

---

**資 料**

---

## 寿命研究と Human Mortality Database

石 井 太

はじめに

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い存在になったということができる。筆者はカリフォルニア大学バークレー校において HMD プロジェクトに参加し、データベース開発の方法論・データ品質の評価や、HMD の今後の方向性などに関する議論に加わる機会を得ることができた。本稿では、この時の経験なども踏まえつつ、改めて、今まさに寿命研究の最前線において活用されている HMD の紹介を行うとともに、HMD の今後の展開について述べることにしたい。なお、HMD の内容や沿革等に関しては HMD ウェブサイト (<http://www.mortality.org/>) に詳細な説明があり、以下でもここから多くの引用を行っていることを最初にお断りしておく。

### I HMD とは

HMD とは、研究者、学生、ジャーナリスト、政策分析者を始めとした、人類の寿命の歴史に関心のある全ての人に対して、詳細な死亡データを提供することを目的とし、国または地域レベルの人口集団について、プロジェクトが独自に計算した死亡率及び生命表を、これらの計算に用いた入力データとともに提供している死亡データベースである。

HMD は、カリフォルニア大学バークレー校 (UCB) とマックスプランク人口研究所 (MPIDR) の研究チームが共同で2000年に立ち上げたプロジェクトであり、UCB の John Wilmoth がプロジェクト責任者、MPIDR の Vladimir Shlinikov が共同責任者となって運営がなされている。HMD の前身となるプロジェクトとして、John Wilmoth による Berkeley Mortality Database (BMD) が1997年に創設され、そこでは HMD と同様の形態での生命表データが提供されていたが、対象となっていた国はフランス、日本、スウェーデン、アメリカの4カ国にとどまっていた。しかしながら、HMD ではデータの収載国 (または地域) を精力的に拡大してきており、2010年7月現在では、オーストラリア、オー

ストリア、ベラルーシ、ベルギー、ブルガリア、カナダ、チリ、チェコ、デンマーク、エストニア、フィンランド、フランス、ドイツ、ハンガリー、アイスランド、アイルランド、イスラエル、イタリア、日本、ラトビア、リトアニア、ルクセンブルグ、オランダ、ニュージーランド、ノルウェー、ポーランド、ポルトガル、ロシア、スロバキア、スロベニア、スペイン、スウェーデン、スイス、台湾、イギリス、アメリカ、ウクライナの37の国または地域となっている。

なお、HMDは1993年に Väinö Kannisto と Roger Thatcher によって創設された、“Kannisto-Thatcher Database on Old Age Mortality” (KTD) にも大きな影響を受けている。KTDは30以上の国に関するデータを収集し、80歳以上の高齢死亡率を対象として、その年齢層における死亡率推定の先進的な方法論を研究したものである。HMDはこれらの先行プロジェクトの成果を継承し、全年齢の詳細な生命表が提供されているBMDの形式を保ちつつも、KTDの成果である高齢死亡率推定に関する新たな方法論を活用し、KTDと同様30~40の国及び地域のデータを収載することを目指してプロジェクトを開始したのである。

## II HMD の設計理念

HMDがこれほどまでの成功を収めることができたのには、生命表の作成法及びその提供法に対する優れた設計理念があったことが要因の一つとなっているというのが筆者の考えである。

Caldwell (2001) は、人口学における典型的なアプローチの例示の一つとして、「多くの労力が数値からすぐに結論を導くことではなく、データを精査し改善することに割かれている。人口学者が「生データはたいいていの場合完全ではない」と想定している点は、多くの社会学・医学・統計学の科学者とは異なる特徴である。」と述べた。生命表作成による死亡の測定も、一次統計として得られる人口・死亡（・出生）という生データそのものではなく、これらのデータを加工し精査することによって、より純粋な死亡力を抽出しようという理念に基づくものであり、このような人口学的アプローチに沿ったものといえる。

しかしながら、データ加工という操作は、異なる生命表間の比較を行う場合に別の問題を引き起こす。人口指標の精度は、その指標の性質や誤差発生要因等によって異なるものであり、さまざまな状況に対応した評価が必要となる（石井 2007）。生命表をその指標の精度を考慮して比較するためには、通常的人口指標と同様、入力データに関する一次統計としての精度評価に加え、一次統計を生命表へと加工する過程が指標にもたらす影響評価も必要となる。例えば、二国間の死亡状況を両国の生命表で比較する問題を考えてみよう。一国のある年齢の生命表上の死亡率が他国より低いことのみをもって、これを直ちに両国の純粋な死亡力格差と結論できるだろうか。先の人口学的アプローチに従えば、その答えには一定の留保が必要である。「データの精査」という観点からは、基となった一次統計の精度評価に加え、生命表においては、その死亡率の推定に用いられたデータ加工の

過程、例えば、死亡率算定の分母の選択（中央人口か生存延べ年数か）、死亡データの年齢の誤報告などの取扱い（データ修正の有無やその方法）、死亡率の平滑化の取扱い、高齢死亡率の数学的関数による当てはめの取扱い、などの様々な工程の影響も考慮が必要である。このように、生命表の国際比較、時系列比較をより厳密に行うためには、生命表作成方法に対しても統一的な取扱いが必要となるのである。

しかしながら、これは同時に非常に困難な問題でもある。まず第一に、各国での人口や死亡に関する統計の取得方法は様々であり、これは必然的に生命表の作成方法に影響を及ぼす。また、各国の死亡状況も多様であり、時系列的にも異なった様相を呈することから、生命表作成にあたっては臨機応変な対応がより望ましいとする考え方もある。したがって、生命表データベース構築にあたっては、各国の統計データ形式や様々な死亡状況の違いに対応しつつ、統一的な生命表作成が行えるよう、設計理念を確立することが重要な課題となる。

そこで、HMD では、まず、データベース作成にあたって以下の四原則に可能な限り従うこととした。それは、比較可能性 (comparability)、柔軟性 (flexibility)、アクセスのしやすさ (accessibility)、および再現性 (reproducibility) である。そして、この四原則に基づき、入力データから最終的な生命表を作成する工程について、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。この方法手順書については、完全な形で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、公表されているのである。

この手順の全貌は大変複雑なものであるが、大きくは以下の6つのデータ収集・作成に対応する工程から成っている。

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. リスク対象生存延べ年数 (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。

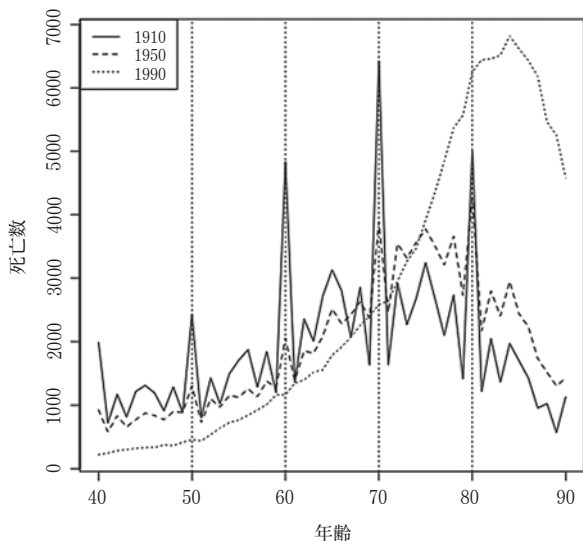
各国・地域の生データは様々な形式になっており、これらにそれぞれアドホックに対応していたのでは統一的な生命表作成は行えない。HMD ではこれらからまず、「満年齢別・

死亡年別・出生年別死亡数」および「各年1月1日現在推計人口」という2つの量を共通に推定するという設計理念を採ることによって、統一的手法による生命表作成を行うこととしたのである。

一方で、HMDでは、明らかな資料の間違いなどを除き、年齢の誤報告やデータのカバー範囲に対する系統的なデータ修正は行わない方針を採っている。これは、全ての国に対して統一的手段でデータ修正を行うことが不可能であり、この種の修正はデータベース構築にあたり適したものではなく、また望ましいものでもないという思想に立っている。これも統一的手法による生命表作成のための重要な設計理念といえる。

HMDでは、上記のようなデータ修正を行わないという観点から、データベースに収載する国や地域は、センサスや人口動態登録システムが比較的良好に整備され、人口・死亡などの入力データの品質が比較的高い人口集団に限定することとしている。しかしながら、それでも国や時代によりある程度の年齢誤報告は観察されることに注意が必要である。年齢誤報告の中で、末尾が0や5の年齢にデータが集積する現象をエイジヒーピング（Age heaping）と呼ぶ。例えば、図1はHMDにおけるスペインの死亡数（女性）を1910年、1950年、1990年について示したものである。これを見ると、1910年では非常に大きなエイジヒーピング、また、1950年においても1910年ほどでないものの、やはりエイジヒーピングが観察される。一方で、1990年においてはエイジヒーピングは全く見られない。HMDデータを各歳で利用するにあたっては、このようなデータの性格に注意が必要だが、一方で、HMDでは5歳階級でのデータ提供も行われているので、こちらを利用することによりエイジヒーピングへの対応も柔軟に行えるようになっている。

図1 死亡数（スペイン、女性、1910、1950、1990年）



資料：Human Mortality Database

なお、HMDの全てのデータは、誰もが無料で完全に自由にダウンロードして使用することができるようになっており、アクセスのしやすさについても完全な対応が図られている。ただし、利用にあたってはユーザー登録が必要となることに注意されたい。

以上に述べてきた通り、HMDは、先に述べた四原則を始め、死亡研究に長年携わってきた経験を持つ研究者から成るチームだからこそ産み出すことのできた、洗練された設計理念を確立することに成功し、これが極めて高品質な生命表データベース構築の達成に結びついたのだということができよう。

### III データベースの内容

現在の HMD ウェブサイトのトップページ (<http://www.mortality.org/>) は図 2 のようになっている。ページ内にある37の国または地域名をクリックすることにより、その国または地域のページへと移動することができる。

図 3 はアメリカのページを例として示したものである。上側の部分では、完成されたデータセットということで、先の手順に示した 6 種類のデータセット、すなわち、期間データに関する出生数、死亡数、人口サイズ、リスク対象生存延べ年数、死亡率、生命表が表側に掲げられているとともに、これらをコーホートデータに組み替えたリスク対象生存延べ年数と死亡率が示されている。一方、表頭は年齢階級と年次階級を示し、1 歳×1 年、1 歳×5 年、1 歳×10 年、5 歳×1 年、5 歳×5 年、5 歳×10 年という多様な年齢・年次区分でのデータが提供され、利用者がその利用法に応じて柔軟に活用できるようになっている。

また、下側では入力データとして、出生・死亡・人口規模に関するデータが収録されている。これらは、上のデータセットの計算に用いられる入力データを生データそのものの形で収録したものである。HMD の 4 原則の一つである再現性の観点から、入力データを完成された生命表とともに示し、このデータ方法手順書の手順を施すことにより、利用者も生命表を再現できるようになっているのである。

次に、表の中に示された「1x1」などの部分ををクリックすることにより、該当するデータがテキスト形式で表示される。図 4 は女性の 1 歳×1 年階級の生命表データを例として示したものである。ここでは、年次別に生命表関数がテキストデータで収録されており、これらを R などの統計ソフトで読み込むことにより、データ分析を行うことができる。

図 2 Human Mortality Database ウェブサイトトップページ

**HMD Main Menu**

- Registration
  - New User
  - Change Password
- Project
  - FAQ
  - Overview
  - History
  - People
  - Acknowledgements
  - Research Teams
  - HMD Publications
  - Methods
    - Brief Summary
    - Full Protocol
    - Special Methods
  - Data
    - What's New
    - Explanatory Notes
    - Data Availability
    - Zippped Data Files
    - Citation Guidelines
  - Links
    - Max Planck Institute
    - UC Berkeley
    - UC Berkeley Demography
    - Human Life Table Database
    - Canadian HMD
  - General
    - Contact us

**The Human Mortality Database**

John R. Wilmoth, Director  
University of California, Berkeley

Vladimir Shkolnikov, Co-Director  
Max Planck Institute for Demographic Research

We have recently upgraded the HMD web site. The new site presents the same information but offers features that should be useful for both data users and our web administrator. From the user's point of view, the most important changes are: (1) a more logical organization of the information on individual country pages, and (2) the ability to download large chunks of the database in zipped data files (see link on the sidebar of this page).

The Human Mortality Database (HMD) was created to provide detailed mortality and population data to researchers, students, journalists, policy analysts, and others interested in the history of human longevity. The project began as an outgrowth of earlier projects in the [Department of Demography at the University of California, Berkeley, USA](#) and at the [Max Planck Institute for Demographic Research in Rostock, Germany](#) (see [history](#)). It is the work of two teams of researchers in the USA and Germany (see [research teams](#)), with the help of financial backers and scientific collaborators from around the world (see [acknowledgements](#)).

We seek to provide open, international access to these data. At present the database contains detailed population and mortality data for the following 37 countries or areas:

Australia	Finland	Lithuania	Spain
Austria	France	Luxembourg	Sweden
Belarus	Germany	Netherlands	Switzerland
Belgium	Hungary	New Zealand	Taiwan
Bulgaria	Iceland	Norway	U.K.
Canada	Ireland	Poland	U.S.A.
Chile	Israel	Portugal	Ukraine
Czech Republic	Italy	Russia	
Denmark	Japan	Slovakia	
Estonia	Latvia	Slovenia	

For more information, please begin by reading an [overview](#) of the database. If you have comments or questions, or trouble gaining access to the data, please write to us ([hmd@mortality.org](mailto:hmd@mortality.org)).

資料：Human Mortality Database

図3 Human Mortality Database U.S.A. のページ

U.S.A.

Background and documentation

Data sources

Complete Data Series [ Explanatory notes ]

Period data	Available dates	Age interval × Year interval					
		1x1	1x5	1x10	5x1	5x5	5x10
Births	1933 - 2006	1-year					
Deaths	1933 - 2006	1x1	1x5	1x10	5x1	5x5	5x10
Deaths by Lexis triangles	1933 - 2006	Lexis					
Population size	1933 - 2007	1-year			5-year		
Exposure-to-risk	1933 - 2006	1x1	1x5	1x10	5x1	5x5	5x10
Death rates	1933 - 2006	1x1	1x5	1x10	5x1	5x5	5x10
Life tables	1933 - 2006						
Females		1x1	1x5	1x10	5x1	5x5	5x10
Males		1x1	1x5	1x10	5x1	5x5	5x10
Total (both sexes)		1x1	1x5	1x10	5x1	5x5	5x10
Life expectancy at birth	1933 - 2006	1-year	5-year	10-year			
Cohort data							
Exposure-to-risk	1833 - 1976	1x1	1x5	1x10	5x1	5x5	5x10
Death rates	1833 - 1976	1x1	1x5	1x10	5x1	5x5	5x10

Input Data [ Explanatory notes ]

	Available dates	Data	Lexis map
Births	1933 - 2006	txt	
Deaths	1933 - 2006	txt	html
Population size	1930 - 2009	txt	html
Notes		pdf	
Reference file		pdf	

資料：Human Mortality Database

図4 生命表データ（アメリカ女性，1歳×1年階級のもの）

The United States of America, Life tables (period lx1), Females Last modified: 29-Jun-2009, MPv5 (May07)

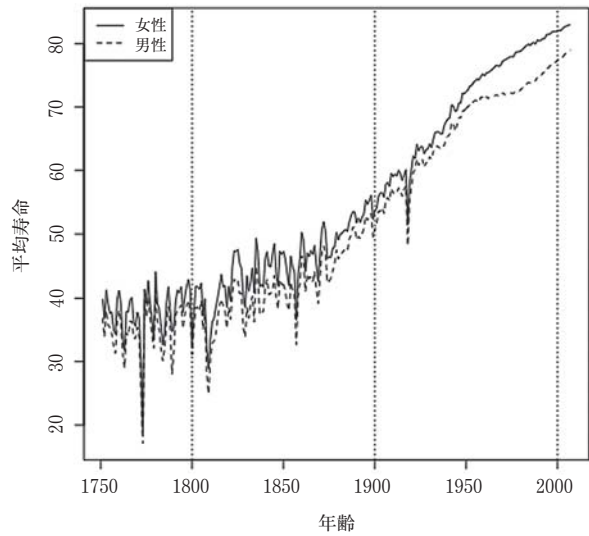
Year	Age	mx	qx	ax	lx	dx	Lx	Tx	ex
1933	0	0.05451	0.05224	0.21	100000	5224	95850	6281628	62.82
1933	1	0.00887	0.00883	0.50	94776	837	94357	6185778	65.27
1933	2	0.00402	0.00402	0.50	93939	377	93750	6091221	64.84
1933	3	0.00287	0.00287	0.50	93562	268	93428	5997671	64.10
1933	4	0.00223	0.00223	0.50	93294	208	93190	5904243	63.29
1933	5	0.00185	0.00185	0.50	93086	172	93000	5811054	62.43
1933	6	0.00164	0.00164	0.50	92913	153	92837	5718054	61.54
1933	7	0.00147	0.00147	0.50	92761	136	92693	5625217	60.64
1933	8	0.00132	0.00132	0.50	92624	122	92563	5532524	59.73
1933	9	0.00119	0.00119	0.50	92502	110	92447	5439961	58.81
1933	10	0.00110	0.00110	0.50	92392	101	92341	5347514	57.88
1933	11	0.00106	0.00106	0.50	92290	98	92242	5255173	56.94
1933	12	0.00110	0.00110	0.50	92193	101	92142	5162931	56.00
1933	13	0.00121	0.00121	0.50	92092	111	92036	5070789	55.06
1933	14	0.00139	0.00139	0.50	91980	128	91916	4978753	54.13
1933	15	0.00163	0.00163	0.50	91852	149	91778	4886836	53.20
1933	16	0.00187	0.00186	0.50	91703	171	91618	4795058	52.29
1933	17	0.00210	0.00210	0.50	91532	192	91436	4703441	51.39
1933	18	0.00233	0.00233	0.50	91340	212	91234	4612005	50.49
1933	19	0.00256	0.00256	0.50	91128	233	91011	4520771	49.61
1933	20	0.00279	0.00278	0.50	90895	253	90768	4429759	48.74
1933	21	0.00299	0.00299	0.50	90642	271	90506	4338991	47.87
1933	22	0.00316	0.00316	0.50	90371	285	90228	4248485	47.01
1933	23	0.00330	0.00329	0.50	90085	297	89937	4158257	46.16
1933	24	0.00340	0.00339	0.50	89789	304	89637	4068320	45.31
1933	25	0.00346	0.00346	0.50	89484	310	89330	3978683	44.46
1933	26	0.00353	0.00353	0.50	89175	314	89018	3889353	43.61
1933	27	0.00360	0.00360	0.50	88860	320	88701	3800336	42.77
1933	28	0.00367	0.00367	0.50	88541	325	88378	3711635	41.92
1933	29	0.00374	0.00374	0.50	88216	330	88051	3623257	41.07
1933	30	0.00381	0.00381	0.50	87886	335	87719	3535206	40.22
1933	31	0.00392	0.00391	0.50	87552	342	87381	3447487	39.38
1933	32	0.00406	0.00406	0.50	87209	354	87033	3360106	38.53
1933	33	0.00426	0.00425	0.50	86856	369	86671	3273073	37.68
1933	34	0.00450	0.00449	0.50	86487	388	86293	3186402	36.84
1933	35	0.00476	0.00475	0.50	86099	409	85894	3100110	36.01
1933	36	0.00497	0.00496	0.50	85690	425	85478	3014215	35.18
1933	37	0.00513	0.00512	0.50	85265	436	85047	2928738	34.35
1933	38	0.00525	0.00524	0.50	84829	444	84607	2843691	33.52
1933	39	0.00536	0.00534	0.50	84385	451	84159	2759084	32.70
1933	40	0.00549	0.00547	0.50	83934	459	83704	2674925	31.87
1933	41	0.00571	0.00569	0.50	83474	475	83237	2591220	31.04
1933	42	0.00603	0.00601	0.50	82999	499	82750	2507984	30.22
1933	43	0.00643	0.00641	0.50	82500	529	82236	2425234	29.40
1933	44	0.00688	0.00686	0.50	81972	562	81690	2342998	28.58
1933	45	0.00739	0.00736	0.50	81409	599	81110	2261308	27.78
1933	46	0.00790	0.00787	0.50	80810	636	80492	2180198	26.98

資料：Human Mortality Database

また、全ての国で同じフォーマットが採用されているので、一つの国のデータに対してプログラムを作成しておけば、他の全ての国のデータに対してもこれをそのまま適用することが可能であり、国際比較分析を簡単に行うことができる。

また、時系列比較分析が容易に行えるのも HMD の特徴であり、中には非常に長期の時系列データが提供されている国もある。最も長期にわたるのはスウェーデンのデータであり、1751年からのデータが収録されている。図5にスウェーデンの平均寿命の推移を示したが、これを見

図5 スウェーデンの平均寿命の推移（1751～2007年）



資料：Human Mortality Database

ると、平均寿命は19世紀初頭から伸長する傾向が見られてはいるものの、19世紀後半以降、とりわけ20世紀における伸長のスピードが著しい事が観察され、長寿革命とも称される、人類の達成した驚異的な死亡率改善の歴史を概観することができる。

#### IV HMD の今後の展開

HMD では死亡データの作成及びその分析に関わる研究者間の議論の活性化と、HMD 自体の今後の方向性に関する議論などを行う目的から、HMD プロジェクトメンバーに加えて外部の死亡研究者を交えたシンポジウムを開催してきている。第1回は2004年、第2回は2008年に行われたが、本年（2010年）6月17～19日、第3回の HMD シンポジウムが、フランスの INED において、14カ国から39名の研究者の参加により開催され、筆者も参加した。今回のシンポジウムでは、死因別データと（国内の）地域別データを HMD において提供することをどのように考えるかがメインテーマとして設定され、8つのセッションと最終日の円卓会議において活発に報告・討議が行われた。

このうち、「HMD に死因データを含める拡張」と題された最初のセッションは、死因データ提供に関する問題点や困難さを冒頭から浮き彫りにし、最終日の円卓会議まで続く通奏低音のような流れを作ったように思われ、象徴的なセッションになったと感じた。最初に、HMD チームで死因分析を担当し、今回のシンポジウムの組織者の一人でもある Magali Barbieri (INED, UCB) と、Carl Boe (UCB) による「HMD に死因データシリーズを含める拡張」という報告が行われた。ここでは、HMD において、100を超えない程度に分けられた新たな死因分類を提案し、この分類に基づいて年齢（階級）別死因別死亡数やその割合などを提供していくことについて提案がなされた。そして、死因データ

に関しては、基本的に修正などを行わない生データに基づいたものを提供するという方向性が示されたが、これがその後の議論の焦点となった。というのは、このセッションのもう一つの報告「死因：時空間比較の困難」が France Meslé (INED) と Jacques Vallin (INED) から行われ、そこでの結論が「死因の生データは、決して直接分析に使えるものではない」とするものであったからだ。この議論は、最終日の円卓会議まで引き続き、様々な意見が出た。HMD では先に述べたように、データ修正は基本的にしないという形で全体の統一性を確保するという設計理念があり、これに照らせば生データに基づくデータ提供が整合的と考えられることから、このような提案が HMD チームからなされたものと考えられる。しかしながら、一方で、死因データにこの理念をどこまで広げられるかという点については、克服すべき様々な問題点があるということがこれらの議論を通じて改めて明らかになったように思われる。

もう一つの議題であった、国内の地域別生命表を HMD データベースに加えることについては、人的なリソースなどの限界などもあり、プロジェクト責任者である John Wilmoth から、かなり困難であるという見通しが示された。以上を総合すると、今後の HMD は、国内の地域別データの収録は行わず、死因データを追加提供するという形でデータベースの拡張を行う方向に展開していくのではないかと考えられる。

おわりに

HMD が掲げている 4 原則の一つに比較可能性がある通り、死亡分析に適したデータベースとして、国際比較可能性・時系列比較可能性という要件は極めて重要である。一般に、統計部局等が作成する生命表は、時宜にかなった国民の死亡状況測定として意義が高いと考えられるものの、必ずしも時系列間で形式が同一とは限らず、また、他国の生命表と比較する場合にも形式や作成方法などについて違いが出ることは避けられない。したがって、死亡研究に最適化された生命表のデータベース化については、HMD が学術研究者から成る研究チームで運営されているように、研究ベースで行われることが望ましいと考えられる。日本においては、未だこのようなデータベースは存在していないが、今後、HMD と整合性の取れた「日本版死亡データベース」開発の検討が望まれるところである。そして、このようなデータベース開発を通じ、わが国の死亡研究の益々の活性化を図ることが重要であろう。

参考文献

- Caldwell, J. C. (2001) "Demographers and the Study of Mortality: Scope, Perspectives, and Theory", *Annals of the New York Academy of Sciences*, Vol.954, pp.19-34.
- Human Mortality Database. University of California, Berkeley (USA) and Max Planck Institute for Demographic Research (Germany). Available at [www.mortality.org](http://www.mortality.org) or [www.humanmortality.de](http://www.humanmortality.de).
- 石井太 (2007) 「人口指標の精度について」、稲葉寿 (編) 『現代人口学の射程』 ミネルヴァ書房, pp.59-76.