

福岡県の地域がん登録事業における地方衛生研究所の役割

篠原志郎, 片岡恭一郎, 甲原隆矢

1. はじめに

我が国のがん死亡数が1981年に死亡原因の第1位となって以来、がん診断、検診、治療、予防等は国民的な関心事となってきた。1984年に国家事業として「対がん10カ年総合戦略」が立てられ、1994年に「がん克服新10カ年戦略」として引き継がれている。がん克服にはがん患者に関する正確な基礎情報が必要であるが、がん登録の歴史をみると、今世紀はじめよりイギリス、ドイツ、オランダ、スペイン等において始められた。当時、がんの拡がりについて病因学的研究に不可欠の資料として、がん患者すべての登録を試みているが成功していない。近代的がん登録は1929年、ドイツのハンブルクにおいて3人の看護婦の出張採録による登録からスタートした。1940年代にはアメリカ、カナダ、デンマーク、イギリス等で普及確立し、今日、世界のさまざまな地域に約200の地域がん登録が存在するといわれている¹⁾。地域がん登録は人口の確定している地域を対象としているので、自治体単位となっている。わが国の地域がん登録は老人保健法の制定(1983年)に伴って、各都道府県が行う事業とされ、がん受療状況の実態把握を通じて罹患率、有病率、生存率等の算出による疫学研究を始め、各市町村が実施しているがん検診事業の評価にも期待されるものである。すでに1990年、25道府県2市において実施²⁾されており、更に、普及拡大していく傾向にある。しかし、一方では、がん登録事業の経験年数を重ねるに従って、データ量の増大とその処理の煩雑化が相乗し、事業としてのがん登録運営の難しさがあるといわれている。福岡県の地域がん登録事業も、このデータ量増大を機会に旧システムの経験を生かし、新たにシステム構築することとなった。その際、運営に当たっては業務分

担を明確化し、相互に協力しながら現在に至っている。地域がん登録事業システムのうち、当研究所はコーディング以後のデータ入力系、コンピュータ処理系及び集計等の出力系を主に分担している。この事業への参画を機に、データ入力系及びコンピュータ処理系における問題点の改善のため、処理中に生じたミスの種類の数量等を検討し、また、地方衛生研究所としての公衆衛生情報活動における、我々の現状と役割等について述べる。

2. 地域がん登録の方法

福岡県の地域がん登録では、まず、県内に発生したすべてのがん患者及びその疑いのある者の実態を把握するため、県保健環境部が(財)福岡県メディカルセンター(以下、メディカルセンター)に委託する。メディカルセンターは地域医師会、病院協会あるいは運営にあたって定められたがん登録委員会を通じて、新生物届出票を県内に所在する医療機関に配布する。該当医療機関の医師は、がん登録対象の疾患を診断したとき、新生物届出票に所要事項を記載のうえ、がん登録委員会宛提出する³⁾。がん患者の死亡情報は目的外使用申請により総務庁の許可を得て、保健所から死亡票が送られてくる。提出された新生物届出票及び死亡票を既に登録されたマスターファイルと照合し、同定作業の後、同一人は情報を結合して再登録していく。もし、死亡していることが確認され新生物届出票がない場合は、この新生物届出票と同じ様式の補充票と呼ばれる補充届出票に記載提出を依頼する。新生物届出票あるいは補充票(以下、届出/補充票)の記載内容は、氏名、生年月日、住所等の基本部事項と診断名、病理診断、診断方法、治療方法、入退院年月日等の詳細部事項からなっている。がん情報は暦年毎に集計し、罹患率、有病率等の基礎情報を求め、医療実態情報として還元するものである。

(福岡県保健環境研究所)

3. 経過及び現状

本県の地域がん登録事業は1984年から開始し、メディカルセンターに事業委託された。当時、メディカルセンターのコンピュータで磁気ディスクにデータファイル化し、端末入力処理による登録であった。年次を重ねるに従い、データ量の増大と処理煩雑化をきたし、1990年より当研究所のコンピュータシステムを使用することを前提に新しくシステム構築が計画された。2年間の基本設計、詳細設計を経て、1993年から本格的にデータベースによる運用を開始した。本システムの処理は入力/検査業務、登録業務及び集計作表業務からなっている。地域がん登録事業の概要については、図1のとおりである。県行政機関、メディカルセンター、保健所医師及び当研究所によって運営されている。福岡県が事業の実施主体であり、その一部をメディカルセンターに委託し、メディカルセンターは県医師会、市郡医師会、病院協会を通じて病院、その他の医療機関等へ届出/補充票の記載及び提出を依頼する。提出された届出/補充票は所在地、診断名等のコード化の後コンピュータに入力され、データベースとして蓄積される。業務分担処理における登録情報の流れと担当機関を図2に示す。この内容は次のとおり

である。がん登録事業全体の予算を始め、進捗状況の把握及び方針、報告書印刷等は健康増進課が主導し、届出/補充票、死亡票等の受付、点検・確認、問合わせ、謝金支払い、原票のファイリング、補充票の封筒詰め及び発送等の事務的業務並びに原票の提出医療機関名、診療科名、住所等の比較的簡単なコード化は、がん登録室（メディカルセンター内）の職員によって処理される。原票の診断名コード化（ICD-9）、組織診及び細胞診のコード化（ICD-O-M）、診断方法、治療方法等の医学的論理チェック、サマリー条件及び報告書作製等はがん登録室において複数の県保健所医師が処理する。当研究所では、コード化された原票のデータ磁気化のための業者委託手続き及びデータチェック、コンピュータ入力、医師修正後の入力、届出/補充票ファイル及び死亡票ファイルの修正・削除、マスターファイルの追加・修正・登録、集計ファイルの作成、基本集計帳表の作成、報告書図表及び関連原稿作成、更に、コンピュータハードウェア、ソフトウェアの維持管理等が業務である。1994年3月、コンピュータシステムをACOS S3500/6に更新し、図3のように研究所内のLANにより、がん登録専用端末として漢字作製、表示可能なS3100/40、N5200モデル98/80及びPC9821Ap2/C9Tを研究室に、メディカルセンターに

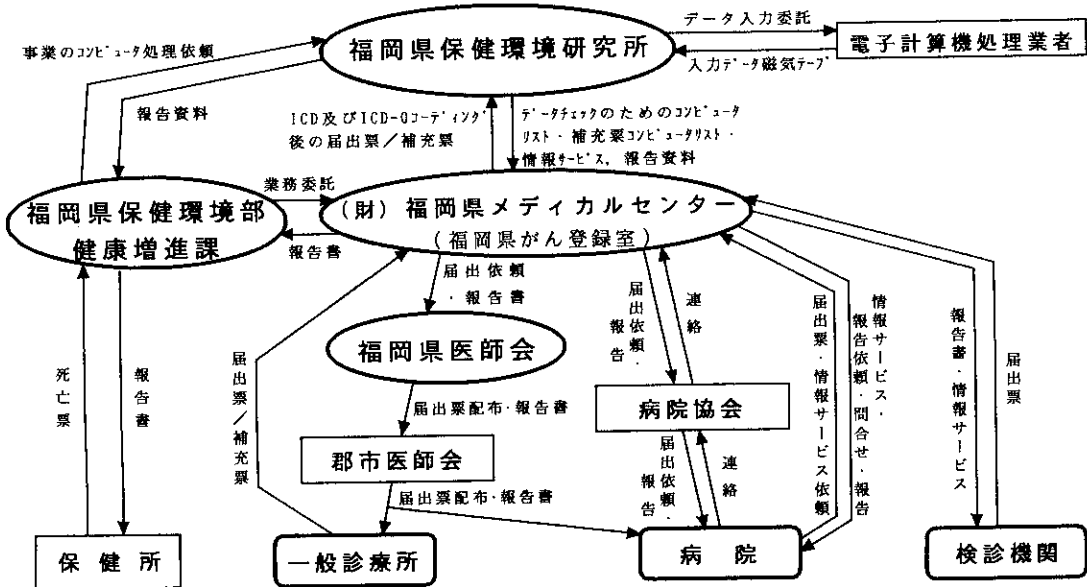


図1 福岡県地域がん登録事業システム概念図

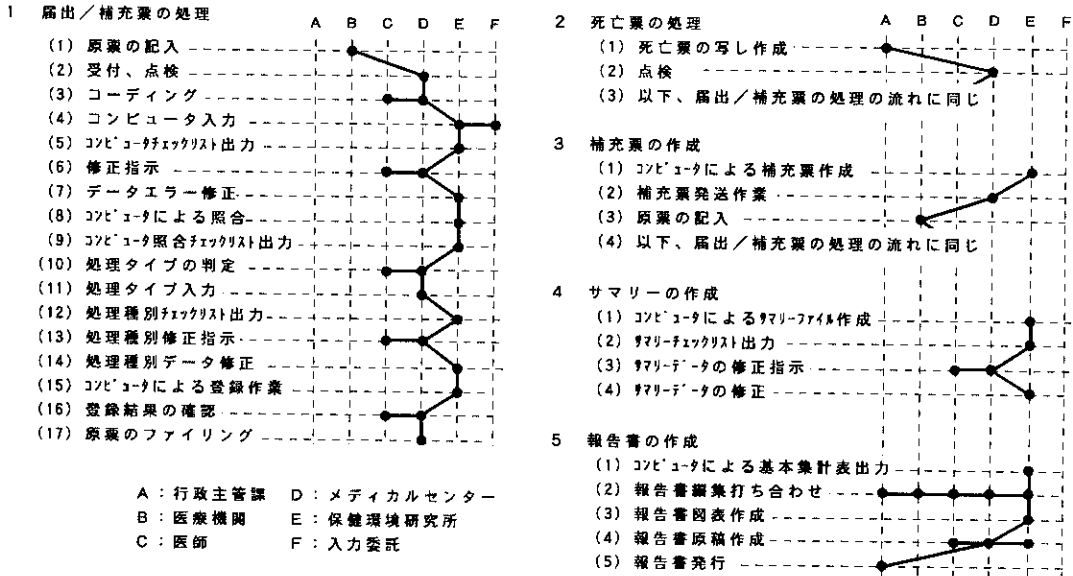


図2 がん登録情報の流れと担当機関

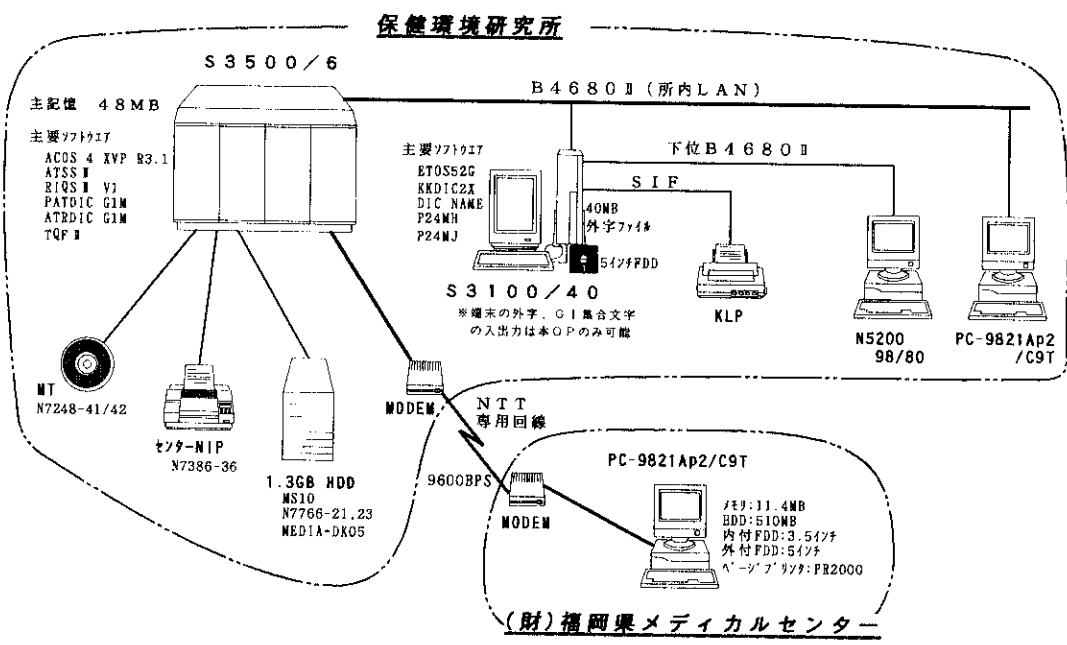


図3 福岡県がん登録システム機器構成

PC9821Ap2/C9Tをそれぞれ設置し、メディカルセンターとはNTT専用回線で結んでいる。図4に示すように、1.3GBのディスク容量をがん登録のために確保し、リレーショナルデータベース(RIQS)にその59.1%

を割り当てている。これまで蓄積されたデータは、そのデータベースの17%を使用している。1994年12月現在、登録済みのデータは基本部事項が94,350件、詳細部事項が147,515件、また、登録が済んでいない途中の

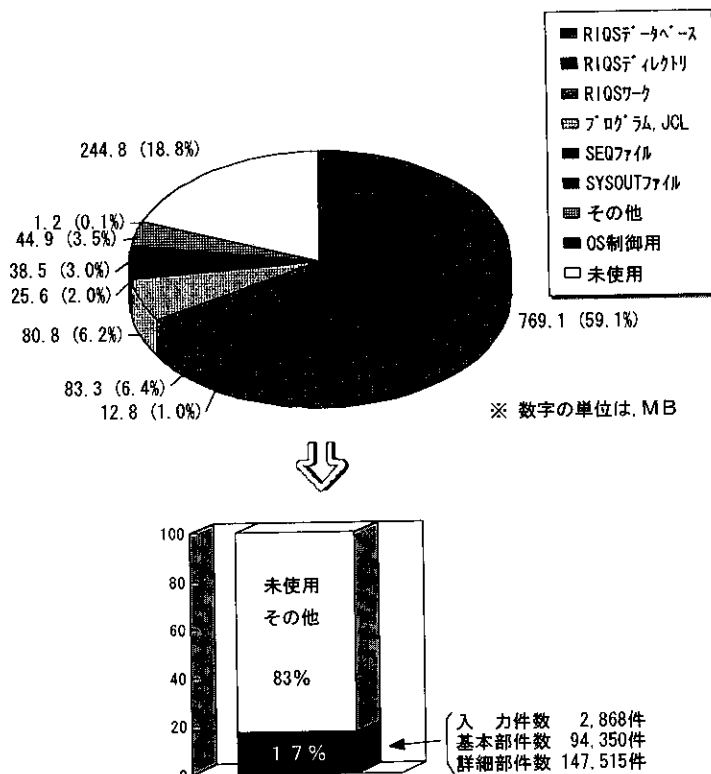


図4 がん登録専用ハードディスク使用状況 (1994年12月現在)

過程にある入力データが2,868件である。当分、確保したディスク容量で処理継続が可能と考えている。

4. 入力データ及び照合データの分析

当研究所のコンピュータシステムで運用開始以来、入力データのチェックからマスターファイルへの追加・登録のうち、特に、入力及び照合を中心に、エラーの種類、発生頻度等を整理検討してきた。これまで1989-1990年の届出/補充票及び死亡票を入力したが、

そのミスの種類と頻度は表1のとおりである。届出/補充票8,949枚のうち総ミス件数は690件(7.7%)である。入力ミスが最も多く、データ磁気化する際の入力タッチミスである。死亡票は2度に分けて入力した。死亡票の1回目は8,228枚のうち総ミスが912件(11.1%)あり、2回目は5,720枚のうちミスは157件(2.7%)に減少している。これは1回目の業者委託時に、データ入力指示が不徹底であったためである。届出/補充票のミスの種類と頻度は表2のとおりである。入力ミス277件のうち、漢字及び平仮名に由来するもの(1の(1), (3))が合わせて160件(57.8%)ある。コーディングミスでは、西暦と和暦の混同不統一によるものが134件(72.8%)存在する。また、記入ミスでは医療機関名が記入されていて、貴院あるいは他院へ紹介の有無に記入がない等いわゆる論理矛盾を含めたミス(3の(1), (3))が83件(36.3%)ある。また、数字の入力間違いも23.8%ある。表3は死亡票について同様に調査した結果である。入力ミスでは2回目に漢

表1 届出/補充票及び死亡票のミスの頻度

| データ処理数 とミスの種類 | 届出/補充票 | 死亡票 | |
|------------------|--------|------|------|
| | | 1回目 | 2回目 |
| 1. データ処理数 | 8949 | 8228 | 5720 |
| 2. ミス総数 | 690 | 912 | 157 |
| (1) 入力ミス | 277 | 806 | 87 |
| (2) コーディングミス | 184 | 27 | 11 |
| (3) 記入ミス | 229 | 79 | 59 |

注) 届出票: 1990.10.1~1991. 3.31受付分け
 補充票: 1990. 9.1~1991. 3.31 "

表2 届出/補充票のミスの種類と頻度

| ミスの種類 | 頻度 | 割合(%) |
|---------------------------|-----|-------|
| 1 入力ミス | 277 | 100.0 |
| (1) 旧字体を新字体で入力, または文字なし | 121 | 43.7 |
| (2) 数字の入力間違い | 66 | 23.8 |
| (3) 漢字あるいは平仮名の入力間違い | 39 | 14.1 |
| (4) 治療方法, 診断方法の読みとり間違い | 29 | 10.5 |
| (5) 県コードの入力間違い | 10 | 3.6 |
| (6) その他 | 12 | 4.3 |
| 2 コーディングミス | 184 | 100.0 |
| (1) 西暦と和暦の混在 | 134 | 72.8 |
| (2) 細胞診あるいは組織診のコード漏れ | 24 | 13.0 |
| (3) 複数記入に対する入力指示なし | 8 | 4.3 |
| (4) 診断名のコード漏れ | 4 | 2.2 |
| (5) 住所のコード漏れ | 4 | 2.2 |
| (6) その他 | 10 | 5.4 |
| 3 記入ミス | 229 | 100.0 |
| (1) 医療機関名記入あり, 貴/他院の有無○なし | 65 | 28.4 |
| (2) 年月日等の記入不明瞭 | 26 | 11.4 |
| (3) 入/退院年月あり, 入院の有無に○なし | 18 | 7.9 |
| (4) 年号間違い (S64年→H元年等) | 18 | 7.9 |
| (5) 性別の記入漏れ | 14 | 6.1 |
| (6) その他 | 88 | 38.4 |

表3 死亡票のミスの種類と頻度

| ミスの種類 | 1回目 | | 2回目 | |
|-------------------------|-----|-------|-----|-------|
| | 頻度 | 割合(%) | 頻度 | 割合(%) |
| 1 入力ミス | 806 | 100.0 | 87 | 100.0 |
| (1) 旧字体を新字体で入力, または文字なし | 542 | 67.2 | 8 | 9.2 |
| (2) 漢字あるいは平仮名の入力間違い | 144 | 17.9 | 60 | 69.0 |
| (3) 病脳期間入力間違い | 49 | 6.1 | 0 | — |
| (4) 住所コード, 死亡場所等の入力間違い | 49 | 6.1 | 6 | 6.9 |
| (5) 年号の入力間違いと入力漏れ | 21 | 2.6 | 3 | 3.4 |
| (6) その他 | 1 | 0.1 | 10 | 11.5 |
| 2 コーディングミス | 27 | 100.0 | 11 | 100.0 |
| (1) 病脳期間の修正ミス | 16 | 59.3 | 2 | 18.2 |
| (2) がん記載がコード化されていない | 8 | 29.6 | 1 | 9.1 |
| (3) レあるいは○印の漏れ | 0 | — | 4 | 36.4 |
| (4) 住所コード間違い | 0 | — | 3 | 27.3 |
| (5) その他 | 3 | 11.1 | 1 | 9.1 |
| 3 記入ミス | 79 | 100.0 | 59 | 100.0 |
| (1) 漢字不明瞭 | 25 | 31.6 | 21 | 35.6 |
| (2) 病脳期間不明瞭 | 12 | 15.2 | 0 | — |
| (3) 医師名未記入 | 9 | 11.4 | 2 | 3.4 |
| (4) 性別, 国籍記入なし, または間違い | 8 | 10.1 | 14 | 23.7 |
| (5) 住所が複数あり, または住所記入漏れ | 7 | 8.9 | 10 | 16.9 |
| (6) その他 | 18 | 22.8 | 12 | 20.3 |

字あるいは読み仮名の指示を徹底したことにより, 大幅な改善がみられた。コーディングミスでは2回目に新たな記入漏れ, 住所コードの間違い(2の(3), (4))がみられた。記入ミスでは1, 2回目とも漢字の不明

瞭さが30%程度あった。

届出/補充票のコンピュータ入力作業の後に, マスターファイル, 死亡票ファイル等との照合作業へ進むが, その照合条件があらかじめ設定されている。これ

は届出／補充票，死亡票及びマスターファイルの間で同一人かどうかを判定する同定照合の基準を定めたものである。コンピュータ処理時に，この照合条件を自動チェックし，1件毎に00～32にコード化されるものあるいは新規登録データとして表示印刷し照合リストを出力する。表4に照合コード00～32の内容を示す。00が完全一致，01は氏名の頭2桁，性別，生年月日及び住所は一致するが，診断名が一致しないものである。届出／補充票8,949件のうち，マスターファイルと照合処理した結果，1,552件（17.3%）が照合コード00～32に判定されたもの，新規登録と判定されたもの7,074件（79.0%），新規登録に結合されたもの323件（3.6%）である（表5）。00～32に判定された1,552件の出現頻度は表6のとおりである。73.8%は完全一致の00である。残り407件の01～32については，照合条件の1～2項目に不一致あるいは部分不一致があるが，記入ミス

表6 マスターファイルの照合コード別頻度と割合

| 照合コード | 頻度 | 割合(%) |
|-------|------|-------|
| 総数 | 1552 | 100.0 |
| 00 | 1145 | 73.8 |
| 01 | 160 | 10.3 |
| 02 | 11 | 0.7 |
| 10 | 73 | 4.7 |
| 11 | 13 | 0.8 |
| 12 | 1 | 0.1 |
| 13 | 41 | 2.6 |
| 14 | 61 | 3.9 |
| 20 | 0 | — |
| 21 | 11 | 0.7 |
| 22 | 6 | 0.4 |
| 23 | 0 | — |
| 24 | 0 | — |
| 25 | 0 | — |
| 30 | 25 | 1.6 |
| 31 | 5 | 0.3 |
| 32 | 0 | — |

表4 照合条件

| 型 | 氏名 | 性別 | 生年月日 | 住所 | 診断名 |
|----|----|----|------|----|-----|
| 00 | ◎ | ○ | ○ | ○ | ○ |
| 01 | ◎ | ○ | ○ | ○ | × |
| 02 | ◎ | × | ○ | ○ | ○ |
| 10 | ○ | ○ | ○ | ○ | ○ |
| 11 | ○ | ○ | ○ | ○ | × |
| 12 | ○ | × | ○ | ○ | ○ |
| 13 | ○ | ○ | ○ | × | ○ |
| 14 | × | ○ | ○ | ○ | ○ |
| 20 | ○ | × | ○ | ○ | × |
| 21 | ○ | ○ | ○ | × | × |
| 22 | × | ○ | ○ | ○ | × |
| 23 | ○ | × | ○ | × | ○ |
| 24 | × | × | ○ | ○ | ○ |
| 25 | × | ○ | ○ | × | ○ |
| 30 | ○ | ○ | △ | ○ | ○ |
| 31 | ○ | ○ | △ | ○ | × |
| 32 | ○ | × | △ | ○ | ○ |

注) 氏名のふりがな ◎：完全一致 ○：頭2桁一致
 ×：不一致
 性別，住所 ○：一致 ×：不一致
 生年月日 ○：一致 △：2/3以上一致
 診断名 ○：頭3桁一致 ×：不一致

表5 届出／補充票とマスターファイルの照合

| 総数 | 8949 |
|-----------------|------|
| (1) マスターファイルに追加 | 1552 |
| (2) 新規登録 | 7074 |
| (3) 新規登録に追加 | 323 |

表7 マスターファイルとの照合結果

| 不一致項目 | 頻度 | 割合(%) |
|-------|-----|-------|
| 氏名 | 67 | 4.3 |
| 性別 | 12 | 0.8 |
| 生年月日 | 30 | 1.9 |
| 住所 | 52 | 3.4 |
| 診断名 | 195 | 12.6 |

注) 届出／補充票：1552件

等で同一人の可能性がある人である。従って，人手により，再度，確認チェックされることになる。この照合結果による項目不一致の状況は表7のとおりである。重複計上されているので全体の数より多いが，氏名不一致は67件，全体の4.3%ある。診断名不一致は重複がん，続発性がん等が含まれるが195件と多く，照合対象の12.6%である。最も少ないのは性別不一致で12件，0.8%である。

5. システムの改善

地域がん登録システムは年数を重ねるに従い，データ量が増大し，大量のデータ処理に対する処理効率が必要になってくる。システム改善を考えるに当たって，このような基礎的データの把握は大切なことである。届出／補充票の入力ミスについては，業者委託時に入力マニュアルを改良することにより，また，単純な指示で済ませられるよう心掛けることで改善される。死

亡票の1回目と2回目の比較で見られるように、明らかに効果が現れている。入力ミスのうち、漢字あるいは平仮名に関連するものが多いが、これは登録事業として、精度の考え方に関わるものであり、高度な精度管理維持のためには必要なことである。また、コーディングミス、記入ミスの改善には届出/補充票そのものの見直し等が検討課題となるが、これは本県の地域がん登録システムの問題であると同時に、他の地域と相互比較する場合に関係する精度上の問題でもある。しかも、罹患率、生存率等の疫学研究にとっては、同一方法による継続性が重要であり、慎重な討議が必要と思われる。コンピュータ処理においては、可能な限り自動化するが、処理量と処理効率が微妙に関係してくる。届出/補充票、死亡票は毎年、どの程度の枚数が集まるか、ほぼ見当がつかず、一度に多量処理すると過去の全登録データ即ちマスターデータと照合、検索するとき、思わぬ処理時間を要することがある。がん登録システムにおいては照合、検索の速度はシステム改善の一つの目安である。

6. ネットワークとしての研究所の役割

地域がん登録事業における我々の果たす役割は、コンピュータ処理周辺の業務であるが、がん登録室とは密接に情報交換している。この業務の目標の1つが精度管理であり、従来から最も重視されてきたもので、真実の意味で確定されたデータといえるものはなく、様々な段階でデータチェックと修正が入るからである。従って、このような分業体制としての業務運営は、我々としても新たな試みである。地域がん登録事業の中核部分はコーディング業務であり、診断名、組織診、細胞診のコーディング作業には特殊な技能と熟練を要し、コンピュータによる自動化は期待できない。従って、地域がん登録において、このような手作業部分の重要性が認識され、当初より中央登録室の必要性が説かれている⁷⁾。それ故に、中央登録室において、コーディング作業が安定して着実に処理される人的システムの確立が先にあって、その上に、この地域がん登録システム全体が効率よく運営されるデータベースによるコンピュータシステムが望ましい姿であると考えている。本県のがん登録室の端末はこの人手作業の手助けとして、過去の履歴等を容易に検索・確認、整理す

るために役立っている。更に、その運営を長期に亘って継続していくことにより、地域がん登録の真の役割である地域サービスが十分果たされると考える。

7. おわりに

届出/補充票、死亡票という個人票の取扱いは、地域がん登録の従事者として道徳的倫理と情報秘密保護は最も重要な要素である。鍵のかかる部屋の確保、コンピュータへのアクセス、がん登録室との情報交換、業者委託契約時の遵守事項等細心の注意をしているところである。これについては、がん登録研究者の間で、地域がん登録事業に携わる関係者の守るべき一般的な事項として明文化⁸⁾されつつある。一方、コンピュータ処理を担当する立場からは、健康増進課とも十分協議しながら、機器及びソフトウェアの改善充実に当たっているところである。更に、この分業体制で重要なことは、全体をよく把握し、強力にリーダーシップを取る能取りが必要であると考えている。コンピュータ化できない人手のかかる作業が多い事業であり、どこかで作業が停滞すると全体的に遅れていく欠点を含んでいる。業務遂行過程で、こうした課題を少しずつ改善していきたいと考えている。

謝 辞

本稿を終えるにあたり、福岡県地域がん登録の創設から今日のデータベースシステムの完成に多くの指導をいただいた福岡大学医学部公衆衛生学教室重松峻夫教授並びに産業医科大学臨床疫学教室吉村健清教授に深く感謝致します。

参考文献

- 1) Jensen, O.M., Parkin, D.M., MacLennan, R., Muir, C.S. and Skeet, R.G., ed.: Cancer Registration Principles and Methods. IARC, Lyon, pp4-6, 1991.
- 2) 厚生省がん研究助成金、地域がん登録の精度向上と活用に関する研究班、ガイドライン作業グループ: 地域がん登録における倫理と情報保護 (試案), pp1-22, 1994.
- 3) 福岡県メディカルセンター、福岡県がん登録委員会: 福岡県におけるがん登録, 第1回標準集計結果 (昭和59年悪性新生物罹患数および罹患率), pp1-6, 1988.