

地域がん登録集計の利用促進に関する試み

三上春夫*

地域がん登録室がデータ利用について受ける依頼中最も多いものが罹患数、罹患率の集計表である。登録室では報告書の形態で集計表を提供しているにも関わらず、集計表のパラメータについて依頼目的に特化した表が得られないため登録室として個別対応作業が必要となる場合が多い。

また報告書では年次の確定値としての精度を高めるため、補充登録（遡り調査）など複数の入力源情報を集めた後に集計を行うこととなり集計の遅れが問題となっている。がん登録においても即時性が求められることから、インターネット（Web）上で定期的にデータを更新できる Web 集計は有力な情報公開の手段と考えられる。

集計の仕組みは以下の通り。

地域がん登録データベースから、性別・罹患時年齢・罹患年・罹患時住所（市区町村）・診断項目を含むデータセットを抽出する。この際 DCO など精度指標データを市区町村単位で集計し、精度指標基準への達不達を個別データに付加する。データセットをデータベースソフト（MySQL）に取り込み、索引を付加する。集計ソフトウェアは Java により開発した。

集計は罹患数と罹患率について行い、本システムが Web ブラウザに対して画面を構成しユーザーに返す。ユーザーは希望するとおりの集計表を構成するために、性・罹患時年齢（5 歳階級）・罹患年（1 年集計、

3 年集計、5 年集計）・住所（市区町村、保健所区分、医師会区分、2 次医療圏区分）・診断（約 20 部位）について集計表の項目構成（縦軸横軸へのセット）と希望する項目内カテゴリーを選択する。

集計結果はカンマ区切り（CSV）ファイルでダウンロード可能で、グラフ化等の加工はエクセル等外部ファイルで行う。現在集計結果のマップ化（県地図の塗りわけ）機能を開発中である。

表示の際、個人情報保護のため規定数（可変）以下の罹患数を非表示にする機能（個人情報マスク機能）がある。また集計表の設計はユーザーに委ねられることから、特定のセルに精度基準不達のデータ割合が規定値（可変）以上占める場合にハッチングや色を変更するなど警告する機能（精度アラート機能）を有する。

今後の課題として

1) データクリーニング

データセット中のコーディングの誤りや入力の誤り、二重登録の確認等クリーニング作業が随時必要である。

2) 集計項目の検討

当初予定項目には病理診断、臨床進行度、発見経路、診断方法等が含まれていないが、集計項目として検討する必要がある。

*千葉県がんセンター研究局疫学研究部

〒260-8717 千葉県千葉市中央区仁戸名町 666-2
