

令和元年度厚生労働科学研究費補助金  
(政策科学総合研究事業(統計情報総合研究事業))  
分担研究報告書

海外における公的死亡データ研究目的利用の実態  
—利用方法・リンケージ可能性・研究実績

研究代表者	伊藤 智子	筑波大学医学医療系	助教
研究分担者	田宮 菜奈子	筑波大学医学医療系	教授
研究分担者	森 隆浩	筑波大学医学医療系	准教授
研究協力者	杉山 雄大	筑波大学医学医療系	准教授

研究要旨

重要な健康アウトカムの一つである死亡を把握するデータとして公的死亡データに着目し、海外主要国における公的死亡データの利用方法や、他データとのリンケージの可能性、研究実績についてレビューした。その結果、公的死亡データを研究目的で活用している海外主要国においては、リンケージデータを用いて多くの研究実績を有していることが分かり、公的死亡データにおけるリンケージ利用の意義を示していると言えた。また、その利用方法においては公的機関によって適切に管理・制限されており、データの不正利用等といった有害事象の予防策が図られているとみられた。

A. 研究目的

海外においては公的データを用いた突合は国によっては活発に行われており、そのデータの種類も公的サービスの利用状況から遺伝子情報など多岐に渡る。その中でも、最終的にかつ最も重要な健康アウトカムの一つである死亡に着目した。生存(あるいは死亡)の正確な観察は、研究コホート内の死亡率の調査に不可欠である。その研究上の活用の可能性を検討することは、今後の我が国の人口動態統計(特に死亡票)のデータ活用に向けて有用な知見が得られると考えられた。

そこで本研究では、海外主要国における公的死亡データにおける利用方法や、他データとのリンケージの可能性、研究実績についてレビューした。

B. 研究方法

公的データの突合が活発に行われている米国、英国、オランダ、デンマーク、スウェーデン、フィンランド、ノルウェー、オーストラリアにおける公的死亡データを挙げ、その突合可能性について方法や突合実績を踏まえてレビューした。

突合実績については、国名+Death+Linkageを検索語としてPubMedで検索したほか、各国の整備するデータベースに関するWebサイトから使用された研究実績を収集した。

(倫理面への配慮)

本研究で用いるデータは、個人情報を含まない。

C. 研究結果

1. 米国

米国では、アメリカ疾病予防管理センタ

ー (Centers for Disease Control and Prevention, CDC) の管轄する全米国民の死亡情報データベースとして National Death Index (NDI) <sup>1</sup> が存在する。NDI は米国で死亡率を確認するために最も頻繁に使用されるデータであり、研究用途の申請を行うことで、特定の対象のデータを取得可能である。データには、氏名、社会保障番号 (Social Security Number, SSN)、生年月日、死亡状況、死亡日などが含まれ、オプションの NDI Plus の利用で死因を把握できる。

NDI の突合 (あるいは個人の特定) においては、少なくとも以下の組み合わせの中の 1 つが必要とされている。

- ① SSN+氏名
- ② SSN+生年月日+性別
- ③ 氏名+生年+生月

そのほか、人種、および婚姻状況なども特定のための情報として用いることができる。

SSN を用いない③の突合は、Probabilistic linkage となる。SSN を使用する①、②の Deterministic linkage とは異なり、個人特定の精度を検討する必要があると述べられており、その精度の検証を目的とした先行研究が行われている。<sup>2</sup> この研究では、大規模な軍事コホートに対して CDC National Program of Cancer Registry (NPCR) アルゴリズムを用いた NDI の一致率が 97.1%であったとし、NDI をリンケージすることによる死亡率や死因の特定は非常に有用であると述べている。

NDI のデータを取得するためには、CDC への用途の申請と CDC に設置された委員会の認可が必要である。<sup>3</sup> 基本料金は 350 ドルであり、1 名の記録の 1 年分の調査につき 0.15 ドルが可算請求される。例えば、10 年間の 1,000 名の死亡記録を調査した場合、 $(\$350 + (\$0.15 * 1,000 * 10)) = 1,850$  ドルとなる。NDI Plus (死因コードを含む) の場合は、さらに 1 人 1 年分につき 0.21 ドルを

加算する。

また Social Security Administration (SSA) が報告を行った死亡者データである Social Security Administration's Death Master File<sup>4</sup> を元にした Social Security Death Index (SSDI) によっても、同様に国民の生存および死亡を確認できる。SSDI は、複数のサイトにおいて無料で利用可能であることから、簡便で安価なデータソースとしても紹介されている。SSDI における個人の特定は、SSN+氏名+生年月日が必要である。すべての識別情報と完全に一致する場合、死亡が確定される。しかし、この SSN+氏名+生年月の組み合わせに誤りがある場合は、個人が特定されない。

死亡を確認するデータソースとして、この NDI と SSDI とのデータの精度や相互関係を検証した研究が行われており、条件の限られたコホートでは SSDI が高い正確性を発揮するものの<sup>5</sup>、包括的な人口を把握する上では NDI が最も正確なデータベースであるとされている。<sup>6</sup> そのため、NDI は死亡の確定のためのデータソースとして用いられることが多い。また、NDI は米国の全国的大規模調査<sup>7,8,9</sup> あるいは臨床データ<sup>10,11</sup> とリンケージされ、調査内容に含まれる曝露因子と死亡との関連を検証する研究を可能にしている。また死亡の確定のためのデータソースとして用いられることがあり<sup>12</sup>、研究の質を担保する上で重要な役割を果たしていると言える。

## 2. 英国

英国では、医薬品・医療製品規制庁 (Medicines and Healthcare products Regulatory Agency) と国立衛生研究所 (National Institute for Health Research) の助成を受けた Clinical Practice Research Datalink (CPRD)<sup>13</sup> が、死亡情報を含む最も大規模な研究用医療データベースとして存在している。CPRD

は英国に在住記録のある累計 5000 万人（現在の居住者では 1500 万人）の医療データを収録する。

CPRD では、英国の一般診療情報等の医療関連データを個人レベルで突合している。突合後、個人を特定する情報は消去される。死亡データは国家統計局（Office for National Statistics、ONS）の全国的な死亡記録情報を取得しており、死亡日のほか、死因（ICD コード）を含む。CPRD を使用することで、研究利用例として、心不全による入院患者と死亡情報の突合研究が行われている。<sup>14</sup>

CPRD と死亡データの突合には、NHS 番号、氏名、性別、誕生日、郵便番号等による 8 段階の個人情報 **Deterministic linkage** アルゴリズムを採用している。

CRPD の死亡データを取得するためには、プロトコル申請書の提出と **Independent Scientific Advisory Committee (ISAC)** による承認、およびライセンス支払いが必要である。<sup>15</sup> ライセンスは、複数の研究に使うことができる 12 ヶ月間の定額ライセンスと、単独研究用のライセンスに分かれており、アクセス可能なデータベースと利用者の性質に応じて種別と料金が異なる。また、データリンクには追加料金を請求する。<sup>16</sup>

一方、ONS が保有する英国の全国民の死亡情報については、国民保健サービス（NHS : National Health Service）内の組織、NHS Digital が、**Primary Care Mortality Database (PCMD)**<sup>17</sup> として、死因などを含めたデータベースとして構築している。しかし PCMD は認可を受けた公的機関にのみ開放されており、一般研究用途には使用できない。

### 3. オランダ

オランダでは、オランダ統計局（Statistics Netherlands、CBS）が、全

国民の死亡データベース、**Cause of Death Registry**<sup>18</sup> を研究用途に公開している。市民サービス番号（BSN）が含まれているため、他の公的データと個人レベルでの突合が可能である。

**Cause of Death Registry** の死亡情報は、性別、年齢、死因（ICD コード）を収録する。研究利用例として、**Income Survey** と死亡データとのリンケージデータを用いて、認知症患者における社会経済的要因（**socioeconomic status, SES**）と死亡との関連を 10 年間のコホートを用いて検証している。<sup>19</sup> また、全国民コホートを使用して、少数派移民のがんによる死亡率をオランダ人と比較している。<sup>20</sup>

データの利用は研究機関や EU の統計機関の研究に限定されており、データの抽出を行うためには CBS のチームに依頼を行う必要がある。料金は研究の性質に応じて決定される。<sup>21</sup>

またオランダには、処方情報をはじめとする国内の 400 万人以上の情報を収めた大規模医療情報リンケージネットワーク、**PHARMO Database Network** が存在する。**PHARMO** は **The Central Bureau of Genealogy** が管理する死亡日データを収録しているが、死因情報が含まれていないため、死因に関する研究には使用できない。<sup>22</sup> **PHARMO** 内部のデータへのアクセスは、大学か研究機関に所属する研究者に限定されている。研究申請はコンプライアンス委員会の審査に合格する必要がある。<sup>23</sup>

### 4. デンマーク

デンマークでは、デンマーク国家保健委員会（**National Board of Health**）が、全住民の死亡データベースである **The Danish Register of Causes of Death**<sup>24</sup> を公開している。**The Danish Register of Causes of Death** の死亡データには国民登録番号（**Civil Registration Number**）が

含まれているため、他の公的データと個人レベルでの突合が可能である。その他に、住所、死亡日、死亡場所、死因（ICDコード）などを収録する。また、The Danish Register of Causes of Deathの個別の死因データの取得には、Forskertjeneste（調査サービス社）へのコンタクトが必要である。<sup>25</sup>

料金は Forskertjeneste のコンサルティングやデータの引き出し等の作業時間に応じて、1時間あたり 1,475 デンマーククローネ（約 226 ドル）が請求される。<sup>26</sup>

研究利用例として、心房細動患者の頭蓋内出血後の抗凝固療法の再開と再春性脳卒中、死亡率、出血への影響について、全国民の罹患情報を含む Patient Registry および Prescription Registry を死亡データとリンケージすることで検証している。<sup>27</sup>

## 5. スウェーデン

スウェーデンでは、スウェーデン社会庁（National Board of Health and Welfare）が管理する National Cause of Death Register（The Swedish Cause of Death Register）<sup>28</sup> が、国内の全死亡データを記録する。死亡データには国民識別コード（PIN）が含まれており、他の公的データと個人レベルでの突合が可能であり、研究目的であれば、個人レベルのデータにアクセスが可能である。他には、生年月日、死亡日、国籍、性別、死因（ICDコード）、死亡場所などを収録する。

この公的死亡データの研究目的利用においては、The National Board of Health and Welfare による審査に合格することが必要である。利用料金は申請受理とデータの引き出しにかかる時間に応じて、1時間あたり 1,100 クローナ（約 122 ドル）が請求される。<sup>29</sup>

利用例として、火災情報と検死情報と突合することで、火災による死亡者と火災の

性質が検証されている。<sup>30</sup>

## 6. フィンランド

フィンランドでは、フィンランド統計局（Statistics Finland）が管理する National Cause of Death Register（Finland）が、国内の全死亡データを記録する。<sup>31</sup> 死亡データには社会保障番号（Finnish Personal Identification Code）が含まれており、他の公的データと個人レベルでの突合が可能である。その他、死因、享年、性別、死因（ICDコード）などを収録する。

個人データへのアクセスは、研究目的に限り可能であり、事前に Statistics Finland による申請書の承認が必要になる。利用料金は調査内容に応じて変動する。

<sup>32,33</sup>

フィンランドを含む北欧諸国では、高い資源配分と公平社会のもと産科ケアへの高いアクセスを背景に、世界で最も低い妊産婦死亡率を誇っている。しかし、改めて国家横断的に妊産婦死亡を正確に把握し、その原因探索を図った。この研究では北欧諸国 5 か国（デンマーク、アイスランド、ノルウェー、スウェーデン、そしてフィンランド）が誇る国民番号による公的データのリンケージが行われ、約 9 年間で妊産婦死亡 168 件と極めて少ない実態を明らかにしている。<sup>34</sup>

## 7. ノルウェー

ノルウェーでは、ノルウェー公衆衛生研究所（Norwegian Institute of Public Health、NIPH）が管理する Norway Cause of Death Registry が、国内の全死亡データを記録している。<sup>35</sup>

死亡データには国民識別番号（Norwegian national identity number）が含まれており、他の公的データと個人レベルでの突合が可能である。<sup>36</sup> その他、死亡

時年齢、性別、死亡場所、死因(ICD コード)を収録する。

Norway Cause of Death Registry の個人レベルの死亡データへのアクセスは、研究目的に限り可能であり、申請書の提出と NIPH による承認が必要である。<sup>37</sup> 料金は調査内容に応じて変動する。下記ページの例では、通常フォーマットの 1 つのデータベースの調査に 10,000~30,000 ノルウェークローネ (約 1,087~3,261 ドル) を請求する。<sup>39</sup>

利用例として、1 型糖尿病の全国的コホートを用いて 1 型糖尿病の長期死亡率、死亡原因、および末期腎疾患の発生を検証している。死亡率は一般集団の 4.4 倍で、特にアルコール関連の死亡が多いことに着目されていた。<sup>38</sup>

## 8. オーストラリア

オーストラリアでは、オーストラリア国立保健福祉研究所 (Australian Institute of Health and Welfare) が管理する National Death Index が存在する。<sup>40</sup> 国内の全死亡データを各州・特別地域から収集、統合している。National Death Index は、氏名、性別、生年月日、死亡日、死因(ICD コード)などを収録する。

オーストラリアには統一の個人番号が存在しないため、研究上の利用では、NDI を用いる個人レベルの突合には専用のソフトウェアを用いた Probabilistic linkage が用いられる。

National Death Index の利用には、研究目的の提出と AIHW の倫理委員会の承認が必要。料金は必要なデータの内容に応じて決定される。<sup>41</sup>

利用例として、国内の地理的に限定されたレジストリベースではあるが、60 年超という長期縦断的なコホートを設定し、重度脳性麻痺における生存期間の推定を行っている。<sup>42</sup>

またオーストラリアには、ニューサウスウェールズ州、オーストラリア首都特別地域の約 6,900 万人を対象とした医療リンケージデータベース、The Centre for Health Record Linkage (CHeReL)<sup>43</sup> が存在する。包括的に複数のデータを統合する都合上、CHeReL は複数の個人情報を組み合わせた Probabilistic linkage を採用する。<sup>44</sup> CHeReL 内部では、公的機関による 2 種類の死亡データを突合している。<sup>45</sup>

CHeReL 内の医療データの突合には、ChoiceMaker Software(Choice-Maker Technologies Inc., New York, US)等の専用のソフトウェアによる Probabilistic linkage が用いられる。CHeReL は研究目的のアクセスが可能である。データへのアクセスには、研究プロトコルと申請書の承認が必要である。<sup>47</sup> 料金はプロジェクトの性質に応じて変動する。公式ページ上の例では、50,000 人分のリンケージデータの取得には約 5,000 豪ドル (約 3,600 米ドル) が請求される。外部データベースとのリンケージには追加料金が求められる。<sup>48</sup>

研究例として、CHeReL に含まれる前立腺がん患者と死因情報を ChoiceMaker で突合し、自殺リスクとの関連性を明らかにしている。<sup>46</sup>

## D. 考察

本研究では、海外 8 か国における公的死亡データの研究利用について調査した。その結果、各国は公的な死亡データの研究目的利用を可能にしており (英国は部分的)、データを用いて出産、小児から高齢者といったあらゆるライフステージの対象について分析を行っていた。こうした公的死亡データを用いる場合は、ほぼ国民の死亡を把握できるという点から、悉皆性のある研究を可能にする。悉皆性のある研究では、選択バイアスをほぼなくすことがで

きるという、疫学上、大変重要な利点を得ることができ、疫学研究で議論されやすい結果の代表性や一般化可能性といった課題をデータの特性によって解消することができる。そして、リンケージするデータによっては、超長期的な観察期間を設定することが可能であり、特定の因子や介入による長期的アウトカムの検証が可能になる。こうした、公的死亡データを他の既存データとリンケージする手法は、長く一般的になっているとみられ、今後もビッグデータの機運に沿って、より重要になるものと考えられる。

#### ・公的データの利用手続き

公的死亡データの出处は基本的には、中央政府に基づく公的機関であったが、その研究目的の利用においては、米国、オランダ、デンマーク、スウェーデン、フィンランド、ノルウェー、オーストラリアとほとんどの国で、利用料金が発生する。また、多くの国でデータ提供に関する倫理的審査を必要とする。こうした手続きは、データ利用を適切に制限することにつながっていると考えられる。

#### ・リンケージ方法

北欧諸国では個人番号が日常的に使用されており、研究上の利用も活発である。個人番号による **Deterministic linkage** は、リンケージされなかった対象を除くといったバイアスの発生を防ぐことに繋がり、研究の質の担保に寄与していると言える。

一方、米国では個人番号（社会保障番号）を用いる方法が第一選択ではあるが、氏名、生年、生月による **Probabilistic linkage** も選択肢として残しており、リンケージの可能性を広くとって、研究活用の可能性の向上や研究の質担保に少しでも貢献できるよう仕組み立てされていると言える。

#### ・本研究の限界

本研究では、公的死亡データを管理する各国の公的機関が公開する Web ページを情報源としており、収集できる情報の量や質は Web ページの整備状況によって左右されうる。また公的死亡データを利用した研究実績については、データ名や国名を検索語とした検索エンジンによる収集であり、完全に網羅できているとは言えず、今回紹介した以外にも、重要な研究実績がある可能性が考えられる。

#### ・今後の展望

本報告者の田宮らは 2017 年に国立研究開発法人科学技術振興機構 研究開発戦略センターの報告書で、我が国の公的データの利用実態について報告している。<sup>49</sup> それによると、我が国では国民のライフステージに沿って健康・医療・介護情報があり、これらには **NDB** といった世界有数規模で集約しているデータベースも含む。また介護保険総合データベースは、保険発足時より電子化され整備されている。その他、国民生活基礎調査といった代表性を有する多くの全国調査データがある。こうした国が保有する貴重なデータは近年の統計法改正に伴って、研究目的利用が可能になっている。しかし、本報告で示したような公的死亡データとのリンケージは許されておらず、最大かつ最終のアウトカムである「死」の検証が、我が国では行き届かないでいる。今後、我が国のデータヘルスをより促進するためには、この重要な「死」のデータである人口動態統計のリンケージ活用が望まれるところであり、そのためには研究者への利用の門戸を開くとともに、データ運営上の安全を担保するための仕組みの整備が求められるところである。

#### E. 結論

公的死亡データを研究目的で活用してい

る海外主要 8 か国（英国では部分的）においては、リンケージデータを用いて多くの研究実績を有しており、公的死亡データにおけるリンケージ利用の意義を十分に示していると言える。また、公的死亡データの研究目的利用への門戸を確実に開いている一方で、その利用方法においては公的機関によって適切に管理・制限されており、データの不正利用等といった有害事象の予防策が図られているとみられた。

#### F. 研究発表

1. 論文発表 なし
2. 学会発表 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 なし
2. 実用新案登録 なし
3. その他 なし

#### 参考文献

1. Centers for Disease Control and Prevention. Data Access - National Death Index. <https://www.cdc.gov/nchs/ndi/index.htm>, Accessed on July 22 in 2020.
2. Nancy A Skopp, et al. Evaluation of a methodology to validate National Death Index retrieval results among a cohort of U.S. service members. *Ann Epidemiol.* 2017 Jun;27(6):397-400/
3. Centers for Disease Control and Prevention. NATIONAL DEATH INDEX USER FEES. [https://www.cdc.gov/nchs/data/ndi/ndi\\_user\\_fees\\_worksheet.pdf](https://www.cdc.gov/nchs/data/ndi/ndi_user_fees_worksheet.pdf), Accessed on July 22 in 2020.
4. Social Security Administration. Requesting SSA's Death Information. <https://www.ssa.gov/dataexchange/reque>st\_dmf.html, Accessed on July 22 in 2020.
5. Justin T Huntington, et al. The Social Security Death Index (SSDI) Most Accurately Reflects True Survival for Older Oncology Patients. *Am J Cancer Res.*2013 Nov 1;3(5):518-22/
6. T L Lash, et al. Comparison of the National Death Index and Social Security Administration Databases to Ascertain Vital Status. *Epidemiology.* 2001 Mar;12(2):259-61/
7. Zhilei Shan, et al. Association of Low-Carbohydrate and Low-Fat Diets With Mortality Among US Adults. *J AMA Intern Med.* 2020 Jan 21;180(4):513-523.
8. Yanping Li, et al. Healthy lifestyle and life expectancy free of cancer, cardiovascular disease, and type 2 diabetes: prospective cohort study. *BM J.* 2020 Jan 8;368:l6669/
9. Peter D Reaven, et al. Intensive Glucose Control in Patients with Type 2 Diabetes - 15-Year Follow-up. *N Engl J Med.* 2019 Jun 6;380(23):2215-2224.
10. Jennifer K Peterson, et al. Postoperative and long-term outcomes in children with Trisomy 21 and single ventricle palliation. *Congenit Heart Dis.* 2019 Sep;14(5):854-863/
11. David J. Malenka et al. Comparing Long-Term Survival of Patients With Multivessel Coronary Disease After CABG or PCI. *Circulation.* 2005 Aug 30; 112:371-376/
12. David B. Hanna, et al. Comparing the National Death Index and the Social Security Administration's Death Master File to Ascertain Death in

- HIV Surveillance. Public Health Rep. 2009 Nov-Dec;124(6):850–860/
13. Clinical Practice Research Datalink. ONS death registration data and CPRD primary care data Documentation (set 18). [https://www.cprd.com/sites/default/files/Documentation\\_Death\\_set18\\_v2.1\\_Final.pdf](https://www.cprd.com/sites/default/files/Documentation_Death_set18_v2.1_Final.pdf), Accessed on July 22 in 2020.
  14. Alex Bottle et al. Frailty and comorbidity predict first hospitalisation after heart failure diagnosis in primary care: population-based observational study in England. *Age Ageing*. 2019 May 1;48(3):347-35/
  15. Clinical Practice Research Datalink. Research applications | CRPD. <https://cprd.com/research-applications>, Accessed on July 22 in 2020.
  16. Clinical Practice Research Datalink. Pricing | CRPD. <https://www.cprd.com/pricing>, Accessed on July 22 in 2020.
  17. NHS Digital. Primary Care Mortality Database. <https://digital.nhs.uk/services/primary-care-mortality-database>, Accessed on July 22 in 2020.
  18. Statistics Netherlands. Causes of death statistics: <https://www.cbs.nl/en-gb/our-services/methods/surveys/korte-onderzoeksbeschrijvingen/causes-of-death-statistics>, Accessed on July 22 in 2020.
  19. Irene E van de Vorst, et al. Socioeconomic Disparities and Mortality After a Diagnosis of Dementia: Results From a Nationwide Registry Linkage Study. *Am J Epidemiol*. 2016 Aug 1;184(3):219-26
  20. Junus M van der Wal, et al. A population-based retrospective study comparing cancer mortality between Moluccan migrants and the general Dutch population: equal risk 65 years after immigration? *BMJ Open*. 2019 Aug 15;9(8):e029288.
  21. Statistics Netherlands. Microdata: Conducting your own research. <https://www.cbs.nl/en-gb/our-services/customised-services-microdata/microdata-conducting-your-own-research>, Accessed on July 22 in 2020.
  22. PHARMO. PHARMO Database Network. <https://pharmo.nl/what-we-have/pharmo-database-network/>, Accessed on July 22 in 2020.
  23. PHARMO. Data request for the PHARMO Database Network. <https://pharmo.nl/what-we-have/data-request-PHARMO/>, Accessed on July 22 in 2020.
  24. Karin Helweg-Larsen. The Danish Register of Causes of Death. *Scand J Public Health*. 2011 Jul;39(7 Suppl):26-9/
  25. Sundhedsdatastyrelsen. Dødsårsagsregisteret (DAR). <https://sundhedsdatastyrelsen.dk/da/registre-og-services/om-denationale-sundhedsregistre/doedsaarsager-og-biologisk-materiale/doedsaarsagsregisteret>, Accessed on July 22 in 2020.
  26. Sundhedsdatastyrelsen. Priser og fakturering.



- <https://sundhedsdatastyrelsen.dk/da/forskerservice/leveringsbetingelser/priser-og-fakturering>, Accessed on July 22 in 2020.
27. Peter Brønnum Nielsen, et al. Restarting Anticoagulant Treatment After Intracranial Hemorrhage in Patients With Atrial Fibrillation and the Impact on Recurrent Stroke, Mortality, and Bleeding. *Circulation*. 2015 Aug 11;132(6):517-25/
  28. Hannah Louise Brooke, et al. The Swedish cause of death register. *Eur J Epidemiol*. 2017 Sep;32(9):765-773/
  29. Socialstyrelsen. Statistics. <https://www.socialstyrelsen.se/en/statistics-and-data/statistics/>, Accessed on July 22 in 2020.
  30. Anders Jonsson, et al. The state of the residential fire fatality problem in Sweden: Epidemiology, risk factors, and event typologies *J Safety Res*. 2017 Sep;62:89-100/
  31. Statistics Finland. Statistics Finland - About statistics - Causes of death. [https://www.stat.fi/meta/til/ksyyt\\_en.html](https://www.stat.fi/meta/til/ksyyt_en.html), Accessed on July 22 in 2020.
  32. Statistics Finland. Application process for microdata | Statistics Finland. [https://www.tilastokeskus.fi/tup/mikroaineistot/hakumenettely\\_en.html](https://www.tilastokeskus.fi/tup/mikroaineistot/hakumenettely_en.html), Accessed on July 22 in 2020.
  33. Statistics Finland. Pricing for research data and services | Statistics Finland. [https://www.tilastokeskus.fi/tup/hinnat/tutkimuspalvelut\\_en.html](https://www.tilastokeskus.fi/tup/hinnat/tutkimuspalvelut_en.html), Accessed on July 22 in 2020.
  34. Siri Vangen, et al. Maternal deaths in the Nordic countries. *Acta Obstet Gynecol Scand*. 2017 Sep;96(9):1112-1119/
  35. Norwegian Institute of Public Health. Norwegian Cause of Death Registry. <https://www.fhi.no/en/hn/health-registries/cause-of-death-registry/>, Accessed on July 22 in 2020.
  36. Anne Gro Pedersen, et al. Data quality in the Causes of Death Registry. *Tidsskr Nor Laegeforen*. 2015 May 5;135(8):768-70/
  37. Norwegian Institute of Public Health. How to apply for access to data – NIPH. <https://www.fhi.no/en/more/access-to-data/applying-for-access-to-data/>, Accessed on July 22 in 2020.
  38. Vibeke Gagnum, et al. Long-term Mortality and End-Stage Renal Disease in a Type 1 Diabetes Population Diagnosed at Age 15–29 Years in Norway. *Diabetes Care*. 2017 Jan;40(1):38-45/
  39. Norwegian Institute of Public Health. Prices for access to data and biological material. <https://www.fhi.no/en/more/access-to-data/prices-for-access-to-data-and-biological-material/>, Accessed on July 22 in 2020.
  40. Australian Institute of Health and Welfare. About National Death Index. <https://www.aihw.gov.au/about-our-data/our-data-collections/national-death-index/about-national-death-index>, Accessed on July 22 in 2020.
  41. Australian Institute of Health and Welfare. Data linkage.

- <https://www.aihw.gov.au/our-services/data-linkage>, Accessed on July 22 in 2020.
42. Eve Blair, et al. Survival and mortality in cerebral palsy: observations to the sixth decade from a data linkage study of a total population register and National Death Index. *BMJ Neurol.* 2019 Jun 4;19(1):111/
  43. The Centre for Health Record Linkage (CHeReL). <https://www.cherel.org.au/>, Accessed on July 22 in 2020.
  44. The Centre for Health Record Linkage (CHeReL). How record linkage works. <https://www.cherel.org.au/how-record-linkage-works>, Accessed on July 22 in 2020.
  45. The Centre for Health Record Linkage (CHeReL). Data dictionaries, NSW Mortality Data. <https://www.cherel.org.au/data-dictionaries#section7>, Accessed on July 22 in 2020.
  46. David P. Smith, et al. Increased risk of suicide in New South Wales men with prostate cancer: Analysis of linked population-wide data. *PLoS One.* 2018; 13(6): e0198679.
  47. The Centre for Health Record Linkage (CHeReL). Apply for linked data – NSW. <https://www.cherel.org.au/apply-for-linked-data> , Accessed on July 22 in 2020
  48. The Centre for Health Record Linkage (CHeReL). Pricing. <https://www.cherel.org.au/pricing>, Accessed on July 22 in 2020
  49. 研究開発の俯瞰報告書 ライフサイエンス・臨床医学分野（2017年）3.

5.4 健康・医療・介護情報. 国立研究開発法人科学技術振興機構 研究開発戦略センター. 2017 Mar. Page 562-576

表 各国の公的死亡データの概要、内容、リンケージ方法、リンケージされた主なデータ

国名	公的死亡データ	データ概要	データ内容	リンケージ方法	リンケージされた主なデータ (末尾の数字は文献番号)
米国	National Death Index (NDI)	Centers for Disease Control and Prevention (CDC, アメリカ疾病予防管理センター)が管轄する全米国民の死亡情報データベース。 研究用途の申請を行うことで、対象人物のデータを取得可能。	氏名、社会保障番号、生年月日、死亡状況、死亡日、死因(ICD コード)など	"eligible"な死亡データ取得のためには、以下の組み合わせの中の少なくとも1つが必要。 ①氏名・社会保障番号(SSN) ②氏名・生年・月 ③社会保障番号・生年月日・性別 統一した個人番号等による個人を完全に特定可能なマッチングを行っていないため、②のリンケージは Probabilistic linkage と考えられる。 ①、③は SSN を使用しているため、Deterministic と言える。	*The US National Health and Nutrition Examination Survey (NHANES) 7 *The Nurses' Health Study (NHS) 8 *the Health Professions Follow-up Study (HPFS) 8 *Centers for Medicare and Medicaid Services (CMS) claims files 9 *The central VA medical files 9 *The dataset of the Pediatric Cardiac Care Consortium (PCCC) 10 *Organ Procurement and Transplantation Network (OPTN) 10 *The Northern New England Cardiovascular Disease Study Group 11 *The NYC HIV Surveillance Registry 12 *NYC DOHMH Vital Statistics Registry 12

表 各国の公的死亡データの概要、内容、リンケージ方法、リンケージされた主なデータ（つづき 1）

国名	公的死亡データ	データ概要	データ内容	リンケージ方法	リンケージされた主なデータ (末尾の数字は文献番号)
米国 (つづき)	Social Security Death Index (SSDI)	1936 年開始。社会保障番号を持つ全国民の死亡データ。 United States Social Security Administration(社会保障局)が公開する Death Master File (Social Security Administration's Death Master File (SSDMF) を元に作成される。複数のサイトで無料で入手可能。	氏名、死亡日、社会保障番号、社会保障番号の発行州、郵便番号	社会保障番号を使うことで他データとのリンケージが可能。	*The NYC HIV Surveillance Registry 12 *NYC DOHMH Vital Statistics Registry 12
英国	Clinical Practice Research Datalink (CPRD)	医薬品・医薬製品規制庁と国立衛生研究所 (NIHR) が共同スポンサーである医療リンケージデータベース。5000 万人の医療データを含む。	Office for National Statistics (国家統計局) の死亡データを引用。死亡日、死因(ICD コード)	8 段階の個人情報マッチングを用いた Deterministic linkage. NHS 番号、氏名、性別、誕生日、郵便番号等を用いる。	*Hospital Episodes Statistics (HES) 14
	Primary Care Mortality Database (PCMD)	NHS Digital が管理。全国の死亡情報を毎月 ONS から収集している。	統計情報、死因、死亡場所、医師 (GP)および検視官の情報	地方政府か NHS の職員にのみ公開されており、一般的研究には使用できない。	

表 各国の公的死亡データの概要、内容、リンケージ方法、リンケージされた主なデータ（つづき 2）

国名	公的死亡データ	データ概要	データ内容	リンケージ方法	リンケージされた主なデータ (末尾の数字は文献番号)
オランダ	Causes of death Registry (Netherlands)	1901 年開始。Statistics Netherlands(CBS) が管理。オランダに在住する全住民の死亡データベース。	性別、年齢層、死因(ICD コード)など	個人番号による他の公的データとの Deterministic linkage が可能。	*Hospital Discharge Register (HDR) 19 *Population Register 19 *Regional Income Survey 19 *The municipal registries 20
デンマーク	The Danish Register of Causes of Death	1871 年開始。2007 年から電子化。National Board of Health が管理。デンマークに在住する全住民の死亡データベース。	個人番号、住所、死亡日、死亡場所、死因(1994 年から ICD-10 を採用)など	Danish Civil Registration Number(Personal identification number)および死亡日データが内包されており、他デンマークの公的データとの Deterministic linkage が可能。	*The Danish Civil Registration System 27 *Danish National Patient Registry 27 *Danish National Prescription Registry 27
スウェーデン	National Cause of Death Register	1952 年開始。スウェーデン社会庁(National Board of Health and Welfare)が管理。国内の全死亡データを記録。	Personal identification numbers (PINs)、生年月日、死亡日、性別、死亡状況・死因の詳細	PIN を使用することでスウェーデンの他の公的データベースとの Deterministic linkage が可能。	*The database on forensic examinations 30 *The database on fatal fires 30

表 各国の公的死亡データの概要、内容、リンケージ方法、リンケージされた主なデータ (つづき 3)

国名	公的死亡データ	データ概要	データ内容	リンケージ方法	リンケージされた主なデータ (末尾の数字は文献番号)
フィンランド	National Cause of Death Register (Finland)	1996 年開始。Statistics Finland が管理。国内の全死亡データを記録。	Finnish personal identification code (個人番号)、死因 (ICD コード)、享年、性別など	Finnish personal identification code (social security number) を使用することで、フィンランドの他の公的データベースとの Deterministic linkage が可能。	*The in-patient register 34
ノルウェー	Norway Cause of Death Registry	1951 年から死因データを記録。Norwegian Institute of Public Health が管理。国内の全死亡データを記録。	国民識別番号、享年、性別、死亡場所、死因 (ICD コード)	Norwegian national identity number(国民識別番号)を使用することで、ノルウェーの他の公的データベースとの Deterministic linkage が可能。	*The National Population Registry 38 *The Norwegian Renal Registry 38
オーストラリア	National Death Index (Australia)	1980 年開始。Australian Institute of Health and Welfare が管理。国内の全死亡データを各州から統合している。	氏名、性別、生年月日、死亡日、死因 (ICD コード) など	個人情報を用いた Probabilistic linkage が可能。	*The Western Australian Register of Developmental Anomalies (WARDA) 42

表 各国の公的死亡データの概要、内容、リンケージ方法、リンケージされた主なデータ（つづき 4）

国名	公的死亡データ	データ概要	データ内容	リンケージ方法	リンケージされた主なデータ (末尾の数字は文献番号)
オーストラリア (つづき)	The Centre for Health Record Linkage (CHeReL)	ニューサウスウェールズ州、オーストラリア首都特別地域の約 6900 万人が対象の医療リンケージデータベース。	死亡データベースには、CHeLeL ID、生年月日、死亡日、享年、死亡した場所、死因などが含まれる。	CHeReL のデータリンケージは複数の個人情報を組み合わせた Probabilistic linkage 。マッチには "ChoiceMaker"等の専用のソフトウェアが用いられる。 外部データベースとのリンケージにも Probabilistic linkage が用いられる。	*NSW Cancer Registry 46 *NSW Registry of Births, Deaths and Marriages 46 *Socio-economic status for each individual was derived from ABS census data 46