

❁ 医療職のための統計シリーズ

医療職のための学び直し—研究デザインから論文報告までの生物統計学の道標—

第19回 既存データの利用

ヨネクラ ユウキ
米倉 佑貴*

I はじめに

国や企業、研究者など、様々な実施主体により調査が行われ、データが収集・記録されている。また、医療機関における診療記録や医療費請求のためのレセプト、企業によって運営されているオンラインサービスやオンラインショッピングの利用記録など、能動的な調査以外にも、業務等で蓄積されているデータは多い。既存の調査のデータや業務等で蓄積されたデータは近年では電子化されたものも多く、研究のためのデータとして比較的容易に利用することができるものもある。今回はこうした、既存のデータの利用について、どのようなデータが利用できるか、利用の仕方などについて解説していく。

II 既存データの研究での利用

研究における既存データ、統計調査結果の用途は、論文や研究計画書の緒言や背景において研究テーマで扱う健康問題等の記述的なエビデンスを示すような文献としての利用と、統計解析、分析の対象としての利用に大別できる。前者については、計画書や論文を執筆する際にほぼ必須と言えるため、主要な指標、調査項目がどのような統計調査にアクセスすれば良いのかを知っておくことは重要である。後者については、冒頭に述べたように近年利用可能なデータが増加してきており、それに伴って既存データを再利用する二次分析の研究は増加してきている。自分の研究テーマに合致する既存データがいつも見つかるとは限らないため、二次分析を行うことができないこともあるが、二次分析が可能であれば、対象者の負担もかからない上、調査にかかる費用や労力、時間を大幅に削減す

ることができるという大きなメリットがある。したがって、新規に調査を計画する前に、既存のデータで自分の研究テーマに合致するものがないか探すと良い。完全に自分のテーマに合致するものが見つからなくても、近いテーマで行われた調査データを見つけることができれば、その調査のデザインや質問項目を参考にすることで、自分の調査の質を高めることもできる。

III 保健医療系の研究で使用できる既存データ

(1) 公的統計

公的統計とは、国の行政機関や地方公共団体が作成する統計のことである。公的統計は、統計調査を行うことにより作成される調査統計、業務データを集計することで作成される業務統計、他の統計が加算することで作成される加工統計に分類される。

国が実施・作成した統計調査・統計は、統計法に基づき調査を所管する行政機関がウェブサイト等で公開している。保健医療に関係する統計調査の例は、表1のとおりである。これらの統計調査結果は、政府統計のポータルサイトであるe-Stat (<https://www.e-stat.go.jp/>) から閲覧・入手することができる。e-Statでは政府統計の結果を検索してダウンロードできるほか、時系列、地域別の集計や、地域別の集計結果の図示など、結果を加工することもできる。

公的統計のデータは集計する前の個票データも一定の条件で利用することができる。公的統計の個票データの利用形態は「匿名データの利用」「オーダーメイド集計の利用」「調査票情報の利用」の3種類がある¹⁾。「匿名データの利用」は個票データを回答者が特定されないように調査票情報を加工した上で提供されるデータである。「オーダーメイド集計」は調査の所管

* 聖路加国際大学大学院看護学研究科准教授