

中央登録室とコンピュータ

味木 和喜子*

はじめに

地域がん登録の中央登録室において、コンピュータは、登録データの保管および集計表の作成に必須である。ところが、中央登録室におけるコンピュータの役割は、それだけにとどまらず、手作業の軽減、質的・量的精度の向上に大きな役割を果たす。この報告では、コンピュータを利用すべき作業として、記録照合および質的精度管理を取り上げ、その役割および方法を示す。また、それ以外の作業について、コンピュータが果たしうる役割を述べる。

1. 登録作業の概要とコンピュータの役割

図1に、登録作業の流れと電算機システムの役割とを示す¹⁾³⁾。

地域がん登録では、主な情報源として、がん患者の診断、治療に関する医療機関からの届出票（あるいは採録票）、および「がん」が記載されている死亡票を用いている。まず、これらの情報を、コンピュータで処理できる形にコード化し、入力する。

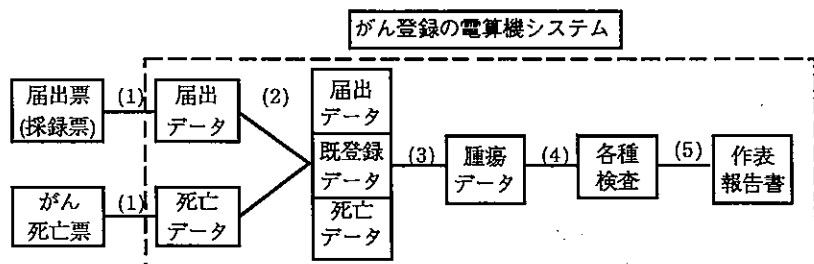
次に、同じ患者に由来するデータをまとめ、登録番号を与える記録照合を行う。単一腫瘍について複数の情報が得られた場合に、それらを別

個の腫瘍として登録すると、罹患数を過大に計測することになる。重複登録を防止するために、患者情報を入力する都度、それらの患者が既に登録されているか、新たな患者であるかを判定する。

単一の腫瘍について、複数の情報が存在するので、それらの情報から、1件の総括情報を作成する必要がある。この1腫瘍1件の情報を、コンピュータで保存している登録室もあれば、1腫瘍複数件の情報を保存しておき、それから、必要時に1腫瘍1件のファイルを作成する登録室もある。

また、情報に、漏れ、誤りが無いかをチェックし、修正すること（精度管理）も、登録室の重要な作業である。この1腫瘍1件の総括情報を用いて、集計し、報告書を作成するとともに、これらの情報を活用していく。

図1. 登録作業の流れと電算機システム



- (1) コード化、入力 : 情報を電算機で処理できるようにコード化し、入力する
- (2) 記録照合 : 同じ患者に由来するデータをまとめ、登録番号を与える
- (3) 情報の統合・要約 : 各々の腫瘍に関する総括情報を作成する
- (4) 質的精度管理 : 総括情報に誤り、矛盾がないか検査する
- (5) 集計、報告書作成

*大阪府立成人病センター調査部集検整合係長

2. 記録照合の方法およびその精度と効率

(1) 記録照合の役割

登録室に届いた情報が、新たな患者の情報か、既登録患者の追加、修正情報かを決定することは、正確な罹患数を計測するために必須の作業である。登録室で処理する票件数が多いほど、また、既登録件数が蓄積されるほど、この作業量は、年々増大する。従って、同一人物に由来する情報を、効率よく、かつ、見落としの無いように、発見する方法を考える必要がある。

(2) 記録照合の方法

記録照合は、機械照合および判定作業の2段階からなる⁴⁾⁶⁾。機械照合では、新着情報と既登録情報とのあらゆる組み合わせについて、生年月日、姓名などの患者同定指標を比較し、その中から、同一人物の可能性のある組み合わせ、すなわち、類似ペアを作成する。ついで、これらの類似ペアについて、他の指標を参考にしながら、同一人物か別人かを判定する。

機械照合の段階で、類似ペアとされなかった組み合わせは、たとえ同一人物に由来する情報であっても、それを判定するチャンスが与えられないため、別人として登録されることになる。従って、機械照合の方法として、同一人物ペアを、もれなく抽出する仕組みが必要である。しかし、類似ペア数が多くなると、判定作業の負荷が大きくなる。従って、この両者の兼合いを見極めながら、機械照合方式を決定する必要がある。

(3) 機械照合方式の例

人口規模の小さい登録室では、照合のためのプログラムを作成せず、情報を生年月日順に並べ変えて、見比べる方法が用いられている。通常は、生年月日が異なる同一人物を見落とさないために、姓名順のリストをも作成し、補足している。これを、A方式とする。

既登録情報が大きくなると、生年月日の一致するものを全て判定することが困難になる。

その場合は、生年月日が一致する患者のうち、さらに、姓名、性別、住所などの他指標が一致する場合に、類似ペアを作成する。これをB方式とする。

B方式に加え、生年月日の一致しない同一人物を見落とさないために、姓名および他指標の一致リストで補足する方式をC方式とする。この方法を採用している登録室が多く、大阪府がん登録でも、C方式を用いている。

このA-C方式は、生年月日、あるいは姓名が一致した場合のみを類似ペアとしているが、これらの指標を分割して、そのうち一定数以上の項目が一致した場合に類似ペアとするD方式もある。例えば、生年月日の年、月、姓、名の4指標のうち3指標以上一致を類似ペアとする方法、あるいは、生年月日の元号、年、月、日、性別、住所、姓、名の8指標のうち6あるいは7指標以上一致を類似ペアとする方法、などがある。

(4) 大阪府がん登録の機械照合

大阪府がん登録では、照合指標として、生年月日、姓名、住所、性別および部位の5指標を用いている。

姓名には、姓の第1字に特有の読みを与えた仮名コードを用いている。この仮名コードが一致している組み合わせでは、姓名の他の漢字が異なっても、コンピュータ上で同じ姓名と判断する。従って、他の姓名の処理方法（完全一致など）よりも、姓名の不一致による同一人物の見落としが少ない。

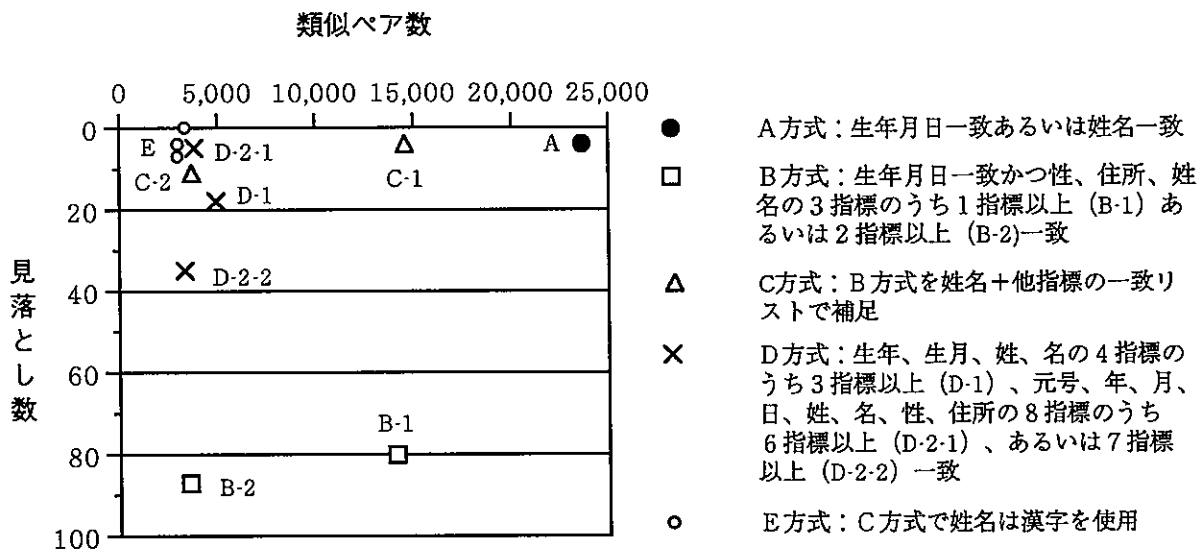
(5) 機械照合の精度と効率の比較

大阪府がん登録の資料を用いて、A-D方式の見落とし数と類似ペア数を比較し、図2に示した⁵⁾⁶⁾。

方法は、大阪府がん登録の1年間の届出票24,648件を用いて、それぞれの機械照合方式で作成される類似ペア数を計測した。さらに、その中で同一人物と判定した2,912件のうち、それぞれの機械照合方式で類似ペアとして作成されない見落とし件数を調べた。

図2. 機械照合方式による類似ペアと見落とし数の比較

—大阪府がん登録の1年間の届出票同士の照合を用いた検討(届出票 24,648 件、同一人物と判定したペア 2,912 件)



A方式では、見落とし数は4件と少なかったが、類似ペア数が23,662件で最も多かった。

B方式では、見落としが、80件以上と大きかった。

C方式では、類似ペア数がB方式よりやや多くなったが、見落とし数は11件以下と小さかった。

D方式は、類似ペア数が5,000件未満であり、効率のよい方法であった。見落とし数は、用いる方法により、5-35件あった。

また、E方式として、大阪府がん登録の方式に、漢字照合をとりいれた方法を試行した。仮名コードのかわりに、姓の第1字の漢字、姓名の第1字と第3字、姓名の全漢字、を用いると、類似ペア数が少なくなり、見落としが0-7件に止まることがわかった。

以上より、人口規模が少ない登録室であっても、照合効率の点からみて、コンピュータによる機械照合を取り入れるべきである。方法としては、C、D、およびE方式を推奨する。生年月日不一致の同一人物を発見する機能が必須であり、B方式を用いることは奨め難い。

3. 質的精度管理の方法

項目の記載漏れ、コード化の間違い、など

を検出する機能は、がん登録の精度向上のために必須である^{2), 3), 7)}。

(1) 精度管理(論理チェック)の内容
表1に、精度管理の方法を示す。

項目単体における妥当性の検査(validation check)として、必須項目がもれていないか、2月30日のような、存在しえない日付はないか、存在するコードであるか、を調べる。部位のように、分類の多い項目では、用語ファイルを管理し、その中に存在するコードが用いられていることをチェックする。

項目単体で問題がなければ、次に項目間の関連を調べる(consistency check)。診断日と生年月日とが逆転していないか、女性の前立腺がん、男性の子宮がん、のように、性別と部位・組織との関係に矛盾がないか、など、項目の関連に矛盾がないかを確認する。また、70歳の神経芽細胞腫のように、発生する可能性が0ではないが、極めてまれな組み合わせも合わせてチェックし、間違いがないことを確認する。特に、部位と組織との関係を確認することは、腫瘍を正しく分類するために必須の作業である。例えば、肝細胞癌が肝以外の部位に分類されていたり、carcinomaが発生しない部位(脳など)に腺がんがあれば、書かれ

表1. 精度管理（論理チェック）の内容

| | |
|---------------------|--|
| 項目単体検査(Validity) | 項目の妥当性 必須項目の漏れ、日付、コードの妥当性 |
| 項目関連検査(Consistency) | 項目間の矛盾 日付の順番、性別と部位・組織、 年齢と部位・組織、部位と組織 |
| 票間検査 | 票間の矛盾 複数の票間での情報の不一致、矛盾 |
| 疑診検査 | 診断の妥当性 疑診あるいは診断を裏付ける検査、治療のない 症例の、集計対象としての妥当性 |

ていた部位が転移先部位でないか、を確認する必要がある。

また、1つの腫瘍について、複数の情報がある場合には、それらの間で不一致、矛盾がないかをチェックする（票間検査）。

さらに、集計前には、疑診のみ、あるいは、診断を裏付ける検査や治療のデータがない症例について、集計対象としての妥当性をチェックする必要がある。

(2) 精度管理のツール

精度管理のためのツールとして、IARCが開発したCHECKプログラムがある⁷⁾。これは、「5大陸のがん罹患」のために収集されたデータをチェックすることを目的に開発されたプログラムで、ICD-Oを用いて部位、組織をコード化している場合に利用することができる。プログラムは英語で、DOS/V機のMS-DOS上で作動する。

ところが、ICD-9あるいはICD-10で部位をコード化している登録室が、このプログラムを用いるためには、あらかじめ部位コードをICD-Oに変換する必要がある。そこで、ICD-10で部位、ICD-O-2で組織をコード化している登録室が、日常的に、IARC-CHECKと同様のチェックを行えることを目的に、地域がん登録研究班がICD-10用チェックプログラムを開発した⁸⁾。これは、MACおよびWindowsのファイルメーカープロ3.0で作動する。

手作業のみでチェックする場合と比較する

と、コンピュータでエラーリストを作成した上で、手作業で確認して修正するという方法は、短時間ですみ、かつ、エラーの見落としも少なくなる。但し、コンピュータで検出できるエラーは、ありえないコードのデータのみであることに留意する必要がある。例えば胃癌を間違えて食道とコード化しても、コンピュータではエラーとならない。このような間違いは、手作業によって初めて確認できる。

(3) 精度管理のタイミング

情報到着時に必須項目の漏れ、間違いを確認し、必要に応じて届出医療機関に問い合わせる。次に、情報を入力する時に、項目単体、項目関連検査を実施する。また、新たな情報が追加されることにより、今まで検出されなかった誤りが発見される場合もあるため、情報を更新する時にもチェックが必要である。さらに、総括作成時には、全情報を通覧して、票間の矛盾を確認する。

4. 今後の方向

(1) 入力支援

大阪府立成人病センター院内がん登録では、カルテから手書きで採録票を作成し、それをパンチ入力し、コンピュータによる論理チェックで内容を確認してから、情報を登録してきた。それを、1997年診断患者から、採録時に直接コンピュータに入力する方法に変更した。入力時に自動的にコード化するとともに、項目単体、項目関連検査をも同時に行うため、採録作業の効率が向上した。さらに、磁気化されている他の診療情報（病理組織診断など）を、採録時に参照できるため、詳細な情報を正確に採録できるようになった。

このような入力支援システムを、地域がん登録でも届出票の入力に活用し、将来的には医療機関に配布して直接入力していただく方向を検討していきたい。

(2) 総括情報作成支援

同一の腫瘍に対する複数情報を、一腫瘍一件にまとめる総括には、手作業で人が判断するという方法と、コンピュータにより自動的

に編集する方法とがある。

手作業では、症例ごとに柔軟に判断できるというメリットがある。一方、判断が人により、時により異なることを防止するため、ルールを記述し、作業者に徹底することが必要になる。また、判断した結果のみをコンピュータに保管していると、ルールを変更した場合に、遡ってそのルールを反映させることができない。件数が増えると、それに伴って作業の効率と精度とが悪くなる可能性がある。

一方、コンピュータによる自動編集では人の判断を正確に分解し、順番に組み立てる必要がある。系統的な誤りが生じないように、あらかじめ情報の誤り、矛盾をチェックし、修正しておく必要がある。一方、件数が増加しても、速くて正確に処理することができる。ルールを変更した場合には、全登録症例にそれを適用し、ルールの変更にもなう集計値の変化を検証することもできる。大阪府がん登録では、このような編集プログラムを用いて、集計用の総括ファイルを作成している。

(3) 外部照合方式の確立にむけて

信頼性のある罹患情報が蓄積されると、それを検診受診者におけるその後のがん発生の把握などに活用することができる。地域がん登録以外で収集された情報と、がん登録の情報との照合を、外部照合と呼ぶこととする。

外部照合では、収集項目およびその様式が、地域登録のそれらと異なる場合があり、登録室内での記録照合方式をそのまま用いると、精度、効率が低下する場合が多い。今後、登録室内記録照合方式を参考に、外部照合方式を検討していく予定である。

5. まとめ

中央登録室の作業量は、毎年増加する。コンピュータの活用による作業効率の改善を、常に考慮しなければならない。中央登録室の全作業過程において、コンピュータの活用が可能であり、有用である。記録照合、精度管理には、コンピュータを用いるべきである。人の判断を正確に記述し、系統的な誤りが混

入しないように配慮すれば、コンピュータによる作業精度は、手作業のそれに劣らない。情報の収集、整理・保管、活用の全ての面で、情報の保護に十二分の注意を払う必要がある。

文献

- 1) 厚生省がん研究助成金「がん予防、医療活動におけるがん登録の役割に関する研究」班：地域がん登録の手引き 改定第3版。同研究班（主任研究者：福間誠吾），1986。
- 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」班訳：がん登録の原理と方法。同研究班（主任研究者：花井彩），大阪，1996。
- 3) Jensen OM et al. eds: Cancer Registration: Principles and Methods. IARC Scientific Publications No. 95. IARC, Lyon, 1991.
- 4) 味木和喜子他：13 地域がん登録室における記録照合方式の実態調査。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」班平成5年度報告書，p. 141-145，同研究班（主任研究者：花井彩），大阪，1994。
- 5) 味木和喜子他：登録室内記録照合の精度向上のための研究—照合方式および姓名の取り扱い方による自動同定での非出現数の比較—。同上研究班平成7年度報告書，p. 238-244，同研究班（主任研究者：花井彩），大阪，1996。
- 6) 味木和喜子他：記録照合での氏名取扱い方式による照合の精度および効率の検討。同上研究班平成6年度報告書，p. 138-144，同研究班（主任研究者：花井彩），大阪，1995。
- 7) Parkin DM et al. : Comparability and Quality Control in Cancer Registration. IARC Technical Report No. 19. IARC, Lyon, 1994.
- 8) 味木和喜子：ICD-10用CHECKプログラム（試用版）。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（主任研究者：大島明），大阪，1997。