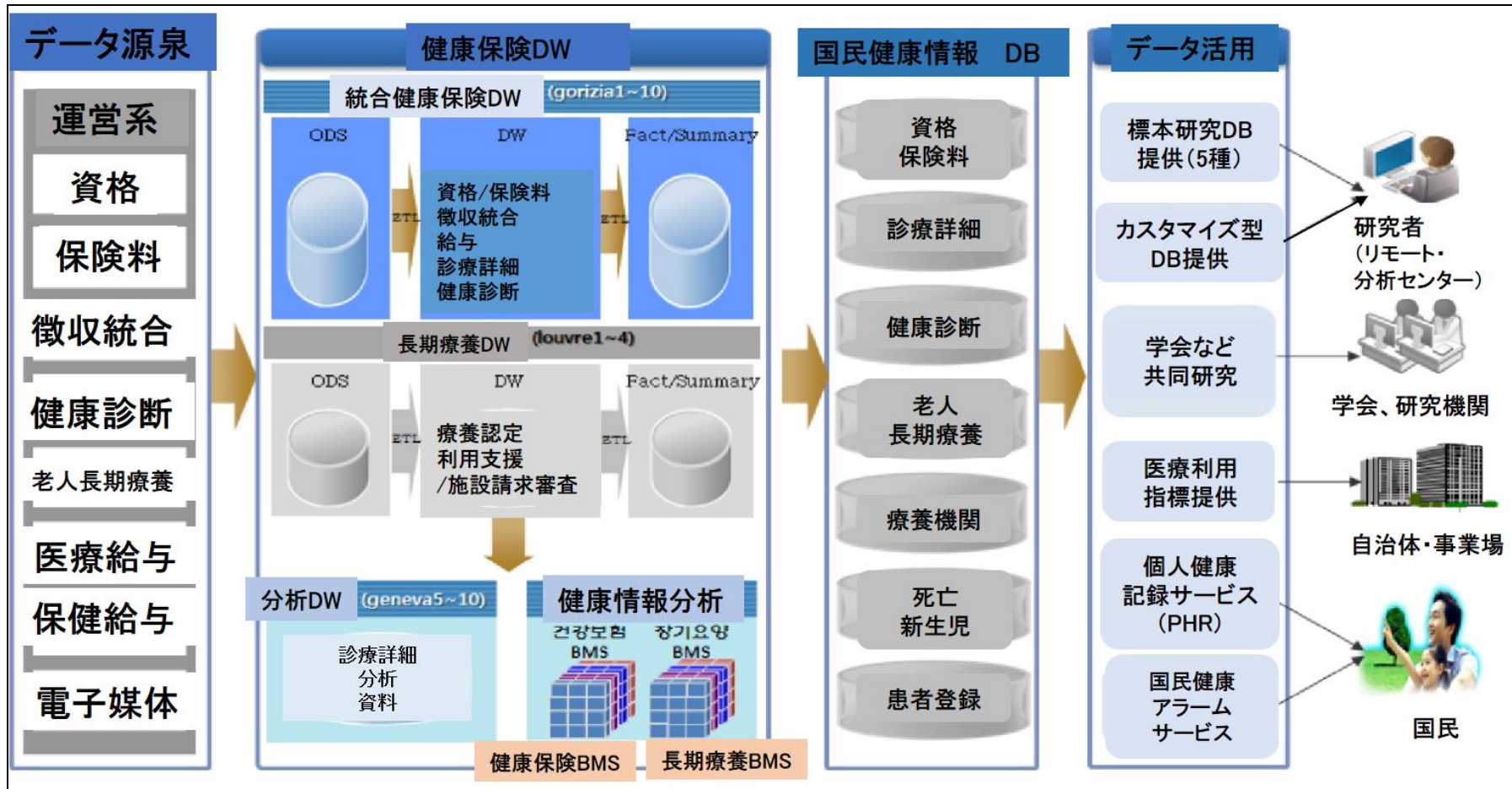


図 1. 国民健康情報 DB の構築プロセス (出典：国民健康保険公団のビッグデータ運営室)

表 1. 国民健康情報 DB 構築及びビッグデータを提供の沿革

年度	主な出来事
2011	・ 国民健康情報 DB 構築
2012	・ 標本コホート DB 構築
2013	・ 標本コホート DB モデル研究事業 (16 機関)
2014	・ 国民健康情報資料提供運営規定の制定 ・ 標本コホート DB 及び「カスタマイズ型 DB」提供
2015	・ 標本研究 DB 2 種 (健康診断および高齢者コホート DB) モデル研究事業を推進 (韓国疫学会、大韓老人病学会) ・ 標本研究 DB 2 種 (健康診断および高齢者コホート DB) を追加提供 ・ 公共データ開放サービスの Web ページオープン及び国家重点開放データ (3 種) を公開
2016	・ 健康保険ビッグデータ分析センターをオープン ・ 健康保険ビッグデータ分析センターを 8 カ所に拡大 ・ 標本研究 DB (働く女性および子ども健診コホート DB) モデル研究事業を推進 ・ 健康保険ビッグデータのプラットフォームの構築
2017	・ 標本研究 DB 2 種 (働く女性および子ども健診コホート DB) を追加提供 ・ 新規標本コホート DB の変更提供 ・ 健康保険ビッグデータのリモート研究分析システムをオープン
2019	・ 国民健康保険ビッグデータのカスタマイズ型分析センターシステムをオープン ・ 国民健康保険ビッグデータカスタマイズ型研究 DB 提供 ・ 国民健康保険ビッグデータ医療利用マップサービスを提供 ・ 国民健康保険健康情報リサーチ協力センターをオープン
2020	健康保険ビッグデータ糖尿病研究 DB を追加提供

(出典：国民健康保険公団。健康保険共有サービス)

表 2. 国民健康保険公団の「国民健康情報 DB」の概要

分類	データの種類	データの特性
標本研究 DB	標本コホート	[基準] 2006 年 1 年間の健康保険加入者及び医療給与受給権者の資格を維持した全国民 [対象者] 100 万人 [年度] 2002～2015 年 [サンプリング] 全国国民母集団の 2%、性別・年齢・加入者区分・保険料分位・地域別層別抽出 [内容] 社会・経済的状況（資格及び保険料、障害及び死亡）、医療利用状況（診療及び健康診断）、療養関状況
	健康診断のコホート DB	[基準] 2002 年の資格維持者のうち、2002～2003 年に、40～79 歳の一般健康診断受検者 [対象者] 約 51 万人 [年度] 2002～2015 年 [内容] 社会・経済的資格変数（障害及び死亡を含む）、医療利用（診療及び健康診断）状況、療養機関状況
	高齢者のコホート DB	[基準] 2002 年の資格維持者のうち 60 歳以上の対象者 [対象者] 約 55 万人 [年度] 2002～2015 年 [内容] 社会・経済的資格変数（障害及び死亡を含む）、医療利用（診療及び健康診断）状況、療養機関現状、老人長期療養サービスの現状
	働く女性のコホート DB	[基準] 2007 年の資格維持者のうち、15～64 歳（生産可能人口）働く女性 [対象者] 約 18 万人 [年度] 2007～2015 年 [内容] 社会・経済的資格変数（障害及び事業場を含む）、医療利用（診療及び健康診断）状況、療養機関の状況
	子ども健診のコホート DB	[基準] 子ども検診第 1～2 次を 1 回以上受けた全受検者のうち、2008～2012 年生まれを抽出し、各出生年度別 5%サンプリング [年度] 2008～2015 年（8 カ年） [内容] 社会・経済的資格変数（障害及び死亡）、医療利用（診療及び健康診断）状況、療養機関の状況

(出典: 国民健康保険公団。健康保険共有サービス)

表 2. 国民健康保険公団の「国民健康情報 DB」の概要（つづき）

カスタマイズ型研究 DB	カスタマイズ型研究 DB	<p>[内容]国民健康保険公団が収集・保有・管理する健康保険と長期療養保険の資料を政策及び学術研究目的のために利用できるように、申請者の研究目的に応じて抽出、要約、加工して、情報主体を見つめることができないように措置したデータセット</p> <p>[データの管理]カスタマイズ型研究 DB を閲覧および研究分析できる PC が設置された国民健康保険公団内の場所であるビッグデータ分析センターで統計分析を実行する必要がある</p>
--------------	--------------	--

（出典：国民健康保険公団。健康保険共有サービス）

表 3. 国民健康情報 DB を通じた研究支援資料の提供状況

(単位：件)

区分	合計	2014 年	2015 年	2016 年	2017 年	2018 年	2019 年
計	3,553	77	201	464	729	802	912
カスタ マイズ型 研究 DB	2065	8	84	186	391	512	623
標本研究 DB	1488	69	117	278	338	290	289

(情報源：国民健康保険公団、出典：青年医師（2020.07.21）)

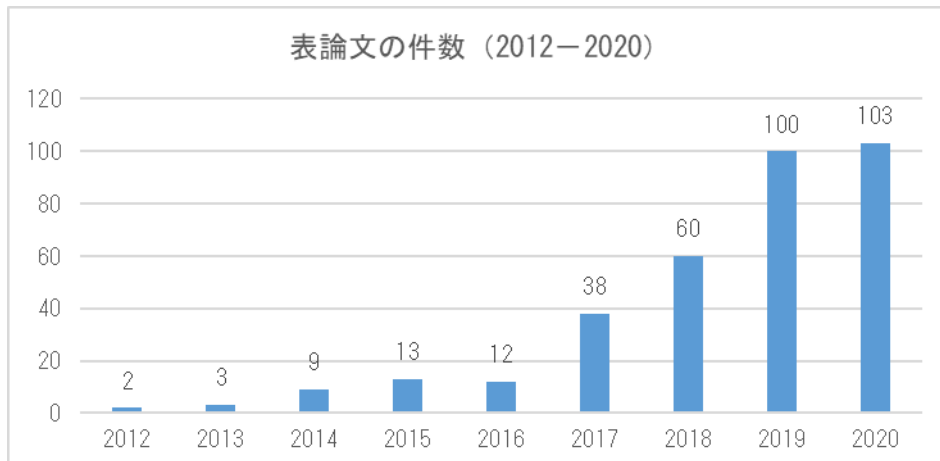


図 2. 「Pubmed」で「Nationwide Population-Based Cohort Study Korea」が入っている論文の件数

(出典: Pubmed)

表 4. データの公開レベルに応じた国民健康情報 DB の分類

データの公開レベル	データの定義	情報源	事例	データ提供の範囲
源泉 DB	別途の加工を施していない資料	国民健康保険公団	「健康保険資格詳細」など ODS(Operational Data Store), DW(data warehouse) 資料	法的根拠に基づいて制限的に提供
制限公開 DB	研究用に新規構築した資料として公開の危険性が高く、まだ検証が進行中の資料	国民健康保険公団及び国家機関	国籍、行政区域、入院のエピソード、医療人など	国家機関の公益目的（政策研究など）でのみ使用可能
公開 DB	研究用に新規構築した資料として公開の危険性が低く、十分な検証が行われた資料	国民健康保険公団	健康保険の資格及び保険料、再構築された診療及び療養機関資料等	公団の資料提供審議委員会の承認に基づいて提供

(出典：国民健康保険公団のビッグデータ運営室)

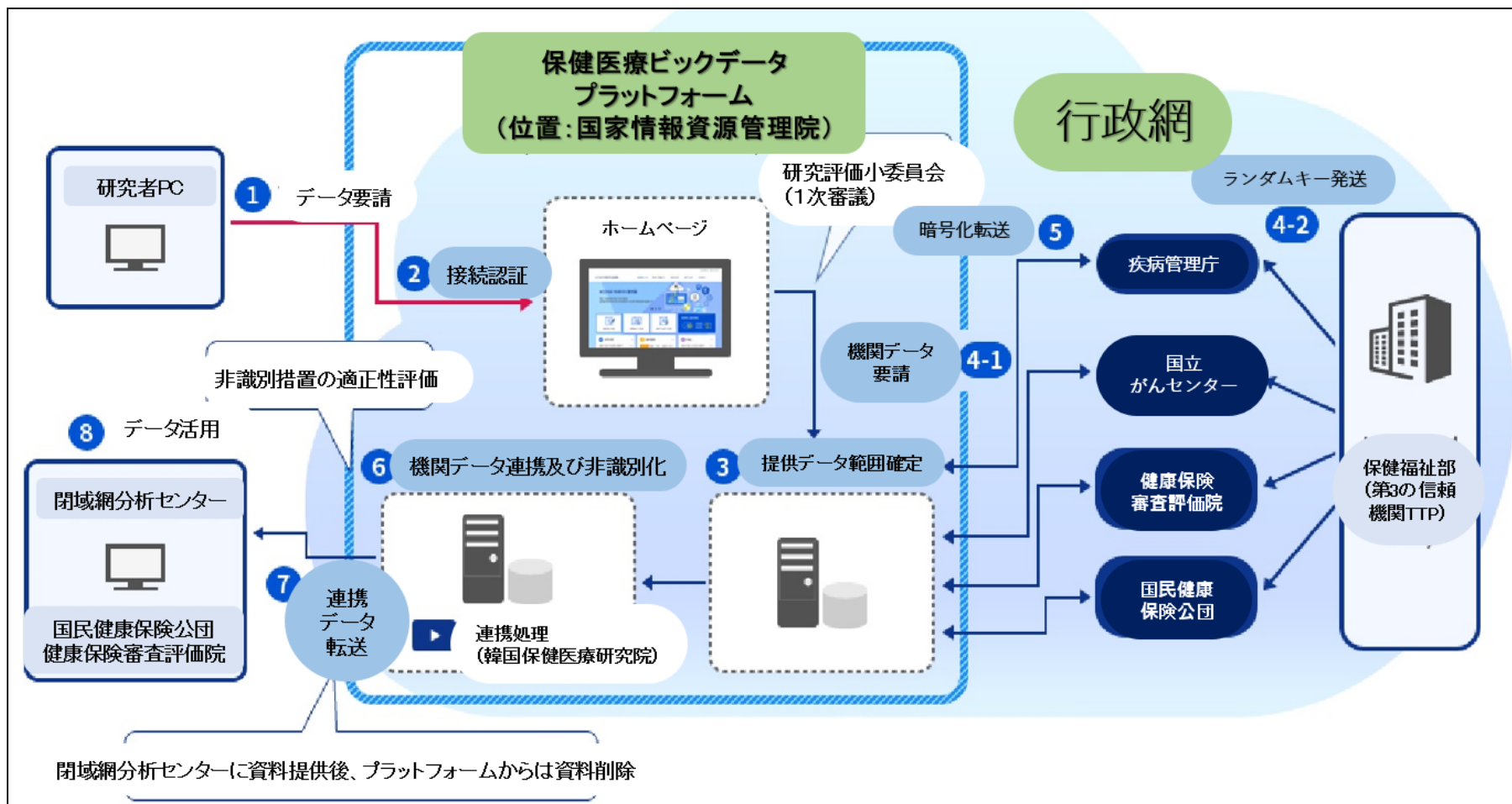


図 3. 保健医療ビッグデータプラットフォーム運営プロセス

(出典: 韓国保健産業振興院。保健医療ビッグデータプラットフォーム)

表 5. 保健医療ビッグデータモデル事業沿革

年度	主な出来事
2015.01- 2016.12 (政策準備の段階)	<ul style="list-style-type: none"> 保健医療ビッグデータ政策決定及びシステム構築のための研究 (2015年) 保健医療ビッグデータの R&D 企画研究及びコア技術の R&D (2016年)
2017.01-2018.06 (政策形成の段階)	<ul style="list-style-type: none"> 保健医療ビッグデータ推進団を構成・運営 国会・市民社会と対話、専門家諮問 保健医療ビッグデータの R&D の公告及び公共性評価 保健医療ビッグデータ実務推進団を構成
2018.07-現在 (モデル事業の段階)	<ul style="list-style-type: none"> 保健医療ビッグデータ政策審議委員会を構成 保健医療ビッグデータモデル事業計画 議決 保健医療ビッグデータプラットフォーム開通 (1次：2019.07、2次：2019.11)

(出典：韓国保健産業振興院。保健医療ビッグデータプラットフォーム)

表 6. 保健医療ビッグデータのデータカタログ

提供機関	データ名
疾病管理庁	国民健康栄養調査の基本 DB (2007-2017) 検疫、結核患者申告状況年報 DB (2013-2018) 予防接種 DB (2012-2018) など
国民健康保険公団	資格及び保険料一般健康診断、療養機関、死亡情報、がん情報など ※データ提供年度：2007-2018
健康保険審査評価院	健康保険明細書一般内訳、診療内訳、受診者傷病内訳、処方内訳等 ※データ提供年度：2007-2018
国立がんセンター	がん登録資料 ※データ提供年度：2002-2017

(出典：韓国保健産業振興院。保健医療ビッグデータプラットフォーム)

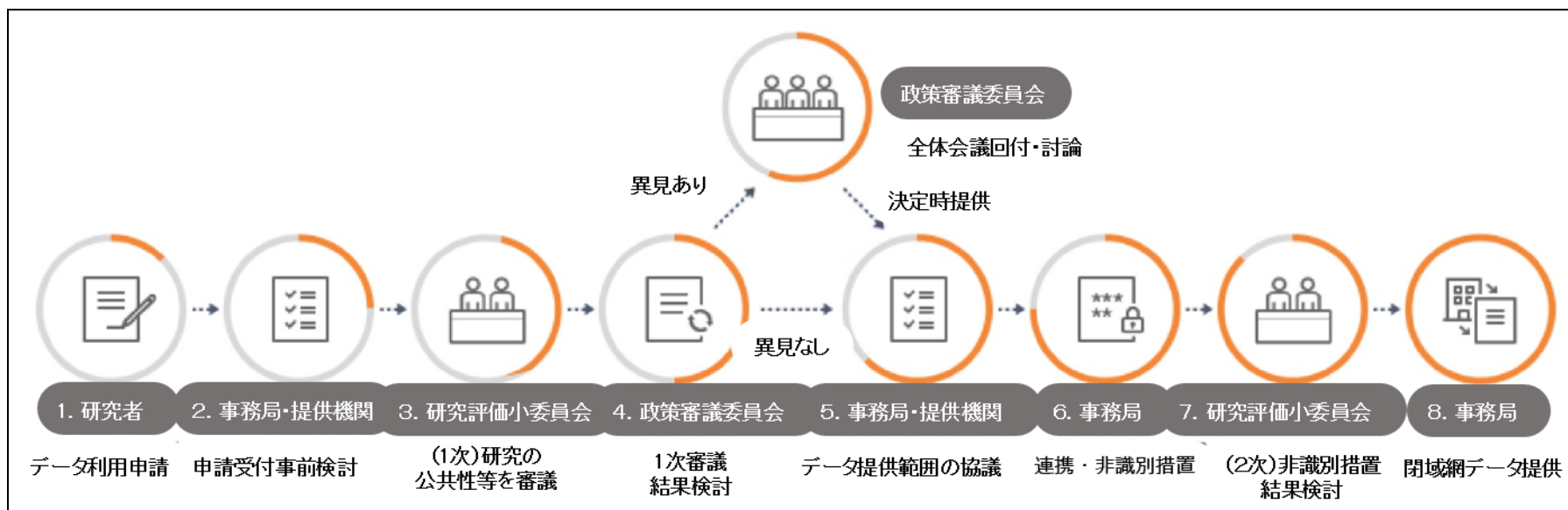


図 4. 保健医療ビッグデータプラットフォームのデータ利用手順

(出典：韓国保健産業振興院。保健医療ビッグデータプラットフォーム)