

令和元年度 厚生労働行政推進調査事業費補助金
(政策科学総合研究事業(政策科学推進研究事業))

健康寿命の延伸に向けた地域保健と職域保健の連携に資する
データ項目及び分析方法の明確化のための研究
分担研究報告書

NDB 特定健診等情報から二次医療圏別に
特定健診調査項目を集計する際の課題について

研究代表者 加藤 源太 (京都大学医学部附属病院 診療報酬センター 准教授)
研究分担者 横山 徹爾 (国立保健医療科学院 生涯健康研究部 部長)
研究協力者 植嶋 大晃 (京都大学医学部附属病院 特定助教)

研究要旨

(研究目的)

NDB 特定健診データを二次医療圏別に集計する際に特有の課題を整理した。

(結果および考察)

2014 年度の NDB 特定健診データにおいて、現在の喫煙状況を問う質問では、26,419,765 もの回答数がみられたものの、二次医療圏別かつ性・年齢階級別の集計では集計結果が 10 未満となる事例が所々に認められ、他の集計値から加減算を行うことで集計結果が類推されないよう、追補的に複数の空欄を作成せざるを得なかった。公共性の高い議論に活用するための集計結果として公開する場合には、こうしたルール的基本的な方針を尊重しつつも、より柔軟な対応が可能となることが望ましい。また、郵便番号情報を元に受診者を二次医療圏別に割り付けたものの、その割り付けに漏れた事例が 685,095 例あり、全体の約 2.6%を占めていた。利用者による集計の精度向上を図る目的でも、こうした基本的な情報は利用者の解釈や努力にゆだねることなく、何らかの形で簡単に参照できるような仕組みがあることが望ましいのではないかと。

A. 研究目的

本研究では、地域保健と職域保健の連携の推進を見据えて、地域・職域の関係者が課題を把握するための基盤となるよ

うな指標の作成を目指し、レセプト情報・特定健診等情報データベース (NDB) のうち、特定健診等情報の調査項目から指標を作成するとともに、その

作業を通じて、NDB の利用可能性およびその限界を明らかにすることを目的としている。

理論的には、特定健診の調査項目であれば、当該データを収集しているわけであるから集計すること自体は容易なはずである。しかし、NDB データはそのデータ規模の大きさや複雑な構造等によって、必ずしも研究者や政策立案者が簡便に利用できるデータであるとは認識されていない。そうした研究や報告は有識者会議や各種学会、研究会等の場でしばしば議論されている¹のでここでは繰り返さないが、今回、NDB 特定健診データを二次医療圏別に集計するという貴重な経験を得たので、その特異性を踏まえて、NDB データを利用するにあたって指摘できる課題を本稿にて記すこととしたい。

B. 研究方法

本研究の目的を遂行するために二次医療圏情報を加えた 2014 年度分 NDB 特定健診データを用い、集計・分析を行った。

(倫理面への配慮)

NDB データの提供を受けるにあたっては、「レセプト情報・特定健診等情報の提供に関するガイドライン」を遵守した運用を明記した提供依頼申出を行い、厚生労働省による承諾を得た。またデータ提供を受けるに際しては所属機関における倫理審査での承認が必要であるため、京都大学医の倫理委員会にて申請を行い、承認を得た (R1119-2)。

C. 研究結果

NDB 特定健診等データを用いて二次医療圏別の集計を行う過程で直面した課題のうち、ここでは以下 2 つの課題について筆者らの対応および見解を述べる。

C1 : 「Cell size suppression」ルールによる限界

NDB データの分析結果を公表する場合には、個人が特定されるリスクを鑑みて、原則として集計値が「10 未満」となるような結果の公表は行ってはならないというルールが定められている²。特定健診は現在、受診率が 50%前後であり、2014 年度の NDB データにおける質問票の問 8「現在、たばこを習慣的に吸っている。(※「現在、習慣的に喫煙している者」とは、「合計 100 本以上、又は 6 ヶ月以上吸っている者」であり、最近 1 ヶ月間も吸っている者)」においては、26,419,765 もの回答数を有している。しかし、二次医療圏は 300 以上に分かれており、また医療圏間での人口差も大きいことから、たとえば「①はい ②いいえ」を回答とするような質問でそもそも回答に偏りが大きい質問の場合など、二次医療圏別かつ性・年齢階級別の集計を行った場合に集計値が非常に小さくなり、この公表ルールに抵触してしまう可能性を否定できない。実際、この問いにおいても、集計結果が 10 未満となる事例が所々に認められた。その結果、他の集計値から加減算を行うことで集計結果が類推されないよう、複数の空欄を追補的に作成せざるを得なかった。結果的に

そのような対応を取らざるを得なかったのは以下の集計区分においてである。

- ・岩手県久慈：65歳～74歳女性
- ・岩手県二戸：65歳～74歳女性
- ・秋田県北秋田：65歳～74歳女性
- ・秋田県横手：65歳～74歳女性
- ・長野県木曾：65歳～74歳女性
- ・長野県大北：65歳～74歳女性
- ・和歌山県有田：65歳～74歳女性
- ・和歌山県御坊：65歳～74歳女性
- ・島根県雲南：65歳～74歳女性
- ・島根県隠岐：65歳～74歳女性
- ・山口県柳井：65歳～74歳女性
- ・山口県長門：65歳～74歳女性
- ・香川県小豆：65歳～74歳女性
- ・香川県三豊：65歳～74歳女性
- ・長崎県上五島：65歳～74歳女性
- ・長崎県対馬：65歳～74歳女性
- ・熊本県宇城：65歳～74歳女性
- ・熊本県芦北：65歳～74歳女性
- ・鹿児島県出水：65歳～74歳女性
- ・鹿児島県熊毛：65歳～74歳女性
- ・沖縄県北部：65歳～74歳女性
- ・沖縄県宮古：65歳～74歳女性

今回はこの集計値を元にZ値を作成したため、いくつかの地域では年齢調整を行うことができず、Z値そのものが導けなくなるなどの影響が認められた。このこと自体は集計値を得る前の時点でz値を作成するなどすれば対応することは可能であったが、作業分担の都合上、そうした対応を取ることが難しかった。

C2：「住所不明事例」の発生

NDB 特定健診等データに含まれている受診者の住所地の郵便番号をもとに、二次医療圏情報に置き換える対応表を独自に作成してNDB 提供事務局に送付し、データ提供時に郵便番号情報を二次医療圏情報に振り替えてもらう作業を行うことで受診者を二次医療圏別に割り付けることを行った。ただし、今回その割り付けに漏れた事例が685,095例あり、受診者全体の26,419,765例の約2.6%を占めていた。

D. 考察

今回、10未満の集計結果が表に出ないよう、複数の個所を空欄にして対応することとしたが、同じような処理はNDBオープンデータにおいても認められる。ただし、今回の場合で言えば、集計結果が公表されず指標が得られない地域の関係者にとっては、このルールがかえって大きな不利益をもたらす恐れがある。本研究に限らず、公共性の高い議論に活用するための集計結果として公開する場合には、より柔軟な対応が可能となることが望ましい。

また、二次医療圏への受診者割り付けについては、こうした割り付けの不具合が発生する原因として、対応表作成時の不具合（郵便番号情報の確認不具合や二次医療圏情報の確認不具合）、あるいは受診者の誤記などが原因として考える。対応表の作成間違いについては、もちろん利用者の責任において細心の注意

を払って準備すべきであり、そうするに越したことはないが、一方で利用者がそれぞれ異なる参照元から異なる手法で準備をした場合には、仮に不具合なく準備ができた場合であっても、集計結果に揺れが発生する可能性がある。利用者による集計の精度向上を図る目的でも、こうした基本的な情報は利用者の解釈や努力にゆだねることなく、何らかの形で簡単に参照できるような仕組みがあることが望ましいと言えよう。

E. 結論

NDB 特定健診データを二次医療圏別に集計する際に特有の課題を整理した。特定健診情報においても NDB における「Cell size suppression」の公表ルールに抵触する可能性があることや、二次医療圏情報を直接利用できないことなど、研究成果や政策利用に直接影響を及ぼしうる課題があることが確認できた。

参考文献

1. 厚生労働省、「レセプト情報・特定健診等情報データの第三者提供の在り方

に関する報告書」

(<https://www.mhlw.go.jp/stf/shingi/2r9852000002s0h8.html>、2020年9月1日確認)。

2. 厚生労働省、「第5回レセプト情報等の提供に関する有識者会議 資料3」(<https://www.mhlw.go.jp/stf/shingi/2r98520000013uoy-att/2r98520000013utz.pdf>、2020年9月1日確認)。

F. 健康危険情報

なし

G. 研究発表

1. 論文発表
なし
2. 学会発表
なし

H. 知的財産権の出願・取得状況

なし