

複数の匿名データの作成・提供について

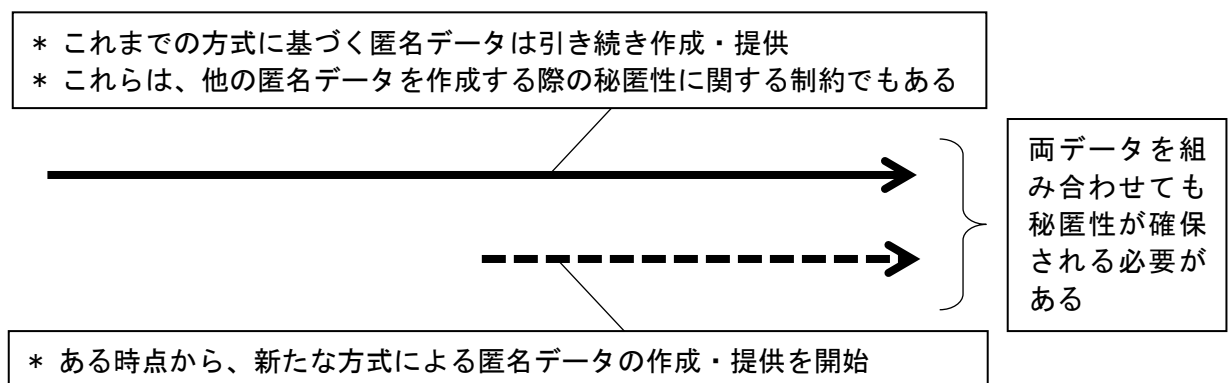
1. 背景・経緯・論点

- ・ 複数ファイルの作成が課題とされている。

※匿名データ作成の際の答申における課題

※匿名データ有識者会議における議論

- ・ アメリカ、カナダ、イギリス等では、個人単位と世帯単位などの複数のデータを作成（伊藤（2017）・伊藤（2018））。
- ・ これまでに作成した匿名データとの関係（イメージ）：



- ・ 2019年度の進め方（案）

⇒複数の方式に基づく匿名データの作成・提供について、諸外国の事例や先行研究などの調査を行ってはどうか。

⇒調査の結果を踏まえ、有識者の御意見も伺いながら、我が国の匿名データへの適用方法を検討

⇒2020年度以降に（※統計局のデータにより）検討・試算

2. 外国事例等の調査に当たっての論点

- ・ 有用性と秘匿性のトレードオフ
- ・ 提供項目の詳細化←→地域情報の詳細化
※地域情報は強力なキー情報となり得る（星野（2010））
- ・ ノイズに基づく方法の導入（地域情報の攪乱を含む）

- ・ 複数のファイルに対する秘匿性の考え方（複数ファイルの共通部分に特殊なレコードが含まれるか、データリンケージの実験など）

※既存の匿名データに新たな地域情報を付加する方法もありうる

（現状では大まかな地域（都市圏等）の情報

← 都市の（人口）規模に関する変数を追加する など）

4. 今後の予定

2019年6月～：匿名データ有識者会議における議論、

諸外国における事例のサーベイ

2020年：複数ファイルの作成方法の検討、試算・分析

参考文献：

- [1] 伊藤伸介（2017）国勢調査マイクロデータにおける匿名化の誤差の評価方法に関する一考察，中央大学経済学論纂，第57巻，第3・4号併号，pp189-209.
- [2] 伊藤伸介（2018）公的統計マイクロデータの利活用における匿名化措置のあり方について，日本統計学会誌，第47巻，第2号，pp77-101.
- [3] 星野伸明（2010）公的統計マイクロデータ提供制度の課題，日本統計学会誌，第40巻，第1号，pp23-45.