

厚生労働科学研究費補助金（循環器疾患・糖尿病等生活習慣病対策総合研究事業）：

「非肥満者に対する保健指導方法の開発に関する研究」分担研究報告書

分担研究名 保健指導効果評価データの効率的な収集に関する研究

研究分担者 岡山 明 所属 生活習慣病予防研究センター

要旨 保健指導効果の効果評価には、指導開始前の情報と指導開始後の情報が必要である。さらに指導対象候補全体の変化を観察することにより、指導を受けた人の特徴を考慮した分析が可能になる。本研究では研究対象者の保健指導効果を明らかにするために、健診結果が同じ人についてすべてのデータを継続的に収集し、比較可能な情報を収集するシステムの開発と運用を行った。

本システムの持つべき基本的な機能と、それをういた対象者選定とその後のデータ収集の在り方について検討し、対象保険者のデータ収集が円滑に行えるシステムの開発と運用を試みた。

A. 目的

特定健診保健指導制度では、健診項目や問診結果の規格が詳細に定められており、多数の保険者を対象としたデータ収集が飛躍的に容易となった。医療保険者は所有するシステムを操作することで対象者の健診データなどを容易に出力することができるようになった。

一方で個人情報保護の観点からは研究で収集するデータを制限し、匿名化したうえで収集するなど個人情報漏洩の可能性が低い方法が求められている。

保健指導を行う対象者は一般的に健診結果を持つすべての対象者ではなく、健診結果に一定の特徴を持つものである。したがって、保健指導効果を評価するにはこうした対象者の特性に基づいて対象者を抽出したうえで、連結可能な匿名化をし、匿名化番号をキーと

して各年度の健診データの出力を行う必要がある。さらに健診結果と保健指導データを結合するには医療保険者内で対象者の匿名化番号を照合したうえで、管理する必要がある。

本研究ではこうしたデータ管理の課題を整理し、研究に必要な情報を簡単な操作で出力可能なシステムを開発し、データ収集を可能とすることを試みた。

B. 研究方法

1. データフローの検討

医療保険者内のデータフローを考えると、その年度の健康診断が終了後、システム内にデータが投入されるまでに1～2か月かかることが多い。そこで、前年度の健診結果を活用し保健指導の対象者をあらかじめ選定することができれば、健診現場での対象者の選定や、保健指導の実施が可能となる。

また保健指導効果を検証するには、保健指導の対象となる疾患を持つ対象者を選定し、その中で研究参加に同意した人に保健指導を実施することになる。比較のためには抽出対象者全員のデータが必要となるが、健診受診者全員の情報は必要ない。

研究に必要なデータを整理すると、前年度の健診データと今年度の健診データ、さらにその翌年度の健診データが必要となる。今年度の健診結果を用いて保健指導の対象者を選定した場合は、当該年度の健診データと翌年度の健診データが必要となる。

2. 匿名化キーの設定方法の検討

医療保険者が用意可能な健診情報は年度別の情報であり、複数年度にわたるすべてのデータに特定の番号をつけて出力するには、対象者に固有の番号を付与し、同一レコード内に保持させることが必要となる。こうすることで、個人ごとに固有の番号を持つデータセットの作成が可能となる。

特定健診におけるデータ管理の方法と匿名化キーの設定方法について検討した。

3. データ収集システムの開発

データ収集システムは保健指導の有効性にかかわる研究班で作成したものをさらに機能追加して作成した。医療保険者内で経年的に健診データが管理できること、データの保管に大規模なデータベースソフトを要せず、ハードディスク内に格納可能なシステムである。PC上にはデータが残らず、すべて USBHDD 上で作業可能なため、データの保管が容易

な点が特徴である。

データ収集開発システムの要点について検討し、開発を行った。

4. データ収集支援

円滑なデータ収集のためには、わかりやすい操作方法と、各施設のサポートが必要となる。

操作手順書の整備や、サポートの体制整備を行って研究の円滑な進捗を支援することを試みた。

C. 研究結果

1. 匿名化キーの設定に関する検討

特定健診では使用されるキーが被保険者記号と番号であり家族単位で同じ番号を持つため、さらに性別と生年月日を用いて照合する必要がある。

この方法の問題点は双子児の場合同一の識別情報となるため健診情報がどちらのものか区別することができなくなることである。そのため同じ識別情報を持つ対象者をあらかじめ選定したうえで収集対象から除外しておく必要がある。研究の対象者は40歳以上であり、婚姻等により同一世帯である可能性はかなり低いため大きな影響はないと考えられた。

複数年度にわたる健診データをまとめて一つのデータセットとして出力することは困難なため、各年度の健診データを出力し、すべてに匿名化キーを付与することとした。匿名化キーは照合表を紛失した場合、大きな問題となるため、一方向関数を用いて匿名化キーを計算することで、万一の匿名化キー紛失の場合でも作成可能な仕組みを構築するこ

ととした。

具体的には記号・番号・生年月日・性別・カナ氏名の文字列からMD5を計算し、HEX文字列から数字情報のみを収集して作成した。こうして作成した文字列は有効数字を十分多くとれば対応する個人情報と一意に対応するランダムな順序となる。本研究では11桁の番号を割り当てた。試験的に行った検証では100万件以上のデータで行っても重複した匿名化キーは作成されなかった。

2. データ収集システムの開発

本研究では収集したすべてのデータを出力回収するのではなく、一定の基準で対象者を抽出し、データを回収可能な仕組みを構築する必要がある。

そこでデータ収集を容易に行うために、データ分析の仕組みを構築した。条件設定の際に、健診項目など複数条件を設定することができ、対象とする年度の健診結果で当てはまる人を抽出できるようにした。さらに任意のデータ項目を出力可能とした。

その結果年度ごとの健診結果の条件別抽出とデータの出力を区分することが可能となった。

最終的に、前年度の健診結果をもとに抽出した人について3年度分、今年度健診結果をもとに抽出した人については2年分をまとめて出力可能とした。健診データは将来のものを含め取り込み可能であり、同じ様式で収集管理されることを考えると、来年度以降のデータであっても出力可能と考えられた。

3. データ収集支援

データ収集のためには、不慣れな職員

や専門職が行う操作を支援する必要がある。またPCのOSが異なると操作方法が異なる、または操作できないなど様々な阻害要因が考えられる。担当者の負担をできるだけ少なくするために、データ管理システムの操作性を向上させるとともに、電話での操作支援の体制を整備した。これらを通じて研究に必要なデータの収集支援体制を構築した。

最終的に操作手順書を作成して各施設の支援を行うこととした（別添操作手順書参照）。

D. 考察

特定健診制度では、すべてのデータの標準様式が定められたことでデータ収集の可能性が飛躍的に高まった。しかし、文字コードが知らない間に変更される場合もある。またキーになる情報が台帳と健診データでは異なっているなど様々な要因があるため、匿名化データを幅広く収集することは困難である。

本研究ではこうした困難を克服するため、健診データを個人ごとに管理する仕組み、匿名化番号を一意に割り付ける仕組みを持つシステムをさらに改善して、特定の対象者について複数年度にわたるデータ収集を行うシステムを構築した。

この際、抽出に用いた情報とは独立して出力データを設定する仕組みを作成した。これにより、健診データを用いた保健指導の効果評価が容易になるものと考えられた。

E. 結論

データ収集のための環境整備の一環としてシステム開発を行い、保健指導効果評価のためのデータ収集体制を整備した。

参考文献

なし

F. 健康危機情報

なし

G. 研究発表

1: Okuda N, Itai K, Okayama A. Usefulness of a Short Dietary Propensity Questionnaire in Japan. *J Atheroscler Thromb.* 2017 Nov 15. doi:10.5551/jat.42226. [Epub ahead of print] PubMed PMID: 29142179.

2: Nakamura K, Watanabe M, Okuda N, Yoshita K, Kabayama M, Torii S, Kuribayashi T, Itai K, Kamide K, Miura K, Okayama A. The Influence of the Japanese Nationwide Cardiovascular Prevention System Health Guidance on Smoking Cessation Among Smokers: A Propensity Score Matching Analysis. *J Atheroscler Thromb.* 2017 Dec 2. doi: 10.5551/jat.42051. [Epub ahead of print] PubMed PMID: 29199202.

3: Koyama T, Yoshita K, Okuda N, Saitoh S, Sakata K, Okayama A, Nakagawa H, Miyagawa N, Miura K, Chan Q, Elliott P, Stamler J, Ueshima H. Overall nutrient and total fat intake among Japanese people: The INTERLIPID Study

Japan. Asia Pac J Clin Nutr. 2017;26(5):837-848.

doi:10.6133/apjcn.072016.11. PubMed PMID:28802293.

4: Turin TC, Okamura T, Rumana N, Afzal AR, Watanabe M, Higashiyama A, Nakao YM, Nakai M, Takegami M, Nishimura K, Kokubo Y, Okayama A, Miyamoto Y. Diabetes and lifetime risk of coronary heart disease. *Prim Care Diabetes.* 2017 Oct;11(5):461-466.

doi: 10.1016/j.pcd.2017.04.007. Epub 2017 May 22. PubMed PMID:28545843.

2. 学会発表

岡山明 健康偏差値による脳卒中・心筋梗塞ハイリスク者のスクリーニング法の開発

第76回日本公衆衛生学会総会（鹿児島）

H. 知的所有権の取得状況

なし