

第一回提供サンプリングデータセットの仕様

2011.05.14

厚生労働省保険局総務課
保険システム高度化推進室

- (注1) 実際に提供するには本仕様に微修正を加える場合がありうるが、空欄化、削除および匿名化の基本的な方針は本仕様に則るものとする。
- (注2) 本仕様書に記載されている空欄化、削除および匿名化等の処理は、レセプト情報・特定健診等情報データベースに格納されている情報に対して行うものである。このため、実際に提供されるサンプリングデータセットにおいては「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」に述べられている内容と比較した場合、本仕様書の記載以外の処理が施されている場合がある。提供されるサンプリングデータセットの具体的な項目やその配置については、

「第一回提供サンプリングデータセット 抽出データレコードフォーマット」

もあわせて参照されたい。

(1. レセプトの種類)

- ・医科入院レセプト（以下、「医科入院」とする）、DPCレセプト（以下、「DPC」とする）、医科入院外レセプト（以下、「医科入院外」とする）、調剤レセプト（以下、「調剤」とする）の4種類とする。
- ・「医科入院外」は、抽出を行った後にハッシュ値1で紐付けられる調剤レセプトの抽出を行い、両者をあわせて提供する。

(2. レセプトの対象時期)

- ・平成23年10月診療分で、かつ平成23年11月、平成23年12月、平成24年1月に審査されたレセプトとする。
- ・「医科入院外」に紐付ける調剤レセプトは、平成23年10月診療分レセプトおよび平成23年11月診療分レセプトとする。このうち、平成23年10月診療分として提供する調剤レセプトは平成23年11月、12月および平成24年1月に審査されたレセプト、平成23年11月診療分として提供する調剤レセプトは平成23年12月、平成24年1月および2月に審査されたレセプトとする。

(3. 抽出手法)

- ・あらかじめ、「医科入院」では700,000点以上、「医科入院外」「調剤」では50,000点以上のレセプトを除外してから抽出を行う（「5. 高額レセプトの抽出」を参照）。

- ・「医科入院」「DPC」は、レセプト全体の10%を抽出して提供する。
- ・「調剤」はレセプト全体の1%を抽出して提供する。
- ・「医科入院外」はハッシュ値1全体の1%を抽出して提供する。
- ・「医科入院」「DPC」「調剤」においては、レセプト総数を母集団として抽出を行い、ハッシュ値は提供しない。
- ・「医科入院外」においては、ハッシュ値1の総数を母集団として抽出を行い、ハッシュ値1も提供する。また、これらのレセプトに紐付けることのできた調剤レセプトを、あわせて提供する。
- ・性別、および5歳刻みの年齢層においては、母集団と構成比率を変えないように抽出して提供する。

(4. 匿名化処理)

- ・診療報酬請求における基本マスターのうち、「傷病名コード」「医科診療行為コード」「医薬品コード」においては出現頻度の低いコード、「診断群分類」においては出現頻度の低い診断群分類番号について匿名化処理を行う。
- ・「傷病名コード」「医科診療行為コード」「医薬品コード」「診断群分類」においては、抽出したレセプトに出現するコードの総数のうち、少ないものから順に0.1%に達するまでのコードを、特定のコードにおきかえる。
- ・ただし、「医科入院」「医科入院外」における「医科診療行為コード」については、抽出したレセプトに出現するコードの総数のうち、少ないものから順に0.01%に達するまで、特定のコードにおきかえる。
- ・「傷病名コード」においては、出現頻度の低いコードは「9999999」に置き換える。同様に、「医科診療行為コード」においては「88888888」に、「医薬品コード」においては「999999999」に置き換える。
- ・「診断群分類」においては、出現頻度の低い診断群分類番号は「01010101010101」に置き換える。
- ・「医科診療行為コード」および「医薬品コード」において匿名化処理を行ったコードについては、同一行に点数が記されている場合は空欄とする。ただし他のコードとまとめて別欄で点数が記載されている場合は特段の処理をほどこさない。
- ・同様に、「DPC」において匿名化処理を行ったコードを有するレセプトについては、「包括評価レコード(レコード識別情報：HH)」における「入院期間区分別点数」「包括小計点数」、および「合計調整レコード(レコード識別情報：GT)」における「包括小計点数合算」を空欄とする。
- ・「記録されている点数から匿名化したコードを推定してはならない」という文言を明記し、匿名化された診療行為や医薬品が他の情報から推定されないようにする。
- ・平成23年10月診療分のうち、平成23年11月審査分のレセプトに記載された情報で集計を行って匿名化の基準を作成し、その基準に基づいて匿名化を行う。
- ・当初、平成23年10月時点のマスターにない「傷病名コード」「医薬品コード」「医科診療行為コード」「診断群分類」についても空欄化して提供することを検討していたが、技術的側面から整備が困難であるため、レセプト情報・特定健診等情報データベースに格納されている情報のうち、空欄化、削除および匿名化処理を終えたものについては、そのまま提供することとする。

(5. 高額レセプトの抽出)

- ・「医科入院」においては、「保険者レコード (レコード識別情報: HO)」における「合計点数」が 700,000 点以上、「医科入院外」「調剤」においては同様に 50,000 点以上のレセプトを最初に削除したうえで、抽出を行う。
- ・「医科入院外」に紐付ける調剤レセプトにおいても、「保険者レコード (レコード識別情報: HO)」における「合計点数」が 50,000 点以上となるものは提供を行わない。

(6. DPC レセプトにおける合計点数の処理)

- ・「DPC」においては、「医療機関別係数」の数値によって医療機関が一意に限定される可能性があるため、「外泊レコード (レコード識別情報: GA)」における「医療機関別係数」および「合計調整レコード (レコード識別情報: GT)」における「包括評価点数」「今月包括合計点数」は空欄とする。
- ・また、「合計調整レコード (レコード識別情報: GT)」における「包括小計点数合算」、および出来高部分の合計などで合計点数を一定程度推定できるため、「保険者レコード (レコード識別情報: HO)」における「合計点数」も空欄とする。

(7. その他処理を行う項目:「医科入院」「医科入院外」)

- ・「医療機関情報レコード (レコード識別情報: IR)」における「都道府県」「医療機関コード」は、空欄とする。
- ・「レセプト共通レコード (レコード識別情報: RE)」における「レセプト番号」「レセプト種別」「給付割合」「一部負担金・食事療養費・生活療養費標準負担額区分」「レセプト特記事項」「割引点数単価」は、空欄とする。
- ・「レセプト共通レコード (レコード識別情報: RE)」における「病床数」は、「1: 0~199」「2: 200~399」「3: 400~599」「4: 600~799」「5: 800~」とカテゴリー化する。空欄の場合は、そのまま空欄とする。
- ・「保険者レコード (レコード識別情報: HO)」における「保険者番号」「職務上の事由」「負担金額」は、空欄とする。
- ・「公費レコード (レコード識別情報: KO)」における「負担者番号」「任意給付区分」「負担金額」は、空欄とする。
- ・「診療行為レコード (レコード識別情報: SI)」「医薬品レコード (レコード識別情報: IY)」「特定器材レコード (レコード識別情報: TO)」「コメントレコード (レコード識別情報: CO)」における「負担区分」は、空欄とする。
- ・「臓器提供医療機関情報レコード (レコード識別情報: TI)」「臓器提供者レセプト情報レコード (レコード識別情報: TR)」「臓器提供者請求情報レコード (レコード識別情報: TS)」についての情報は、すべて削除する。

(8. その他処理を行う項目:「DPC」)

- ・「医療機関情報レコード (レコード識別情報: IR)」における「都道府県」「医療機関コード」は、空欄とする。
- ・「レセプト共通レコード (レコード識別情報: RE)」における「レセプト番号」「レセプト種別」「給

付割合」「一部負担金・食事療養費・生活療養費標準負担額区分」「レセプト特記事項」「割引点数単価」は、空欄とする。

- ・「保険者レコード（レコード識別情報：HO）」における「保険者番号」「合計点数」「職務上の事由」「負担金額」は、空欄とする。
- ・「公費レコード（レコード識別情報：KO）」における「負担者番号」「任意給付区分」「負担金額」は、空欄とする。
- ・「コメントレコード（レコード識別情報：CO）」における「負担区分」は、空欄とする。
- ・「患者基礎レコード（レコード識別情報：KK）」における「入院時年齢」では、5歳刻みで年齢を区分し、85歳以上は一括りにして提供する。
- ・「患者基礎レコード（レコード識別情報：KK）」における「出生時体重」では、100g単位で体重を四捨五入して提供する。
- ・「患者基礎レコード（レコード識別情報：KK）」における「Burn Index」では、10%単位で数値を四捨五入して提供する。
- ・「外泊レコード（レコード識別情報：GA）」における「医療機関別係数」は、空欄とする。
- ・「包括評価レコード（レコード識別情報：HH）」における「負担区分」は、空欄にする。
- ・「合計調整レコード（レコード識別情報：GT）」における「負担区分」「包括評価点数」「今月包括合計点数」は、空欄にする。
- ・「診療行為レコード（レコード識別情報：SI）」「医薬品レコード（レコード識別情報：IY）」「特定器材レコード（レコード識別情報：TO）」における「負担区分」は、空欄とする。
- ・「臓器提供医療機関情報レコード（レコード識別情報：TI）」「臓器提供者レセプト情報レコード（レコード識別情報：TR）」「臓器提供者請求情報レコード（レコード識別情報：TS）」についての情報は、すべて削除する。

（9. その他処理を行う項目：「調剤」）

- ・「薬局情報レコード（レコード識別情報：YK）」における「都道府県」「薬局コード」は、空欄とする。
- ・「レセプト共通レコード（レコード識別情報：RE）」における「レセプト番号」「レセプト種別」「給付割合」「レセプト特記事項」「医療機関名称・所在地」「一部負担金区分」は、空欄とする。
- ・「保険者レコード（レコード識別情報：HO）」における「保険者番号」「職務上の事由」「負担金額」は、空欄とする。
- ・「公費レコード（レコード識別情報：KO）」における「負担者番号」「任意給付区分」「一部負担金額」「公費給付対象一部負担金」は、空欄とする。
- ・「処方基本レコード（レコード識別情報：SH）」における「第一公費」「第二公費」「第三公費」「第四公費」は、空欄とする。
- ・「調剤情報レコード（レコード識別情報：CZ）」「医薬品レコード（レコード識別情報：IY）」「特定器材レコード（レコード識別情報：TO）」「基本料・薬学管理料レコード（レコード識別情報：KI）」における「負担区分」は、空欄とする。

(10. 提供形式)

- ・「医科入院」「DPC」「医科入院外」「調剤」のサンプリングデータはそれぞれ1枚のDVDに複写され、承諾された申出それぞれに対して順次提供する。
- ・上記のうち、複数種類のサンプリングデータ提供の申出があれば、複数枚のDVDにて提供する。1枚にまとめて提供することを行わない。
- ・「国保連固有情報レコード」(レコード識別情報：KH)は削除する。
- ・コード番号については、平成23年10月時点の各種マスターとの比較対照を行うことはせず、データベースに格納されている情報をそのまま提供することとする。