

## データ分析業務の効率化～がん検診事業のプロセス指標分析システムの開発～

高崎 光浩<sup>1</sup>、佐々木 和美<sup>1</sup>、北島 健一<sup>2</sup>、川原 康義<sup>2</sup>、古川 修一<sup>2</sup>、田中 丈晴<sup>2</sup>、池上 愛子<sup>2</sup>、松田 智大<sup>3</sup>

1：佐賀大学、2：佐賀県健康増進課、3：国立がん研究センター

### 背景

【背景】データに基づく政策形成への期待が高まる一方で、分析業務は煩雑である上、非定型的業務が多いため、担当者の手を煩わせている。ICTの利活用はデータの利活用を進める一助となり得る。

がん検診の精度管理については、厚生労働省の「がん検診に関する検討会（平成15年開始）」、「がん検診事業の評価に関する委員会（H19年開始）」等で検討され、がん検診に関わる機関（都道府県、市区町村、検診機関）の役割及び各々がとるべき手法が決定され、精度管理の基本的な考え方として位置づけられるようになった。がん検診の精度管理指標には、「技術・体制指標」「プロセス指標」「アウトカム指標」がある。

直接的な評価につながるアウトカム指標が判明するには相当な時間を要するので、「技術・体制指標」や「プロセス指標」によって評価されている。がん検診事業は大まかに「受診者の募集」「スクリーニング」「精密検査への誘導」「精密検査」「事業評価」のプロセスに分かれており、プロセス指標は、各プロセスが適切に行われているかを評価するための指標である。

（プロセス指標分析作業の現状）

プロセス指標分析に用いるデータは各市、町の担当者が入力する「地域保健・健康増進事業報告」（Excel形式、以下 報告用ファイル）で佐賀県に集約される。

報告用ファイルは地域保健・健康増進事業全般に関する報告ファイルであり、がん検診以外のデータも含まれている。がん検診のプロセス指標分析に用いるデータは複数のシートに分散して存在する。

各指標の算出方法は定式化されているが、県の担当者は報告用ファイルのバラバラのシートに散在

しているデータの必要な部分を手作業でコピー&ペーストで計算できる形に整理し直す作業に相当の労力と時間を要している。また、類似した操作の繰り返しも多いため誤操作も生じやすいなどの課題がある。

### 目的

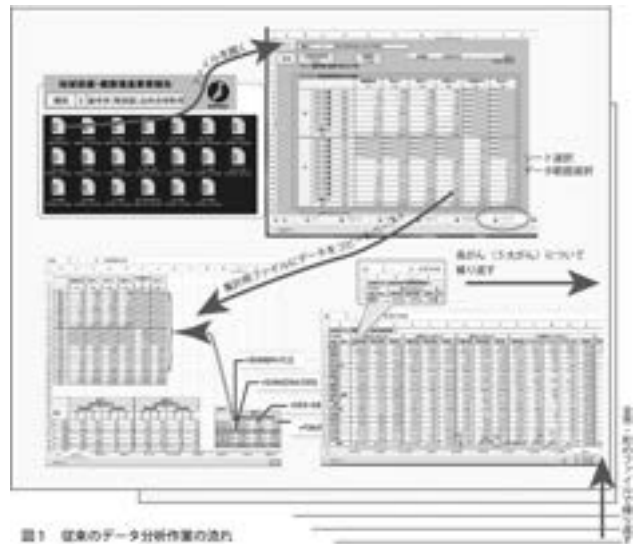
がん検診事業のプロセス指標分析業務を効率化するため、地域保健・健康増進事業報告ファイル（Excel形式）を取り込みデータベース化し、受診率、プロセス指標を簡便に作表、図示できるシステムを構築する。

### 方法

現行業務分析を行い、システムの到達目標を①システム導入による担当者の追加作業を最小限に抑え、②現行業務を効率化し、③必要であるが現状で実行できていない分析機能を追加する、の3項目として開発した。

### 結果

（業務の分析）



指標分析に用いるデータは、「地域保健・健康増進事業報告」（以下 報告ファイル）として各市、町の担当者が入力し佐賀県健康増進課に集約されたExcel形式のファイルである。報告ファイルは59のシートで構成され、がん検診のプロセス指標分析には34シートが関係していた。その中から必要なデータを担当者が手作業で別のExcelシートに

貼り付けて、Excel の計算機能や関数を用いて計算、作表、作図していた。指標算出に当たっては、データ範囲選択、コピー&ペースト、他シートのデータ参照等類似の操作の繰り返しが多く、操作間違い等が生じやすい状況であった（図1）。

（システム開発方針）これらの課題を解決するため、がん検診データを可視化するシステムを構築した。システムは報告ファイルをデータベース化する機能と、データベースを用いて各種指標を計算し、作表、作図する機能からなる web ベースのシステムとした。

プログラミング言語は php、フレームワークとして Laravel を用いた。

（報告用ファイルのデータベース化機能）報告用ファイル（Excel 形式）からがん検診関連のデータのみを抽出し、データベース化するシステムを構築した。対象となるシートやシート内のデータ位置（どこにデータが入力されているか）が年度によって異なっているため、一般的なデータベースシステム構築で行われる CSV ファイルのインポートでは対応が困難であった。そのため、一旦、xlsx 形式（Excel 2007 以降で用いられている形式）にファイル形式を変換してそれを読み込んでデータベース化した。この形式のファイルは内部構造は XML 形式に準拠しているため、Excel 以外の外部プログラムでも容易にデータを読み出すことができる。データ取り込みの際、シート内に入力されている年度や市町名のデータをプログラムで確認することにより、取り込みの際に別の市町や年度のデータファイルを間違えて取り込んでしまうのを防ぐことができる。

（分析機能）

データ参照（確認）、受診率、プロセス指標を地域、年度、がん種等を指定して作表グラフ表示ができるように開発した。

基本的な操作手順は、必要に応じて、「年度」、「がん種」、「市・町」を選択して表示ボタンをクリックすると希望の数表やグラフが表示される。

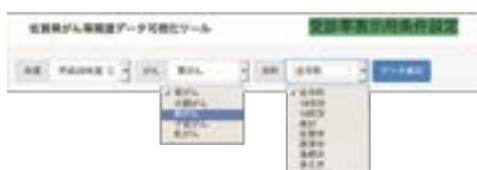


図 2 a 表示条件設定例

図 2 b 数値表の表示例

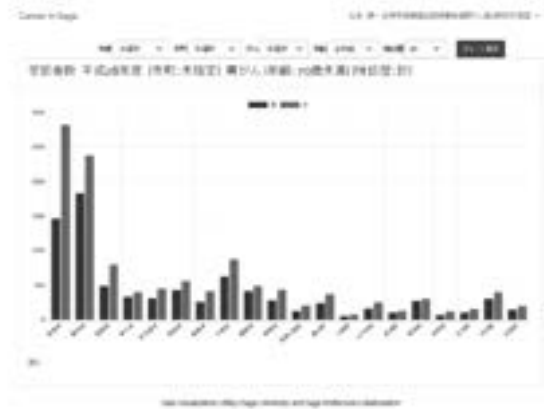


図 2 c グラフの表示例

### 考察とまとめ

各指標の算出方法は定式化されているが、複数のシートに散在しているデータの必要な部分を手作業でコピー&ペーストで計算できる形に整理し直す作業に相当の労力と時間を要していた。本システムにより全 20 市町の 1 年分の報告用ファイルが 2 分弱でデータベース化でき、表示条件設定後数秒で結果が得られるようになった。手作業での分析は時間を要するだけでなく、使用データの範囲設定、多くのミスが発生する可能性が存在する。システム化することによって元データの正しさを確認するだけでよく、担当者は分析データの評価・考察に専念できる。

当演題発表に関し、開示すべき COI はありません。