

## 疫学追跡終了後コホートデータの共通利用(アーカイブ化)の際の死因データ利用に関する検討

研究分担者 大橋 靖雄 中央大学理工学部人間総合理工学科  
研究協力者 原田亜紀子 東京大学大学院医学系研究科公共健康医学専攻

### 研究要旨

死因情報を付与した形でのデータアーカイブ化が難しい現状を鑑みて、現制度下での運用案を提案した。提案の運用案を複数のコホート研究から構成される国内循環器疫学研究(JALS)に適用したところ、死因照合率は99.7%であり、照合作業の技術的側面、作業手順化の面で問題はなく、運用案の一つになりうると考えられた。想定するアーカイブセンター等で実際に運用を行う上では、データ利用の規約、データ利用の物理的環境についてあらかじめ検討すべき事項があると考えられた。

### A. 目的

米国では、National Heart, Lung, and Blood Institute (NHLBI)などの公的研究資金で実施した臨床試験や疫学観察研究のデータは、Biologic Specimen and Data Repositories Information Coordinating Center (BioLINCC) に対して追跡終了後一定の期間、条件を充足した後に研究データを登録し公開するよう進められている(<https://biolincc.nhlbi.nih.gov>)。2010年時点で、BioLINCCには、56臨床試験、20疫学観察研究のデータが収集されており、使用申請を行うことにより、研究利用できるとともに、生物統計学系の学生などに対してはテストデータの提供なども実施されている(BioLINC HP: FAQs <https://biolincc.nhlbi.nih.gov/faqs/#q2>)。2000～2010年の間で、約500件の利用申請に対し、出版物が224件、68本の論文が出版されており、研究データの二次利用により新たな研究や若手研究者の教育・研究機会が創出されているといえる。

翻って我が国の現状をみると、厚生労働省科学研究費補助金を受けた研究では、National Bioscience Database Center (NBDC) 等へのデータ提供が求められ、バイオサイエンス基礎研究でのデータアーカイブ化には進展はみられるが、個人の健康情報を含む疫学観察(コホート)研究領域では、未だデータの二次利

用環境は整っていない。このように疫学観察研究において、データの共有化、オープンデータ化が進まない背景には、いくつかの要因が考えられる。一点目は、コホート研究が様々な研究費で維持されており、米国 NHLBI などに代表される大型の研究費を受けた先へのデータ提供という単純な構図になりえないことである。二点目は、多くのコホート研究では主要なアウトカムとなる死因情報を人口動態統計調査の二次利用申請によって得ており、この情報を付与した状態でのデータ公開が難しい点である。三点目としては、研究企画時点から、データアーカイブ化を見据えた倫理的諸問題の対応、調査票やデータベース構造の標準化の準備されていない点にあると考えられる。

そこで本報告では、死因情報を付与した形でのデータアーカイブ化が難しい本邦の現状を鑑みて、現制度下で研究基盤づくりを進めていく上での運用案を提案する。

### B. 方法

1. 死因情報を外したアーカイブ環境を想定し、必要時に死因を人口動態二次利用申請し、アーカイブセンター(データセンター)にて照合・集計・解析を行う運用例を提案する。

2. 国内大規模循環器疫学研究 (JALS) を例に、研究を統括する中央研究事務局が人口動態二次利用申請し、上記1で提案する運用例になって死因を照合する。この際に生じる問題点を考察し、提案例をアーカイブセンターで運用する際に想定される課題を検討する。

### 1) 人口動態統計の二次利用申請

現在、我々が行っている国内の循環器疫学研究 Japan Arteriosclerosis Longitudinal Study (JALS) では、国内 58 市町村、8 職域の計 11 万人を追跡している。この JALS コホートのうち、職域コホートを除き、研究開始 (2002 年 ただし一部コホートは 1999 年) から 2012 年 12 月までの異動状況が、住民基本台帳 (住民票) 情報により確認されている追跡対象者で期間中に死亡した 7,137 件を死因照合の対象とした。総務省による「統計法第 33 条の運用に関するガイドライン」を参照した上で、人口動態調査二次利用の申請書類を作成し、厚生労働省大臣官房統計情報部企画課審査解析室に人口動態調査二次利用の申請を行った。

申請作業と並行し研究事務局において、研究開始から 2012 年 12 月までの死亡例について、性別、生年月日、死亡日、死亡時の居住地 (市町村コード) のリストを作成した。

### 2) 提案例に従った照合作業の検証

人口動態調査の使用許可がおりた後、性別、生年月日、死亡年月日、死亡時の居住地 (市町村コード) を照合変数とし、人口動態調査データと照合し原死因を確定した。

## C. 結果

### 1. 死因情報を外したアーカイブデータを利用する研究基盤の提案

#### 1) 統計法 33 条下での個票データの提供

本邦の多くのコホート研究では、死因情報は人口動態統計調査の二次利用申請を行い入手している。コホート研究を行うにあたっては、個別 (研究対象者個人) の死因が必要となることから、公的統計の二次的な利用方法のうち、「統計法 33 条に基づく個票データの利用を申請する方法」がとられている。この利用申請は、公的な研究費の補助を受けた研究などに限定され、研

究者は利用したい統計の担当省庁の事前相談窓口を利用の申し出を行い、事前審査と本審査を経て許可される仕組みとなっている。利用申請、実際の利用にあたっては、統計法 33 条により、データを利用する者、利用する場所、利用する目的の範囲が厳しく制限されており、本報告において議論の対象としている死因情報を付加したデータのアーカイブ化 (オープンデータ化、一定の制限を設けた下での二次利用) は、不可能な状態にある。

### 2) 死因データを必要時に連結させる方法の提案

上述の通り現行の統計法下では、ローカル (各研究グループ) が行った二次利用申請で得た死因を付与したデータをアーカイブすること (中央に集めること) はできないので、死因を連結した形でのデータセット構築はあきらめて、必要時に中央 (アーカイブデータを保持するセンター等) で死因照合作業を行って解析用データセットを作成する方法を提案する。

JALS では、ローカルコホートから追跡調査データ (生存・死亡 (死因はなし)、発症) の提供があり、死因の同定については中央事務局 (データセンター) で人口動態の二次利用申請を行い一括で照合作業を行っている。この JALS で採用しているシステムを例にして、JALS を国内疫学研究のコンソーシアム (アーカイブ事業に参加する研究グループ) に、中央事務局をアーカイブセンターに置き換え考えることで、研究コンソーシアムに属する研究者が、このデータを利用した研究を計画し、人口動態の二次利用申請で死因情報を得ることによって、センターのスタッフ (必ずしも専門知識を必要としない) により容易に照合が行えるような環境を検討する。

### データ提供からアーカイブ化まで

提案するデータ利用基盤の概略を [図 1] に示した。研究コンソーシアムに参加する各研究が、基本データ (生活習慣、検査データなど) と死因を除いた追跡データをアーカイブセンターに提供する ( **データ提供** )。アーカイブセンターでは、基本データベースと追跡データベースを分けて構築しておき、基本データベースは原則登録時から修正なしの状態、追跡データベースは、今後の追跡継続に応じて更新できる構造と

する（**アーカイブ化**）。追跡データは、今後の死因照合作業で必要となる「死亡地（市町村）」、「死亡日」、「生年月日」、「性別」を含むように設計する。

### 死因情報の申請

コンソーシアム内の研究者（あるいは一定の条件を設け、研究グループ外の研究者も可能とするか）が、このデータベースを使用する研究を計画し、死因情報を得るために厚生労働省に対して人口動態調査二次利用申請を行うとする。この際に、データの利用場所およびその利用者に、研究申請者とアーカイブセンターを含む形で申請を行う（**死因情報申請**）。

### 照合作業から研究データセット作成

承認後に提供を受けた死因情報をアーカイブセンター内で、「死亡地（市町村）」、「死亡日」、「生年月日」、「性別」をキー変数として、保有する追跡情報と照合し（**データ照合**）、死因を付与した一時的な解析データセットを作成し、研究計画に基づいた解析に使用する（**解析用データセットの作成**）。研究終了後は死因情報を削除（抹消）し、厚生労働省に利用後報告を行うという流れである（**利用報告、死因情報の削除**）。

死因を付与したデータをアーカイブせずに、研究を計画するたびに、この から の作業をアーカイブセンターで手順化し対応するという運用方法の提案で

## 2. 必要時に死因を照合する方法の実際（JALS 事例）

JALS で現在実施している方法をもとに、照合手順について具体的に説明する。

### 1) 追跡情報の収集

コホート（各参加研究）から中央事務局（アーカイブセンター）に、死因照合の際にキーとなる「死亡地（市町村）」、「生年月日」、「死亡年月日」、「性別」を含めた追跡情報を提出する。死亡地（市町村）については、市町村コードに加え、都道府県コード、保健所コードなどを併記した対応表を作成した。

### 2) 二次利用申請

必要とする統計情報の変数、上記1)でリストアップされた市町村、必要とする調査対象年などをもとに、死因情報の申請範囲をとりまとめる。その他の申請書類を合わせて作成し申請した。

### 3) 承認後、データの受領

2015年1月29日付で利用許可があり、JALS対象地域の市町村で1999年1月1日から2012年12月31日までに発生した死亡の調査票情報を受領した。提供を受けたデータ（ファイル）の形式は、CSVファイル形式で、1年1ファイルで提供された。

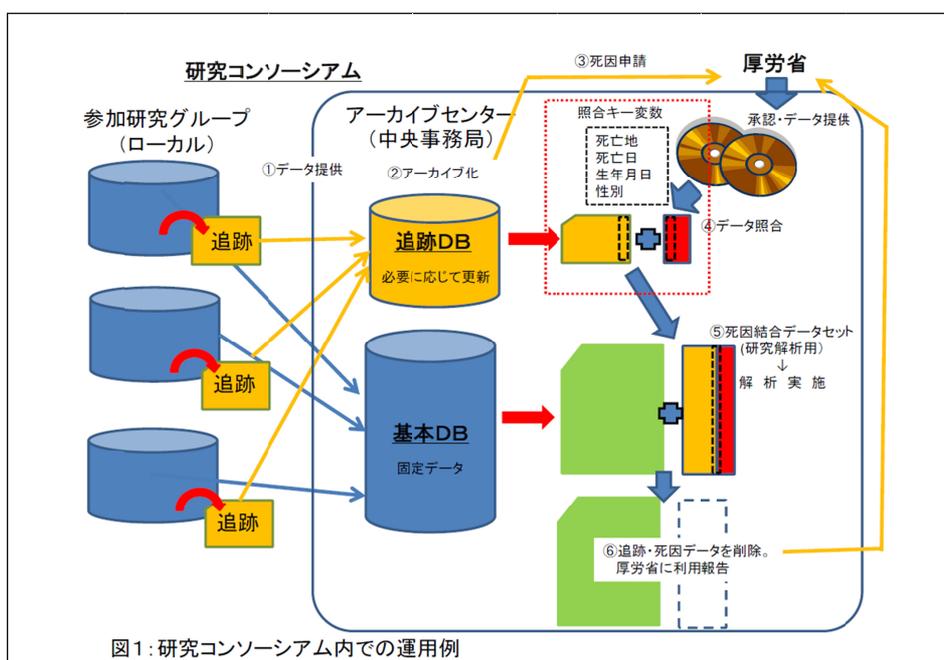


図1：研究コンソーシアム内での運用例

ある。

## 4) 照合作業

### 提供を受けた人口動態データの読み込み

提供を受けるデータの構造(提供変数)に応じて変更は必要となるが、データの構造と読み込みプログラム例を付録に示した(SAS プログラム#1)。

### 研究全体の市町村(コード)をリストアップした一覧表を読み込み(SAS プログラム#2)

### 追跡調査データベースの対象データから死因照合を行いたい死亡者を抽出する作業(SAS プログラム#3)

### 照合作業(SAS プログラム#4)

上記 と で作成されたデータセットについて、「死亡地(市町村コード)」、「死亡日」、「生年月日」、「性別」でマッチマージしていく。

JALS の対象者で、職域コホートと死亡調査データが確定していないコホートを除き、死亡が特定できていたのは7,137件であった。性別、生年月日、死亡年月日、死亡時の居住市町村名を照合変数とし、人口動態調査データと一致がみられたのは7,099件(99.5%)であった。

### 照合例、未照合例のリスト化(SAS プログラム Excel ファイル#5)

死因一致例の中に複数の候補例(一意に決まらない例)が存在する対象がないか、また未照合例の情報確認を行うために該当者のリストアップを行う必要がある。これらを SAS から Excel ファイルに出力するプログラムを作成した(帳票例を付録に示した)。

複数の候補例が見つかったのは5件あり、いずれも東日本大震災の被災地域での死亡者で、死因が同じであったためその死因で照合を行った。死因が照合できなかった例は33件であり、このうち14件については、以前に JALS が行った死因照合作業において既に未照合が判明しており、各コホートに対して死亡時情報を確認したがいずれも情報に誤りのなかった例で、人口動態統計作成の過程で入力間違い等が発生した事例と判断した。このため、今回照合出来なかった例は、実質として19件(0.27%)であった。

## D. 考察

本検討では、死因情報を外したアーカイブ環境を想定し、必要時に死因を人口動態二次利用申請し、アーカイブセンターにて照合・集計・解析を行う運用例を提案した。この提案法を国内大規模循環器疫学研究(JALS)で、実際に運用し死因照合を行った。JALS は国内コホート研究データを中央事務局で統合して実施している研究であることから、この研究で行う死因照合作業の実際が、本検討で想定している研究データアーカイブの環境づくりに対して参考になる点が多いように思われた。

死因データの提供を受けてから実施する照合作業については、JALS の研究進捗にあわせて、統計プログラム(SAS)を用い、データの読み込み、Excel などへの帳票出力に至るまで一連の作業の手順化をすすめてきた。多くの作業をルーチン化してきており、大部分の課題は解決してきているが、本提案例をアーカイブセンターで運用する際に予想される課題について検討する。

### 1. 厚生労働省から受領する統計フォーム(データ形式)に大幅な変更がないことを前提とする

本検討で提案した仕組みは、人口動態二次利用申請により提供される死因データの構造に大幅な変更がないことを前提としている。構造の変更を伴わない小規模の変数追加等の変更であれば、受領後の確認作業を手順化すること(変数表の確認作業やプログラム修正箇所の明示)で、アーカイブセンタースタッフ、疫学研究に従事する研究者等で十分対応できる作業と考えられる。

### 2. 死因照合率はコホートから提出される照合キー変数の精度に依存している

死因照合の一致率については、各コホート研究から提出される照合キー変数(死亡地、性別、死亡年月日、生年月日)の情報精度に依存している。JALS では、自治体の住民基本台帳などの確認後の情報に基づき照合作業を実施しているが、未照合となる例は、提供情報に何らかの誤りがあり、情報の再確認後に照合可能となる例が大部分である。一方で、提供情報に誤りがないにもかかわらず照合例を検索できない例が少なからず存在しており、これらは性別、生年月日、死亡日につ

いて、人口動態統計作成の過程で発生した何らかの入力ミス等が原因として考えられる(例:性別が男女逆、1日ずれている、生年月日と死亡日が同じ日付となっているなど)。しかし、本検討で想定しているのは追跡終了後コホート研究のデータアーカイブ化であるので、これまでの研究実施の過程(各研究での死因照合作業などの過程)でこれらの変数は十分確認されてきており、情報の精度については、進行中の研究に比べて十分高く、こうした問題の発生は少ないと考えられる。

### 3. 市町村合併の問題

JALSの例では、研究実施期間内に平成の市町村大合併の時期を含んでおり、実際、市町村リストの作成作業が煩雑で、データ受領後の照合作業において工夫が必要となっている。合併年月日のリスト化、合併前後の自治体の包含関係(新規設置であるか、吸収合併であるか等)について注意を払ってリストを作成しておかねばならない。しかし、これらの対応については一度その対応手順を検討してしまえば、以後は無理なく対応できるものと考えられる。

### 4. 研究成果報告に関連した対応

#### 1) 死因つきデータセットの保持可能な期間

昨今の臨床研究実施に関連した不祥事に関連し、研究の質保証の観点から(結果の再現性など)、研究成果報告後に、データ再検証・再解析対応が必要なことも考慮しておく必要があるといえる。また、このような事例に限らず、論文化手続き(査読課程)などが長期化して再解析実施の可能性が考えられる場合に、死因付きデータをいつまで保持できるかは課題であるといえる。公的研究費を継続できるような場合には、死因情報を保持することは可能であるとはいえ、研究の質や結果の再現性を保証するという点では、解析に使用した死因付きのデータセットが長期に保持できることが望ましいといえる。

#### 2) 死因つきデータセットが使用できる環境が限定され

### る

提案する方法では、死因つきデータセットの利用場所がアーカイブセンター(あるいは申請書に記載した研究者の所属する機関)に限られることになる。したがって、研究の解析もこの利用場所に限られることになる。アーカイブデータの利用規定も合わせて、アーカイブセンターで対応する場合は、データ解析を行える環境(物理的な環境、統計家の配置等)について検討する必要がある。また、死因データの申請者の所属機関で実施する場合には、アーカイブデータの外部利用の規約等の整備も必要があるといえる。

### E. 結論

死因情報を付与した形でのデータアーカイブ化が難しい本邦の現状を鑑みて、現制度下での運用案を提案した。提案した運用案を複数コホート研究から構成される国内の循環器疫学研究に適用したところ、死因照合率は99.7%であり、照合作業の技術的側面、作業手順化の面で問題はなく、運用案の一つになりうると考えられた。想定するアーカイブセンター等で実際に運用を行う上では、データ利用の規約、データ利用の物理的環境についてあらかじめ検討すべき事項があると考えられた。

### F. 研究発表

1. 論文発表
  2. 学会発表
- いずれもなし

### G. 知的財産権の出願・登録状況

(予定を含む。)

1. 特許取得
  2. 実用新案登録
  3. その他
- いずれもなし

## 付録

### 作業1.人口動態統計データの読み込み (SAS プログラム#1)

<データセット構造>

以下のようなデータセットが1年1ファイル提供されているとする(ただし変数は申請内容により異なる)。

例)

	調査年	都道府県	保健所	市町村	性別	生年月日				死亡年月日				原死因
						元号	年	月	日	元号	年	月	日	
1	1999													
2	1999													
3	1999													
4	1999													
5	1999													
6	1999													
7	1999													
8	1999													
9	1999													
10	1999													
11	1999													
12	1999													
13	1999													
14	1999													
15	1999													
16	1999													
17	1999													
18	1999													

H11(1999)年から H24 年までのデータを読み込む

Cドライブの「C:¥死因照合¥照合\_hxx」の階層に、shibo\_hxx.csv ファイルが保存されている状態 (xx は年数)。

<SAS プログラム例 #1>

```
%MACRO CSVREAD (YEAR)
```

```
data JINKO.H&YEAR ;
```

```
infile "C:¥死因照合¥照合_h&year.¥sibo_h&year.csv" dlm=',' MISSOVER DSD lrecl=32767
```

```
firstobs=1 ;
```

```
input
```

```
COL1 $2.
```

```
COL3 $1.
```

```
COL14 $2.
```

```
COL15 $2.
```

```
COL16 $3.
```

```
COL17 $1.
```

```
·
```

```
·
```

```
·
```

```
;
```

CSV ファイル形式で提供された下記の人口動態データを読み込み、調査年毎に SAS データセットとして格納する。

申請内容に応じて変数は変更

調査年,都道府県,保健所,市区町村,  
性別,生年月日\_\_元号,生年月日\_\_年,  
生年月日\_\_月,生年月日\_\_日,死亡日\_\_元号,  
死亡日\_\_年,死亡日\_\_月,死亡日\_\_日,原死因,

```
%MEND ;
%CSVREAD(11);%CSVREAD(12);%CSVREAD(13);%CSVREAD(14);
%CSVREAD(15);%CSVREAD(16);%CSVREAD(17);%CSVREAD(18);
%CSVREAD(19);%CSVREAD(20);%CSVREAD(21);%CSVREAD(22);
%CSVREAD(23); %CSVREAD(24);
```

DATA JINKO ;

SET JINKO.H11(in=A) JINKO.H12(in=B) JINKO.H13(in=C) JINKO.H14(in=D) JINKO.H15(in=E)

JINKO.H16(in=F) JINKO.H17(in=G) JINKO.H18(in=H) JINKO.H19(in=I) JINKO.H20(in=J)

JINKO.H21(in=K) JINKO.H22(in=L) JINKO.H23(in=M) JINKO.H24(in=N) ;

run;

調査年ごとにセットして 1 つの SAS データセットとして格納する。

 データセット“JINKO”

## 作業 2. 研究全体の市町村(コード)をリストアップした一覧表を読み込む

追跡データに市町村コードを事前に付与しておくか、各研究と該当する市町村の対応リストを作成し読み込む。

《データイメージ》

研究識別コード	都道府県	都道府県コード	コホート名	保健所	保健所コード	市町村
A01	北海道	1	a	aa	11	〇〇町
A01	北海道	1	a	bb	22	〇△町
A03	秋田	5	b	cc	33	〇×町
A05	秋田	5	c	dd	44	△×町
	.					
	.					
	.					
A30	沖縄					

## 作業 3. 追跡調査データベースの対象データから照合を行いたい死亡者のみを抽出

市町村コードが事前に入力されていない場合は、作業 2 のリストから付与し読み込む

研究識別コード	市町村	ID	性別	生年月日	死亡日	生死(死亡=2)	市町村コード
A01	〇〇町	1000222	1	1906-02-25	2009-11-07	2	01548
A01	〇〇町	1000200	1	1906-11-28	2001-11-21	2	01548
A01	〇〇町	1000030	1	1907-04-02	2008-12-31	2	01548
A01	〇〇町	1004000	2	1911-01-14	2006-04-29	2	01548
A01	〇〇町	1004020	1	1911-06-21	2008-07-13	2	01548
A01	〇△町	1005030	2	1911-06-29	2001-05-13	2	01575
A01	〇△町	1105030	1	1909-01-01	2000-02-22	2	01575
A01	〇△町	1105044	1	1909-12-01	2000-04-30	2	01575
A01	〇△町	1106048	1	1939-03-09	2000-10-25	2	01575
A01	〇△町	1126048	1	1924-09-19	2011-12-25	2	01575
A03	〇×町	9987	2	1934-09-20	2003-05-11	2	05324
A03	〇×町	6989	2	1952-07-30	2006-03-17	2	05324
A03	〇×町	7878	1	1922-11-19	2006-06-17	2	05324
A03	〇×町	9900	1	1953-01-04	2011-03-11	2	05324
A03	〇×町	6677	1	1910-01-07	2005-09-30	2	05324
A03	〇×町	6789	1	1928-07-18	2008-05-03	2	05324
A05	△×町	99993737	2	1944-10-24	2005-06-13	2	05367
A05	△×町	98883799	1	1955-11-28	2010-06-23	2	05367
A05	△×町	99994444	2	1910-10-24	2012-08-09	2	05367
A05	△×町	98896677	1	1966-01-02	2008-07-13	2	05367

 読み込んだ照合用死亡者データセットを “S\_SICHO” とする。

## 作業4 照合対象データと人口動態データをキー変数で結合

データセット JINKO と S\_SICHO とを性別、生年月日、市町村コード、死亡日の情報でマージする

<SAS プログラム例 #4>

```
(1) data S_SICHO;  
    set S_SICHO ;
```

```
mkey=compress(put(sex,best.)||"_"||put(birth,yymmdd10.)||"_"||shicyou_c||"_"||put(shibou,yymm  
dd10.));
```

```
run ;
```

(1) 研究対象データで、照合時のキー変数を作成、付与を行う。  
照合時のキー変数とは、  
性別、生年月日、市町村コード、死亡日の情報を結合し、文字列としたものである。  
(プログラム中"mkey"として記載)

```
(2) data JINKO ;  
    set JINKO ;
```

```
mkey=compress(put(sex,best.)||"_"||put(birth,yymmdd10.)||"_"||shicyou_c||"_"||put(shibou,yymm  
dd10.));
```

```
run ;
```

(2) 人口動態データで、照合時のキー変数を作成、付与を行う。

```
(3) data S_SHOGO S_ERR JINKO_ERR ;  
    merge S_SICHO(in=a) JINKO(in=b) ;  
    by MKEY ;  
    if a and b then output S_SHOGO ;  
    else if a and b^=1 then output S_ERR ;  
    else if a^=1 and b=1 then output JINKO_ERR ;  
run ;
```

(3) 作成したキー変数を使用し、  
研究の死亡データと、人口動態データを結合する。  
結合した対象者は死因照合例としてデータセットに格納する。  
結合しなかった対象者は、照合不能例として別のデータセットに格納する。

```
(4) data DUPTAI ;  
    set S_SHOGO ;  
    by mkey ; if first.mkey then count=0 ; COUNT+1 ;  
        IF COUNT=2 ;  
        KEEP MKEY;  
run ;
```

```
data DUPlist ;  
merge S_SHOGO duptai(in=a) ;  
by mkey ;  
if a ;  
run ;
```

(4) 照合候補例が複数発生した場合のデータセットを格納する。

## 作業5 死因照合

上記作業4で作成したデータセットをもとにして、(1)照合例、(2)未照合例、(3)重複例(候補例が複数存在したものを)をリストアップする

- (1) 照合例 データセット … “S\_SHOGO”  
 (2) 未照合例 … “S\_ERR”  
 (3) 重複例(候補例が複数存在) … “DUPLIST”

JALS では、これらのデータセットから EXCEL へ出力するプログラムを作成済みである。

### a) 人口動態統計 照合状況集計表

受領人口動態データ件数(～2012)= XXXXXX		重複削除後人口動態データ件数(～2012)= XXXXXX		集計			率	
統合研究死亡者数	人口動態統計対象者数	結合件数	結合者数	結合中結合キー重複者数	統合研究対象者のみ数	人口動態統計対象者のみ数	結合率 【結合者数÷統合研究死亡者数×100】	結合中結合キー重複率 【結合中結合キー重複者数÷結合者数×100】
7,137	XXXXX	7,104	7,099	5	XX	XXXXXXX	99.47	0.07

### b) コホート別照合・未照合数表

コホート毎不一致対象者数		
コホート	JALS死亡者数	不一致対象者数
1	126	1
2	181	4
3	2120	4
4	76	0
5	127	0
6	252	4
7	520	6
8	118	0
9	0	0
10	112	0
11	295	3
12	0	0
13	899	2
14	93	1
15	0	0
16	0	0
17	356	1

・  
 ・  
 ・  
 ・

c) 未照合 (不一致) 対象者一覧

受領人口動態データ件数 (~2012) = XXXX 重複削除後人口動態データ件数 (~2012) = XXXX 統合研究対象者のみ数 = XXXX				統合研究死亡者数 = XXXX 結合者数 = XXXXX					
コホート番号	ローカルコード	ID	性別 男=1,女=2	生年月日	都道府県・保健所・市町村結合	死亡日	生存状態コード	死因	市町村名
1		aaaaaaa	1	19XX-XX-X	0xxxx	20XX-XX-X	2		〇〇市
2		bbbbbbb	1	19XX-XX-X	0xxxx	20XX-XX-X	2		〇△町
3		ccccccc	1	19XX-XX-X	0xxxx	20XX-XX-X	2		△△町
5		ddddddd	2	19XX-XX-X	0xxxx	20XX-XX-X	2		△×町
6		eeeeeee	1	19XX-XX-X	0xxxx	20XX-XX-X	2		××村
6		fffffffff	1	19XX-XX-X	0xxxx	20XX-XX-X	2		〇〇町
7		sssssss	2	19XX-XX-X	0xxxx	20XX-XX-X	2		△△市
8		hhhhhhh	2	19XX-XX-X	0xxxx	20XX-XX-X	2		〇〇町
9		iiiiiiii	2	19XX-XX-X	0xxxx	20XX-XX-X	2		××町
9		jjjjjjjj	1	19XX-XX-X	0xxxx	20XX-XX-X	2		△△〇市××区
11		kkkkkkk	2	19XX-XX-X	0xxxx	20XX-XX-X	2		△△〇市××区
13		lllllllll	2	19XX-XX-X	0xxxx	20XX-XX-X	2		△△〇市××区
13		mmmmmm	2	19XX-XX-X	0xxxx	20XX-XX-X	2		△△〇市××区
20		nnnnnnn	1	19XX-XX-X	0xxxx	20XX-XX-X	2		〇△町
21		oooooooo	2	19XX-XX-X	0xxxx	20XX-XX-X	2		××市
31		ppppppp	2	19XX-XX-X	0xxxx	20XX-XX-X	2		××市
32		qqqqqqq	2	19XX-XX-X	0xxxx	20XX-XX-X	2		××市
33		rrrrrrrr	1	19XX-XX-X	0xxxx	20XX-XX-X	2		××市

d) 重複対象者一覧

受領人口動態データ件数 (~2012) = XXXX 重複削除後人口動態データ件数 (~2012) = XXXX 重複者数 = 3				統合研究死亡者数 = XXXXX 結合者数 = XXXXX										
コホート番号	ローカルコード	ID	JALS情報					人口動態統計情報						
			性別 男=1,女=2	生年月日	都道府県・保健所・市町村結合	死亡日	生存状態コード	死因	市町村名	性別 男=1,女=2	生年月日	都道府県・保健所・市町村結合	死亡日	原死因
3		AAAAA	1	1935-XX-XX	0XXXX	20XX-XX-X	2		〇△町	1	1935-XX-XX	0XXXX	20XX-XX-X	aaa
3		AAAAA	1	1935-XX-XX	0XXXX	20XX-XX-X	2		〇△町	1	1935-XX-XX	0XXXX	20XX-XX-X	aaa
3		ZZZZZZZ	1	1940-XX-XX	0XXXX	20XX-XX-X	2		××市	1	1940-XX-XX	0XXXX	20XX-XX-X	bbbb
3		ZZZZZZZ	1	1940-XX-XX	0XXXX	20XX-XX-X	2		××市	1	1940-XX-XX	0XXXX	20XX-XX-X	bbbb
7		WWWWW	2	1930-XX-XX	0XXXX	20XX-XX-X	2		〇〇市	2	1930-XX-XX	0XXXX	20XX-XX-X	cccc
7		WWWWW	2	1930-XX-XX	0XXXX	20XX-XX-X	2		〇〇市	2	1930-XX-XX	0XXXX	20XX-XX-X	dddd