

日本版死亡データベースの構築に関する研究 (1)

石井 太

はじめに

わが国の平均寿命は 20 世紀後半に著しい伸長を遂げ、2010 年には、男性 79.64 年、女性 86.39 年と、現在、世界有数の長寿国として国際的に見てもトップクラスの水準を誇っている。そして、「日本の将来推計人口（平成 24 年 1 月推計）」によれば、平均寿命は 2060 年には男性 84.19 年、女性 90.93 年（死亡中位仮定）に達すると推計されており、今後も長寿のフロントランナーとして走り続けるものと見込まれる。

このような、世界にも類を見ない長寿化のメカニズムと背景、また、これらが日本社会に与える影響を的確に捉えるためには、死亡研究に最適化したデータベースの構築とそれに基づいた人口学的分析を中心としつつ、社会・経済面や医学・生物学的視点などに基づく学際的アプローチが不可欠であると考えられる。しかしながら、わが国において、死亡研究に最適化した生命表データベースとして構築されたものは現在のところない。

本研究は、このような問題意識の下、長寿化に関する多角的かつ学際的なアプローチに基づく総合的研究のための基礎とするべく、海外の先進的な死亡データベースの事例を参照し、わが国の生命表を人口分析の目的から総合的に再編成した「日本版死亡データベース (Japanese Mortality Database、以下 JMD)」を構築することを目的としたものである。

1 Human Mortality Database について

近年、死亡・寿命分野の研究論文や学会報告において、Human Mortality Database(以下、HMD と略す)が引用されているのを目にする機会が多くなってきている。いまや、HMD は死亡研究者の共通言語といっても過言ではないくらい、寿命研究になくはならない存在になったといえることができる。本節は、海外の先進的な死亡データベースの代表例であり、また、JMD の構築に当たって基礎とした HMD の作成方法についてレビューする。

HMD では、入力データから最終的な生命表を作成する行程について、全ての国・地域または時代に統一的に適用する方法論を確立し、これを方法手順書 (methods protocol) の形に取りまとめている (Wilmoth et al. 2007)。この方法手順書は、HMD サイトにおいて公開されており、この手順書に従って全ての生命表が作成され、公表されている。以下、この方法手順書に従って、その作成方法の概要を述べる。

1.1 生命表算定の概略と死亡データの調整

生命表は、おおまかに以下の6段階の手順に基づいて作成されている。

1. 出生数: 可能な限り長期間の性別年間出生数を収集する。
2. 死亡数: 可能な限り詳細なレベルの死亡数を収集し、生データがまとめられている場合、統一的な手法によって、満年齢別・死亡年別・出生年別の死亡数を推定する。
3. 人口規模: 各年1月1日現在推計人口を、統計データから得るか、またはセンサスと出生・死亡数に基づいて推計する。
4. リスク対応生存延べ年数 (Exposure-to-risk): ある年齢×時間区間において、死亡リスクにさらされる生存延べ年数を推計する。
5. 死亡率: 死亡率は、ある年齢×時間区間に属する死亡数の、対応する区間の生存延べ年数に対する比として計算する。
6. 生命表: 生命表を作成するため、まず死亡確率が死亡率より計算され、これにより生命表が求められる。

HMD では様々な国の様々な時代の生命表を作成しているが、その基となっているデータの形式は当然のことながら多種多様である。基データが HMD の作成に必要となる分類よりも詳細なものであれば特別な調整を行わずに HMD が必要とする形式に変換することが行えるが、一般的には何らかの調整が必要となる。また、詳細なデータが存在する場合であっても、不詳データの取扱いなど調整が必要となることもある。HMD では、死亡データについて、以下の4種類の調整法に基づき、基データの調整を行うこととしている。

1. 年齢不詳死亡数のあん分
2. 1歳×1年死亡数のレキシストライアングルへの分割
3. 5歳×1年死亡数の1歳×1年死亡数への分割
4. 開放区間死亡数のレキシストライアングルへの分割

ここで、「レキシストライアングル」とは、死亡者数を横軸に時間、縦軸に年齢を取ったレキシス図で表した際、ある暦年の特定の満年齢の死亡数はレキシス図内の正方形内の死亡点に対応するが、これをさらに生年別に分割した直角二等辺三角形領域のことを指す(図1)。具体的には、図1において、暦年 t 、満 x 歳の死亡数は、 $\{(T, X) | T \in [t, t+1), X \in [x, x+1)\}$ という黒枠の正方形領域で与えられるが、このうち、左上の直角二等辺三角形の領域は $t-x-1$ 年生まれコーホート、右下の直角二等辺三角形の領域は $t-x$ 年生まれコーホートであり、生年が違うコーホートとなっている。レキシストライアングルはこれらを区別したものであり、左上の領域を上方トライアングル (upper-triangle)、右下の領域を下方トライアングル (lower-triangle) と呼ぶ。

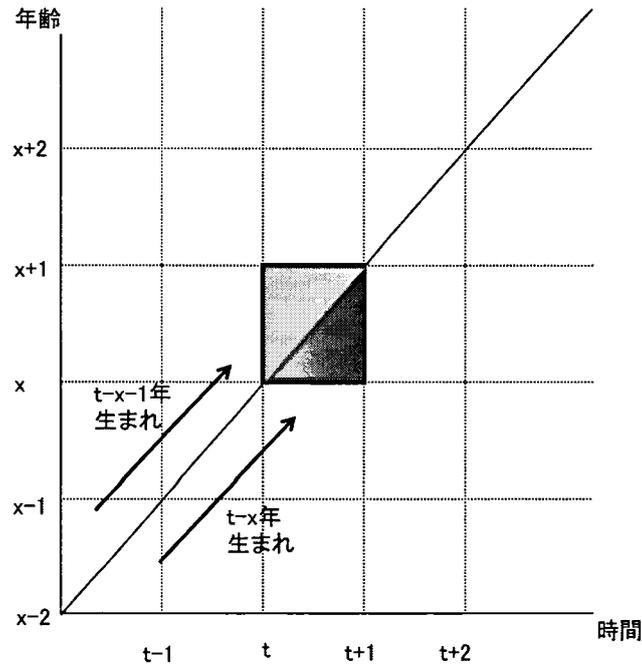


図1 レキシストライアングル

本年度、JMDにおいて開発の対象としたわが国の全国規模での死亡データに関しては、基データの段階でレキシストライアングルレベルのデータが存在することから、2~4の調整は必要ないため、ここでは、1の調整についてのみ述べる。

年齢不詳の死亡データについては、生年も不明であることから、上方・下方トライアングルとも同じ率で拡大することにより不詳あん分を行う。すなわち、

$D_L(x, t)$: 暦年 t , 年齢区間 $[x, x + 1)$ の下方トライアングルの死亡数

$D_U(x, t)$: 暦年 t , 年齢区間 $[x, x + 1)$ の上方トライアングルの死亡数

$D_{UNK}(t)$: 暦年 t の年齢不詳死亡数

$D_{TOT}(t)$: 暦年 t の死亡総数

として、

$$D_L^*(x, t) = D_L(x, t) \cdot \frac{D_{TOT}(t)}{D_{TOT}(t) - D_{UNK}(t)}$$

$$D_U^*(x, t) = D_U(x, t) \cdot \frac{D_{TOT}(t)}{D_{TOT}(t) - D_{UNK}(t)}$$

となる。ただし、 $D_L^*(x, t)$ 、 $D_U^*(x, t)$ はそれぞれ不詳あん分後の下方、上方トライアングルの死亡数である。

1.2 1月1日現在人口の推計

1月1日現在人口の推計については、

1. 線形補間 (linear interpolation)
2. センサス間生残者推計 (intercensal survival method)
3. 死滅コーホート推計 (extinct cohort method)
4. 生残比推計 (survivor ratio)

の4つの推計法が用いられている。

1.2.1 線形補間

センサス以外に利用可能な推計人口が1月1日現在でなく、データの間隔が1年またはそれよりも小さい場合には、線形補間によって1月1日現在人口を推計する。ただし、センサス以外のデータを用いる場合には、利用可能であること、かつ、その信頼性が高いと考えられることが条件となる。

1.2.2 センサス間生残者推計

センサス間生残者推計は、センサスとセンサスの間の期間の各年の1月1日人口を推計するための簡便で信頼性の高い方法である。わが国ではセンサスが10月1日に実施されていることから、センサスが1月1日以外の場合にも対応している方法についてのみ述べる。推計は、(1) 既に存在するコーホート、(2) 0歳のコーホート、(3) 新たに生まれるコーホートの3グループに分けて行われる。

(1) 既に存在するコーホート

t と $t+N$ をセンサス間に含まれる最初と最後の1月1日の年次とする。 f_1 を年次 $t-1$ における1番目のセンサス以前の期間の割合、 f_2 を年次 $t+N$ における二番目のセンサス以前の期間の割合とする。 t 年1月1日に満 x 歳であるコーホートに着目する。このコーホートの i 番目のセンサスでの人口を C_i 、また i 番目のセンサスにおける満 x 歳の人口を $C_i(x)$ と表すと、センサスにおいて人口が各年齢内は一様に分布していると仮定すると、

$$C_1 = (1 - f_1)C_1(x-1) + f_1C_1(x)$$

$$C_2 = (1 - f_2)C_2(x+N) + f_2C_2(x+N+1)$$

が成立する。

また、死亡についてもレキシストライアングル内では一様に分布していると仮定すると、当該コーホートの $t-1$ 年の1番目のセンサス後における死亡数は、以下の2つの項から成る。

$$D_a = (1 - f_1^2)D_L(x, t-1)$$

$$D_b = (1 - f_1)^2D_U(x-1, t-1)$$

同様に、当該コーホートの $t+N$ 年の2番目のセンサス前における死亡数は、

$$D_c = f_2^2D_L(x+N+1, t+N)$$

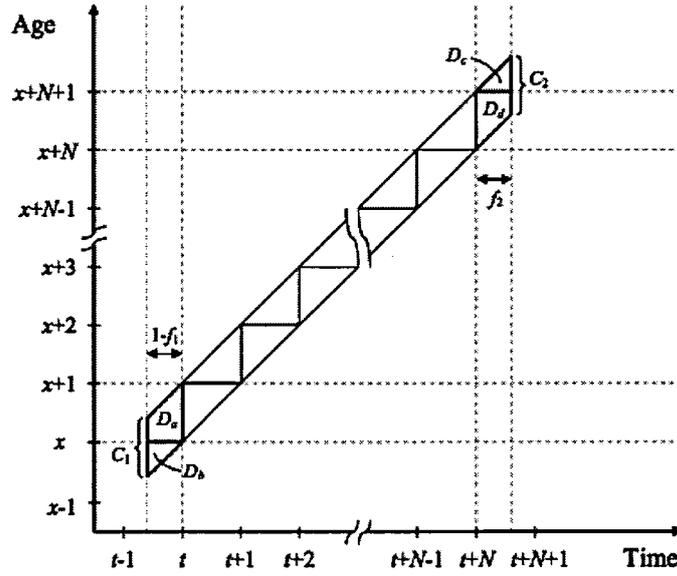


図2 センサス間生残者推計 (既に存在するコーホート)

$$D_d = (2f_2 - f_2^2)D_U(x + N, t + N)$$

C_1 とこれらの死亡数を用いて C_2 を推計すると、

$$\hat{C}_2 = C_1 - (D_a + D_b) - \sum_{i=0}^{N-1} [D_U(x + i, t + i) + D_L(x + i + 1, t + i)] - (D_c + D_d)$$

$\Delta_x = C_2 - \hat{C}_2$ はセンサス間における移動と誤差による違いとなるので、これを均等に配分することにより、各年の1月1日現在人口を以下のように推計する。

$$P(x + n, t + n) = C_1 - (D_a + D_b) - \sum_{i=0}^{n-1} [D_U(x + i, t + i) + D_L(x + i + 1, t + i)] + \frac{1 - f_1 + n}{N + 1 - f_1 + f_2} \Delta_x$$

ただし、 $n = 0, \dots, N$ である。

(2) 0 歳のコーホート

(1) はセンサス間の最初の1月1日時点で満1歳以上のコーホートのみに有効である。これより年少のコーホートについては別の取扱いが必要となる。まず、最初の1月1日時点で0歳のコーホートについては、(1) で述べた式について、以下の修正が必要となる。

$$C_1 = (1 - f_1)B_{t-1} + f_1C_1(0)$$

$$\hat{C}_2 = C_1 - D_a - \sum_{i=0}^{N-1} [D_U(i, t + i) + D_L(i + 1, i)] - (D_c + D_d)$$

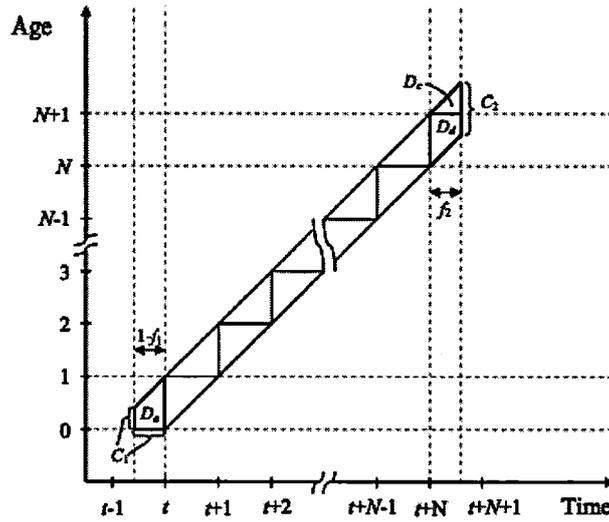


図3 センサス間生存者推計 (0歳のコーホート)

$$P(n, t+n) = C_1 - D_a - \sum_{i=0}^{n-1} [D_U(i, t+i) + D_L(i+1, t+i)] + \frac{\frac{1}{2}(1-f_1^2) + n}{N + \frac{1}{2}(1-f_1^2) + f_2} \Delta_0$$

ただし、 $n = 0, \dots, N$ 、 $\Delta_0 = C_2 - \hat{C}_2$ 、 B_{t-1} は1番目のセンサスの暦年における出生数である。

(3) 新たに生まれるコーホート

最後に、センサス間において新たに生まれるコーホートについて考える。

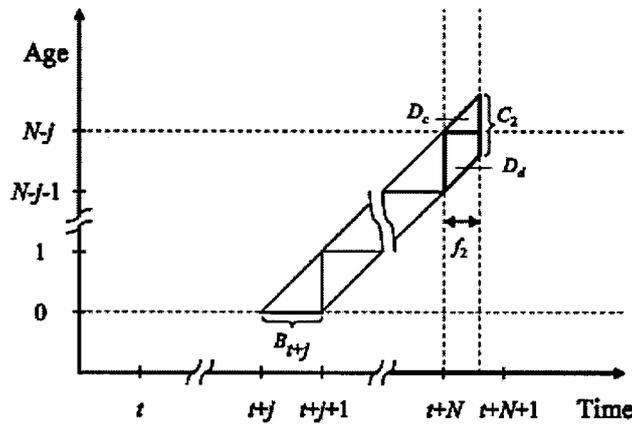


図4 センサス間生存者推計 (新たに生まれるコーホート)

$t+j$ 年に生まれたコーホートは、2番目のセンサスの直前の1月1日に $K = N-j-1$ 歳となる。このとき、

$$\hat{C}_2 = B_{t+j} - D_L(0, t+j) - \sum_{i=1}^K [D_U(i-1, t+j+i) + D_L(i, t+j+i)] - (D_c + D_d)$$

$$P(k, j+k+1) = B_{t+j} - D_L(0, t+j) - \sum_{i=1}^k [D_U(i-1, t+j+i) + D_L(i, t+j+i)] \\ - (D_c + D_d) + \frac{2k+1}{2K+1+2f_2} \Delta'_{i+j}$$

ただし、 $k=0, \dots, K$ 、 $\Delta'_{i+j} = C_2 - \hat{C}_2$ である。

センサス間の1月1日現在人口については以上のようにして推計が行われるが、最初のセンサスより前、および最後のセンサスより後の短い期間においては、この方法を利用して、センサス人口に単純に死亡数を加えたり減じる事により1月1日人口を推計する。これは、移動と誤差による修正分を含まない事から、直近の期間における推計人口は暫定的なものとして捉える必要があり、それ以降に新たなセンサスが実施された場合にはこれに基づいてセンサス間生残者推計を行って数値を改定することとなる。

このような推計の例を図5に示す。

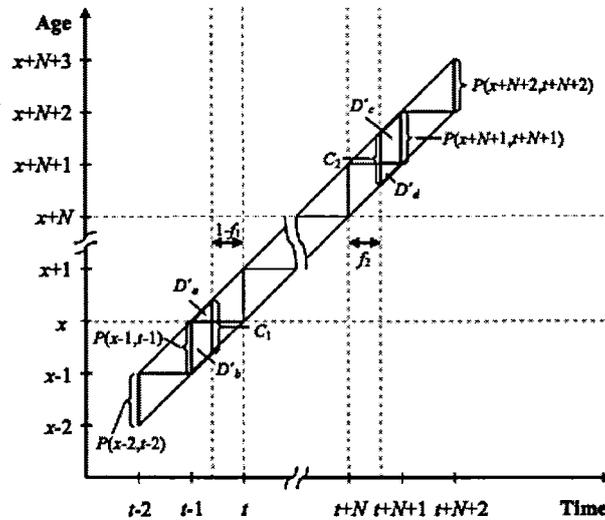


図5 センサス前・後生残者推計

$t-x-1$ 年生まれコーホートの $t-1$ 、 $t-2$ 年1月1日現在人口は、

$$P(x-1, t-1) = C_1 + D'_a + D'_b$$

$$P(x-2, t-2) = C_1 + D'_a + D'_b + D_L(x-1, t-2) + D_U(x-2, t-2)$$

同じコーホートの $t+N+1$ 、 $t+N+2$ 年1月1日現在人口は、

$$P(x+N+1, t+N+1) = C_2 - D'_c - D'_d$$

$$P(x+N+2, t+N+2) = C_2 - D'_c - D'_d - D_U(x+N+1, t+N+1) + D_L(x+N+2, t+N+1)$$

ここで、 D'_a 、 D'_b 、 D'_c 、 D'_d は、それぞれ、 D_a 、 D_b 、 D_c 、 D_d が属するレキシストライアングルの残りの部分を示す。すなわち、 $D_a + D'_a = D_L(x, t-1)$ などを満たす。

1.2.3 死滅コーホート推計

この方法では、あるコーホートの x 歳の人口を、それ以降の全ての死亡数を足し上げる
ことによって推計する。すなわち、

$$P(x, t) = \sum_{i=0}^{\infty} [D_U(x+i, t+i) + D_L(x+i+1, t+i)]$$

この方法では、 x 歳以降の国際人口移動がない事を仮定しており、具体的には、80 歳
以上の人口の推計のみに用いられる。この方法の適用にあたってはどのコーホートが
死滅しているかを定める必要がある。これにあたり、Väinö Kannisto により提案され、
Kannisto-Thatcher oldest-old mortality database で用いられた方法を使うこととした。
この方法では、あるコーホートが死滅しているとは、観察期間の最後 (t_n 年の 1 月 1 日)
において、ある年齢 ω に達していることと定義される。また、この年齢 ω は、

$$\omega = \arg \min_x \{ \tilde{D}(x, t_n, l) \leq 0.5 \}$$

で定義される。ただし、 $l = 5$ であり、

$$\tilde{D}(x, t_n, l) = \frac{1}{l} \sum_{j=1}^l \sum_{i=0}^{j-1} [D_U(x+i, t_n-j+i) + D_L(x+i+1, t_n-j+i)]$$

とする。

1.2.4 生残比推計

生残比推計は 80 歳以上人口のうち、概ね死滅しているコーホートに適用される方法で
ある。これは、観察期間の最後で少なくとも 90 歳以上であり、まだ死滅していないコー
ホートに適用する。これまでに、この方法に関する多くのやり方が提案・研究されてきて
いるが、ここでは、比較研究において最も信頼されると示された方法を用いる。

生残比を、 t 年 1 月 1 日に x 歳のコーホートが、 k 年前に生存していたコーホートに対
して残存している者の割合として定義する。すなわち、

$$R = \frac{P(x, t)}{P(x-k, t-k)}$$

ここで、人口移動がこの間なかったと仮定すると、

$$R = \frac{P(x, t)}{P(x, t) + \dot{D}}$$

と書ける。ただし、 $\dot{D} = \sum_{i=1}^k [D_U(x-i, t-i) + D_L(x-i+1, t-i)]$ である。これを
 $P(x, t)$ について解くと、

$$P(x, t) = \frac{R}{1-R} \dot{D}$$

となる。

仮に、生残比が対象とするコーホートとその前の m コーホートについて概ね一定であったとする。すなわち、

$$R(x, t, k) = \frac{P(x, t)}{P(x-k, t-k)} \approx \frac{P(x, t-1)}{P(x-k, t-k-1)} \approx \dots \approx \frac{P(x, t-m)}{P(x-k, t-k-m)}$$

を仮定する。このとき、 R を m コーホート分をプールしたものとして以下のように推定することができる。

$$R^*(x, t, k) = \frac{\sum_{i=1}^m P(x, t-i)}{\sum_{i=1}^m P(x-k, t-k-i)}$$

R^* と \dot{D} を使って、 $P(x, t)$ は、

$$\tilde{P}(x, t) = \frac{R^*}{1-R^*} \dot{D}$$

で推定できる。

最も単純な形の生残比推計では、この方法によりまず $P(\omega-1, t_n)$ を推定し、ここから死滅コーホート推計法と同様に死亡数を足し上げて戻ることによって、このコーホートの 80 歳までの人口を推定する。すると、 $P(\omega-2, t_n)$ 、 $P(\omega-3, t_n)$ と、次々により若いコーホートに対して同様の方法を適用することができる。

しかしながら、一般的には死亡率改善によって生残比は増加するのが普通であり、この場合、 R^* は R を過小推計し、その結果として \tilde{P} は P を過大推計することとなる。

そこで、生残比推計に以下の修正を行う。

$$\hat{P}(x, t) = c\tilde{P}(x, t) = c \frac{R^*}{1-R^*} \dot{D}$$

ここで c は定数であり、

$$\sum_{x=90}^{\omega-1} \hat{P}(x, t_n) = P(90+, t_n)$$

を満たすように選ぶこととする。ただし、 $P(90+, t_n)$ は観察期間の最後における 90 歳以上の人口推計値である。

1.3 死亡率の推計

1 歳 \times 1 年の場合について考える。死亡率は、死亡数を事象対応延べ年数で割ることにより計算されるので、

$$M_{xt}^p = \frac{D_{xt}^p}{E_{xt}^p}$$

となる。ここで、

$$D_{xt}^p = D_L(x, t) + D_U(x, t)$$

であり、事象対応延べ年数は、

$$E_{xt}^p = \frac{1}{2} [P(x, t) + P(x, t + 1)] + \frac{1}{6} [D_L(x, t) - D_U(x, t)]$$

で推定される。

1.4 生命表の推計

死亡率は 1.3 節で述べた方法で推計されるが、高齢部については統計的安定性を欠くことから、80 歳以上の死亡率を用いてロジスティック曲線への当てはめを行うことにより平滑化された死亡率を求めている。

いま、年齢 $x = 80, 81, \dots, 110+$ に関する D_x, E_x が推計されたとする (簡単のため、 $110+$ は 110 と表す)。このとき、観測された死亡率 $M_x = \frac{D_x}{E_x}$ を、死力 μ_x が Kannisto モデル、すなわち、

$$\mu_x(a, b) = \frac{ae^{b(x-80)}}{1 + ae^{b(x-80)}}$$

に従うとして関数あてはめを行う ($a \geq 0, b \geq 0$)。

パラメータの推定にあたっては、 $D_x \sim \text{Poisson}(E_x \mu_{x+0.5}(a, b))$ を仮定し、以下を最大にする最尤法によって推定値 \hat{a}, \hat{b} を求める。

$$\log L(a, b) = \sum_{x=80}^{110} [D_x \log \mu_{x+0.5}(a, b) - E_x \mu_{x+0.5}(a, b)] + \text{constant}$$

このパラメータ推定値 \hat{a}, \hat{b} を用いて、平滑化された死亡率 \hat{M}_x を、 $\hat{M}_x = \hat{\mu}_{x+0.5} = \mu_{x+0.5}(\hat{a}, \hat{b})$ とする。また、このようにして当てはめられた死亡率は Y 歳以上について用いる。ただし、 Y は死亡数が 100 未満となる最も低い年齢とし、 $80 \leq Y \leq 95$ を満たすものとする。したがって、最終的に得られる死亡率は、

$$M_0, M_1, \dots, M_{Y-1}, \hat{M}_Y, \dots, \hat{M}_{109}, \infty \hat{M}_{110}$$

となる。次に、生命表の死亡率 m_x が上の系列と等しいと仮定する。すなわち、 $m_x = M_x, (0 \leq x \leq Y-1), m_x = \hat{M}_x, (Y \leq x \leq 109), \infty m_{110} = \infty \hat{M}_{110}$ である。

次に、この m_x を q_x に変換する。 a_x については、0 歳以外については $a_x = \frac{1}{2}$ を仮定する。そして、 $x = 0, 1, \dots, 109$ に対して

$$q_x = \frac{m_x}{1 + (1 - a_x)m_x}$$

によって q_x を計算し、開放区間については、 $\infty a_{110} = \frac{1}{\infty m_{110}}, \infty q_{110} = 1$ とする。

また、0 歳については、Coale-Demeny life tables で使用された式を用いた。すなわち、 $m_0 \geq 0.107$ に対して、

$$a_0 = \begin{cases} 0.350 & \text{for females} \\ 0.330 & \text{for males} \end{cases}$$

$m_0 < 0.107$ に対して、

$$a_0 = \begin{cases} 0.053 + 2.800m_0 & \text{for females} \\ 0.045 + 2.684m_0 & \text{for males} \end{cases}$$

である。

生命表を完成させるためには、まず、 p_x を

$$p_x = 1 - l_x$$

とし、 $l_0 = 100000$ として、

$$l_x = l_0 \prod_{i=0}^{x-1} p_i$$

により、 l_x を計算する。また、 $x = 0, \dots, 109$ について、

$$d_x = l_x q_x$$

とし、 ${}_{\infty}d_{110} = l_{110}$ とする。次に、 $x = 0, \dots, 109$ について、

$$L_x = l_x - (1 - a_x)d_x$$

とし、 ${}_{\infty}L_{110} = l_{110} a_{110}$ とする。同様に $x = 0, \dots, 109$ について、

$$T_x = \sum_{i=x}^{109} L_i + {}_{\infty}L_{110}$$

で、 $T_{110} = {}_{\infty}L_{110}$ である。最後に、

$$e_x = \frac{T_x}{l_x}$$

となる。

なお、以上の全ての工程において、実際には推計期間内で対象となっている領域の大きさが変わることがあり、その場合にはこの影響を考慮する「領域調整」を行う必要がある。しかしながら、本節では、領域調整がない場合についてのみの記述を行った。

2 日本版死亡データベースの構築に関する検討点

本年度は、日本版死亡データベースの構築にあたり、まず、全国版の生命表の作成を行うこととした。そこで、前節における HMD の生命表作成方法のレビューに基づき、全く同じ方法を用いて日本全国の生命表の作成を試みた。その結果として、死亡率のロジスティック曲線への当てはめにおける数値計算上の誤差を除き、概ね同じ生命表が再現できることが確認された。

HMD の利点は、全ての生命表が基本的に同一の方法で作成され、国際比較可能性に優れていることであり、わが国の固有の状況を優先させることは適当ではないと考えられる

が、一方で、HMD の生命表の作成方法に関する詳細なレビューを通じ、一定の比較可能性を担保しつつも、わが国の生命表という観点から作成方法に改善が可能である点がいくつかあることが明らかとなった。そこで、JMD では、単に HMD の方法論をそのまま適用するだけでなく、いくつかの点について、比較可能性を阻害しない範囲で、日本の死亡状況により適合させるための変更を行うこととした。主な変更点は以下の通りである。

- 人口の不詳あん分について、総務省統計局の現在推計人口で用いられているのと同様の基準人口を用いた。
- 1月1日人口の推計の際に、国勢調査の出生月別人口のデータを利用した。また、直近の期間について HMD で用いられている線形補間を用いず、他の方法から推計を行った。
- 沖縄の返還に関する領域調整を、わが国の人口動態統計と整合性がとれるようにした。
- a_0 の推定には、日本の近年のデータを用いて別途推計式を作成した。

以下、これらの方法論の改善点についてより詳細に述べることとする。

2.1 人口の不詳あん分

一般的に、人口動態統計はある領域の中で起きる事象を対象とすることが多い。しかしながら、従来から、わが国の人口動態統計で最もよく用いられているのは、「日本における日本人」に関する統計であり、日本における外国人のデータは含まれていない。一方、人口静態統計である国勢調査は外国人を含めた日本に常住する人口となっている。このため、国勢調査では、外国人を含めた「総人口」だけではなく、日本人のみを対象とした「日本人人口」に関する統計表も作成を行っており、人口動態率を作成する際には、日本人人口を用いて行うこととなる。

2005年までの国勢調査においては、人口動態率を作成するための基本的な属性である性・年齢別人口について、性別のみわかるが、年齢・国勢が不詳である人口のみが集計されていた。したがって、国勢調査の統計表上では、総人口の中に性別の年齢不詳人口があるものの、日本人人口には年齢不詳が計上されていない。しかし、これは日本人の年齢不詳がないということではなく、年齢不詳人口の国籍別内訳が集計されていなかったということによるものである。総務省統計局においては、国勢調査と国勢調査の間の期間に関する推計人口を作成するために、国勢調査人口について不詳按分を行った基準人口を作成している。これは、2005年までの国勢調査では、性・都道府県別に年齢不詳人口を日本人・外国人にあん分し、これをさらに年齢構成比であん分することにより基準人口を作成している。

ところが、2010年の国勢調査では、不詳あん分の精度を向上させる観点から、世帯の家族類型（単独世帯・単独世帯以外）別に、性・国籍・年齢別人口を集計し、また、国

籍や年齢の不詳についても、国籍のみ不詳、年齢のみ不詳、国籍・年齢とも不詳に分けた集計結果を基礎データとしてあん分が行われている。具体的には、都道府県・世帯類型・性別に、まず国籍のみ不詳人口を各年齢で日本人・外国人にあん分し、次に、国籍・年齢とも不詳の人口を日本人・外国人に按分して年齢のみ不詳人口に加える。最後に、年齢不詳人口を年齢構成比によってあん分して全てを合計し、不詳あん分人口を作成する。

このように、わが国の国勢調査においては、推計人口に用いるためのあん分の方法が存在しており、各種の人口動態率等もこれを基準として算出されている。ところが、HMDでは日本人人口を基礎としているものの、国籍・年齢とも不詳人口を全て日本人人口にあん分することにより国勢調査を利用している。特に、近年、国勢調査の国籍・年齢不詳人口は増加をしてきており、この方法によると死亡率を過小に評価する可能性がある。そこで、JMDにおいては、HMDのあん分方法によらず、総務省統計局の現在推計人口で用いられているのと同じの基準人口を用いることとした。

2.2 1月1日人口の推計方法

わが国の国勢調査は10月1日現在で調査が行われているため、1.2.2節で述べた通り、センサス間生残者推計では、 f_1 :年次 $t-1$ における1番目のセンサス以前の期間の割合、 f_2 :年次 $t+N$ における二番目のセンサス以前の期間の割合を用いて推計が行われる。これらの割合は複数の目的で使われているが、そのうちの 하나가10月1日現在の満年齢人口を生まれ年に分けることである。しかしながら、国勢調査では各年齢について出生月別人口の集計があることから、 f_1 、 f_2 によらずに実績の割合を使うことが可能である。また、わが国ではひのえ午の年など、ある年の出生数とその前後の年とギャップを持つことなどもあることから、単純に f_1 、 f_2 を用いると年齢別の人口動態率が不自然となることがある。そこで、JMDでは、10月1日現在の満年齢人口を生まれ年に分けるための割合については、国勢調査の出生月別人口のデータを利用することとした。

また、HMDでは1995年以降の期間について、センサス間生残者推計・死滅コーホート推計・生残比推計によらず、総務省統計局の現在推計人口を線形補間することによって1月1日現在人口を算出している。しかしながら、JMDにおいては、国勢調査間の人口ギャップをより滑らかにつなぐ観点から、線形補間を用いず、センサス間生残者推計等の方法を用いることとした。

2.3 沖縄の返還に関する領域調整

1節では簡単な記述のみにとどめたが、生命表の推計期間内で対象となっている領域の大きさが変わることがあり、そのような場合、HMDではこの影響を考慮した「領域調整」を行っている。HMDにおける日本の生命表については、1972年5月15日に沖縄県が米国から返還されたことから、これに関する領域調整が行われている。

領域調整については、領域が変化する時点の前後の年齢別人口・出生数の増減割合を領域調整係数として各種統計から推定し、これを用いて各種の推計法に修正を加える仕組みとなっている。HMDでは、1972年1月1日に領域が変化すると仮定して領域調整係数を算出している。しかしながら、わが国の人口動態統計では、1972年は沖縄県の事象を含まず、1973年から含む整理とされている。そこで、JMDでは人口動態統計と整合性がとれるよう、領域が変化する時点を1973年1月1日とし、これにあうように領域調整係数を独自に算出することとした。JMDとHMDの領域調整係数を比較したものが表1である。

表1 領域調整係数

	JMD(1973年)		HMD(1972年)	
	男性	女性	男性	女性
出生	1.01124	1.01106	1.01113	1.01118
0～4歳	1.01086	1.01106	1.01135	1.01078
5～9歳	1.01239	1.01266	1.01281	1.01217
10～14歳	1.01440	1.01468	1.01509	1.01433
15～19歳	1.01214	1.01238	1.01228	1.01166
20～24歳	1.00696	1.00775	1.00732	1.00695
25～29歳	1.00609	1.00642	1.00630	1.00599
30～34歳	1.00783	1.00791	1.00817	1.00776
35～39歳	1.00842	1.00835	1.00860	1.00816
40～44歳	1.00787	1.00784	1.00803	1.00763
45～49歳	1.00614	1.00757	1.00708	1.00673
50～54歳	1.00776	1.00782	1.00800	1.00760
55～59歳	1.00697	1.00817	1.00780	1.00740
60～64歳	1.00649	1.00737	1.00720	1.00684
65歳～	1.00740	1.00959	1.00896	1.00851

2.4 a_0 の推定

1節で述べた通り、HMDでは a_0 についてCoale-Demeny life tablesで使用された式を用いている。

表2に1955年以降について、日本の公式生命表とHMDの a_0 を比較したものを示した。これを見ると、HMDの a_0 は1955年以降単調に減少しているのに対し、公式生命表では1980年以降は反転上昇している。この結果として、2005年では公式生命表の a_0 が0.2前後なのに対して、HMDでは0.05前後と極めて低い値となっている。

表2 a_0 の比較

年	公式生命表		HMD		JMD	
	男性	女性	男性	女性	男性	女性
1955	0.18	0.18	0.16	0.16	0.18	0.19
1960	0.18	0.19	0.14	0.13	0.17	0.18
1965	0.16	0.17	0.10	0.10	0.16	0.17
1970	0.15	0.17	0.09	0.09	0.15	0.16
1975	0.14	0.15	0.07	0.08	0.15	0.16
1980	0.14	0.16	0.07	0.07	0.15	0.16
1985	0.16	0.16	0.06	0.07	0.17	0.18
1990	0.18	0.18	0.06	0.06	0.18	0.19
1995	0.20	0.20	0.06	0.06	0.19	0.19
2000	0.19	0.20	0.05	0.06	0.20	0.20
2005	0.21	0.21	0.05	0.06	0.21	0.21
2010	0.22	0.23			0.21	0.21

そこで、近年の公式生命表における a_0 と m_0 の関係を調べるため、両者の関係をグラフに表したものが図 6、7 である。×印で表したのが 1955 年以降の実績値であり、破線で表されているのが、HMD で用いられている関係式である。これによれば、 m_0 が小さくなると次第に実績値と破線の乖離が大きくなっていることがわかる。

そこで、JMD においては、 m_0 が小さい場合には、実績値に基づき、実線で表されるような新たな関係式を用いることとした。JMD で用いている関係式を表すと以下の通りとなる。

男性については、

$$a_0 = \begin{cases} 0.330 & (m_0 \geq 0.107) \\ 0.045 + 2.684m_0 & (0.0612 \leq m_0 < 0.107) \\ 0.132 + 1.264m_0 & (0.00869 \leq m_0 < 0.0612) \\ 0.242 - 11.373m_0 & (m_0 < 0.00869) \end{cases}$$

女性については、

$$a_0 = \begin{cases} 0.350 & (m_0 \geq 0.107) \\ 0.053 + 2.800m_0 & (0.0557 \leq m_0 < 0.107) \\ 0.152 + 1.015m_0 & (0.00637 \leq m_0 < 0.0557) \\ 0.239 - 12.537m_0 & (m_0 < 0.00637) \end{cases}$$

Relation between a_0 and m_0 (Males)

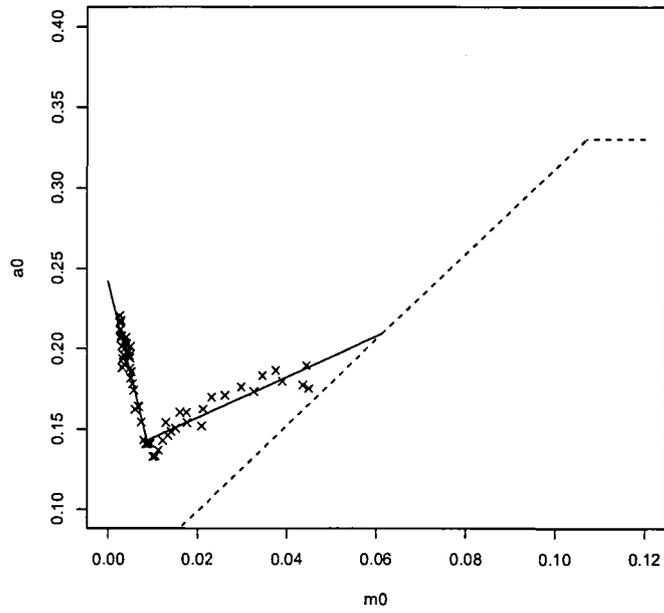


図6 a_0 と m_0 の関係 (男性)

Relation between a_0 and m_0 (Females)

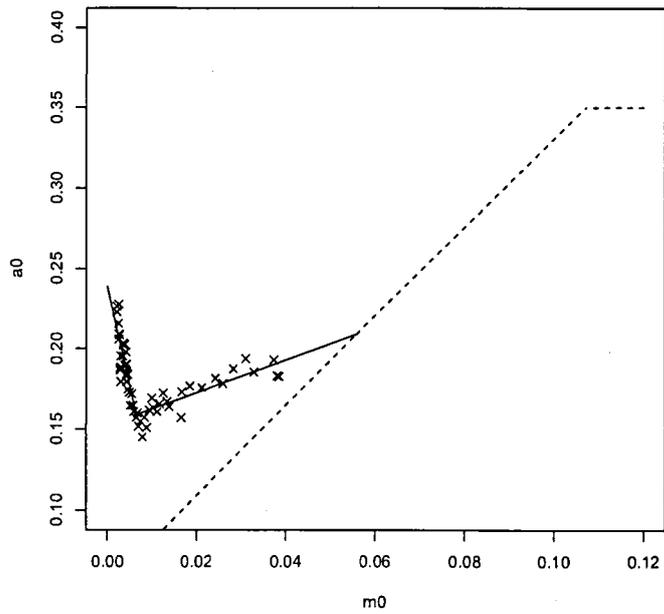


図7 a_0 と m_0 の関係 (女性)

3 全国版生命表の推計結果

前節までに述べた方法に従い、全国版の生命表の作成を行った。本年度の研究においては、男女別、1歳×1年について、1947～2010年の生命表を作成した。詳細な生命表等の結果については別途資料編に掲載するが、ここでは主要な結果として平均寿命の比較について述べる。

図8及び表3は、公式生命表、JMD、HMDによる平均寿命を比較したものである。図8を見ると、1950年代においては公式生命表では簡易生命表が5歳階級で作られていたことなどもあって、JMD・HMDによる平均寿命に乖離が見られるものの、それ以降は三者とも概ね同様の傾向を示していることがわかる。また、1972年の平均寿命は、公式生命表では男性70.50年、女性75.94年、JMDでは男性70.52年、女性75.92年であるのに対し、HMDでは男性70.63年、女性76.02年と0.1年程度高い値を示している。これは、HMDでは沖縄返還に関する領域調整が1972年となっているためであると考えられる。一方、1995年以降では、JMDの値が他の2つに比べてやや異なる傾向を示す部分が見られるが、これは公式生命表とHMDは国勢調査と国勢調査の中間年では総務省の推計人口を用いているのに対し、JMDではこれを用いていないことに原因があると考えられる。

このように、JMDによる平均寿命は、HMDの方法論を基礎とし、時系列比較が可能となるよう同一形式の生命表を提供することを可能にしたとともに、国際比較可能性を阻害しない範囲で、日本の死亡状況により適合させる改善にも成功している。したがって、わが国の死亡研究に特化した生命表として、非常に使いやすいデータ系列を提供することを可能にしたといえることができるであろう。

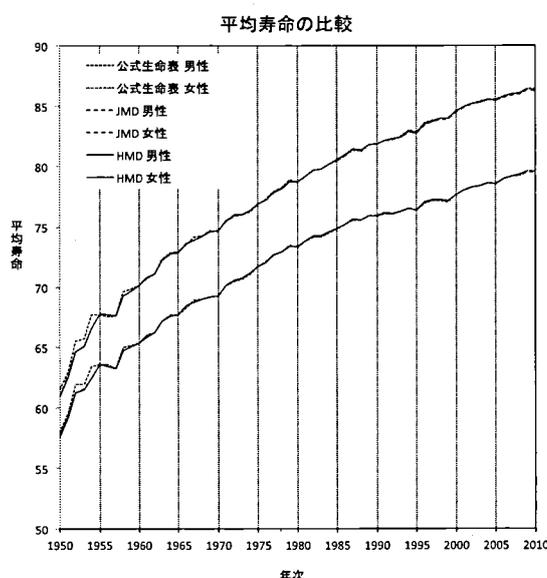


図8 e_0 の比較

表3 e_0 の比較

年	公式生命表		JMD		HMD	
	男性	女性	男性	女性	男性	女性
1947			49.82	53.60	49.78	53.65
1948	55.60	59.40	55.05	58.64	55.02	58.63
1949	56.20	59.80	56.01	59.49	55.97	59.44
1950	58.00	61.50	57.68	60.99	57.58	60.90
1951	59.57	62.97	59.36	62.66	59.26	62.57
1952	61.90	65.50	61.28	64.69	61.25	64.66
1953	61.90	65.70	61.55	65.08	61.54	65.06
1954	63.41	67.69	62.49	66.58	62.49	66.58
1955	63.60	67.75	63.68	67.83	63.62	67.76
1956	63.59	67.54	63.52	67.75	63.44	67.67
1957	63.24	67.60	63.26	67.60	63.26	67.61
1958	64.98	69.61	64.76	69.31	64.76	69.31
1959	65.21	69.88	65.09	69.75	65.09	69.76
1960	65.32	70.19	65.36	70.18	65.31	70.14
1961	66.03	70.79	65.96	70.87	65.91	70.83
1962	66.23	71.16	66.19	71.12	66.19	71.13
1963	67.21	72.34	67.18	72.26	67.19	72.27
1964	67.67	72.87	67.63	72.76	67.64	72.76
1965	67.74	72.92	67.69	72.85	67.67	72.84
1966	68.35	73.61	68.46	73.65	68.46	73.65
1967	68.91	74.15	68.73	73.93	68.74	73.95
1968	69.05	74.30	69.06	74.28	69.06	74.29
1969	69.18	74.67	69.19	74.60	69.20	74.61
1970	69.31	74.66	69.32	74.67	69.32	74.67
1971	70.17	75.58	70.13	75.47	70.13	75.47
1972	70.50	75.94	70.52	75.92	70.63	76.02
1973	70.70	76.02	70.74	76.01	70.75	76.01
1974	71.16	76.31	71.23	76.35	71.24	76.35
1975	71.73	76.89	71.74	76.86	71.74	76.86
1976	72.15	77.35	72.12	77.27	72.12	77.28
1977	72.69	77.95	72.66	77.85	72.66	77.85
1978	72.97	78.33	72.95	78.20	72.95	78.21
1979	73.46	78.89	73.42	78.74	73.42	78.75
1980	73.35	78.76	73.37	78.74	73.38	78.75
1981	73.79	79.13	73.79	79.15	73.80	79.15
1982	74.22	79.66	74.24	79.71	74.25	79.71
1983	74.20	79.78	74.22	79.80	74.23	79.81
1984	74.54	80.18	74.58	80.24	74.59	80.25
1985	74.78	80.48	74.87	80.51	74.88	80.53
1986	75.23	80.93	75.25	80.95	75.26	80.96
1987	75.61	81.39	75.62	81.41	75.63	81.41
1988	75.54	81.30	75.57	81.31	75.57	81.31
1989	75.91	81.77	75.94	81.78	75.95	81.78
1990	75.92	81.90	75.92	81.84	75.93	81.85
1991	76.11	82.11	76.13	82.15	76.15	82.16
1992	76.09	82.22	76.11	82.27	76.12	82.29
1993	76.25	82.51	76.24	82.43	76.26	82.44
1994	76.57	82.98	76.57	82.88	76.58	82.90
1995	76.38	82.85	76.39	82.76	76.41	82.79
1996	77.01	83.59	77.01	83.49	77.03	83.52
1997	77.19	83.82	77.23	83.73	77.23	83.75
1998	77.16	84.01	77.20	83.92	77.20	83.95
1999	77.10	83.99	77.16	83.92	77.16	83.95
2000	77.72	84.60	77.68	84.53	77.70	84.58
2001	78.07	84.93	78.01	84.86	78.05	84.91
2002	78.32	85.23	78.29	85.16	78.32	85.21
2003	78.36	85.33	78.34	85.26	78.36	85.30
2004	78.64	85.59	78.63	85.52	78.65	85.57
2005	78.56	85.52	78.53	85.44	78.55	85.49
2006	79.00	85.81	78.95	85.73	79.00	85.79
2007	79.19	85.99	79.14	85.91	79.20	85.98
2008	79.29	86.05	79.25	85.98	79.31	86.04
2009	79.59	86.44	79.55	86.37	79.61	86.42
2010	79.64	86.39	79.54	86.28		

おわりに

本年度においては、わが国の生命表を人口分析の目的から総合的に再編成した日本版死亡データベースを構築することを目的とし、Human Mortality Database で用いられている方法論を詳細にレビューするとともに、この方法論の評価に基づいた改善法を提案し、新たな方法論を用いて全国版の生命表の作成を行った。

HMD の方法論については、公開されている方法手順書に従ってレビューを行った。その概要は、死亡数や人口に対して年齢不詳あん分等の必要な調整を行い、1月1日現在人口を、線形補間、センサス間生残者推計、死滅コーホート推計、生残比推計の4つの方法により推計するとともに、これに基づいて事象対応延べ年数を推計する。そして、ここから死亡率を関数あてはめなどを利用することにより推定し、生命表の作成を行っている。

一方、HMD の生命表の作成方法に関する詳細なレビューを通じ、一定の比較可能性を担保しつつも、わが国の生命表という観点から作成方法に改善が可能である点があることが明らかとなったことから、JMD では比較可能性を阻害しない範囲で、日本の死亡状況により適合させるための変更を行った。主な変更点は以下の通りである。

- 人口の不詳あん分について、総務省統計局の現在推計人口で用いられているのと同じの基準人口を用いた。
- 1月1日人口の推計の際に、国勢調査の出生月別人口のデータを利用した。また、直近の期間についてHMD で用いられている線形補間を用いず、他の方法から推計を行った。
- 沖縄の返還に関する領域調整を、わが国の人口動態統計と整合性がとれるようにした。
- a_0 の推定には、日本の近年のデータを用いて別途推計式を作成した。

そして、この方法に基づいて、男女別、1歳×1年について、1947～2010年の全国版生命表の作成を行い、得られた平均寿命を公式生命表やHMDと比較した。その結果の考察を通じ、JMDによる生命表は、HMDの方法論を基礎とし、時系列比較が可能となるよう同一形式の生命表を提供することを可能にしたとともに、国際比較可能性を阻害しない範囲で、日本の死亡状況により適合させる改善にも成功しており、わが国の死亡研究に特化した生命表として、非常に使いやすいデータ系列を提供することを可能にしたということができる。

今後、全国版の生命表だけではなく、都道府県単位などの地域別の生命表の作成、及び、全国版については死因に関するデータベースの構築を進めていく必要がある。また、全国版の生命表についても、本年度一つの結果を得ることができたが、方法論について、まだ改善の余地がないとはいえない。最終的にどこまで日本の状況を反映するかについては国際比較可能性とのバランスの観点から判断する必要があるが、比較可能性を保ちつつも、

できるだけわが国の死亡状況をよく表している生命表作成に関する方法論を研究することについても、引き続き検討が必要な課題である。

参考文献

Human Mortality Database. University of California, Berkeley (USA) and Max Planck Institute for Demographic Research (Germany). Available at www.mortality.org or www.humanmortality.de.

Wilmoth, J. R., K. Andreev, D. Jdanov, and D. A. Glej (2007) “Methods Protocol for the Human Mortality Database (Version 5)”, available at <http://www.mortality.org/Public/Docs/MethodsProtocol.pdf>.