

【意見を表示したものの全文】

レセプト情報・特定健診等情報データベースシステムにおける収集・保存データの不突合の状況等について

(平成27年9月4日付け 厚生労働大臣宛て)

標記について、会計検査院法第36条の規定により、下記のとおり意見を表示する。

記

1 レセプト情報・特定健診等情報データベースシステム（NDBシステム）等の概要

(1) 医療費適正化計画と生活習慣病予防対策の概要

貴省は、高齢者の医療の確保に関する法律（昭和57年法律第80号。以下「高齢者医療確保法」という。）第8条の規定に基づき、医療に要する費用の適正化（以下「医療費適正化」という。）を総合的かつ計画的に推進するために、医療費適正化に関する施策についての基本的な方針（以下「基本方針」という。）を定めるとともに、5年ごとに、5年を1期として、医療費適正化を推進するための計画（以下「全国医療費適正化計画」という。）を定めることとなっている。また、都道府県は、高齢者医療確保法第9条の規定に基づき、基本方針に即して、5年ごとに、5年を1期として、医療費適正化を推進するための計画（以下「都道府県医療費適正化計画」といい、全国医療費適正化計画と合わせて「医療費適正化計画」という。）を定めることとなっている。

平成20年3月に定められた当初の基本方針及び24年9月改正後の基本方針では、医療費の急増を抑えていくための重要な政策の一つとして、生活習慣病予防対策が掲げられている。そして、都道府県は、これを踏まえて、都道府県医療費適正化計画において、医療保険制度の下における全国健康保険協会、健康保険組合、市町村（特別区を含む。以下同じ。）、共済組合等の保険者（以下「保険者」という。）が生活習慣病予防対策として実施する特定健康診査及び特定保健指導（以下、特定健康診査と特定保健指導を合わせて「特定健診等」という。）に関する目標を定めることとなっている。

また、貴省は、健康保険法（大正11年法律第70号）等の規定に基づき、保険者のうち全国健康保険協会、健康保険組合、市町村及び国民健康保険組合（以下「補助対象保険者」という。）に対し、全国健康保険協会特定健康診査・保健指導補助金等の国庫補助金（以下「国庫補助金」という。）を交付して、特定健診等の実施を助成している。補助対象保険者に対する国庫補助金の交付額は、20年度から26年度までの合計で1257

億余円（見込み）となっている。

(2) 医療費適正化計画の実績に関する評価

貴省及び都道府県は、高齢者医療確保法第12条の規定に基づき、医療費適正化計画の期間の終了の日の属する年度の翌年度において、当該計画に掲げる目標の達成状況及び施策の実施状況に関する調査及び分析を行い、当該計画の実績に関する評価を行うとともに、その結果を公表することとなっている。

また、貴省及び都道府県は、これらの評価を行うに当たっては、高齢者の医療の確保に関する法律施行規則（平成19年厚生労働省令第129号）第3条及び第4条の規定に基づき、医療費適正化計画に掲げる目標の達成状況及び施策の実施状況のほか、当該施策に要した費用に対する効果に係る調査及び分析を行うことなどとなっている。

(3) レセプト情報・特定健診等情報データベースシステム（NDBシステム）の概要

貴省は、「医療サービスの質の向上等のためのレセプト情報等の活用に関する検討会」が20年2月7日に公表した報告書を踏まえて、全ての保険者から全ての特定健診等の内容等に関する電子情報（以下「特定健診等データ」という。）を、また、全ての保険者等（保険者のほか、都道府県の区域内の全ての市町村が加入する後期高齢者医療広域連合を含む。）から全てのレセプトの内容等に関する電子情報（以下「レセプトデータ」という。）を収集・保存した上で、生活習慣病予防対策として実施されている特定健診等が医療費適正化に及ぼす効果等について適切な分析を行うことなどを目的として、レセプト情報・特定健診等情報データベースシステム（以下「NDBシステム」という。）を構築し、21年4月からその運用を開始している。

20年9月に作成されたNDBシステムの入札仕様書（以下「入札仕様書」という。）によれば、NDBシステムに収集・保存するデータを活用して行われる分析の内容は、①レセプトデータに関する分析、②特定健診等データに関する分析及び③特定健診等データとレセプトデータとの突合データに関する分析とされている。このうち③の突合データに関する分析は、特定健診等を受診した全ての被保険者（被扶養者を含む。以下同じ。）の特定健診等データと同一人の全てのレセプトデータを突合した上で、追跡調査を行うなどして詳細な分析等を行い、生活習慣病予防対策が医療費適正化に及ぼす影響等について評価するものである。

そして、貴省は、30年度に予定されている第2期医療費適正化計画の実績に関する評価においては、生活習慣病予防対策として実施されている特定健診等が医療費適正化

に及ぼす効果について、NDBシステムに収集・保存されているデータの十分な突合・分析等により得られる詳細な分析データに基づき適切な評価を行うことにしている。

(4) 特定健診等データ及びレセプトデータの収集方法等

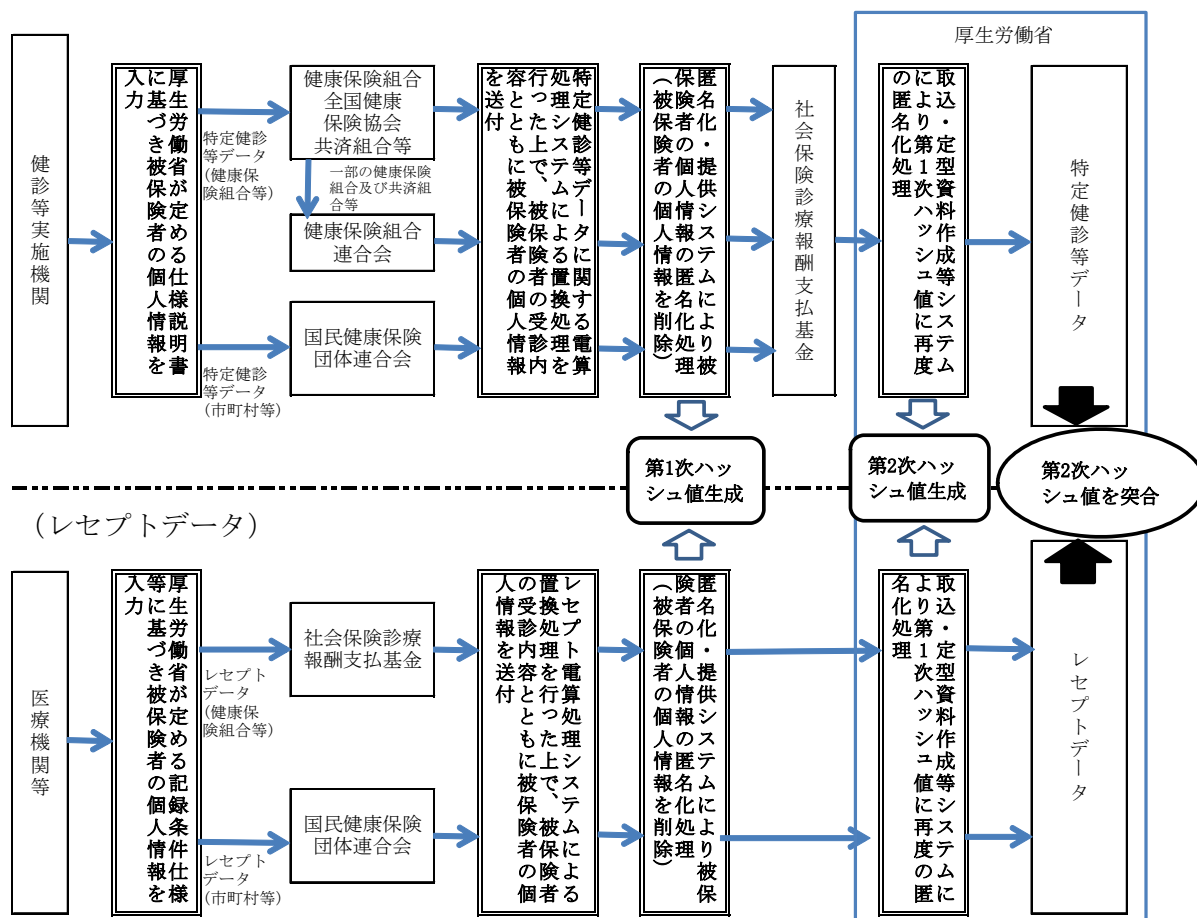
図1のとおり、NDBシステムに収集・保存する特定健診等データについては全国の特定健診等の実施機関（以下「健診等実施機関」という。）が、レセプトデータについては医療機関又は薬局（以下「医療機関等」という。）が、被保険者（特定健診等については受診の対象となる被保険者。以下同じ。）の受診内容等のデータとともに、被保険者の加入する保険者の番号、被保険者証等記号・被保険者証等番号、生年月日、氏名等（以下「被保険者の個人情報」という。）を入力している。

そして、特定健診等データについては、健診等実施機関が「特定健診の電子的なデータ標準様式特定健診情報ファイル仕様説明書」（以下「仕様説明書」という。）に基づき入力することとなっている。一方、レセプトデータについては、従来、医療機関等が「オンライン又は光ディスク等による請求に係る記録条件仕様」等（以下「記録条件仕様等」という。）に基づき入力しているものである。

また、貴省は、NDBシステムの運用に当たり、保険者、健康保険組合連合会、社会保険診療報酬支払基金（以下「支払基金」という。）及び国民健康保険団体連合会（以下「国保連合会」といい、国保連合会と支払基金を合わせて「審査支払機関」という。）に対して、特定健診等データ及びレセプトデータにそれぞれ含まれる被保険者の個人情報については、これを匿名化する処理（以下「匿名化処理」という。）を行うよう通知するとともに、匿名化処理を行うためのシステムである匿名化・提供システムを提供している。匿名化処理は、被保険者の個人情報を疑似乱数化された別の文字列（以下「ハッシュ値」という。）に置き換えるものである。

保険者、健康保険組合連合会及び審査支払機関は、健診等実施機関又は医療機関等から送付された特定健診等データ及びレセプトデータについて匿名化処理を行い、生成された第1次ハッシュ値（ハッシュ値を生成した後、元の被保険者の個人情報を削除したもの。以下同じ。）をNDBシステムに送付している。そして、貴省は、送付された第1次ハッシュ値に再度の匿名化処理を行って第2次ハッシュ値を生成し、これをNDBシステムに保存している。

図1 NDBシステムの概念図
(特定健診等データ)



注(1) 健康保険組合等には、全国健康保険協会、共済組合等を含む。
注(2) 市町村等には、国民健康保険組合を含む。

(5) NDBシステムの構築等に係る経費及び保存データ

NDBシステムの構築、運用等に係る経費は、20年度から26年度までの間の累計で27億9734万余円となっている。

そして、NDBシステムには、27年2月末現在で特定健診等データ約1億2000万件及びレセプトデータ約87億8900万件が保存されている。

2 本院の検査結果

(検査の観点、着眼点、対象及び方法)

本院は、有効性等の観点から、NDBシステムは所期の目的どおり運用されているか、特に、第2期医療費適正化計画の実績に関する評価において、生活習慣病予防対策として実施されている特定健診等が医療費適正化に及ぼす効果について、NDBシステムに収集・保存されているデータの突合・分析等により得られる詳細な分析データに基づき適

(注1)
切な評価を行うことができる運用状況となっているかなどに着眼して、貴省、22都県の155市区町（155保険者）、2国民健康保険組合、全国健康保険協会本部及び2健康保険組合において、関係書類の提出を受けるとともに、関係者から説明等を徴取するなどして会計実地検査を行った。

(注1) 22都県 東京都、青森、秋田、栃木、群馬、神奈川、富山、石川、山梨、岐阜、愛知、三重、兵庫、和歌山、広島、山口、徳島、愛媛、福岡、佐賀、熊本、沖縄各県

(検査の結果)

検査したところ、次のような事態が見受けられた。

(1) NDBシステムにおける特定健診等データとレセプトデータの突合状況

NDBシステムの運用における特定健診等データとレセプトデータの突合状況について、23、24両年度における各保険者の特定健診データ（特定健診等データから特定保健指導に関するデータを除いたもの。以下同じ。）のレセプトデータとの突合率に(注2)より確認したところ、次のとおり、多数の保険者について、特定健診データをレセプトデータと突合できない状況となっていた。

すなわち、23年度では全3,420保険者のうち1,943保険者（56.8%）、24年度では全3,403保険者のうち1,667保険者（49.0%）の特定健診データをレセプトデータと突合することが全くできない状況となっていた。特に、これらの保険者のうち、全国健康保険協会については、23、24両年度ともに、特定健診データをレセプトデータと突合することが全くできない状況となっていた。また、健康保険組合についても、23、24両年度ともに、ほぼ全ての保険者について、すなわち、23年度では全1,429保険者のうち1,418保険者について、24年度では全1,412保険者のうち1,403保険者について、特定健診データをレセプトデータと突合することが全くできない状況となっていた（表1及び表2参照）。

このほか、国民健康保険の保険者である市町村及び国民健康保険組合についてみると、23年度では全1,903保険者のうち488保険者（409市町村及び79国民健康保険組合）、24年度では全1,903保険者のうち227保険者（194市町村及び33国民健康保険組合）の特定健診データをレセプトデータと突合することが全くできない状況となっていた（表1及び表2参照）。

以上のことから、NDBシステムにおいて、収集・保存されている多数の保険者の特定健診等データをレセプトデータと突合できない事態が生じていることが判明した。

(注2) 突合率 突合率は次の算式により算出されるものである。

$$\text{突合率 (\%)} = \frac{\text{②}}{\text{①}} \times 100$$

①：全ての特定健診データ（第2次ハッシュ値）の件数

②：①と突合できる全てのレセプトデータ（第2次ハッシュ値）について、同一の保険者における同一の被保険者を名寄せして個人単位のデータとした上で、これを集計したものの件数

表1 保険者別の特定健診データとレセプトデータの突合率の分布状況

平成23年度

(単位：保険者数)

保険者		全体	全国健康保険協会	健康保険組合	市町村	国民健康保険組合	共済組合等
突合率	80%以上	1,162	0	9	1,044	58	51
	60%以上80%未満	251	0	1	235	15	0
	40%以上60%未満	6	0	0	2	4	0
	20%以上40%未満	4	0	0	1	3	0
	0%以上20%未満	1,997	1	1,419	457	84	36
	うち、0%	1,943	1	1,418	409	79	36
計		3,420	1	1,429	1,739	164	87

平成24年度

(単位：保険者数)

保険者		全体	全国健康保険協会	健康保険組合	市町村	国民健康保険組合	共済組合等
突合率	80%以上	1,512	0	8	1,345	108	51
	60%以上80%未満	103	0	0	91	12	0
	40%以上60%未満	3	0	0	2	1	0
	20%以上40%未満	7	0	0	5	2	0
	0%以上20%未満	1,778	1	1,404	296	41	36
	うち、0%	1,667	1	1,403	194	33	36
計		3,403	1	1,412	1,739	164	87

注(1) 「共済組合等」の欄は、国家公務員共済組合、地方公務員共済組合、警察共済組合、公立学校共済組合、日本私立学校振興・共済事業団等の合計である。

注(2) 特定健康診査を受診しても医療機関等を受診しない者等がいることから、突合率が100%になることはない。

表2 保険者別の特定健診データとレセプトデータの不突合の状況
平成23年度

(単位：件)

保険者	全体	全国健康保険協会	健康保険組合	市町村	国民健康保険組合	共済組合等
①：全ての特定健診データ（第2次ハッシュ値）の数	23,615,176	4,990,532	7,922,629	7,370,368	628,401	2,703,246
②：①と突合できる全てのレセプトデータ（第2次ハッシュ値）の数	4,483,229	0	29,978	3,713,843	160,078	579,330
不突合の件数（①－②）	19,131,947	4,990,532	7,892,651	3,656,525	468,323	2,123,916
突合率（②／①）	19.0%	0.0%	0.4%	50.4%	25.5%	21.4%

平成24年度

(単位：件)

保険者	全体	全国健康保険協会	健康保険組合	市町村	国民健康保険組合	共済組合等
①：全ての特定健診データ（第2次ハッシュ値）の数	24,656,823	5,451,153	8,230,619	7,600,042	646,004	2,729,005
②：①と突合できる全てのレセプトデータ（第2次ハッシュ値）の数	6,149,288	0	23,619	5,179,655	351,101	594,913
不突合の件数（①－②）	18,507,535	5,451,153	8,207,000	2,420,387	294,903	2,134,092
突合率（②／①）	24.9%	0.0%	0.3%	68.2%	54.3%	21.8%

注(1) 「共済組合等」の欄は、国家公務員共済組合、地方公務員共済組合、警察共済組合、公立学校共済組合、日本私立学校振興・共済事業団等の合計である。

注(2) 特定健康診査を受診しても医療機関等を受診しない者等がいることから、突合率が100%になることはない。

(2) NDBシステムにおける収集・保存データの不突合の重要な要因

ア 特定健診等データ及びレセプトデータの入力形式等

前記の事態を踏まえて、特定健診等データ及びレセプトデータのそれぞれの被保険者の個人情報の入力文字の種類・形式（英数、漢字等の別及び全角・半角文字の別）や数字等の入力方法（以下、これらを合わせて「入力形式等」という。）について確認したところ、例えば、健診等実施機関では、特定健診等データの入力の際に、「被保険者証等記号」を全角文字で入力するなどしているのに、医療機関等では、レセプトデータの入力の際に、これを半角英数文字で入力するなど、被保険者の個人情報の入力形式等が両データで異なるものとなっていた。

そこで、仕様説明書及び記録条件仕様等のそれぞれにおける被保険者の個人情報等の入力形式等に関する指示内容の記載について確認したところ、例えば、「被保険

者証等記号」及び「被保険者証等番号」等の入力形式等に関する両者の指示内容は異なるものとなっており、しかも、レセプトデータについては「半角英数文字又は全角文字（漢字・カナ・数字等を含む。）」で入力することとなっているなど、指示内容が選択可能なものとなっていた（表3参照）。

このため、同一の被保険者の個人情報であるのに、両データで被保険者の個人情報の入力形式等が異なるものとなっていて、異なるハッシュ値が生成される場合があることが判明した。

イ レセプト電算処理システム等における置換処理

レセプトデータについては、従来、審査支払機関のレセプト電算処理システムにおいて、例えば、半角英数文字又は全角文字で入力された「被保険者証等記号」については全て全角文字に置き換えるなど、被保険者の個人情報の入力形式等を同じものとする処理（以下「置換処理」という。）が行われていた。

一方、特定健診等データについては、一部の保険者、健康保険組合連合会又は国保連合会の特定健診等データに関する電算処理システムにおいて、例えば、半角英数文字又は全角文字で入力された「被保険者証等記号」については全て半角英数文字に置き換えるなど、レセプト電算処理システムとは異なる置換処理が行われていた。

このため、両データで被保険者の個人情報の入力形式等が同じ場合であっても、異なるハッシュ値が生成される場合があることも判明した。

ウ 匿名化・提供システムにおける置換処理

匿名化・提供システムも、特定健診等データとレセプトデータとで被保険者の個人情報の入力形式等（置換処理が行われた後のものを含む。以下同じ。）が異なっている場合に、匿名化処理の過程で「被保険者の生年月日のうち和暦のものは西暦に置き換える」など一定の置換処理を行う機能を有している。

しかし、この置換処理機能は、「被保険者証等記号」等の入力形式等が異なっている場合には置換処理を行うことになっていないなど、限定的な機能となっていて、両データで被保険者の個人情報の入力形式等が異なる全ての場合に置換処理を行うものとなっていなかった（表3参照）。

以上のように、同一の被保険者であっても、特定健診等データとレセプトデータとでは被保険者の個人情報の入力形式等が異なっている場合があり、その場合には、そ

それぞれの匿名化処理の過程で異なるハッシュ値が生成され、NDBシステムには異なる第2次ハッシュ値が保存されることになる。そして、このことがNDBシステムにおいて、多数の保険者について両データを突合できない事態を生じさせている重要な要因となっていると認められた。

表3 レセプトデータと特定健診等データの入力形式等の相違及び置換処理の内容等

種類 項目	レセプトデータの入力形式等	特定健診等データの入力形式等	匿名化・提供システムにおける 置換処理
保険者番号	8桁以内で設定された保険者番号については、右詰めに記録し残りは“スペース”を記録する。	8桁に満たない場合には先頭の0をつけて8桁とする。	半角スペースは全て0に置換する。
被保険者証等記号 ・ 被保険者証等番号	半角英数文字又は全角文字（漢字・カナ・数字等を含む。）	全角だけからなる文字列又は半角だけからなる文字列のどちらかとする。	全角・半角スペースは全て削除する。
	健康保険被保険者証、船員保険被保険者証、受給資格者票及び国民健康保険被保険者証等の「記号及び番号」欄の記号（番号）を左詰めに記録する。		
		ただし、電子化レセプトを作成している医療機関においては、これにかかわらず電子化レセプトでの記述形式と同一とすること（注）	
生年月日	和暦	西暦	西暦に変換する。
性別	男=「1」、女=「2」	男=「1」、女=「2」	
氏名	半角英数文字又は全角文字（漢字・カナ・数字等を含む。）	全角カナ氏名	
	姓と名の間に“スペース”を1桁記録する。	空白を含まない。	全角・半角スペースは全て削除する。

（注）平成24年10月改定の仕様説明書では、この記述は削除されている。

上記の事態について事例を示すと、次のとおりである。

<事例>

全国健康保険協会は、NDBシステムの構築に合わせて、特定健診等データを支払基金に送付するためのシステム（以下「特定健診情報報告システム」という。）を構築し、平成21年4月から運用している。そして、特定健診情報報告システムでは、健診等実施機関から送付された特定健診等データに含まれる「被保険者証等記号」及び「被保険者証等番号」については、全て半角英数文字に置き換えるという置換処理が行われている。

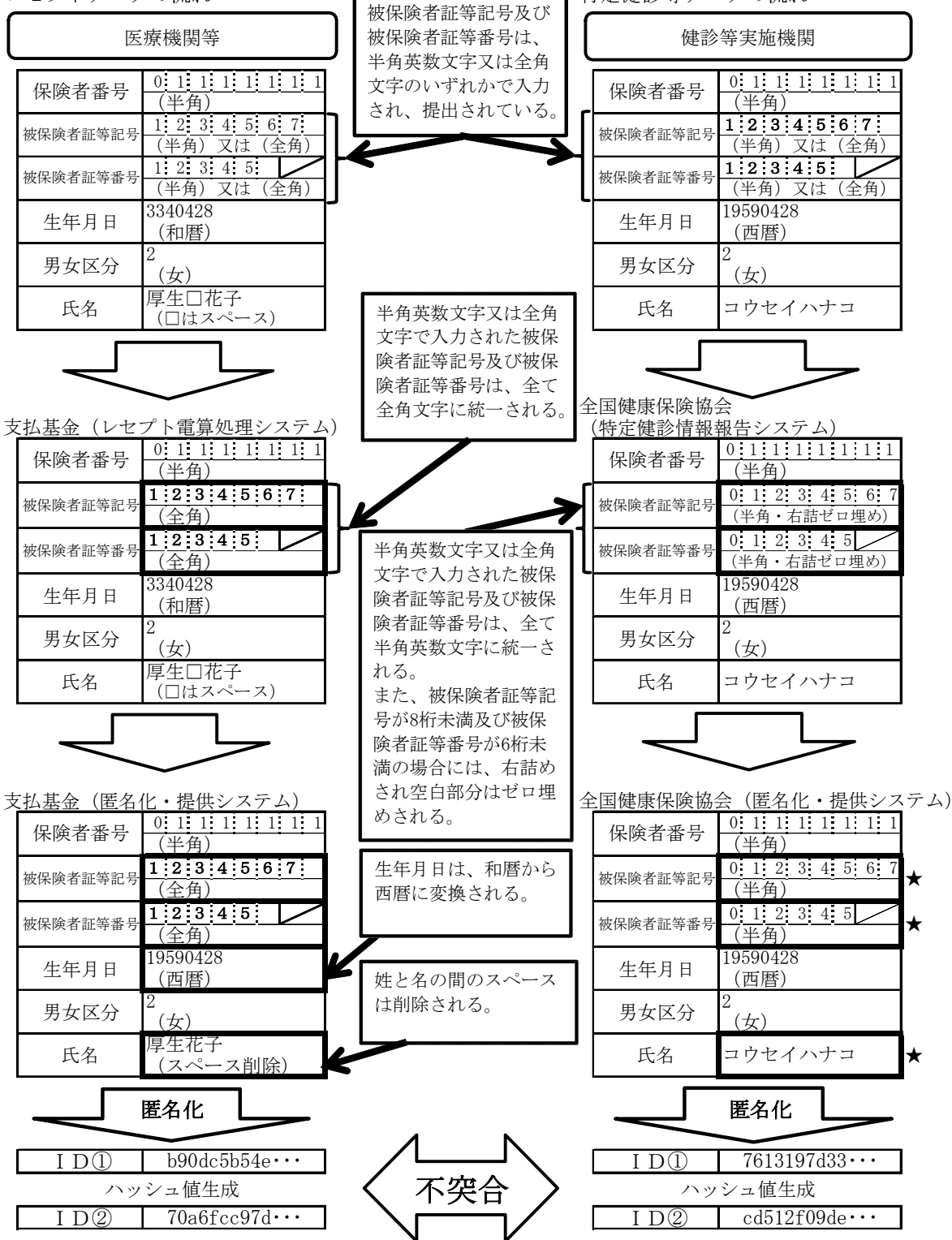
しかし、従来、支払基金のレセプト電算処理システムでは、同協会の特定健診情報報告システムにおける置換処理とは逆に、レセプトデータとして入力された被保険者の個人情報のうち半角英数文字で入力された「被保険者証等記号」、「被保険者証等番号」及び「氏名」については、全て全角文字に置き換えるという置換処理が行われていた。

また、特定健診情報報告システムでは、健診等実施機関から同協会に送付された特定健診等データについては、「被保険者証等記号」が8桁未満の場合及び「被保険者証等番

号」が6桁未満の場合には右詰めを入力とし、空白部分については0を入力するという置換処理が行われている。しかし、支払基金におけるレセプト電算処理システムでは、医療機関等から支払基金に提出されたレセプトデータについて、同様の置換処理は行われていなかった。

以上のようなことから、同協会の被保険者の個人情報については、特定健診等データとレセプトデータのそれぞれの匿名化処理の過程で異なるハッシュ値が生成され、NDBシステムには異なる第2次ハッシュ値が保存されていたことになり、NDBシステムにおいて、同協会の特定健診等データをレセプトデータと突合することは全くできない状況（突合率0%）となっていた（図2参照）。

図2 レセプトデータと特定健診等データの不突合の状況（全国健康保険協会の例）
レセプトデータの流れ



(3) 特定健診等データとレセプトデータの突合に関する検証等

入札仕様書によれば、NDBシステムにおいて実際に特定健診等データとレセプトデータを突合することができるかどうかについての検証（以下「突合検証」という。）は、設計・開発段階において、受託業者が模擬データを作成して行うこととされている。そして、貴省は、模擬データの作成については、受託業者が仕様説明書及び記録条件仕様等に基づき行ったとしており、突合検証は21年3月に行われている。

しかし、貴省は、被保険者の個人情報を入力形式等が特定健診等データとレセプトデータとで異なる場合があるなどのため、同一の被保険者の個人情報であっても両データで異なるハッシュ値が生成される場合があることなどについては想定しておらず、突合検証に当たり、受託業者と十分な協議を行っていなかった。このため、受託業者により作成された模擬データは、被保険者の個人情報を入力形式等が両データで異なる場合があるなどの実態を適切に反映したものとなっていなかったと認められる。

そして、貴省は、NDBシステムの運用開始された同年4月以降も、実際にNDBシステムに収集・保存されている特定健診等データ及びレセプトデータを用いた突合検証を一切行っていないかった。

また、貴省の「保険者による健診・保健指導等に関する検討会」（24年2月24日開催）の資料には、22年4月から23年3月までの間のNDBシステムにおける特定健診等データとレセプトデータとの突合率は男性で9.8%、女性で15.7%と記載されている。したがって、貴省は、遅くともこの24年2月の時点では、NDBシステムにおいて両データの不突合が生じている事態を把握していたことになる。

しかし、貴省は、当時、上記のような不突合が生じている原因について、被保険者の加入する保険者の変更等によるものではないかと推測したり、突合の精度には限界があると説明したりするなどして、事態の原因究明と改善に向けた調査等を速やかに実施していなかった。

（改善を必要とする事態）

NDBシステムは、27億9734万余円の経費を投じて構築されたものであるのに、被保険者の個人情報を入力形式等が特定健診等データとレセプトデータとで異なるものとなっていたり、電算処理システムにおける置換処理が特定健診等データに関する電算処理システムとレセプト電算処理システムとで異なるものとなっていたり、匿名化・提供システムにおける置換処理機能が限定的なものとなっていたりするなどしているため、多

数の保険者の特定健診等データをレセプトデータと突合できない状況となっている事態は適切ではなく、改善の要があると認められる。

(発生原因)

このような事態が生じているのは、貴省において、次のようなことなどによると認められる。

ア 従来のレセプトデータの入力形式等に対する理解が十分でないため、仕様説明書の作成に当たり、被保険者の個人情報と特定健診等データとして入力する際には、その入力形式等をレセプトデータの入力形式等と同一のものとするなどの指示を十分に行っていないこと

イ 審査支払機関のレセプト電算処理システム等における置換処理に関する理解が十分でないため、特定健診等データとレセプトデータとで被保険者の個人情報の入力形式等が同じ場合であっても、それぞれ異なる内容の置換処理が行われる結果、生成されるハッシュ値は異なったものとなる場合があることについての認識が欠けていること

ウ 匿名化・提供システムの設計等に当たり、特定健診等データとレセプトデータとで被保険者の個人情報の入力形式等が異なる場合があることを踏まえた十分な置換処理機能を用意していないこと

エ NDBシステムの運用に当たり、実際に収集・保存された特定健診等データとレセプトデータを用いた突合検証を全く行っていないこと

3 本院が表示する意見

貴省は、生活習慣病予防対策として特定健診等を実施している全国の補助対象保険者に対して多額の国庫補助金を交付しているのに、NDBシステムにおいて収集・保存されている多数の保険者の特定健診等データをレセプトデータと突合できない事態が生じており、このまま推移すれば、30年度に予定されている第2期医療費適正化計画の実績に関する評価において、当該施策が医療費適正化に及ぼす効果について、収集・保存されているデータの十分な突合・分析等により得られる詳細な分析データに基づき適切な評価を行うことは困難であると見込まれる。

については、貴省において、NDBシステムの運用状況を大幅に改善し、第2期医療費適正化計画の実績に関する評価に当たっては、生活習慣病予防対策として実施されている特定健診等が医療費適正化に及ぼす効果について、収集・保存されているデータを十分に活用した適切な評価を行うことができるようにするために、データの不突合の原因等

を踏まえたシステムの改修等を行うなどの措置を講ずるよう意見を表示する。