

等)を扱う臨床研究情報データベースサーバ、および、Web コンテンツを提供する Web サーバにより、サーバ群は構成されている。個人特定情報データベースについては、検索可能暗号化されており、データセンタ側で一切復号ができない暗号方式を採用している。検索可能暗号化とは、従来の検索可能なデータベースがデータセンタ側で復号が可能なのに対し、暗号化したまま検索が可能な技術であり、鍵を持ったユーザからのみデータにアクセスできる方式である。サーバ、外部ネットワーク間は、ファイアウォールやルータを介し、それぞれネットワークセグメントが分離されている。

2) システム運用を実施している企業(統合 DB の運用企業と各施設から統合 DB へのデータ抽出システムを運営する企業等複数ある場合には、複数企業名とその担当業務)

日立製作所

5. データの個人情報・匿名情報の取り扱いに関する情報

1) データの匿名化の有無、匿名化を行う者

データは匿名化されている。匿名化を行うものは、個人特定情報を行うものであり、個人、事務局、医師等が該当する。システム運用者は関与しない。

2) 匿名化の場合、匿名加工のためのデータ加工方法

(現行システム)

連結可能匿名化を行い、登録対象者に新たに付された符号または番号(コード)と対象者氏名の対応表(コード表)を管理する。

(今後導入予定のシステム)

Web アプリからサーバ送信時に個人特定情報と、臨床研究情報を分離して匿名化を行う。

3) 匿名化の場合、匿名個人データを連結可能な符号処理や情報源の異なる特定個人の情報を突合可能な仕組みの有無とその方法

匿名個人データを連結可能な符号処理：有

特定個人の情報を突合可能な仕組み：有

個人特定情報と臨床研究情報は、入力の際で分離され、個人特定情報は個人特定情報データベースサーバへ、臨床研究情報は臨床研究情報データベースサーバへ分離して送信される。分離される際に符号処理が行われる。個人特定情報と臨床研究情報を結合する ID の組み合わせが生成され、個人特定情報データベースサーバにて保持される。

6. データベースに関する情報

1) インデックス DB の採用有無

(現行システム)

個人特定情報データベースサーバ:無、臨床研究情報データベースサーバ:有
(今度投入予定のシステム)

個人特定情報データベースサーバと臨床研究情報データベースサーバのいずれもインデックス DB を採用している。なお、個人特定情報データベースサーバについては、インデックスも暗号化されており、データセンタ側で復号することはできない。

2) 解析用 DB で採用している DB 記述 (cache, Oracle, SQLserver, Netezza, Hadoop など)

個人特定情報データベースサーバ:MySQL (ただし、独自に検索可能暗号化機能を実装したものを
用い、日立製作所のサービスシステムであり、情報管理サービスとして提供されている)。臨床研究情報
データベースサーバ:HiRDB (ただし、日立製作所のサービスシステムであり、情報管理サービスとし
て提供されている)。

※インデックスデータベースとは

SS-MIX2「標準化ストレージ」は患者 (患者 ID) を特定してから、当該患者の診療情報を検索するこ
とに特化した物理構造を採用している。したがって、例えば特定の診療日に該当する診療情報や、特定
の診療行為に該当する診療情報等、複数の患者に跨った参照を行う場合や大量の診療情報を参照する場
合にはシステムに多大な負荷を与えることとなる。このような負荷を軽減するとともに、標準化ストレ
ージを利活用するアプリケーションが上記のような検索機能を容易に実装できるようにするため、リレ
ーションデータベースシステム等を利用し、物理構造を構成する値をデータベースに保持する、イン
デックスデータベースを規定する。

(「SS-MIX2 標準化ストレージ構成の説明と構築ガイドライン Ver.1.2c」より抜粋)

SS-MIX2 を利用していない事業においても、同様のインデックスデータベースを装備している場合は、
「有」ときさいしてください。

7. データベースの現状と課題

1) データ収集の方法について

- ・患者からの郵送での登録に基づいて内容確認を行うしくみとなっている。
- ・患者の約 1 / 3 が登録している。

2) データの内容 (項目) および精度 (正確さ) に関して

- ・主治医に内容確認を行っている。
- ・状態が変わったら患者が登録を変更するしくみである。

3) データの活用に関して

- ・臨床試験登録等に活用できると考えられる。

1.4 難病データベース（仮称）

1. 担当部局

健康局疾病対策課

2. 事業の位置づけ、データ利用に関する情報

1) 根拠規定

難病の患者に対する医療等に関する法律施行規則第 18 条（平成 26 年 11 月 12 日省令第 121 号）

2) 事業の目的

難病の患者の QOL 向上を図ることを目的として、難病研究の推進を図るため。

2)～5)については検討段階。

3) 事業におけるデータ取得対象者への同意取得の有無と在りの場合の同意内容

特定医療費支給認定申請書を申請する際に、臨床調査個人票（患者データ）の研究利用についての同意を求めている。

4) 当該事業におけるデータ利用者（記載例：〇〇局〇〇課、事業参加医療機関、等）

都道府県における特定医療費支給認定

5) 当該事業以外での第三者データ利用の可否とその場合の利用法

可。患者データを疫学研究、臨床開発に活用。

3. 事業の規模に関する情報

1) 事業（収集）開始時期（予定含む）

未定

2) データ保有期間（記載例：2010 年から現在までの 5 年分）

3) 保有するデータ量（患者数、データ件数）

4) 参加医療機関

難病指定医が臨床調査個人票（患者データ）を記載する。

4. 事業システムに関する情報

1) システム基本構成

未定

2) システム運用を実施している企業(統合 DB の運用企業と各施設から統合 DB へのデータ抽出システムを運営する企業等複数ある場合には、複数企業名とその担当業務)

未定

5. データの個人情報・匿名情報の取り扱いに関する情報

1) データの匿名化の有無、匿名化を行う者

未定

2) 匿名化の場合、匿名加工のためのデータ加工方法

未定

3) 匿名化の場合、匿名個人データを連結可能な符号処理や情報源の異なる特定個人の情報を突合可能な仕組みの有無とその方法

未定

6. データベースに関する情報

1) インデックス DB※の採用有無

2) 解析用 DB で採用している DB 技術 (cache, Oracle, SQLserver, Netezza, Hadoop など)

※インデックスデータベースとは

SS-MIX2「標準化ストレージ」は患者(患者 ID)を特定してから、当該患者の診療情報を検索することに特化した物理構造を採用している。したがって、例えば特定の診療日に該当する診療情報や、特定の診療行為に該当する診療情報等、複数の患者に跨った参照を行う場合や大量の診療情報を参照する場合にはシステムに多大な負荷を与えることとなる。このような負荷を軽減するとともに、標準化ストレージを利活用するアプリケーションが上記のような検索機能を容易に実装できるようにするため、リレーショナルデータベースシステム等を利用し、物理構造を構成する値をデータベースに保持する、インデックスデータベースを規定する。

(「SS-MIX2 標準化ストレージ構成の説明と構築ガイドライン Ver.1.2c」より抜粋)

SS-MIX2 を利用していない事業においても、同様のインデックスデータベースを装備している場合は、

「有」ときさいしてください。

7. データベースの現状と課題

1) データ収集の方法について

- ・平成27年1月から新法のもとで306疾患が対象となる。
- ・医師の診断に基づくデータ収集である。
- ・データ収集開始時期は未定である。
- ・患者の医療費助成申請書に同意文書が付与される。
- ・国の機関でデータ管理を行う。

2) データの内容（項目）および精度（正確さ）に関して

- ・登録は難病指定医に限定する。

3) データの活用に関して

- ・56疾患のデータベースについては限定的に利用している。

1.5 国保データベース(KDB)システム

1. 担当部局

保険局 国民健康保険課、高齢者医療課

2. 事業の位置づけ、データ利用に関する情報

1) 根拠規定

- ・国民健康保険法に基づく保健事業の実施等に関する指針（平成 16 年厚生労働省告示第 307 号）
- ・高確法に基づく保健事業の実施等に関する指針（平成 25 年厚生労働省告示第 141 号）

2) 事業の目的

国保連合会が保険者の委託を受けて行う審査支払業務等を通じて管理する、「特定健診・特定保健指導等情報」「レセプト情報」「介護保険情報」を利活用し統計情報等を保険者向けに情報提供することで、保険者の効率的かつ効果的な保健事業の実施をサポートすることを目的として構築。

3) 事業におけるデータ取得対象者への同意取得の有無と在りの場合の同意内容

保険者：「国保データベースシステムを利用した帳票データの作成業務に関する委託契約」を各国保連合会と保険者間にて締結。

（内容）別添契約書（案）参照

被保険者：無し

4) 当該事業におけるデータ利用者（記載例：〇〇局〇〇課、事業参加医療機関、等）

国保保険者、後期高齢者医療広域連合

5) 当該事業以外での第三者データ利用の可否とその場合の利用法

第三者利用不可

3. 事業の規模に関する情報

1) 事業（収集）開始時期（予定含む）

平成 25 年 10 月稼働開始

2) データ保有期間（記載例：2010 年から現在までの 5 年分）

2012 年から現在までの約 3 年分（MAX5 年分）

3) 保有するデータ量 (患者数、データ件数)
患者数 : 約 5,450 万人、レセプト : 約 23 億件

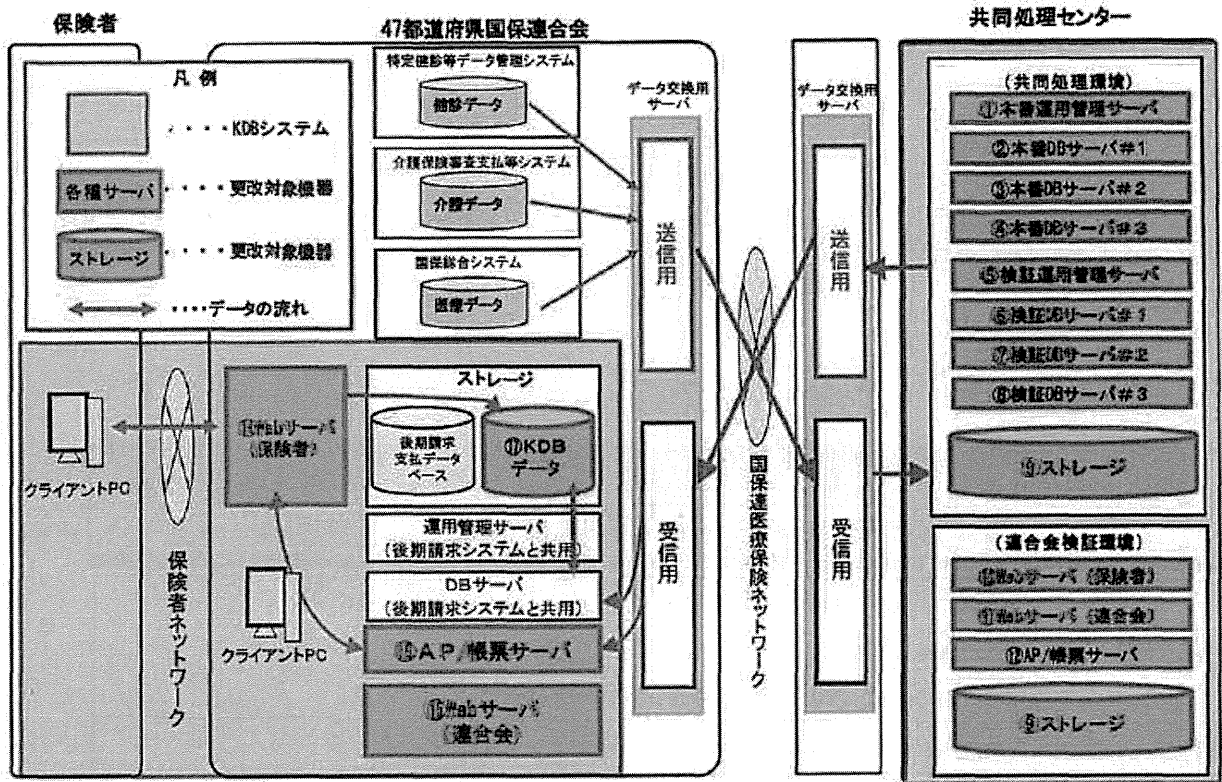
4) 参加医療機関
なし

4. 事業システムに関する情報

1) システム基本構成
共同処理環境

構成要素	内容 (バージョン)
ハード・CPU	PRIMERGY RX300 S6 / Xeon® E5640 2.66G(2P8C)
FW、ミドル	Waha! Transformer Ver4.2 JPI/Base V9.5 JPI/Automatic Job Management System3 Agent V9.5
OS	Windows Server 2008 R2 SP-1
DBMS	SQL Server 2012 Enterprise Edition

システム概要図



2) システム運用を実施している企業(統合DBの運用企業と各施設から統合DBへのデータ抽出システムを運営する企業等複数ある場合には、複数企業名とその担当業務)

安川情報システム株式会社

5. データの個人情報・匿名情報の取り扱いに関する情報

1) データの匿名化の有無、匿名化を行う者

データの匿名化(暗号化)有、47都道府県国保連合会

2) 匿名化の場合、匿名加工のためのデータ加工方法

ビット操作による独自の暗号アルゴリズム

3) 匿名化の場合、匿名個人データを連結可能な符号処理や情報源の異なる特定個人の情報を突合可能な仕組みの有無とその方法

情報源の異なる特定個人の情報を突合している。

暗号化前に特定個人の情報のクレンジング処理を行い、情報源の異なる情報を統一化する方法を採用している。

6. データベースに関する情報

1) インデックス DB※の採用有無

有 (SS-MIX2 は利用していませんが、毎月連携される大量データを処理するため、リレーショナルデータベースシステム等を利用し、物理構造を構成する値をデータベースに保持する)

2) 解析用 DB で採用している DB 技術 (cache, Oracle, SQLserver, Netezza, Hadoop など) SQLserver

※インデックスデータベースとは

SS-MIX2「標準化ストレージ」は患者（患者 ID）を特定してから、当該患者の診療情報を検索することに特化した物理構造を採用している。したがって、例えば特定の診療日に該当する診療情報や、特定の診療行為に該当する診療情報等、複数の患者に跨った参照を行う場合や大量の診療情報を参照する場合にはシステムに多大な負荷を与えることとなる。このような負荷を軽減するとともに、標準化ストレージを利活用するアプリケーションが上記のような検索機能を容易に実装できるようにするため、リレーショナルデータベースシステム等を利用し、物理構造を構成する値をデータベースに保持する、インデックスデータベースを規定する。

（「SS-MIX2 標準化ストレージ構成の説明と構築ガイドライン Ver.1.2c」より抜粋）

SS-MIX2 を利用していない事業においても、同様のインデックスデータベースを装備している場合は、「有」ときさいしてください。

7. データベースの現状と課題

1) データ収集の方法について

- ・特定健診・医療・介護のデータが網羅されている。
- ・システムは国保中央会で作成しているが、データは各連合会単位となっている。
- ・参加市町村は約 95%である（条例等の関係で参加していない自治体がある）。
- ・平成 25 年 10 月から運用している。
- ・過去のデータを取り込んでいない市町村が多い（システムとしては過去 5 年分は登録可）。
- ・転居等で保険者が変わると追跡不可となる。

2) データの内容（項目）および精度（正確さ）に関して

3) データの活用に関して

- ・データ分析は各連合会で実施し、フィードバックする
- ・第三者提供は想定していない。

- ・データの所有権は各保険者なので、利用には同意が必要である。

4) データの相互利用によるメリット

- ・保険加入時点での情報があると、さらに利用価値が高まるのではないか。

1.6 DPC データベース管理運用システム

1. 担当部局

保険局医療課

2. 事業の位置づけ、データ利用に関する情報

1) 根拠規定

「日本再興戦略」改訂 2014」（平成 26 年 6 月閣議決定）により DPC データの第三者提供の本格的な運用を目指すこととされたことを受け、システム構築に着手したものを。

2) 事業の目的

保険局医療課が収集する DPC データをデータベース化し、データの一元管理及びその利活用を可能とすることを目的とする。

3) 事業におけるデータ取得対象者への同意取得の有無と在りの場合の同意内容

無し

4) 当該事業におけるデータ利用者（記載例：〇〇局〇〇課、事業参加医療機関、等）

保険局医療課

5) 当該事業以外での第三者データ利用の可否とその場合の利用法

平成 26 年度からの試行期間中は、国の行政機関に所属する者が研究目的で利用する場合は、「レセプト情報等の提供に関する有識者会議」において認められた場合に限り利用可能としている。第三者提供の本格運用の時期及び利用範囲等については未定。

3. 事業の規模に関する情報

1) 事業（収集）開始時期（予定含む）

平成 29 年度からシステム稼働開始

2) データ保有期間（記載例：2010 年から現在までの 5 年分）

4 年分を保有することを想定

3) 保有するデータ量（患者数、データ件数）

800 億件程度を想定

4) 参加医療機関

1,804 病院（平成 25 年度）

4. 事業システムに関する情報

1) システム基本構成

未定

2) システム運用を実施している企業（統合 DB の運用企業と各施設から統合 DB へのデータ抽出システムを運営する企業等複数ある場合には、複数企業名とその担当業務）

未定

5. データの個人情報・匿名情報の取り扱いに関する情報

1) データの匿名化の有無、匿名化を行う者

医療機関において匿名化されたデータが提出される。

2) 匿名化の場合、匿名加工のためのデータ加工方法

医療機関において患者毎に独自の「データ識別番号」を割り当てている。

3) 匿名化の場合、匿名個人データを連結可能な符号処理や情報源の異なる特定個人の情報を突合可能な仕組みの有無とその方法

無し。

6. データベースに関する情報

1) インデックス DB※の採用有無

未定

2) 解析用 DB で採用している DB 技術（cache, Oracle, SQLserver, Netezza, Hadoop など）

未定

※インデックスデータベースとは

SS-MIX2「標準化ストレージ」は患者（患者 ID）を特定してから、当該患者の診療情報を検索することに特化した物理構造を採用している。したがって、例えば特定の診療日に該当する診療情報や、特定の診療行為に該当する診療情報等、複数の患者に跨った参照を行う場合や大量の診療情報を参照する場

合にはシステムに多大な負荷を与えることとなる。このような負荷を軽減するとともに、標準化ストレージを活用するアプリケーションが上記のような検索機能を容易に実装できるようにするため、リレーショナルデータベースシステム等を利用し、物理構造を構成する値をデータベースに保持する、インデックスデータベースを規定する。

(「SS-MIX2 標準化ストレージ構成の説明と構築ガイドライン Ver.1.2c」より抜粋)

SS-MIX2 を利用していない事業においても、同様のインデックスデータベースを装備している場合は、「有」ときさいしてください。

7. データベースの現状と課題

1) データ収集の方法について

- ・データベースを構築中で、平成29年度からの運用を予定している。
- ・入院については平成18年度からデータがある。平成26年度からは外来データもある。
- ・3ヶ月に1度、病院から調査事務局にデータが送られるしくみである。
- ・全DPC対象病院のデータが網羅されている。

2) データの内容(項目)および精度(正確さ)に関して

- ・データチェックのためのソフトを配布している。
- ・データ提出時に事務局でもチェックしている。

3) データの活用に関して

- ・第三者提供を検討している。
- ・利用に関するガイドラインがある。
- ・患者単位での他のデータとのリンクは難しい。
- ・病院が変わると同じ患者を同定することはできない。

1.7 レセプト情報・特定健診等情報データベース（NDB:ナショナルデータベース）

1. 担当部局

保険局医療介護連携政策課保健システム高度化推進室

2. 事業の位置づけ、データ利用に関する情報

1) 根拠規定

高齢者の医療の確保に関する法律第 16 条

2) 事業の目的

全国医療費適正化計画及び都道府県医療費適正化計画の作成、実施及び評価に資するため

3) 事業におけるデータ取得対象者への同意取得の有無と在りの場合の同意内容

同意取得なし。

4) 当該事業におけるデータ利用者（記載例：〇〇局〇〇課、事業参加医療機関、等）

保険局医療介護連携政策課医療費適正化計画推進室。

5) 当該事業以外での第三者データ利用の可否とその場合の利用法

医政局指導課、医政局地域医療計画課、医薬食品局安全対策課、健康局疾病対策課、保険局医療課、国土交通省、文部科学省、岐阜県、神奈川県、千葉県、川崎市、国立がん研究センター、国立感染症研究所、国立健康・栄養研究所、PMDA、各大学・大学附属病院等。

3. 事業の規模に関する情報

1) 事業（収集）開始時期（予定含む）

<レセプトデータ>平成 21 年度から

<特定健診データ>平成 20 年度から

2) データ保有期間（記載例：2010 年から現在までの 5 年分）

<レセプトデータ>平成 21 年度から現在まで 5 年分

<特定健診データ>平成 20 年度から現在まで 6 年分

3) 保有するデータ量（患者数、データ件数）

<レセプトデータ>平成 21 年度から平成 27 年 4 月診療分までで 97 億 2000 万件。

<特定健診データ>平成 20 年度から平成 25 年度分までで、1 億 4200 万件。

4) 参加医療機関

<レセプトデータ>全国保険医療機関（病院、診療所、薬局、歯科医院）

<特定健診データ>特定健診等実施機関

4. 事業システムに関する情報

1) システム基本構成

- ・レセプトについては平成 21 年 4 月より収集開始し、平成 27 年 4 月診療分まで合計およそ 97 億 2000 万件収集している。
- ・特定健診・特定保健指導については平成 20 年度より収集開始し、平成 25 年度分まで合計約 1 億 4200 万件収集している。
- ・本来目的である、医療費適正化計画の作成・評価に用いるのみならず、第三者提供として公益性の高い申出者に対して有識者会議での審査を経て、データの一部を提供する枠組みも整備されており、平成 27 年 4 月時点で 58 件に対してデータ提供を完了している。
- ・保険者及び後期高齢者医療広域連合から必要な情報を受領（支払基金、国保連合会経由）。

2) システム運用を実施している企業（統合 DB の運用企業と各施設から統合 DB へのデータ抽出システムを運営する企業等複数ある場合には、複数企業名とその担当業務）

データベースの保守・運用、データの抽出は（株）富士通が行っている。

5. データの個人情報・匿名情報の取り扱いに関する情報

1) データの匿名化の有無、匿名化を行う者

匿名化している。

<レセプトデータ>審査支払機関および（株）富士通が実施している。

<特定健診データ>保険者が実施している。

2) 匿名化の場合、匿名加工のためのデータ加工方法

レセプト上の個人情報（被保険者番号、生年月日等）から固定長の疑似乱数（ハッシュ値）を生成する。この操作を 2 度繰り返す。

3) 匿名化の場合、匿名個人データを連結可能な符号処理や情報源の異なる特定個人の情報を突合可能な仕組みの有無とその方法

有り。個人情報（被保険者番号、生年月日等）を基にしてハッシュ値を生成し、それを ID として用

いることで個人情報削除したレセプト情報等について、同一人物の情報として特定することが可能。

6. データベースに関する情報

1) インデックス DB※の採用有無

有。

2) 解析用 DB で採用している DB 技術 (cache, Oracle, SQLserver, Netezza, Hadoop など)

Oracle.

※インデックスデータベースとは

SS-MIX2「標準化ストレージ」は患者（患者 ID）を特定してから、当該患者の診療情報を検索することに特化した物理構造を採用している。したがって、例えば特定の診療日に該当する診療情報や、特定の診療行為に該当する診療情報等、複数の患者に跨った参照を行う場合や大量の診療情報を参照する場合にはシステムに多大な負荷を与えることとなる。このような負荷を軽減するとともに、標準化ストレージを活用するアプリケーションが上記のような検索機能を容易に実装できるようにするため、リレーショナルデータベースシステム等を利用し、物理構造を構成する値をデータベースに保持する、インデックスデータベースを規定する。

（「SS-MIX2 標準化ストレージ構成の説明と構築ガイドライン Ver.1.2c」より抜粋）

SS-MIX2 を利用していない事業においても、同様のインデックスデータベースを装備している場合は、「有」ときさいしてください。

7. データベースの現状と課題

1) データ収集の方法について

- ・一次審査のレセプトであるため、返戻が反映されていない。
- ・個人 ID をハッシュ化しているため、他のデータとのリンクが難しい。

2) データの内容（項目）および精度（正確さ）に関して

- ・主疾患等の傷病名の扱いについては課題がある。

3) データの活用に関して

- ・データ活用を推進するオンサイトセンターを設置し、今後拡大の予定である。
- ・第三者提供によるデータ利用を促進しており、既に 66 件の実績がある。
- ・民間からの要請に対してはオープンデータ（集計済みデータ）で対応する。

4) データの相互利用によるメリット

- ・DPC データでは入院中の医療行為がわかるため、相補的な関係にある。
- ・他のデータとリンクするためには新たなハッシュ値等のリンクできるしくみが必須となる。
- ・介護レセプトや死亡票などとリンクできると、医療提供の後の状況が把握できる。

1.8 小児慢性特定疾病登録管理データ運用事業

1. 担当部局

雇用均等・児童家庭局母子保健課

2. 事業の位置づけ、データ利用に関する情報

1) 根拠規定

小児慢性特定疾病の治療方法その他小児慢性特定疾病その他の疾病にかかっていることにより長期にわたり療養を必要とする児童等の健全な育成に資する調査及び研究を推進するため（児童福祉の一部を改正する法律第21条の4第1項）。

2) 事業の目的

小児慢性特定疾病に係るデータベースを構築し、研究者等に当該データを提供することにより小児慢性特定疾病の治療研究の推進を図るとともに、小児慢性特定疾病児童等の治療・療養生活の改善や小児慢性特定疾病に係る理解促進等に資する情報の入口となるポータルサイトを構築し、運用することにより、児童の健全育成に寄与することを目的とする。

3) 事業におけるデータ取得対象者への同意取得の有無と在りの場合の同意内容

あり（別添同意書参照）

4) 当該事業におけるデータ利用者（記載例：〇〇局〇〇課、事業参加医療機関、等）

成育医療研究センター

5) 当該事業以外での第三者データ利用の可否とその場合の利用法

研究機関

3. 事業の規模に関する情報

1) 事業（収集）開始時期（予定含む）

平成27年度中

2) データ保有期間（記載例：2010年から現在までの5年分）

永年

3) 保有するデータ量（患者数、データ件数）

14.8 万人／年（試算）

4) 参加医療機関

- ・ 指定小児慢性特定疾病医療機関
- ・ 成育医療研究センター

4. 事業システムに関する情報

1) システム基本構成

- ・ 医療機関において、小児慢性特定疾病の医療費助成の申請のために医師が患者の臨床データ等を入力すると、中央機関（成育医療研究センター）のデータベースに登録される。
- ・ 小児慢性特定疾病は 704 疾病（包括的疾患含めると 760 疾病）あり、登録される情報はそれぞれの疾病に合わせたものになっている。
- ・ 登録された情報は、管理され研究機関等に提供される。
- ・ 平成 27 年 1 月 1 日施行の児童福祉法の一部を改正する法律と合わせて、システムが構築され、平成 28 年度から本格稼働する予定である。
- ・ 年間 14.8 万件の登録を試算している。

2) システム運用を実施している企業（統合 DB の運用企業と各施設から統合 DB へのデータ抽出システムを運営する企業等複数ある場合には、複数企業名とその担当業務）

- ・ 今後入札により決定。

5. データの個人情報・匿名情報の取り扱いに関する情報

1) データの匿名化の有無、匿名化を行う者

あり。システム上で匿名化される。

2) 匿名化の場合、匿名加工のためのデータ加工方法

医療機関において入力された患者情報のうち、都道府県等には患者の個人情報を含んだ情報が登録される。一方個人情報を除いた患者の臨床データの登録管理を行う中央機関（成育医療研究センター）には、患者の氏名、生年月日などの個人情報を除いた情報が登録される仕組みになっている。

3) 匿名化の場合、匿名個人データを連結可能な符号処理や情報源の異なる特定個人の情報を突合可能な仕組みの有無とその方法

登録するデータの中に含まれる、出生地、出生体重等いくつかの登録項目を組み合わせることで、氏名や生年月日を除いた状態での匿名個人データを連結可能としている。

6. データベースに関する情報

1) インデックス DB※の採用有無

2) 解析用 DB で採用している DB 技術 (cache, Oracle, SQLserver, Netezza, Hadoop など)

※インデックスデータベースとは

SS-MIX2「標準化ストレージ」は患者（患者 ID）を特定してから、当該患者の診療情報を検索することに特化した物理構造を採用している。したがって、例えば特定の診療日に該当する診療情報や、特定の診療行為に該当する診療情報等、複数の患者に跨った参照を行う場合や大量の診療情報を参照する場合にはシステムに多大な負荷を与えることとなる。このような負荷を軽減するとともに、標準化ストレージを利活用するアプリケーションが上記のような検索機能を容易に実装できるようにするため、リレーショナルデータベースシステム等を利用し、物理構造を構成する値をデータベースに保持する、インデックスデータベースを規定する。

（「SS-MIX2 標準化ストレージ構成の説明と構築ガイドライン Ver.1.2c」より抜粋）

SS-MIX2 を利用していない事業においても、同様のインデックスデータベースを装備している場合は、「有」ときさいしてください。

7. データベースの現状と課題

1) データ収集の方法について

- ・自治体でデータ入力し、厚生労働省で収集し、国立成育医療研究センターのデータベースに登録されるしくみとなっている。
- ・平成28年1月からは、14疾患群の704疾患が対象となる。

2) データの内容（項目）および精度（正確さ）に関して

- ・旧システムでは手書きによるデータの提出であったため、誤りが多かったが、フォームを定める方式に移行したことにより誤りが減少できる。

3) データの活用に関して

- ・9割以上の患者からは同意を取得している。
- ・将来的には難病データベースとの統合を検討している。