

いただいた御質問に対する回答について
 (平成 24 年 3 月 21 日開催 サンプルングデータセットの提供に関する事前説明会でいただいたご質問)

	質 問	回 答
1	<p>申出書作成に当たり、データの抽出方法を記載する必要はありますか。</p>	<p>サンプルングデータセットの場合は、あらかじめ用意したデータセットをそのまま提供することとなるので、詳細な抽出条件の記載までは必要ない。 ただし、サンプルングデータセットを用いることで可能となる研究内容かどうかがわかるように、研究内容・方法を記載することは必要である。</p>
2	<p>サンプルングデータセットは、今回説明のあったデータが全て提供されると考えてよいでしょうか。サンプルングデータセットから、さらに申出内容に基づいて抽出したものが提供されるわけではないと考えてよいですか。</p>	<p>同上</p>
3	<p>全数データ同様、コードの指定が必要か。</p>	<p>同上</p>
4	<p>DPC レセプトだけを抽出してもらえるか。</p>	<p>基本的には「医科入院」「DPC」「医科入院外」「調剤」それぞれのデータセットをそのまま提供することとしており、DPC レセプトだけを提供することは可能である。</p>
5	<p>サンプルングデータの分析結果を学術成果として公表することを期待されているでしょうか。</p>	<p>活用され、公表していただくことを期待しており、学会や学術論文をはじめ、HP 掲載等、何らかの形で成果が公表される必要がある。</p>
6	<p>セキュリティ要件について、同一の部屋(入退室管理などは行う)で別々の研究者が行う複数の研究で使用することはセキュリティ要件を満たさないと考えられるか。</p>	<p>利用が認められる利用者は、研究毎に申し出られた者に限定する必要があることから、仮にある研究のためにサンプルングデータセットの提供を受けた者であっても、あらかじめ利用者として申し出られていない別の研究目的のためにそのデータを活用することは認められない。したがって、同一の部屋で別々の研究者が行</p>

		<p>う複数の研究で使用する場合にも、それぞれのデータの利用者は、その研究毎にレセプト情報等の利用をあらかじめ認められた者のみに限定することが担保されている必要がある。この点が担保されているのであれば、一律にセキュリティ要件を満たさないこととはならない。</p>
7	<p>厚生労働省科学研究費での分担主任等でレセプトデータを使用している場合はサンプリングデータセットの提供を受けると、当該の研究からは離脱する必要があるのでしょうか。</p>	<p>サンプリングデータセットの利用がセキュリティ要件を満たす環境のもとで適切に使用される限りにおいては、離脱の必要はない。ただし、ナショナルデータベース由来のレセプト情報を他の情報と照合することは原則認めていないので、情報の取り扱いは、慎重に行っていただく必要がある。</p> <p>また、本データベースの提供を同時に受けられるのは1件としているため、仮に本データベースから提供を受けたレセプトデータの研究に従事している場合には、そもそもサンプリングデータセットの提供は受けられない。</p>
8	<p>抽出元母集団についての必要な情報（性、年齢層別の人数）などや、抽出の手順の詳細については予め、もしくはこちらからの求めに応じて公表されるのか教えてください。</p>	<p>抽出の手順についてはホームページに公表する予定としているので、研究成果を公表される場合等に活用されたい。ただし、母集団についての情報提供は現在のところ考えていない。</p>
9	<p>匿名化作業のために削除されたコード一覧を公表していただくことは可能か（可能なら、さらに何件のレセプトが該当したかの情報の公表がなされれば有用だが、可能か）。</p>	<p>削除されたコードの公表は現在のところ考えていない。</p>
10	<p>サンプリングデータ内の医科入院、DPC、調剤は、ひも付けされないとのこと説明でしたが、ハッシュ値など、研究者がデータ入手後に自分でひも付けすることのできるデータは提供されますか。</p>	<p>当初よりサンプリングデータセットにおいては紐付けを想定していないため、抽出はレセプト単位で行い、抽出されたレセプトにはハッシュ値を入れないこととする。ただし、医科入院外レセプトにおいて</p>

		<p>は調剤レセプトの情報を紐付けられる形式で提供するため、抽出はレセプト単位では行わずハッシュ値単位で行い、ハッシュ値を付与して提供する。</p> <p>すなわち、同じハッシュ値を持つレセプトはすべて抽出される。またこのハッシュ値に紐付けられる調剤レセプトも、ハッシュ値を残して提供する。したがって、以下の整理となる。</p> <p>医科入院レセプト：レセプト単位での抽出、ハッシュ値なし</p> <p>DPC レセプト：レセプト単位での抽出、ハッシュ値なし</p> <p>調剤レセプト：レセプト単位での抽出、ハッシュ値なし</p> <p>医科入院外レセプト：ハッシュ値単位での抽出、ハッシュ値あり、紐付けられる調剤レセプトにもハッシュ値あり</p>
11	入院と入院外をひも付けした上での提供は不可ということであったが、それぞれのデータにハッシュ値を付与されたデータは提供されるのか（申請者自身でハッシュ値でひも付けしてよいのかを知りたい）。	同上
12	入院外のサンプル患者のひも付いた入院レセを抽象して頂くことは可能でしょうか。又、同一患者が複数医療施設を受診した場合（つまり同一患者が複数のレセプトを発生させている）ひも付けることは可能でしょうか。	同上
13	「医科入院」「医科入院外」「DPC」「調剤」レセプトどしはひも付けできない、とのことだったが、「医科入院外」の場合はひも付けを考慮して「調剤」が2か月分付与されることになって	同上

	いるが、ひも付けできるのでしょうか。できないなら「調剤」レセプトはどう使用するのですか。	
14	サンプリングデータセットの審査評価項目は、全数抽出と同じ①～⑤ですか。探索的な研究を認められる場合、③は想定されないのではないのでしょうか。また、④、⑤は全申出で同じ条件となるので、①、②のみが審査対象となると考えてよいのでしょうか。	以下は、質問における「審査評価項目」を「（参考）優先順位付けのイメージ」に記載されている基準を指すものと考えての回答である。サンプリングデータセットであっても具体的な政策への反映が全く想定されないわけではないので、政策への提言（③）も基準となる。加えて、もし今後サンプリングデータセットから何らかの抽出を行ったうえでデータ提供を行う形式が整備された場合には、④、⑤も基準となりうる。
15	診療行為マスターにおける匿名化処理の基準について、頻度の少ない外科治療の出現回数は、必要に応じて緩和できる可能性はありますか。	今回の提供では、さらに匿名化の基準を緩和することは考えていない。
16	研究計画は全く変えても問題ないという説明もあったが、薬剤疫学的研究をデータの特性に鑑み、医療経済学的テーマに変えるなど領域を変えることも可能か。	研究計画を変更することは可能である。
17	「サンプリングデータセット」と「第二回全数」の両方の申し出を行う予定であるが、優先して行いたい分析が「第二回全数」の場合、「サンプリングデータセット」申し出において、その旨を記載することが考慮してもらえるか。サンプリングデータセットの提供内容が限られており、行いたい分析に取り組みないため、できれば「全数」の提供を希望したいと考えている。	現時点での予定では、サンプリングデータセットの審査を全数データよりも先行させることとしているが、仮に全数データの申出を優先したいとの意向が申出者にある場合には、サンプリングデータセットの審査を了した後、全数データの審査が終了するまでデータ提供を留保し、全数データの審査が承諾となればサンプリングデータセットの申出を取り下げ、全数データ審査が不承諾となればサンプリングデータセットの提供を行う、といったことも可能と考えている。

18	匿名化処理に関して質問します。たとえば A、B 医薬品を匿名化した場合、コードは同じコードが付与されるのか。同じコードが付与された場合、処方医薬品数が、元々は5種類あっても、2種類が同じコードになると3種類の処方とカウントされるようになるのか。	同じコードが付与されるが、処方されたという情報自体は匿名化したことを示すコードによって残されるため、処方医薬品数は3種類および匿名化した医薬品が同一のコードで2つカウントされることになる。
----	--	--