

## 第1節 地域がん登録標準データベースシステム

標準データベースシステムは、都道府県で実施される地域がん登録事業において、登録業務の標準化、効率化、品質管理、ならびに登録資料の有効活用を実現するために設計されたデータベースシステムである。「がん罹患・死亡動向の実態把握の研究」班において決定された地域がん登録の標準方式に則り、放射線影響研究所情報技術部にて開発をおこない、山形県地域がん登録をモデル地区として導入、試験、評価を繰り返し、同研究班が定めた「目標と基準」の実現に必要なデータベースシステムとしての基本機能の実装を完了した（表13）。

平成17年度1県（山形県）、平成18年度5県（愛知県、福井県、滋賀県、青森県、広島県）、平成19年度5県（熊本県、山口県、愛媛県、山梨県、兵庫県）、平成20年度2県（栃木県、群馬県）、平成21年度5県（茨城県、香川県、北海道、徳島県、沖縄県）、平成22年度8県（大阪府、石川県、新潟県、高知県、京都府、長野県、島根県、福島県）、平成23年度8県（佐賀県、秋田県、静岡県、三重県、和歌山県、奈良県、福岡県、大分県）、平成24年度3県（鹿児島県、埼玉県、東京都）と、運用が開始され、5県（岩手県、宮城県、鳥取県、岐阜県、宮崎県）が導入準備を進めている。

### 1. 開発方針と特徴

開発方針として、操作が容易であり、かつ、専従の情報システム技術者がいない場合でも、データベースシステムの保守・運用が簡単にできることを第一とした。登録実務担当者の声を聞きながら、操作性の改善が常に行われている。システムの更新、保守、管理についても、特別な知識・経験を必要とすることはなく、地域がん登録室の責任者が容易に実施できる。

地域がん登録が取り扱う登録情報の特性を考慮し、限られた予算の中で、安全性、信頼性の高いシステムの実現を基本的条件とした。地域がん登録業務専有の閉じられたネットワークの中で、サーバー・クライアント型の構成とし、サーバー本体は登録室の運用体制や登録件数に応じて数種類の機器構成から選択可能である。

標準方式の検討・決定が進むにつれ、システムの機能は常に拡張・拡大される。機能の追加・更新には、モデル地区において十分に検証され、あわせて、システムの機能に則した運用手順が検討される。標準データベースシステムの機能追加や各種定義表の更新に伴う費用追加等は発生しない。

がん登録データの集約作業や統計ファイルの作成において、標準データベースシステムでは国際がん研究機関・国際がん登録学会（IARC/IACR）の基準に完全に準拠して行なわれる。従って、このシステムを導入する地域においては、データの品質管理および集約作業が統一された基準に則って行なわれる。わが国の地域がん登録を通して、品質の高いがん罹患・死亡動向の実態を把握する上でこの意義は極めて大きい。標準データベースシステムを用いた登録作業手順に関する教育・指導が提供されることも、標準データベースシステムを採用した地域がん登録事業にとっては大きな利点となることであろう。

### 2. 導入までの手順

「がん罹患・死亡動向の実態把握の研究」班では、第2期（標準化推進期：平成19～21年度）においては「標準データベースシステムの普及と適正な運用」を最優先課題として支援していく方針を定めた。それを受けて、2006年8月に「標準データベースシステムの導入要件」を決定した。第3次対がん総合戦略研究事業の終了を見据え、安定したシステム維持体制の構築のため、国立がん研究センターは、平成22年9月に研究班より標準データベースシステムの譲渡を受け、運営費交付金を利用してシステムの運用と導入を支援することとなった。各県からの標準データベースシステム導入依頼を受け、標準データベースシステムを適正に導入し、きちんと運用していくために、必要な体制・準備について技術支援をしている。すなわち、標準データベースシステムを導入するために、必要な準備と手順を相談する「導入支援」、導入後は標準データベースを適正に運用・管理をおこなう「運用支援」の二つに重点が置かれている。

### 第3章 標準データベースシステム

導入準備状況調査では、国立がん研究センター及び標準データベースシステム導入・運用支援機関から現地を訪問し、導入準備の支援をおこない、導入時研修、集約時研修を同じく現地で実施する。導入後も、メーリングリスト、WEB サイト、研修、導入後現地調査による指導、蓄積されたがん登録情報の精査による品質管理の状態などを基に、適正な運用支援を実施する。

図7に導入までの手順を示した。移行するがん情報がある場合は、現在進行しているがん情報の登録作業の停止や、過去のがん情報の原票へ戻っての確認作業や修正作業等が発生するため、導入計画の作成やその実施にあたって、より長期間に渡る綿密な計画を策定する必要がある。

表 13. 標準データベースシステムの開発方針と特徴

開発方針	
1	操作が容易であり、情報システム技術者がいなくても保守・運用できる。
2	限られた予算の中で、安全性、信頼性の高いシステムを実現する。
3	機能の拡張が容易に行なえ、機能追加や各種定義表の更新に伴う追加の費用を要しない。
特徴	
1	「がん罹患・死亡動向の実態把握の研究」班により定められた標準登録票項目及びモニタリング項目を完全に満たしている。
2	入力間違いを防ぐために2回入力（ダブルエントリー）を必須としている。
3	不明、推定の日付と完全日付の区別ができる。
4	姓名、住所、生年月日については、同一人物でも複数のデータを持つ。
5	姓名を漢字で持ち、異字体を類似辞書として持つので、個人照合の際に候補の幅が広がる。
6	一腫瘍一件への集約作業はその殆どが機械的に処理される。
7	集約作業や統計ファイルの作成は、IARC/IACRの基準に準拠して行なわれる。
8	国立がん研究センターから提供される人口動態統計資料を使用し、標準集計表を自動作成する機能を持つ。
9	集約されたデータだけでなく、集約に使用した全ての票データも保持している。
10	入力時のデータも履歴として保持している。
11	各登録で更新可能な定義テーブルメンテナンス機能を持つ。（各登録室で更新可能な定義テーブルは、住所、医療機関、診療科に限定）
12	定義テーブルはアプリケーションプログラムからでも参照可能であり、確認のために使用できる。
13	実務者間での取決め事や伝達事項を標準DBS内で交換する掲示板機能を持つ。
14	端末IDで入力データが管理され、端末機器が複数でも入力されたデータが競合しない。
15	実務者のパソコンからWebを介してデータベースサーバに接続し、手動によるデータベースのバックアップやリストア、またバックアップファイルを外部メディアへ保存することなどが可能である。
16	研究班で決定された最新情報への更新に対応し、導入した登録室での同レベルでのシステム保守が可能である。
17	機能の追加・更新には、モデル地区において十分に検証されたシステムが、運用手順とともに提供される。
18	標準データベースシステムを用いた登録作業手順に関する教育・指導が提供される。

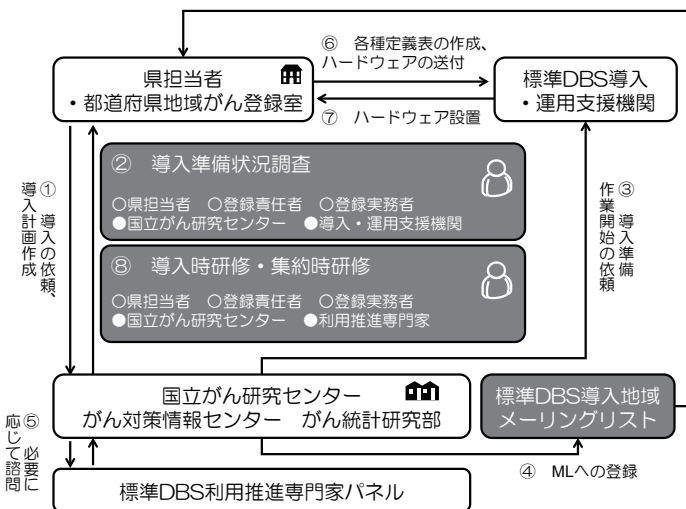


図7. 標準データベースシステム導入までの手順

## 第2節 標準データベースシステムによる標準作業手順

### 1. 標準データベースシステムによる標準作業手順

登録室に収集された罹患情報類（届出票による届出、標準データベースシステムは、これまで地域独自に行われてきた、情報の収集から登録・集計までの地域がん登録に関わる作業手順・工程について標準を示し（標準システムの構築）、結果として、地域比較可能な品質の高いがん罹患登録の実現を支援することを目的として、第3次対がん「がんの実態把握研究」班によって開発された。

標準データベースシステムの特徴は、がん患者についての個人識別情報、腫瘍情報をコンピューターに入力し、データの有効性と整合性について自動チェックを行い、実務者が確認、修正したあとに最終的にデータ登録されること、多重がんやその他の重複情報の集約は半自動で行われ、最終的な罹患集計用のデータセットは自動作成されることである。ただし、地域がん登録の実作業には完全自動化に適さない部分も存在し、その点は標準データベースシステムの運用方法自体を標準化することで補う必要がある。「がんの実態把握研究」班では、各県における運用から、効率的で精度が高く、比較的人的・物的資源の少ない県においても実現可能な方法を標準作業手順としてまとめ、平成24年3月に冊子として刊行した。

### 2. 罹患情報の受理

登録室に収集された罹患情報類（届出票、出張採録、院内がん登録電子ファイルなど）および、死亡転写票は、コンピューターに入力する前に点検、準備が必要である。コンピューターに入力した後から容易に原票にさかのぼれるように、一件毎の情報に対してシリアル番号を付与する。シリアル番号を届出票などの一件ごとの情報に付与してからコンピューター内に登録するまでに時間差を生じることがあるので、複数の実務者が作業している登録室ではシリアル番号を重複して発行しないように、新

たにシリアル番号を付与した時点で最大の番号とその日付を別途記録しておく必要がある。罹患情報類については、実務者が目で見て内容の点検を行い、漢字姓名、生年月日、診断時住所、罹患部位、診断日のような重要な欠落がある場合には届出医療機関に問い合わせする。また、罹患情報類が登録室に届いた日付を記録しておく。また、届出票は内容的にデータベースに登録しないこともあるので、届出件数とデータベース内の登録件数は異なってくる。届出票件数や死亡転写票の件数は、事業報告用として別に管理する必要がある。

### 3. コーディング

届出票には実務者が対応するコードを書き込んでいく。電子ファイルデータについては、データインポート可能な状態か、データ型、コードについて点検を行う。死亡転写票については、悪性腫瘍の記載のある票（がん死亡転写票）と記載のない票（非がん死亡転写票）に分類し、がん死亡転写票については対応する部位コード、組織形態コードを書き込む。がん死亡と非がん死亡の区分は、手引き詳細版「死亡転写票の収集と確認」に従って行なう。

### 4. データ入力

各々の票の入力画面に従って情報をコンピューターに入力していく。誤入力の防止のために別人による2回入力（ダブルエントリー）を基本とする。がん死亡転写票は、死亡転写票の記載内容をそのまま入力した後、記載内容からがん情報として登録できる部分をコーディングして入力する。非がん死亡転写票は既に登録されている患者ががん以外で死亡した場合を想定し、患者識別指標のみ入力する。それぞれ2回目入力後、1回目に入力した内容と比較し、確認、修正を促す点検プログラムを実行する。入力した情報について票内における項目単体チェック、項目間チェックプログラムを実行する。電子ファイルデータをインポートする場合は、この票内チェックからデータベース作業が開始する。

#### 5. コンピュータによる照合

入力されたデータの中から照合に必要な項目をサーバーに転送するプログラムを実行する。転送後、コンピューターによる照合プログラムを実行する。照合結果が同一人物候補リスト（同定リスト）として返されるので、実務者が目で見て同一人物か否かを判定する。判定した結果をコンピューターに入力する。この作業を通じて、既登録患者の場合は既に付与されている個人識別番号が今回登録された情報にも付与され、新規患者の場合は新規に個人識別番号が付与される。

#### 6. データ登録

照合によって付与された個人識別番号を元に、データベース内に保管されている既登録データとの間で項目間チェックを行うプログラムを実行する。確認、修正した後、データ登録を実行する。今回入力した情報が、個人識別番号をキーとして患者単位でデータベース内に登録保管される。

#### 7. データ集約

標準データベースシステムでは、登録室に収集した情報を原則としてすべてコンピューターに入力登録保管するので、データ登録した状態では一患者について一腫瘍一登録の状態ではない。任意の時点で多重がんの判定を行い、重複情報の取捨選択をして、一腫瘍一登録のデータセットを別途作成する必要がある。標準データベースシステムではこの作業がルールに基づいて半自動的に行われる。現時点の集約ルールでは、約90%前後の登録が自動的に処理される。自動的に集約されなかったデータについては、実務者が目視集約画面に従って、いずれかの情報に確定するか、登録内容を修正することで最終的にすべてのデータを集約する。（集約の対象は、前回集約のプログラムを実行した後から新規に登録されたデータと、保管されたデータの中で変更が加えられたデータのみ選択される。）多重がんの判定ルールは、IARC/IACRのRecording ruleに従っている。データを集約したあとは、集約したことによってデータに項目間矛盾が生じていないかを点検するプログラムを実行する。

#### 8. 遡り調査

一患者について一腫瘍一登録に集約したデータセットの中で、少なくとも罹患集計対象年で、人口動態死亡(小)票からの情報でのみ登録されている腫瘍について、死亡診断書を作成した医療機関に情報提供を促す遡り調査を実施する。標準データベースシステムは、遡り調査対象者の抽出、遡り調査票の作成などの遡り調査を支援する機能を備えている。調査終了後、結果をデータベースに登録し、再び集約プログラムを実施する。

#### 9. 生存確認調査

死亡日が確認されていないデータについて、罹患年から5年後等の生死状況について確認する調査を実施する。生存確認調査のうち、最も信頼性の高い方法である住民票照会は、調査対象者について市町村に住民票（除票）の閲覧依頼し、生死、死亡日、転出、転出年月日の情報を返してもらう調査である。標準データベースシステムは、生存確認調査対象者の抽出から結果登録までの作業を支援する機能を備えている。また、非がん死亡との照合にも対応している。

#### 10. 統計資料用データセットの作成

地域がん登録のデータは日々更新されるため、年に一度一定の時期にデータセットを確定し、その確定したデータを罹患集計や統計解析に用いる（統計値の固定のため）。標準データベースシステムでは、年間利用する統計資料用データセットは、集約済みのデータセットから多重がんの判定をIARC/IACRのReporting ruleに従った腫瘍を自動的に抽出するプログラムを実行して作成され、それを外部ファイルとしてアウトプットできる。登録室では、そのファイルを統計ファイルとして保管、利用する。

#### 11. 罹患集計の作成、外部提供用データの作成

これらの地域がん登録室の機能は、標準データベースシステムでは統計ファイルを用いて自動作成される。主たる外部提供用データ（全国がん罹患モニタリング集計用、国際がん罹患調査用）については、提出項目と区分に自動変換したファイルが作成される。



## 第3節 標準データベースシステムにおける ロジカルチェック

### 1. はじめに

地域がん登録の作業では、情報入手時の受付整理、コーディング、登録、追加、修正、データ更新等のいずれの作業段階においても人手による作業が必要で、作業者の単純ミス（記載漏れ、入力ミス）や知識・経験不足・誤解によるコード漏れが生じやすい。これらのミスを完全に防ぐことは不可能だが、標準データベースシステムでは、登録作業の各段階において、その段階に応じたロジカルチェックを組み込み、より正確な登録データを蓄積し、作業者による目視や読み合わせ等による確認作業を軽減することを目標としている。標準データベースシステムのチェックに用いる有効なコード範囲やルールについては、標準データベースシステム導入モデル地域（山形県地域がん登録）による実データの検証に基づき、第3次対がん「がんの実態把握研究」班で検討・決定された定義を採用している。

ここでは、標準データベースシステムで実装されているチェックの種類、そして実際の登録作業の流れに沿って標準データベースシステムにおけるチェックが、どのように行われているかを説明する。

### 2. ロジカルチェックの種類

#### (1) 項目単体チェック

個々の項目について、必須項目の漏れ、データ型や桁数、有効コード範囲かどうか、存在する日付かどうか等をチェックする。

標準データベースシステムには個々の項目のデータ型や桁数等の属性を格納するテーブルマップ情報テーブルがあり、チェックに使用される。テーブルマップ情報テーブル内の「データ型コード」では、数値、英数字、漢字、日付型かどうかを定義し、個々の項目におけるコードミス・入力ミスを防ぐ。また、「値チェックフラグ」が、「1:コードテーブルでチェック」ならば、コード定義を別テーブルでもち、「コードテーブル名」でチェックに使用する定義テーブル名を指定し、定義テーブルによるチェッ

クを可能にする。

項目単体チェックで用いる定義テーブルには、データコード情報、住所コード、医療機関コード、医療機関診療科コード、ICD-O-3 局在 (T) コード定義テーブルがある。

#### (2) 項目間チェック (票内・テーブル間)

複数の項目の関係において矛盾がないかをチェックする。日付項目間の順番が妥当か（生年月日 ≤ 診断日 ≤ 死亡日）、進展度と性状が妥当か、性別と部位・組織型に矛盾がないか、部位と組織型の組み合わせに矛盾がないか等のチェックを行う。

マスタ登録後、性別は基本情報テーブルに、部位・組織のがん情報は登録マスタテーブルに保管される。マスタ登録後のデータ更新時のチェック、もしくは、集約データを作成する前のタイミングでかけるテーブル単位の一括チェックは、テーブル間で項目の矛盾をチェックする。標準データベースシステムにおいてどのような項目チェックがかっているかの詳細については、定義テーブル同様アプリケーション内から随時参照可能である。

がん情報における部位、組織型に関する事項は、IARC/IACR のロジカルチェックに準拠した定義テーブルを採用している。項目間チェックで用いる定義テーブルには、項目単体チェックにも用いる ICD-O-3 局在 (T) コードの他、ICD-O-3 形態 (M) コード、ICD-O-3 組み合わせ定義テーブルがある。

### 3. 登録作業内におけるロジカルチェックの タイミング (図8)

### 第3章 標準データベースシステム

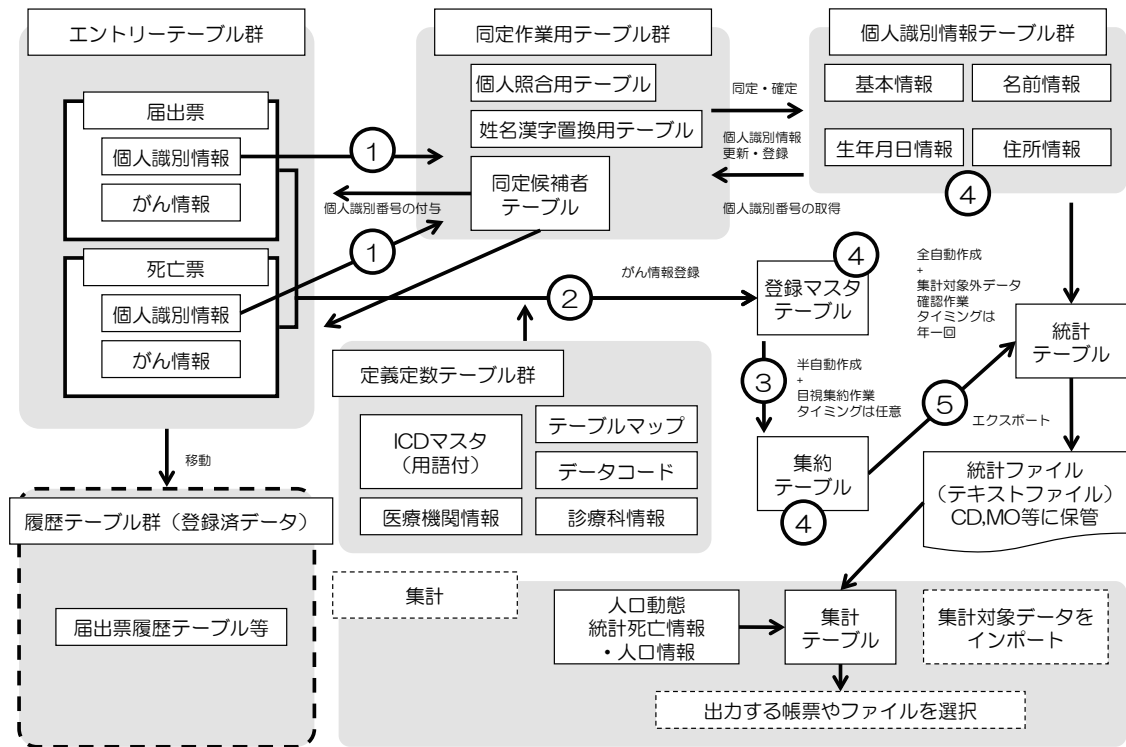


図 8. 標準データベースシステムにおけるロジカルチェックのタイミング

チェック①	届出票入力後（照合前）の票内における項目単体チェック、項目間チェック
チェック②	照合後の既登録データと入力データとの項目間（テーブル間）チェック
チェック③	登録マスタに保管された票データを元に、半自動（自動集約 + 目視集約）で 1 腫瘍 1 件に集約されたデータの項目単体チェック、項目間チェック
チェック④	（更新時） 随時、項目単体チェック、項目間（テーブル間）チェック （一括チェック） （1）集約データ作成前において、登録マスタテーブル、個人識別情報テーブル群をそれぞれの単位で、チェック①②と同じチェック （2）統計データ作成前においては、集約テーブル全体をチェック⑤と同じチェック
チェック⑤	統計テーブル作成前の集約データの最終的なチェック