

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）

健康格差対策に必要な公的統計のあり方に関する研究

平成27年度 総括・分担研究報告書

研究代表者 伊藤 ゆり

平成28（2016）年 3月

目 次

I. 総括研究報告	
健康格差対策に必要な公的統計のあり方に関する研究	
	伊藤ゆり----- 1
II. 分担研究報告	
1. 市区町村別地理的剥奪指標を用いた全死亡・主要死因別 年齢調整死亡率の社会経済格差の推移	
	伊藤ゆり・近藤尚己・中谷友樹・米島万有子・福井敬祐----- 15
2. 地理的剥奪指標データセットの構築	
	中谷友樹・米島万有子----- 51
3. 市区町村別性年齢階級別人口の線形補間について	
	福井敬祐----- 61
4. Probabilistic linkage を用いた大規模公的統計データベースの活用に関する研究	
	福井敬祐・近藤尚己----- 69
5. がん進行度別罹患率における社会経済格差	
	伊藤ゆり・近藤尚己・中谷友樹・米島万有子・福井敬祐・中山富雄----- 77
III. 研究成果の刊行に関する一覧表	
	----- 91
IV. 研究成果の刊行物・別刷	
	----- 95

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))
総括研究報告書

健康格差対策に必要な公的統計のあり方に関する研究

研究代表者 伊藤 ゆり 大阪府立成人病センターがん予防情報センター 主任研究員

研究要旨

国民皆保険制度下の日本において、近年社会経済状況により死亡や疾病発症などの格差が生じ始めている。国民の経済格差が拡大する中、健康格差をモニタリングし、対策を講じる必要がある。そこで、本研究では現状で利用可能なデータを用いて健康格差指標の分析を行うとともに、現行の公的統計での限界や課題を抽出し、健康格差を測るために必要な公的統計のあり方について検討する。

平成 27 年度は、現状で利用可能なデータによる分析・課題抽出を行っている。

空間疫学的手法を用いた全死亡における社会経済格差

人口動態統計の二次利用申請を行い、平成 27 年 11 月に資料を入手した。市区町村別地理的剥奪指標(ADI)による全死亡および主要死因別の年齢調整死亡率の社会経済格差についての基礎的解析を行った。

主要死因ごとの職業・産業別死亡率の時系列分析の地域差

人口動態特殊統計は死亡時の職業・産業となっているため、正確な職業・産業別死亡率の分析が行えない。そこで、諸外国において用いられている Probabilistic Record Linkage の手法が我が国の国勢調査データと人口動態統計データのリンケージに適用可能かどうか調べた。

がんを事例とした社会経済格差およびその要因分析

大阪府がん登録資料を用い、小地域 ADI に基づくがん進行度別罹患率の社会経済格差についての時系列分析を行った。

～ のさらなる詳細分析を通して、健康格差指標のモニタリングにおける今後の公的統計のあり方について引き続き検討を行う。

分担研究者

中谷 友樹 立命館大学 文学部 教授
近藤 尚己 東京大学大学院医学系研究科 准教授

研究協力者

米島万有子 立命館大学衣笠総合研究機構 専門研究員
福井 敬祐 大阪府立成人病センターがん予防情報センター 研究員
中山 富雄 大阪府立成人病センターがん予防情報センター 課長

A . 研究目的

国民皆保険の体制下にあるわが国においても、収入や職業などの社会経済状況により、各種疾患の死亡率や生存率をはじめとした健康指標において、格差が生じていることが報告されている。このように健康格差の問題が顕在化する中、第 2 次健康日本 21 の計画においては、「健康寿命の延伸と健康格差の縮小」が目標に掲げられたが、

我が国の公的統計は健康格差のモニタリング体制は十分に整っていない。

健康格差対策を行うためには公的統計を用いた定期的な健康格差指標のモニタリングが必要である。現状で利用可能なデータを用いた健康格差指標の分析を通して、現行の公的統計での限界や課題を抽出し、健康格差を測るために必要な公的統計のあり方について提言する。

本研究班においては以下の3つの課題について、取り組んでいる。

空間疫学的手法を用いた全死亡における社会経済格差
主要死因ごとの職業・産業別死亡率の時系列分析の地域差
がんを事例とした社会経済格差およびその要因分析

そのうち、研究結果および考察に関しては、～の進捗状況について、報告する。

B. 研究方法

市区町村別社会経済指標を用いた全死亡および死因別死亡率の格差

人口動態統計の二次利用申請を行い、平成27年11月に資料を入手した。Nakayaraの市区町村別地理的剥奪指標(Areal Deprivation Index: ADI) [1]を人口動態統計の二次利用データに付与し、ADI別全死因死亡率の分析を行った。ADIは数値が大きいほど地域の剥奪度が高い、つまり社会的に不利な経済状況にある人々の割合が高いという指標である。これを各地域の人口で重み付けし、5分位に分けた(第1分位が最も剥奪度が低く裕福な地域、第5分位が最も剥奪度が高い地域)。ADIの詳細については、中谷らの分担報告書に詳細が記載

されている。市区町村別性・年齢階級別人口は国勢調査(e-stat)より入手した。国勢調査実施年以外の年については、線形補間により内挿(2011~2014年は外挿)した(詳細は福井の報告書を参照)。市区町村別人口とADIと連結し、人口重み付きADIを5分位および100分位でグループ化した。これを人口動態データに付与し、ADI分位群別の全死亡および主要死因別年齢調整死亡率の分析を行った。これらの手順および結果については伊藤らの分担報告書に記載した。

また死亡率の社会経済指標による格差の指標としては絶対指標としてSlope Index of Inequality (SII)および相対指標としてRelative Index of Inequality (RII)を用いた[2]。トレンドの評価にはJoinpoint regression modelを用いた[3]。

人口動態統計の二次利用申請時に、オンライン届出分の資料には詳細住所が含まれ、入手可能と判明したため、申請手順を経て入手した。これにより、小地域情報に基づくADIを用いた全死亡率(生命表)・各死因別死亡率の計測の可能性が出てきたが、ジオコーディングおよび国勢調査の調査区への対応作業は非常に労力と時間を要することが推察される。将来的にはこの情報を用いて健康格差モニタリングに移行していくため、利用可能性や作業手順などを本研究班でまとめる意義は大きい。平成28年度に一部の都道府県のデータを用いて、分析を行い、利用可能性を検討する。

また、空間的階層ベイズモデルを用いて、職業・産業の分布の違いを補正した市区町村別死亡率の検討に関しては、分析資料を整備し解析が可能な状況となった。

主要死因ごとの職業・産業別死亡率の時系列分析の地域差

本研究課題を実施する上で、留意が必要な点として、人口動態特殊統計は死亡時の職業・産業となっている点が挙げられる。そこで、より正確な職業・産業別死亡率の分析を行うために、諸外国において用いられている Probabilistic Record Linkage の手法が我が国の国勢調査データと人口動態統計データのリンケージに適用可能かどうか調べた。

がん進行度別罹患率における社会経済格差

大阪府がん登録資料より、1993-2004 年に診断された胃・大腸・肺・乳房・子宮頸・前立腺がんの患者データを用い、小地域(町字単位) ADI を人口重み付き 5 分位でグループ化し、がん進行度別罹患率の社会経済格差についての時系列分析を行った。

がん罹患率は検診受診等の予防行動に影響を受けるため(検診受診率の高い地域の罹患率が高い)、診断時の進行度別(早期がん: 上皮内 + 限局 / 進行がん: 領域 + 遠隔転移)の年齢調整罹患率を ADI ごとに算出した。部位・性別・診断時期別に、分散重み付き最小二乗法により ADI と進行度別罹患率の関連を分析し、ADI 第 5 分位地域と第 1 分位地域の罹患率差を推定した。

(倫理面への配慮)

本研究は、既存の厚生労働統計や地域がん登録資料を用いた記述疫学研究であり、本人同意取得の原則は適用されないが、職業・産業・市区町村コードを含む人口動態統計の分析においては、詳細の集計により、個人が同定されないよう注意を払う必要が

ある。また、地域がん登録資料と国勢調査の小地域統計をもとにした社会経済因子の突合に際しては、詳細住所を用いるため、スタンドアロン環境の PC において作業を行い、突合後のデータは個人同定が不可能な状態に変換する。

平成 27 年度に二次利用申請をして入手した人口動態統計資料の一部には詳細住所情報が含まれているため、本データを利用する可能性のある研究者の所属機関における倫理審査委員会に申請を行い、承認を得る(研究代表者の所属機関においては承認済)。

また、研究の進行によっては、自治体において実施する健康事業データの利用をする可能性があるが、その際には申請者所属機関の倫理審査委員会を始め、各自治体における法令を遵守した形で利用を申請する。

C . 研究結果

市区町村別社会経済指標を用いた全死亡および死因別死亡率の格差

1995-2014 年の人口動態統計を用いて、市区町村別 ADI5 分位別に 0-84 歳の年齢調整全死亡率の推移を 5 年ごとに示した(図 1)。男女とも ADI が高いほど(社会経済指標が低いほど)死亡率が高かった。全死亡における年齢調整死亡率の絶対格差(SII)は男性では 2000-2004 年で最大の人口 10 万対 102.5 (95% CI: 100.5-104.5)であった。女性では、2010-2014 年における SII が最大で 30.2 (95% CI: 29.1-31.3)であり、男性よりも絶対格差は小さかった(図 2)。相対的指標 RII でみると、男女ともやや拡大傾向にあった(図 3)。

主要死因別にみた SII では男性では、がんが最も格差が大きかった。男女とも、2011

年の東日本大震災の影響で不慮の事故による格差が絶対指標、相対指標ともに2010-2014年の死亡において大きくなった(図4、図5)。

主要死因ごとの職業・産業別死亡率の時系列分析の地域差

現状の国勢調査および人口動態統計における二次利用データにおいて、リンケージが可能かどうかについて、諸外国で用いられている Probabilistic Record Linkage の手法についてまとめた。Probabilistic record linkage においては、リンケージを行うデータベース間での共通のマッチング変数が重要となるが、本研究で想定している国勢調査および人口動態統計の二次利用データで入手可能な変数では、あまり正確にリンケージできない可能性が示唆された。

がん進行度別罹患率における社会経済格差

分散重み付き最小二乗法により推定された ADI 第1分位(Q1)と第5分位(Q5)における進行度別罹患率の絶対格差(Q5 - Q1)について性別、部位別に示した(図6、図7)

前立腺がん以外の全ての部位で、進行がんの罹患率は ADI の高い地域ほど高かった。ADI 第5分位と第1分位間の進行がんの罹患率差は、男女とも1993-1998年診断の肺がんが最大で、それぞれ人口10万人対8.0(95%信頼区間:6.3-9.7)、5.4(3.6-7.3)であった。しかし、その差は1999-2004年には統計的有意に減少し、他の部位の進行がん罹患率においても、統計的有意には拡大しなかった。一方、早期がんの罹患率は、男性で前立腺、胃、大腸において、ADIの

高い地域ほど罹患率が低かった。その傾向は前立腺で顕著であり、1999-2004年では ADI 第5分位と第1分位間の罹患率差は -7.3 と拡大した。女性では子宮頸がんにおいて、ADI の高い地域の罹患率が高かった。格差の縮小が最も大きかったのは、男性の胃・大腸の進行がんであったが、この傾向は女性ではみられなかった。

D . 考察

市区町村別社会経済指標を用いた全死亡および死因別死亡率の格差

絶対指標において、女性より男性の死亡率格差が大きいのは、男性の平均余命に比べ女性の平均余命が長いことが影響していると考えられるが、相対指標における違いについては、死因別の結果と合わせて検討する必要がある。女性の死亡率格差が拡大している点については、震災の影響を除去しても残るため、要因を調べる必要がある。

東日本大震災の年やその年が含まれる時期の死亡率において格差が大きかったのは、最も剥奪されている地域(Q5)に震災被害に遭った東北地方の市区町村が多く含まれていたためと考えられる。被害地域を除外した分析との比較や、阪神・淡路大震災の影響とを比較するなど、さらなる分析・考察が必要である。

今回の分析は単純な ADI 分位群ごとの層別分析であるため、都道府県などより大きな単位での地域の影響などは考慮できていない。階層構造を考慮したマルチレベル分析や、高度な空間疫学的手法を適用し、全死亡・死因別死亡における社会経済格差をよりクリアに表現していく予定である。

主要死因ごとの職業・産業別死亡率の時系列分析の地域差

人口動態統計および国勢調査の二次利用データにおいて氏名や生年月日などの利用が困難である我が国の現状において、Probabilistic Record Linkage によるリンケージデータの精度が低い可能性がある。将来的には、北欧諸国や英国、米国のように、個人識別番号の整備を経て、各種公的統計のリンケージを公的機関が行い、個人識別可能な情報を削除した匿名化データを利用者に提供する仕組みなどを検討していく必要がある。

がん進行度別罹患率における社会経済格差

進行度別がん罹患率の社会経済格差において、特に進行がんの罹患率の格差には、喫煙やハイリスクな性行動など、がん発症のリスクとなりうる行動の違いやがん検診の受診率の違いなどが影響していると考えられる。男性において観測された早期がんにおける罹患率の逆方向の格差に関しては、企業などの検診提供体制の違いなどに起因する可能性がある。がんのリスク要因や検診受診率などの情報と合わせ、要因を分析しておく必要がある。

E . 結論

人口動態統計および地域がん登録資料を用いて、現状で分析可能な全死亡・主要死因別死亡率およびがん進行度別罹患率における社会経済格差のモニタリングを行った。格差の生じる原因に関しては、死因ごとや年齢区分（小児・若年・中年・高齢者など）に時系列の詳細分析を行うことで、全死亡における社会経済格差のトレンドを

明らかにし、関連する要因を探索し、健康格差対策に役立てたい。また、研究計画当初に比べて、人口動態統計の住所情報に関し、詳細データの利用可能性が高まった。利用手順及び、現段階での限界を検討し、健康格差評価における今後の公的統計のあり方について、まとめていく。

F . 健康危険情報

なし

G . 研究発表

1 . 論文発表

Ito Y, Nakaya T, Ioka A, Nakayama T, Tsukuma H, Uehara S, et al. Investigation of spatial clustering of biliary tract cancer incidence in Osaka, Japan: neighbourhood effect of a printing factory. *Journal of Epidemiology*. 2015; [in press].

2 . 学会発表

伊藤ゆり, 中谷友樹, 近藤尚己, 福井敬祐, 中田佳世, 井岡亜希子, et al. 大阪府におけるがん進行度別罹患率の社会経済格差：1993-2004年における格差の変化. 第74回日本公衆衛生学会総会. 長崎; 2015. P0802-10. Poster

Ito Y, Nakaya T, Kondo N, Fukui K, Nakaya K, Ioka A, et al. SOCIO-ECONOMIC DIFFERENCES IN STAGE-SPECIFIC CANCER INCIDENCE IN OSAKA, JAPAN: 1993-2004. Mumbai, India; 2015. Oral

H . 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

引用文献

1. Nakaya T, Honjo K, Hanibuchi T et al. Associations of all-cause mortality with census-based neighbourhood deprivation and population density in Japan: a multilevel survival analysis. PLoS

One 2014; 9: e97802.

2. Harper S, Lynch J. Selected Comparisons of Measures of Health Disparities: A Review Using Databases Relevant to Healthy People 2010 Cancer-Related Objectives. In NCI Cancer Surveillance Monograph Series. Bethesda, MD,: National Cancer Institute. 2007.

3. Statistical Research and Applications Branch, National Cancer Institute. Joinpoint Regression Program, Ver. 4.2.0.2. In 3.3 Edition. 2015.

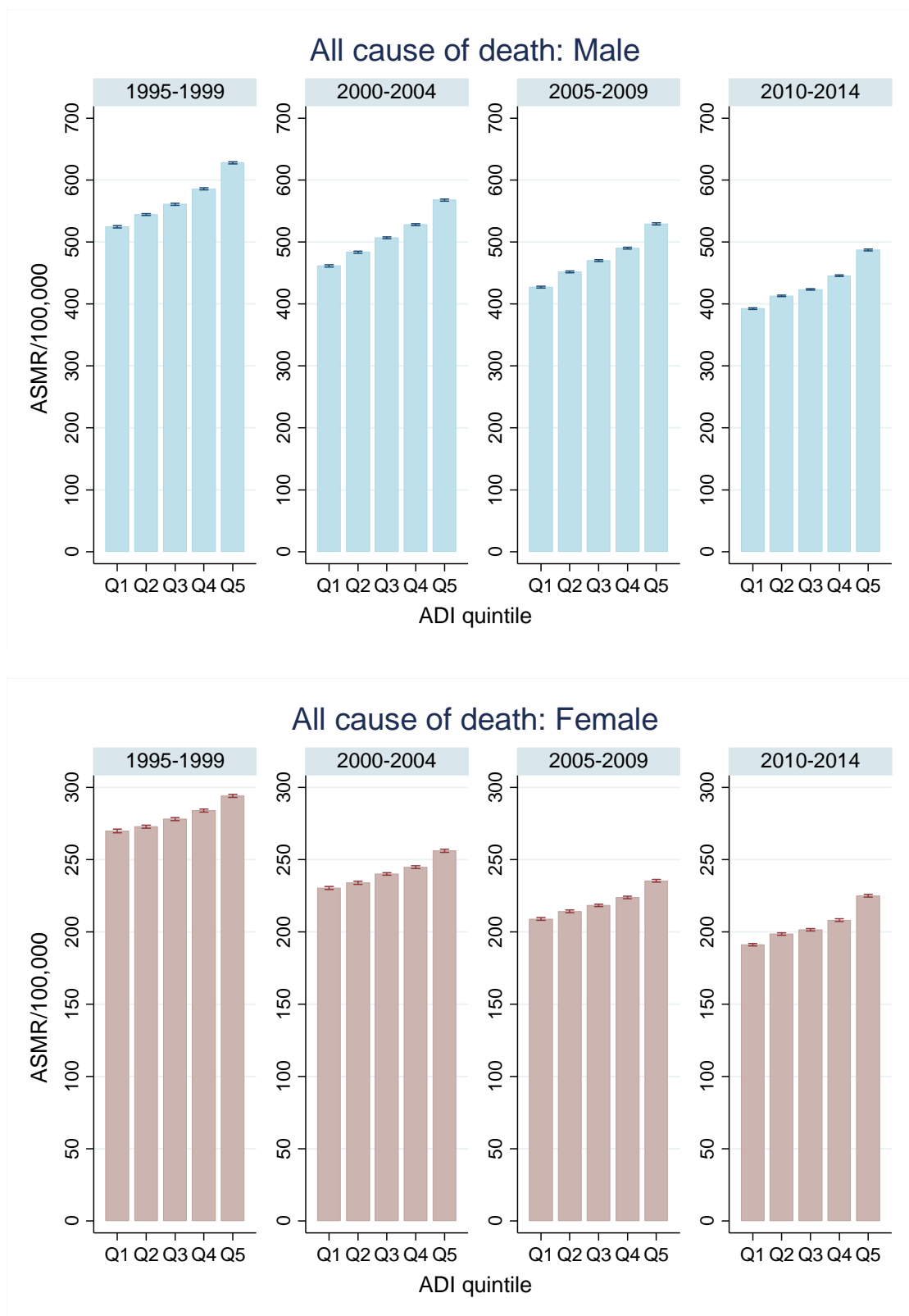


図1. 市区町村別 ADI5 分位ごとの年齢調整死亡率(0-84 歳)の推移:全死亡, 上・男性, 下・女性

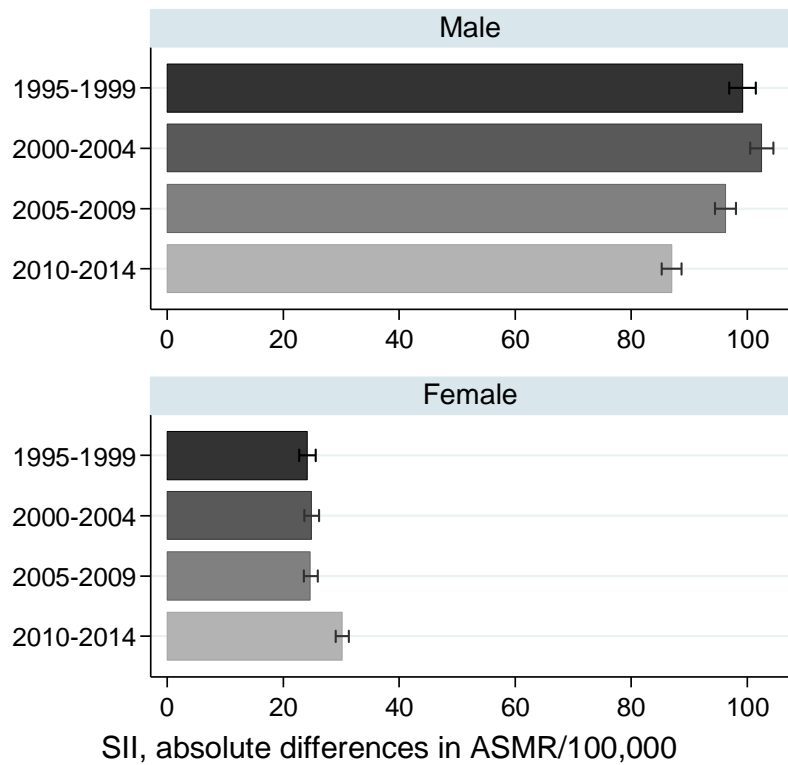


図 2 . 全死亡における年齢調整死亡率の格差の絶対指標 (Slope Index of Inequalities: SII): Q5 の年齢調整死亡率と Q1 の年齢調整死亡率の差

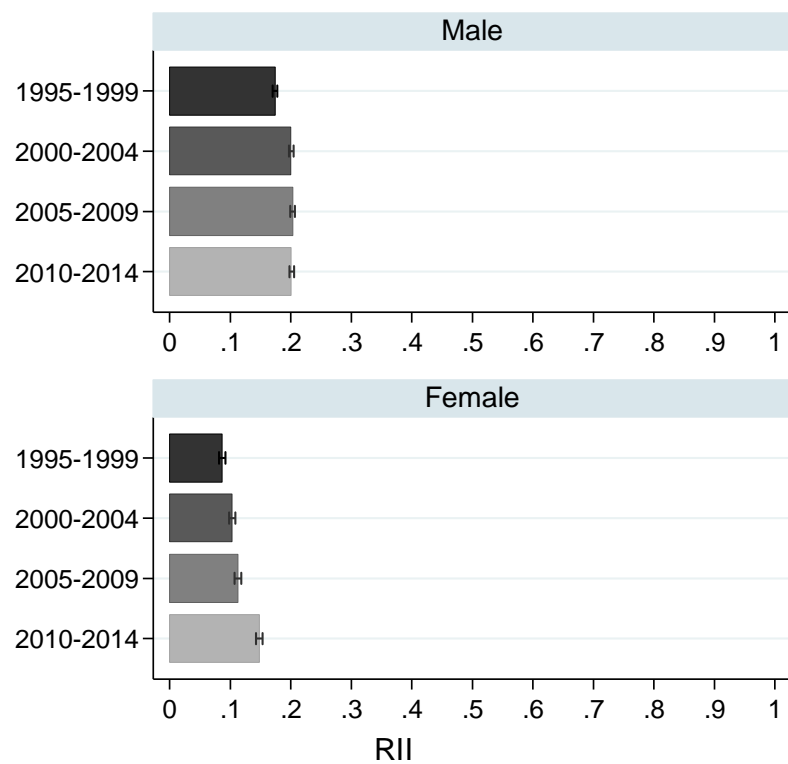


図 3 . 全死亡における年齢調整死亡率の格差の相対指標 (Relative Index of Inequalities: RII): SII を Q1-Q5 全体の年齢調整死亡率で除したものの

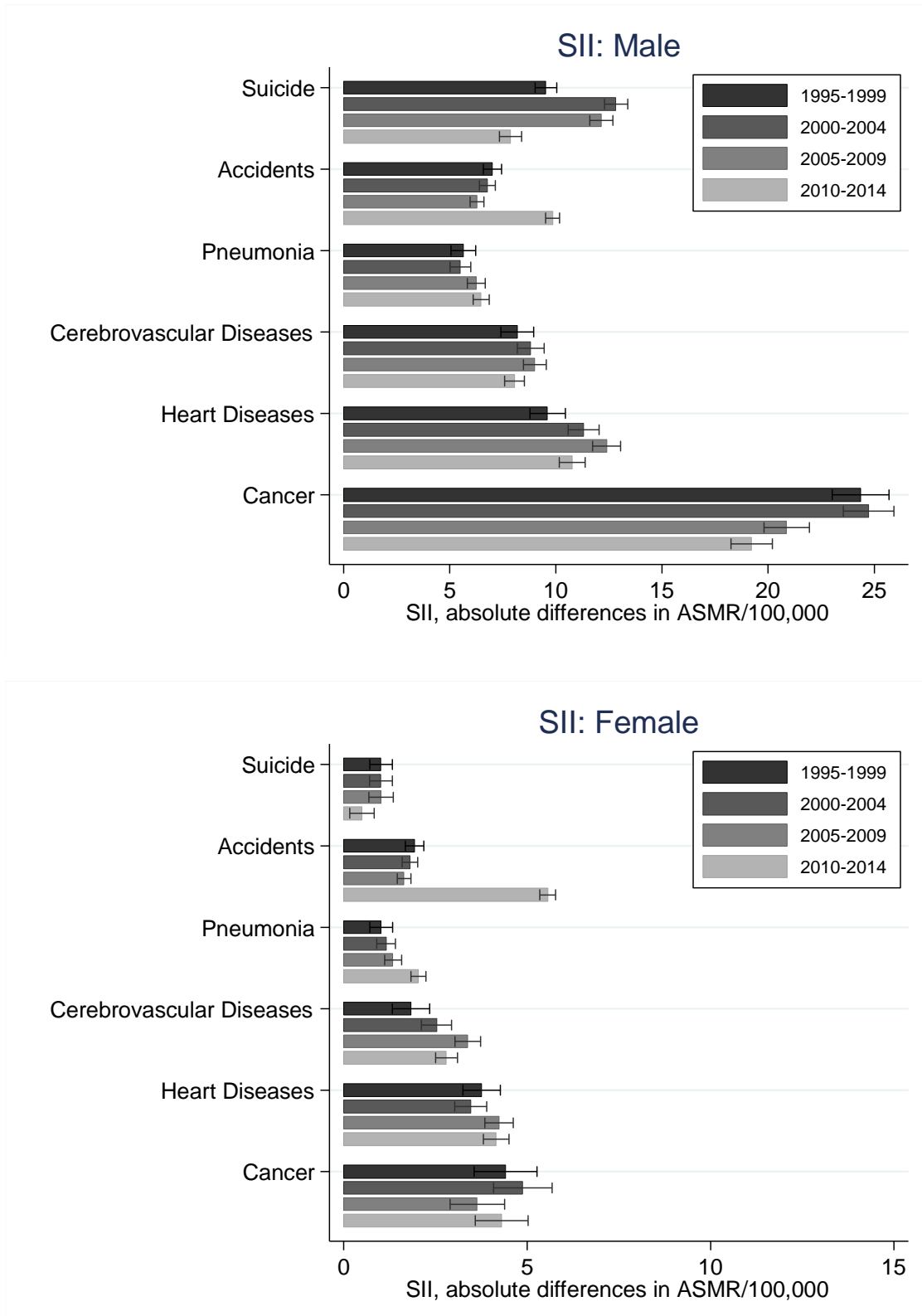


図 4 . 主要死因別年齢調整死亡率における格差の絶対指標 (Slope Index of Inequalities: SII): Q5 の年齢調整死亡率と Q1 の年齢調整死亡率の差, 上・男性, 下・女性

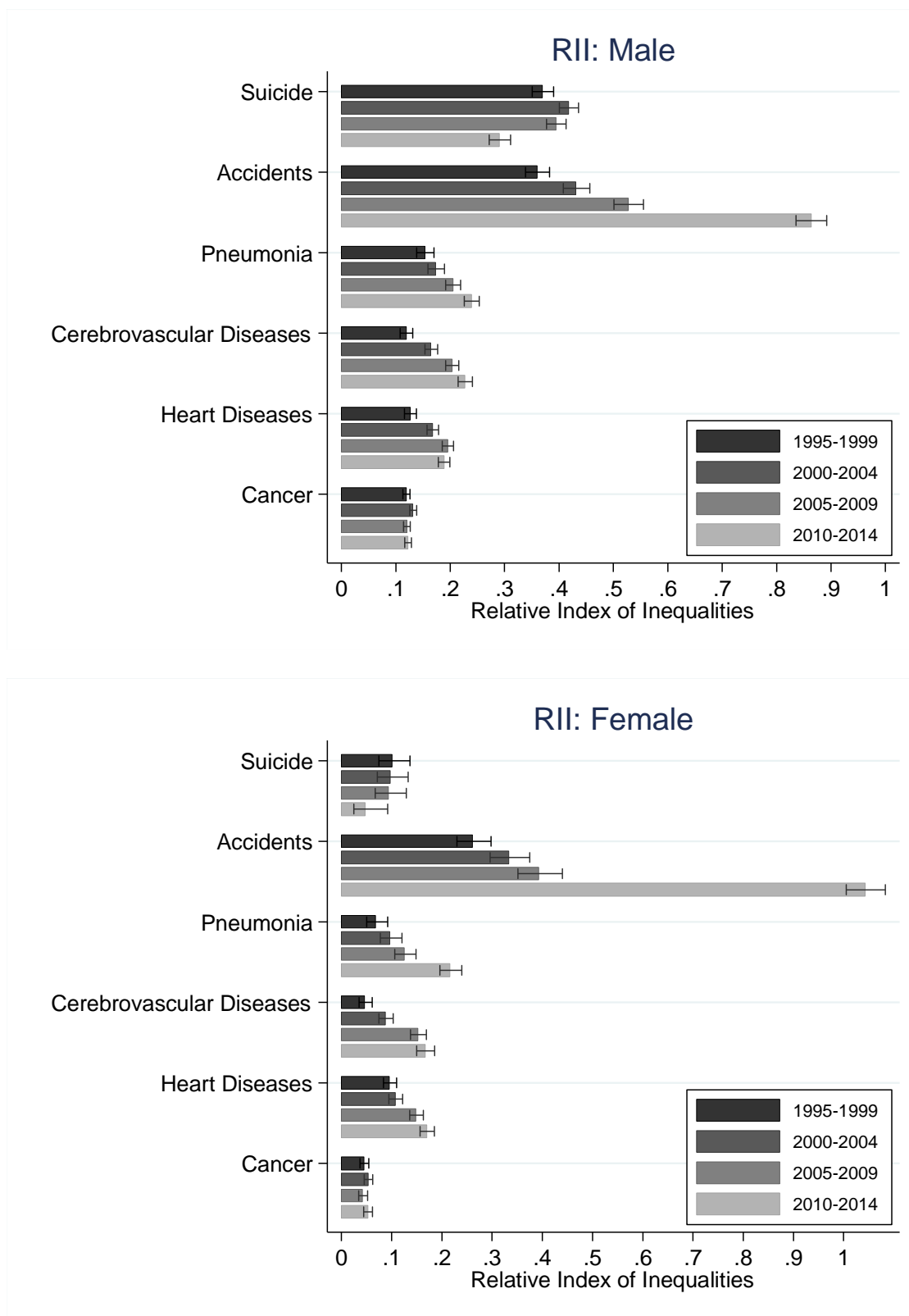


図 5. 主要死因別年齢調整死亡率の格差の相対指標 (Relative Index of Inequalities: RII): SII を Q1-Q5 全体の年齢調整死亡率で除したもの, 上・男性, 下・女性

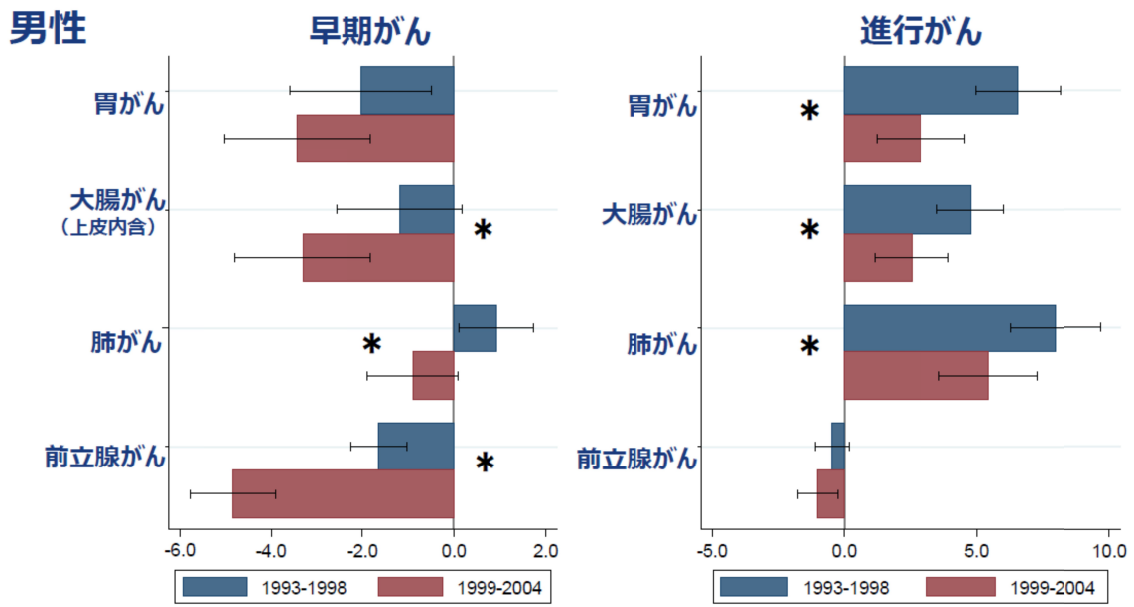


図 6 . 大阪府における進行度別がん罹患率の社会経済格差 (第 5 分位 - 第 1 分位): 男性

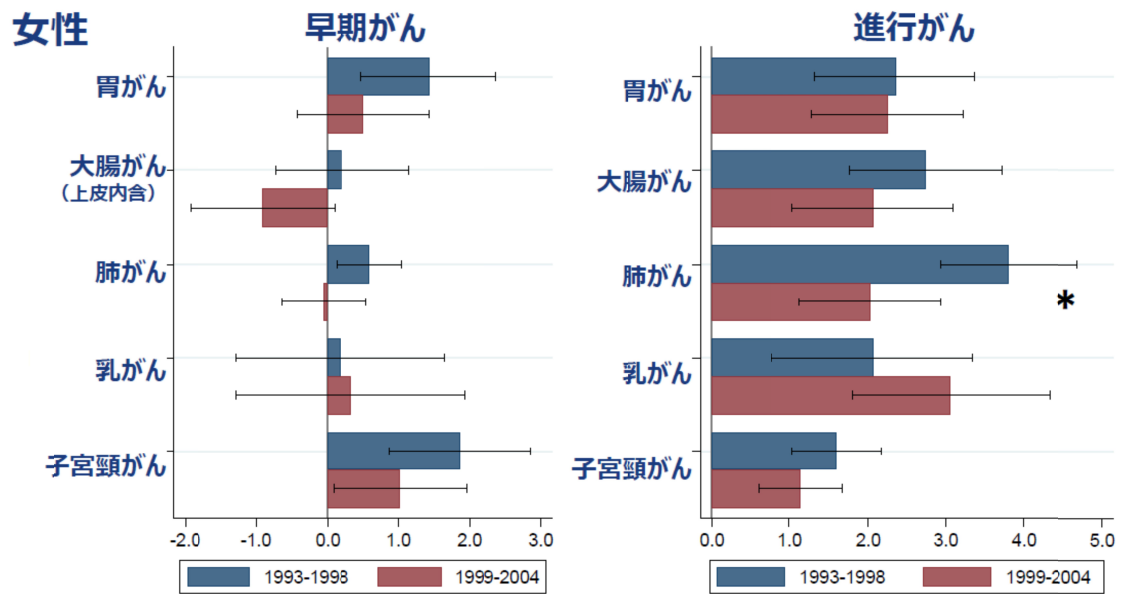


図 7 . 大阪府における進行度別がん罹患率の社会経済格差 (第 5 分位 - 第 1 分位): 女性

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))
分担研究報告書

**市区町村別地理的剥奪指標を用いた全死亡・主要死因別年齢調整死亡率の
社会経済格差の推移**

研究代表者 伊藤ゆり 大阪府立成人病センター がん予防情報センター 主任研究員
研究分担者 近藤尚己 東京大学大学院医学系研究科 准教授
研究分担者 中谷友樹 立命館大学文学部(立命館大学歴史都市防災研究所 兼任) 教授
研究協力者 米島万有子 立命館大学衣笠総合研究機構 専門研究員
研究協力者 福井敬祐 大阪府立成人病センター がん予防情報センター 研究員

研究要旨

人口動態統計の二次利用データを用いて、市区町村別の地理的剥奪指標(Areal Deprivation Index: ADI)に基づき、全死亡および主要死因別の年齢調整死亡率の社会経済格差の年次推移を分析した。分析にあたり、準備したデータの加工手順を本報告に記す。市区町村合併・分割による市区町村の変化の影響を除去するために、2010年時点の市区町村区分をベースとし、分割した市区町村は分割前のものに再統合した1840の市区町村区分を用いた。ADIは国勢調査データを用いるため、1990~2010年の国勢調査実施年のものを使用した。市区町村別・性・年齢階級別人口は、国勢調査実施年以外の年は線形回帰により内挿補間した(2011~2014年は外挿補間)。各国勢調査実施年のADIはもっとも近い国勢調査実施年となるように人口、死亡データに割り当てた。

市区町村別の0~84歳の年齢調整死亡率をADIの国勢調査年別に算出し、市区町村別ADIと死亡率の相関関係をみた。また、ADIを人口で重みづけした5分位・100分位の群に分け、単年ごとに全死亡および主要死因の年齢調整死亡率を算出し、格差の指標をSlope Index of Inequalities (SII)およびRelative Index of Inequalities (RII)で示した。また格差の年次推移をJoinpoint regression modelにより分析した。全死亡では絶対指標(SII) 相対指標(RII)ともに、男性の格差が大きかった。女性では男性に比べ格差が小さいものの、拡大傾向にあった。死因別では、絶対指標(SII)でみると、男性ではがん死亡における格差が最も大きく、女性では東日本大震災のあった時期では不慮の事故における死亡の格差が大きかった。相対指標(RII)でみた場合、男女とも不慮の事故における死亡率格差が最も大きく、東日本大震災の年が含まれる時期では突出していた。男性では、不慮の事故に次いで自殺における死亡率格差が大きく、特に2000年代で顕著であった。今後、年齢区分毎の分析や都道府県の影響など、詳細の分析を通して、格差の生じる要因の探索を行い、健康格差縮小に向けた基礎資料としたい。

A. 研究目的

第2次健康日本21の計画において、「健康寿命の延伸と健康格差の縮小」が目標に掲げられたが、我が国の公的統計を用いた健康格差のモニタリング体制は十分ではない。そこで、本研究班では、現状で利用可能な公的統計を用いた健康格差指標の分析を通して、現行の公的統計での限界や課題を抽出する。本課題では、市区町村レベルの地理的剥奪指標（Areal Deprivation Index: ADI）と1985～2014年の人口動態統計データおよび、国勢調査より得られた人口を用いて、地理的剥奪指標別の全死亡および主要死因別の0～84歳の年齢調整死亡率を算出する。

B. 研究方法

<使用したデータ>

死亡データ：1985～2014年人口動態統計データを使用した。ただし、主要死因別の分析はICD-10への変更以降の1995～2014年データを使用した。

人口データ：1985～2014年の市区町村別、性別・年齢階級別人口を国勢調査年ごとに入手し、国勢調査年以外の年は線形回帰により内挿（2011～2014年は外挿）した。詳細は福井の報告を参照。

市区町村別地理的剥奪指標：中谷らがJGSS調査データより推定した式を用いて、国勢調査に基づき、1990～2010年の国勢調査対象年ごとに作成したものを使用した。ADIは数値が大きいほど、社会的に剥奪されている人が多く住む地域を意味し、5分位で言えば、Q1が最も裕福な人が多く住む地域で、Q5が最も社会経済指標の低い人が住む地域といえる。この指標の詳細は中谷らの

報告書を参照。

～の対応年は表1に示した。

<解析方法>

市区町村別 ADI5 分位ごとの年齢調整死亡率の推移

市区町村別 ADI5 分位ごとに0-84歳の年齢調整死亡率を全死亡および主要死因別に算出し、その経時変化を検討した。経時変化は単年と5年ごとの分析を行った。

格差の指標

市区町村別 ADI に対応する国勢調査年ごとに、人口で重みづけした5分位(Q1～Q5)および100分位で分け(Q1～Q100)、1年ごとに年齢調整罹患率を算出した。Q1～Q5またはQ1～Q100の年齢調整死亡率を分散重み付き最小二乗法により回帰し、その回帰係数をもとに、格差の絶対指標として Slope Index of Inequalities (SII) および相対指標として Relative Index of Inequalities (RII) を算出した。

$$ASMR_i = \beta_{ADI} a_i$$

$$i = 1, \dots, x \text{ (} x \text{ は分位数)}$$

$$SII = \beta_{ADI} \times (x - 1)$$

SII は、分散重み付き最小二乗法により推定された Qx と Q1 の年齢調整死亡率の推定値の差に等しい。

$$RII = SII / ASMR_{all}$$

RII は、SII を Q1-Qx 全グループの年齢調整死亡率 (ASMR_{all}) で除することで、どの年齢層・死因における格差が大きいのか、

などサブグループ間での格差の大小の比較が可能となる。

ADI 別年齢調整死亡率及び格差指標の推移

1995-2014 年死亡の推移を 5 年ずつまとめた結果の推移と単年の年次推移について分析した。

5 分位ごとの年齢調整死亡率の年次推移および格差指標 SII・RII の単年の年次推移は Joinpoint regression model を用いて分析した。[1, 2]

年齢調整死亡率は昭和 60 年の日本人モデル人口を標準人口とし、直接法により算出した。分析には Stata Ver. 13.1 を用いた。[3]

(倫理面への配慮)

本研究は、人口動態統計および国勢調査の市区町村別集計データを用いた記述疫学研究であり、本人同意取得の原則は適用されないが、市町村コードを含む人口動態統計の分析においては、結果の提示においては個人が同定されないよう 3 人未満の集計結果に関しては表として提示しないなどの配慮を行う。

C . 研究結果

市区町村別 ADI5 分位別全死亡および主要死因別年齢調整死亡率の推移(5 年ごと)を図 1-1 ~ 1-7 および表に示した。全死亡における年齢調整死亡率の絶対格差(SII)は男性では 2000-2004 年で最大の人口 10 万対 102.5 (95% CI: 100.5-104.5) となり、2010-2014 年では、87.0 (95% CI: 85.3-88.7) と最も小さくなった。女性では、2010-2014 年が最大で 30.2 (95% CI:

29.1-31.3) であり、男性よりも絶対格差は小さかった(図 2-1、表 2)。相対的指標 RII でみると、男女ともやや拡大傾向にあった(図 3-1、表 2)。主要死因別にみた SII では男性では、がんが最も格差が大きかった。男女とも、2011 年の東日本大震災の影響で不慮の事故による格差が絶対指標、相対指標ともに 2010-2014 年の死亡において大きくなった(図 2-2、図 3-2、表 2)。

市区町村別 ADI5 分位別の年齢調整死亡率の年次推移を死因別に図 4-1 ~ 4-7 に示した。男女とも、どの死因および ADI 群においても、近年、年齢調整死亡率は減少傾向にあった。自殺は特に男性で特異なトレンドを示し、1990 年代半ばにピークがあり、2000 年代後半から減少傾向がみられた(図 4-7)。また、不慮の事故は東日本大震災があった 2011 年で Q4 および Q5 において、高い死亡率を示した(図 4-6A)。全体の傾向を見るために阪神淡路大震災のあった 1995 年および東日本大震災のあった 2011 年を削除したデータでの Joinpoint regression model の結果も示した(図 4-6B)。

格差指標 SII, RII の年次推移も Joinpoint regression model により分析した(図 5-1-1A ~ 図 5-7-2, 表 3-1, 表 3-2)。全死亡の SII(図 5-1-1A), RII(図 5-1-2A)ともに、2011 年突出して高くなっていた。そのため、この影響を除外した Joinpoint regression model による分析結果についても提示した(図 5-1-1B、図 5-1-2B)。2011 年の影響を除外すると、男性で SII は 2006 年までは横ばいであったが、2006 年以降減少傾向に転じた。RII でみると、2005 年まで格差は拡大傾向にあったが、2005 年から

格差が縮小傾向に転じた。女性は SII でみると、1985-2014 年まで横ばいであるが、RII では統計的に有意な格差拡大傾向が見られた（図 5-1-1B, 表 3-1, 表 3-2）。

D . 考察

市区町村別 ADI を用いて全死亡・主要死因別の 0-84 歳の年齢調整死亡率において、社会経済格差を観察した。全死亡では絶対指標（SII） 相対指標（RII）ともに、男性の格差が大きかった。女性では男性に比べ格差が小さいものの、拡大傾向にあった。死因別では、絶対指標（SII）でみると、男性ではがん死亡における格差が最も大きく、女性では東日本大震災のあった時期では不慮の事故における死亡の格差が大きかった。相対指標（RII）でみた場合、男女とも不慮の事故における死亡率格差が最も大きく、東日本大震災の年が含まれる時期では突出していた。不慮の事故に次いで、男性では自殺における死亡率格差が大きく、特に 2000 年代で顕著であった。

絶対指標において、女性より男性の死亡率格差が大きいのは、男性の平均余命に比べ女性の平均余命が長いことが影響していると考えられるが、相対指標における違いについては、死因別の結果と合わせて検討する必要がある。女性の死亡率格差が拡大している点については、震災の影響を除去しても残るため、要因を調べる必要がある。

東日本大震災の年やその年が含まれる時期の死亡率において格差が大きかったのは、最も剥奪されている地域（Q5）に震災被害に遭った東北地方の市区町村が多く含まれていたためと考えられる。被害地域を除外した分析との比較や、阪神・淡路大震災の影響とを比較するなど、さらなる分析・考

察が必要である。

男性の自殺における特異なトレンドはバブルの崩壊やリーマンショックなど経済状況の変化と合わせて検討を行う必要がある。格差のトレンドにおける都道府県間の違いなども他の地域経済指標等と合わせた分析を行う必要がある。

今回の分析は単純な ADI 分位群ごとの層別分析であるため、都道府県などより大きな単位での地域の影響などは考慮できていない。階層構造を考慮したマルチレベル分析や、高度な空間疫学的手法を適用し、全死亡・死因別死亡における社会経済格差をよりクリアに表現していく予定である。

日本における全死亡や主要死因別の死亡率の社会経済格差については、1990 年代までの人口動態統計によるもの[4, 5]や、近年のコホート研究によるもの[6-9]が報告されている。これらの研究結果とは、対象とする年齢区分や死因（疾患）、社会経済指標の違いがあるため単純な比較はできないものの、日本における社会経済格差の存在を示す意味では大きく異なる結果ではなかった。

市区町村という単位は人口規模がかなり大きいいため、先行研究におけるコホート研究に使用したもの[9]や、地域がん登録資料に適用した事例[10]に用いた地理的剥奪指標が町字単位であることと比べると、地理的剥奪指標としては影響度が薄まり、格差を過小評価している可能性がある[11]。今後、人口動態統計のオンライン届出分の住所情報を用いて、より小地域の地理的剥奪指標を用いた分析が可能かどうかの検討を行っていく。

また、死因ごとや年齢区分（小児・若年・中年・高齢者など）に時系列の詳細分析を行うことで、全死亡における社会経済格差

のトレンドを明らかにし、関連する要因を探索し、健康格差対策に役立てたい。

E . 結論

国勢調査データに基づく市区町村別地理的剥奪指標を人口動態統計に付与し、市区町村単位での全死亡・主要死因別の年齢調整死亡率の格差とそのトレンドについて、その集計結果を報告した。

F . 健康危険情報

なし

G . 研究発表

1 . 論文発表

なし

2 . 学会発表

なし

H . 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

引用文献

1. Kim HJ, Fay MP, Feuer EJ, Midthune DN. Permutation tests for joinpoint regression with applications to cancer rates. *Stat Med* 2000; 19: 335-351.
2. Statistical Research and Applications Branch, National Cancer Institute. Joinpoint Regression Program, Ver. 4.2.0.2. In 3.3 Edition. 2015.
3. StataCorp. Stata Statistical

Software: Release 13. College Station, TX: StataCorp LP. 2013.

4. Fukuda Y, Nakamura K, Takano T. Municipal socioeconomic status and mortality in Japan: sex and age differences, and trends in 1973-1998. *Soc Sci Med* 2004; 59: 2435-2445.

5. Fukuda Y, Nakamura K, Takano T. Cause-specific mortality differences across socioeconomic position of municipalities in Japan, 1973-1977 and 1993-1998: increased importance of injury and suicide in inequality for ages under 75. *Int J Epidemiol* 2005; 34: 100-109.

6. Fujino Y, Tamakoshi A, Iso H et al. A nationwide cohort study of educational background and major causes of death among the elderly population in Japan. *Prev Med* 2005; 40: 444-451.

7. Honjo K, Iso H, Fukuda Y et al. Influence of municipal- and individual-level socioeconomic conditions on mortality in Japan. *Int J Behav Med* 2014; 21: 737-749.

8. Honjo K, Iso H, Nakaya T et al. Impact of neighborhood socioeconomic conditions on the risk of stroke in Japan. *J Epidemiol* 2015; 25: 254-260.

9. Nakaya T, Honjo K, Hanibuchi T et al. Associations of all-cause mortality with census-based neighbourhood deprivation and population density in Japan: a multilevel survival analysis. *PLoS One* 2014; 9: e97802.

10. Ito Y, Nakaya T, Nakayama T et al. Socioeconomic inequalities in cancer survival: a population-based study of adult patients diagnosed in Osaka, Japan, during

the period 1993-2004. *Acta Oncologica* 2014; [in press].

11. Woods LM, Rachet B, Coleman MP. Choice of geographic unit influences

socioeconomic inequalities in breast cancer survival. *British Journal of Cancer* 2005; 92: 1279-1282.

表 1 . 使用した各種データの対応年

死亡:人口動態統計	人口:国勢調査年	ADI:国勢調査年
1985	1985	1990
1986	1986	
1987	1987	
1988	1988	
1989	1989	
1990	1990	
1991	1991	
1992	1992	
1993	1993	1995
1994	1994	
1995	1995	
1996	1996	
1997	1997	
1998	1998	2000
1999	1999	
2000	2000	
2001	2001	
2002	2002	
2003	2003	2005
2004	2004	
2005	2005	
2006	2006	
2007	2007	2010
2008	2008	
2009	2009	
2010	2010	
2011	2011	
2012	2012	
2013	2013	
2014	2014	

グレーは線形補間

表 2 . 全死亡・主要死因別年齢調整死亡率（0-84 歳）における絶対的・相対的格差指標の
トレンド（5 年単位）

Cause of death	Period	Male						Female					
		SII	95%CI	RII	95%CI	SII	95%CI	RII	95%CI				
All causes	1985-1989	98.6	95.8	101.4	0.149	0.145	0.154	25.6	23.7	27.4	0.071	0.066	0.076
	1990-1994	99.3	96.8	101.8	0.161	0.157	0.165	25.6	24.0	27.2	0.081	0.076	0.086
	1995-1999	99.2	97.0	101.4	0.174	0.170	0.178	24.2	22.7	25.6	0.086	0.081	0.092
	2000-2004	102.5	100.5	104.5	0.201	0.197	0.205	24.9	23.6	26.2	0.103	0.098	0.109
	2005-2009	96.3	94.5	98.1	0.203	0.199	0.207	24.8	23.6	25.9	0.112	0.107	0.118
	2010-2014	87.0	85.3	88.7	0.201	0.198	0.205	30.2	29.1	31.3	0.148	0.142	0.153
Cancer	1995-1999	24.4	23.0	25.7	0.119	0.113	0.126	4.4	3.6	5.3	0.045	0.037	0.055
	2000-2004	24.7	23.5	25.9	0.132	0.125	0.138	4.9	4.1	5.7	0.053	0.045	0.063
	2005-2009	20.9	19.8	21.9	0.120	0.114	0.126	3.6	2.9	4.4	0.042	0.034	0.052
	2010-2014	19.2	18.2	20.2	0.123	0.117	0.129	4.3	3.6	5.0	0.053	0.044	0.062
Heart Diseases	1995-1999	9.6	8.8	10.4	0.126	0.116	0.138	3.8	3.3	4.3	0.096	0.084	0.110
	2000-2004	11.3	10.6	12.0	0.168	0.157	0.179	3.5	3.0	3.9	0.107	0.095	0.122
	2005-2009	12.4	11.7	13.1	0.196	0.185	0.206	4.2	3.8	4.6	0.149	0.136	0.163
	2010-2014	10.8	10.2	11.4	0.189	0.178	0.200	4.2	3.8	4.5	0.170	0.156	0.185
Cerebrovascular Disease	1995-1999	8.2	7.4	9.0	0.119	0.108	0.131	1.8	1.3	2.3	0.046	0.035	0.061
	2000-2004	8.8	8.2	9.4	0.165	0.153	0.177	2.5	2.1	2.9	0.088	0.075	0.103
	2005-2009	9.0	8.5	9.5	0.203	0.191	0.216	3.4	3.0	3.7	0.153	0.138	0.169
	2010-2014	8.1	7.6	8.5	0.227	0.214	0.241	2.8	2.5	3.1	0.167	0.150	0.186
Pneumonia	1995-1999	5.6	5.1	6.2	0.154	0.139	0.170	1.0	0.7	1.3	0.068	0.050	0.092
	2000-2004	5.5	5.0	6.0	0.174	0.159	0.190	1.2	0.9	1.4	0.097	0.077	0.121
	2005-2009	6.2	5.8	6.7	0.205	0.191	0.219	1.3	1.1	1.6	0.125	0.106	0.148
	2010-2014	6.5	6.1	6.9	0.239	0.226	0.253	2.0	1.8	2.2	0.217	0.196	0.240
Accidents	1995-1999	7.0	6.6	7.4	0.360	0.338	0.383	1.9	1.7	2.2	0.262	0.230	0.298
	2000-2004	6.8	6.4	7.2	0.432	0.408	0.456	1.8	1.6	2.0	0.333	0.296	0.375
	2005-2009	6.3	6.0	6.6	0.527	0.501	0.555	1.6	1.5	1.8	0.393	0.352	0.440
	2010-2014	9.8	9.5	10.2	0.863	0.836	0.892	5.6	5.3	5.8	1.043	1.005	1.083
Suicide	1995-1999	9.5	9.0	10.0	0.370	0.351	0.390	1.0	0.7	1.3	0.101	0.075	0.136
	2000-2004	12.8	12.3	13.4	0.418	0.401	0.436	1.0	0.7	1.3	0.098	0.072	0.133
	2005-2009	12.1	11.6	12.7	0.395	0.377	0.413	1.0	0.7	1.4	0.093	0.068	0.129
	2010-2014	7.9	7.3	8.4	0.291	0.272	0.311	0.5	0.2	0.8	0.047	0.025	0.092

表 3-1. 全死亡・主要死因別の年齢調整死亡率における絶対格差指標（SII）の年次推移：Joinpoint regression model による分析結果

SII	Sex	Trend 1	APC	95% CI	Trend 2	APC	95% CI	Trend 3	APC	95% CI	Trend 4	APC	95% CI		
Allcause of death	Male	1985 2011	0.0	-0.4 0.3	2011 2014	-10.5	-17.5	-2.9 *							
	Female	1985 2014	0.8	-0.2 1.8											
Allcause of death (omit year of 2011)	Male	1985 2006	0.2	-0.1 0.5	2006 2014	-3.9	-5.0	-2.8 *							
	Female	1985 2014	-0.3	-0.7 0.1											
Cancer	Male	1995 2014	-1.7	-2.3 -1.0 *											
	Female	1995 2014	-0.6	-2.3 1.0											
Heart Diseases	Male	1995 2009	2.3	1.1 3.6 *	2009 2014	-6.0	-10.5	-1.3 *							
	Female	1995 2014	0.9	-0.3 2.1											
Cerebrovascular Diseases	Male	1995 2011	0.7	0.0 1.5	2011 2014	-11.5	-18.4	-4.1 *							
	Female	1995 2014	0.6	-1.3 2.5											
Pneumonia	Male	1995 2014	1.1	0.0 2.3											
	Female	1995 1997	-29.4	-67.4 53.0	1997 2005	1.1	-5.0	7.5	2005 2012	10.3	5.2	15.6 *	2012 2014	-13.7	-29.8
Accidents	Male	1995 2014	0.2	-2.4 2.9											
	Female	1995 2014	1.8	-1.6 5.2											
Accidents (omit years of 1995 & 2011)	Male	1996 2009	-1.1	-2.4 0.1	2009 2014	-7.4	-12.9	-1.6 *							
	Female	1996 2014	-0.8	-2.4 0.9											
Suicide	Male	1995 2004	7.6	4.7 10.6 *	2004 2014	-8.1	-10.4	-5.7 *							
	Female	1995 2014	0.2	-3.8 4.4											

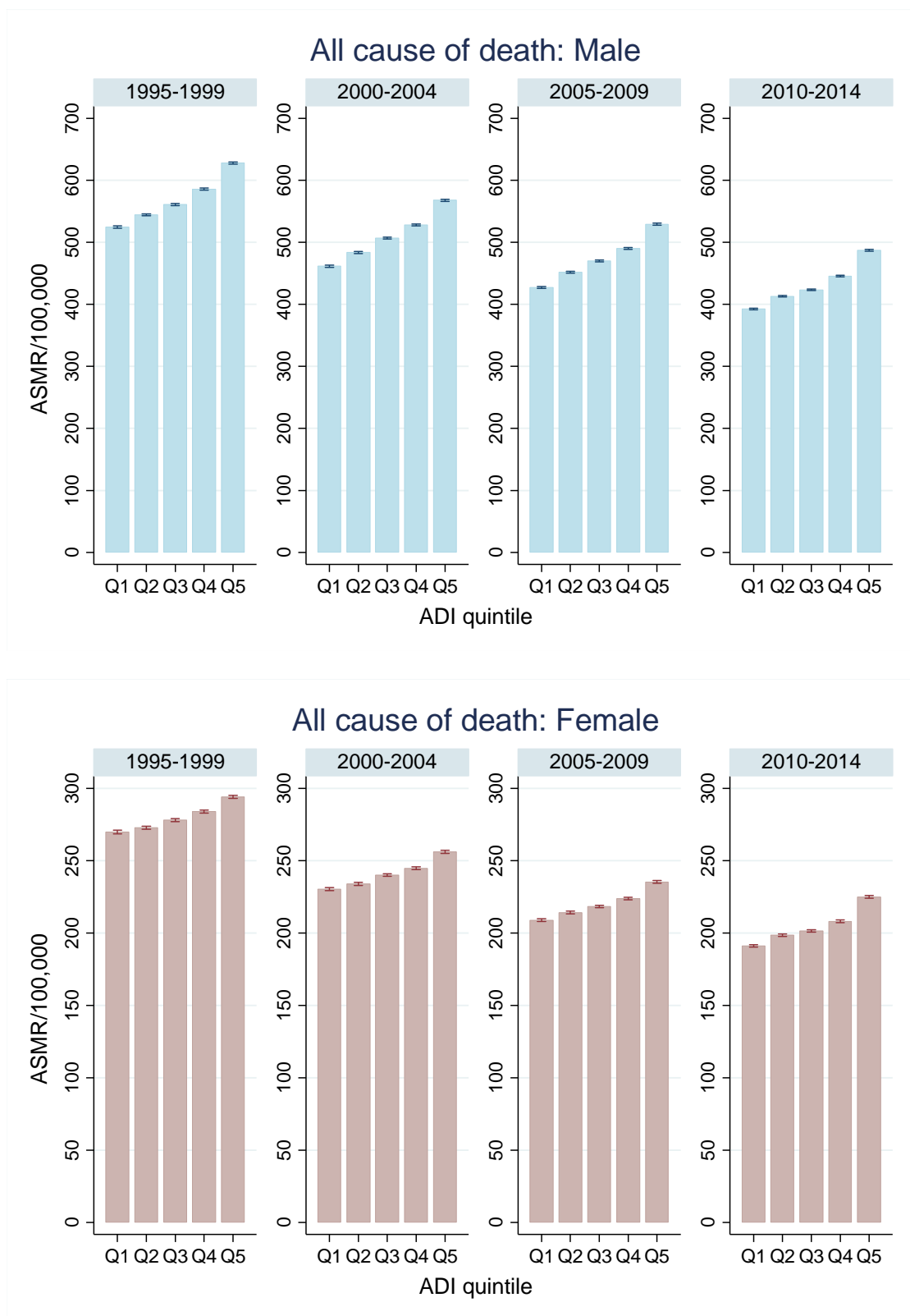


図 1-1. 市区町村別 ADI5 分位ごとの年齢調整死亡率（0-84 歳）の推移：全死亡，上・男性，下・女性

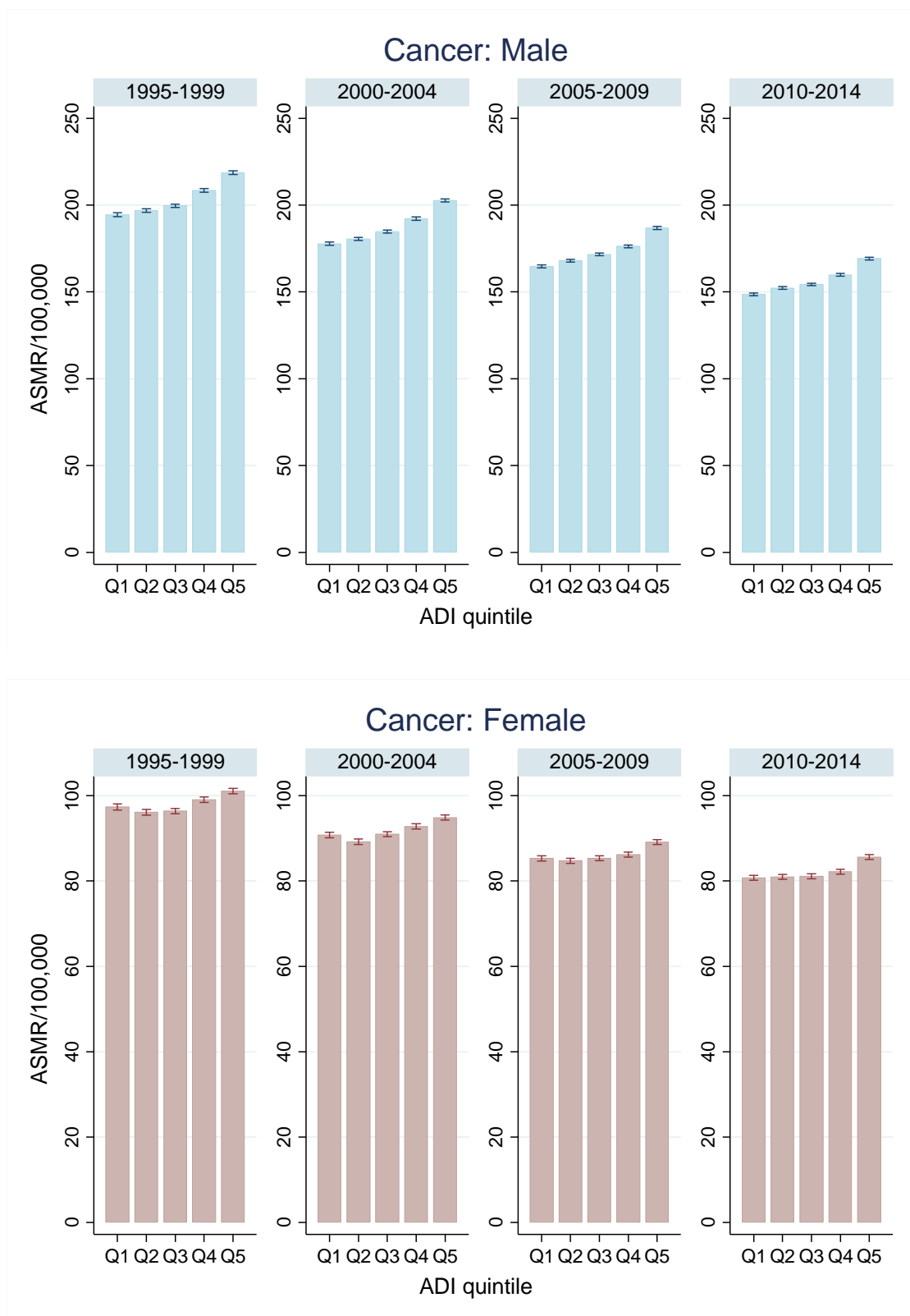


図 1-2. 市区町村別 ADI5 分位ごとの年齢調整死亡率（0-84 歳）の推移：悪性新生物，上・男性，下・女性

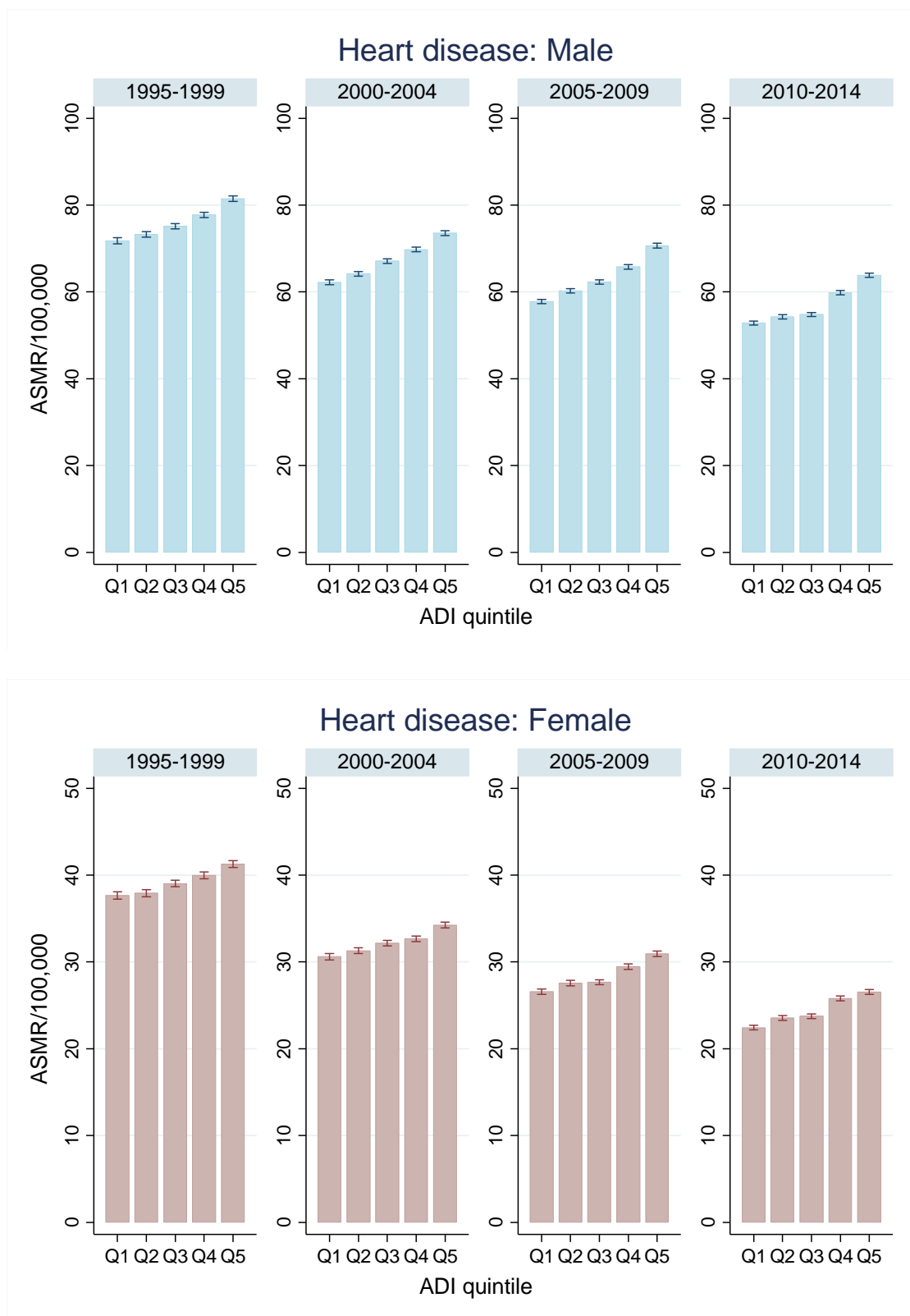


図1-3. 市区町村別 ADI5 分位ごとの年齢調整死亡率（0-84 歳）の推移：心疾患（高血圧性を除く），上・男性，下・女性

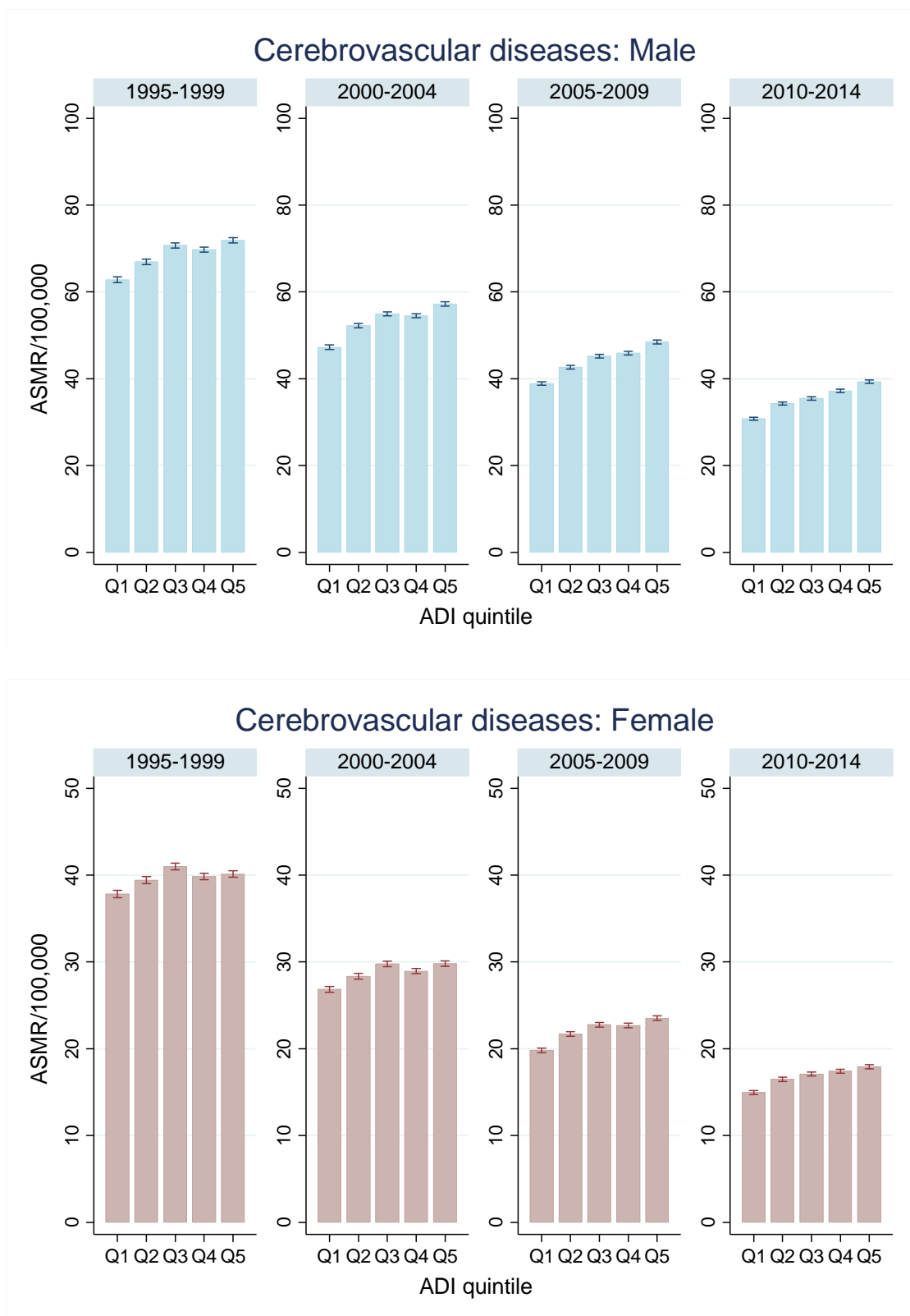


図1-4. 市区町村別 ADI5 分位ごとの年齢調整死亡率（0-84 歳）の推移：脳血管疾患，上・男性，下・女性

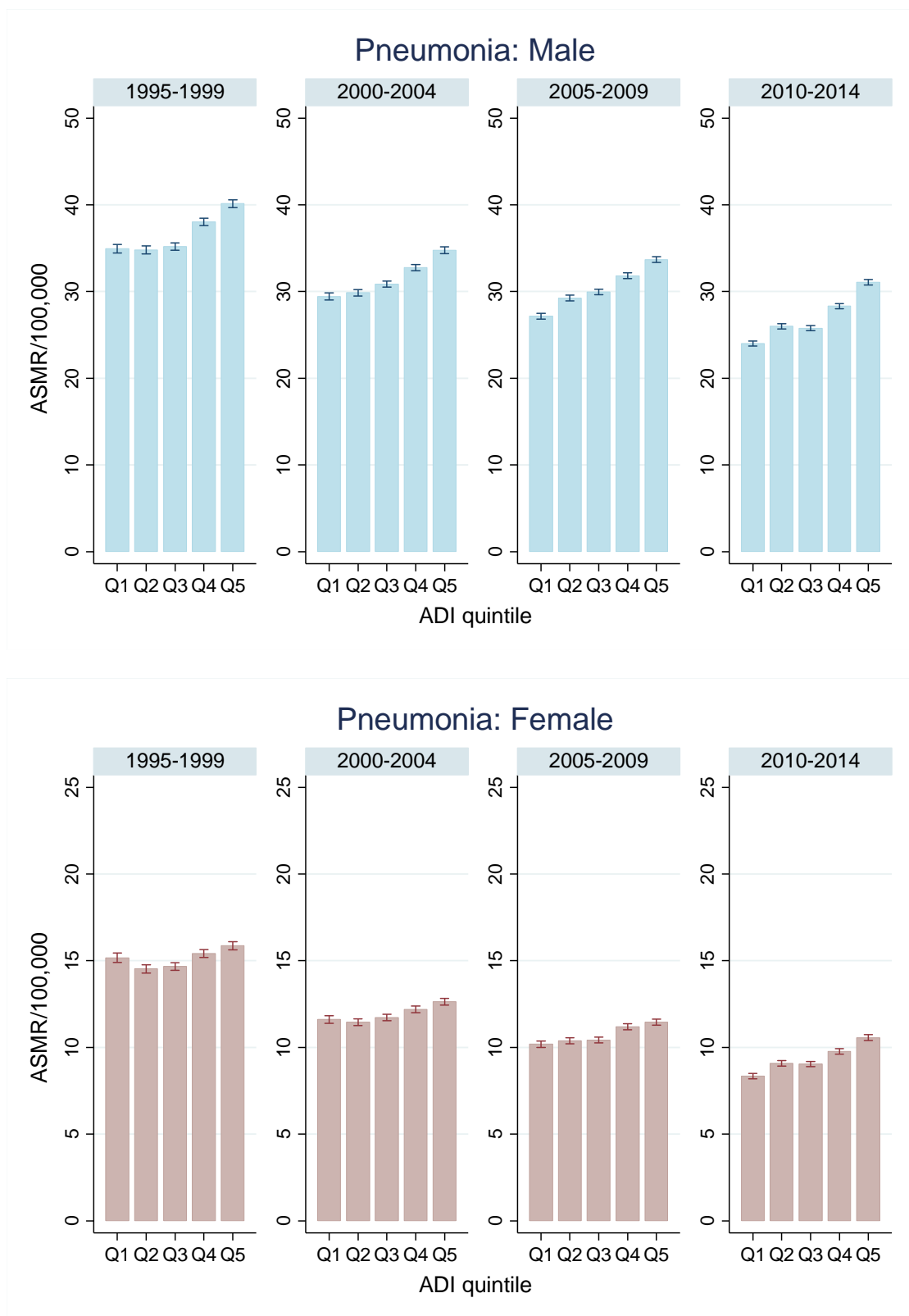


図 1-5. 市区町村別 ADI5 分位ごとの年齢調整死亡率(0-84 歳)の推移:肺炎, 上・男性, 下・女性

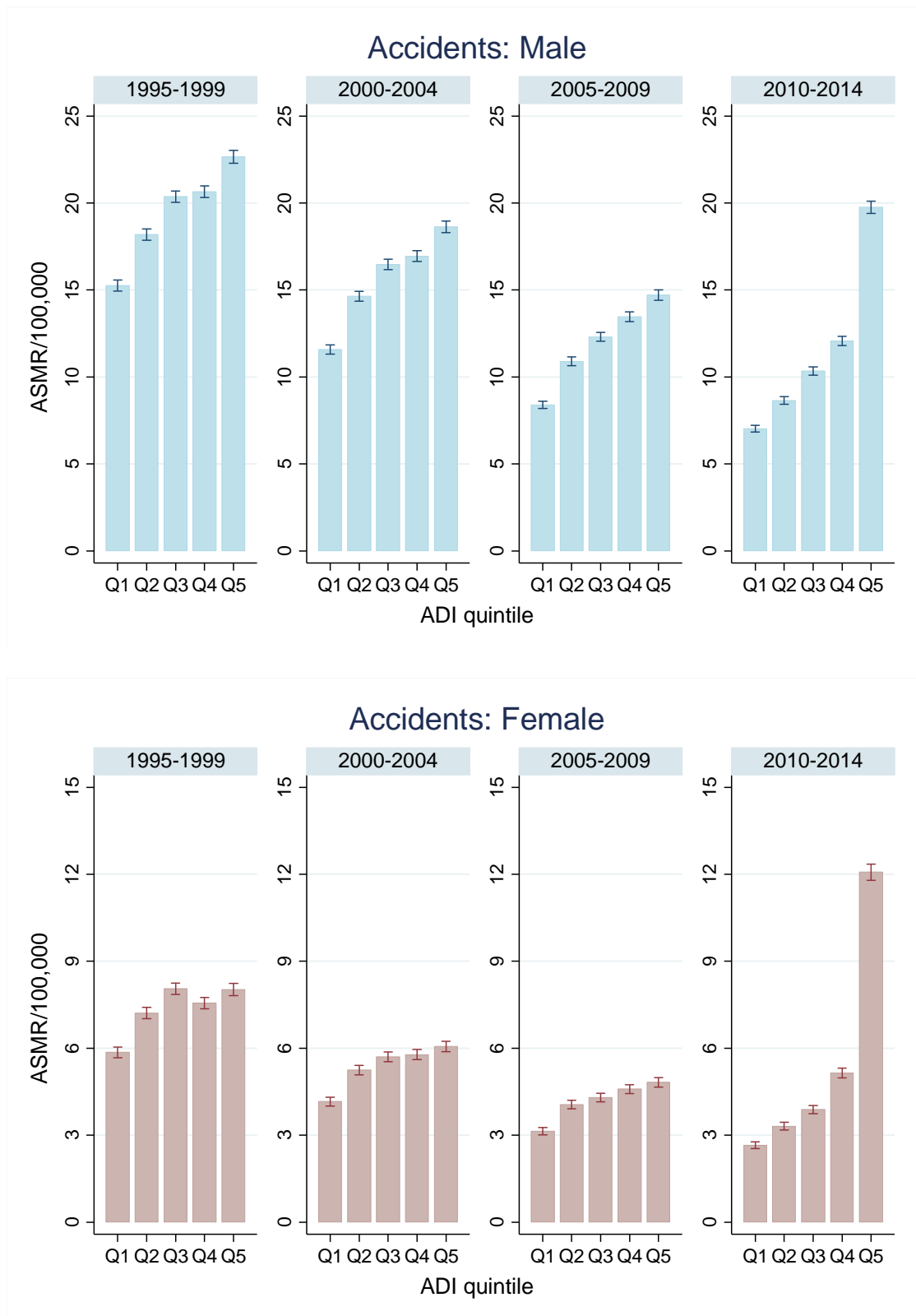


図 1-6. 市区町村別 ADI5 分位ごとの年齢調整死亡率（0-84 歳）の推移：不慮の事故，上・男性，下・女性

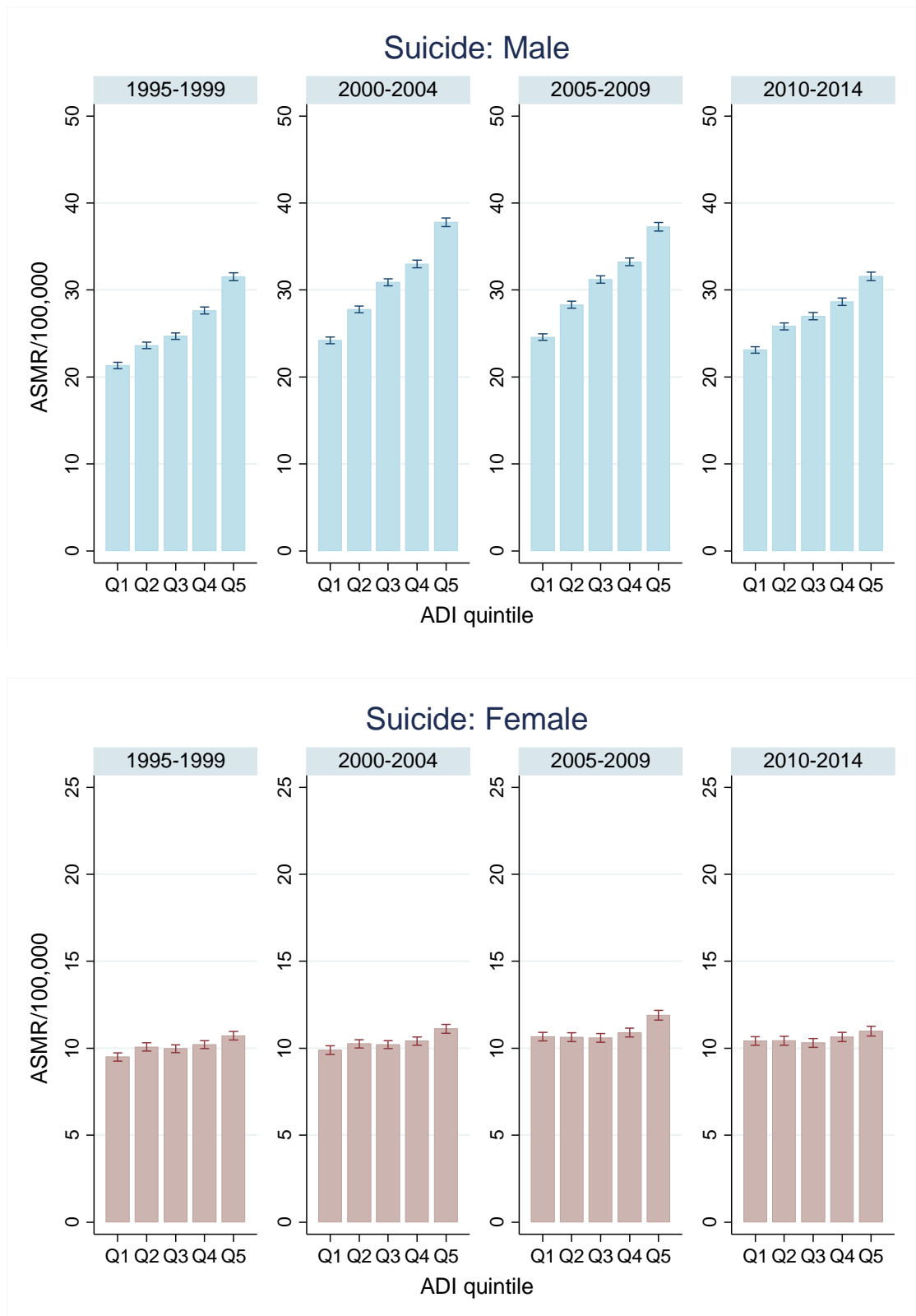


図 1-7. 市区町村別 ADI5 分位ごとの年齢調整死亡率(0-84 歳)の推移:自殺, 上・男性, 下・女性

Slope Index of Inequalities: All cause of death

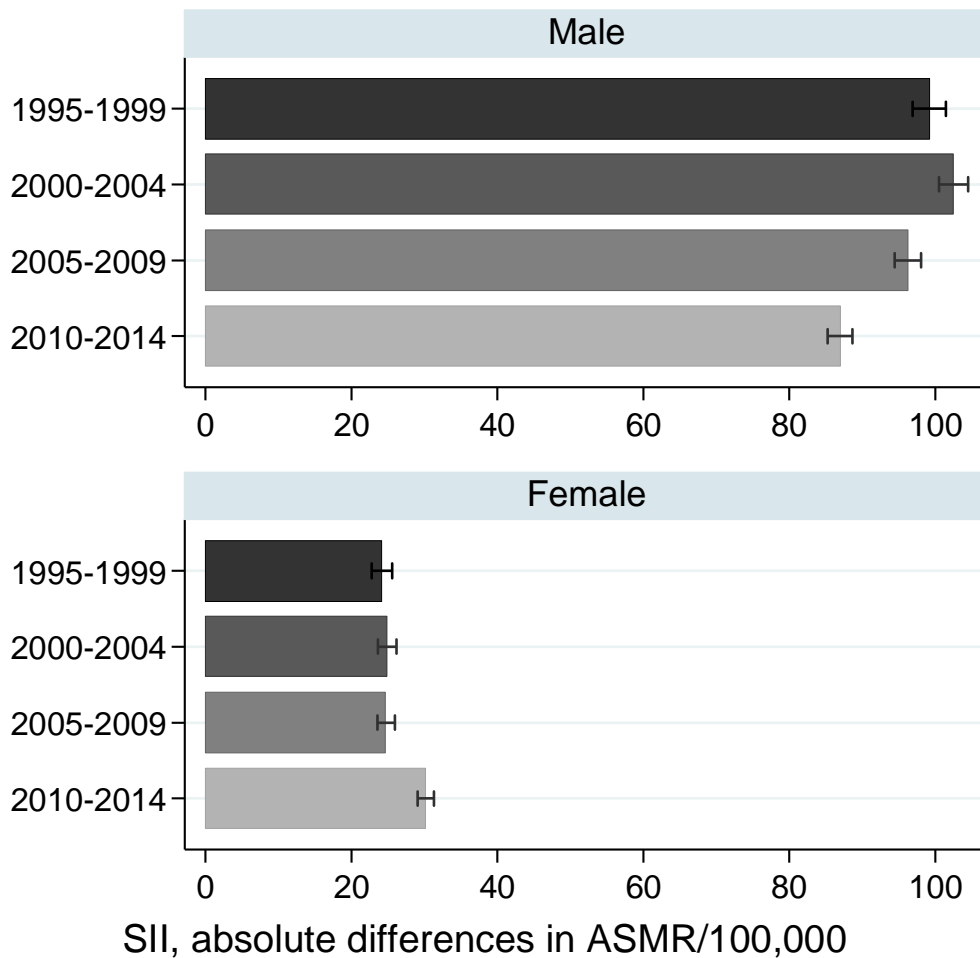


図 2-1 . 全死亡における年齢調整死亡率の格差の絶対指標 (Slope Index of Inequalities: SII): Q5 の年齢調整死亡率と Q1 の年齢調整死亡率の差

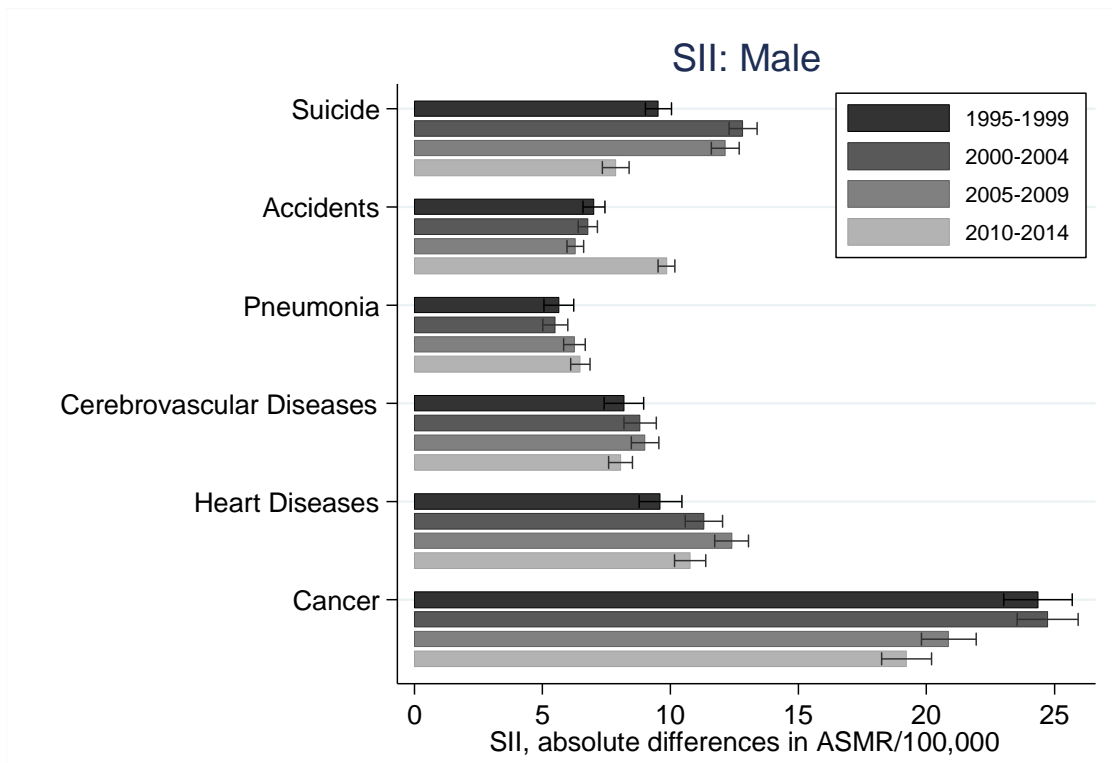


図 2-2 . 主要死因別年齢調整死亡率における格差の絶対指標 (Slope Index of Inequalities: SII): Q5 の年齢調整死亡率と Q1 の年齢調整死亡率の差, 上・男性, 下・女性

Relative Index of Inequalities: All cause of death

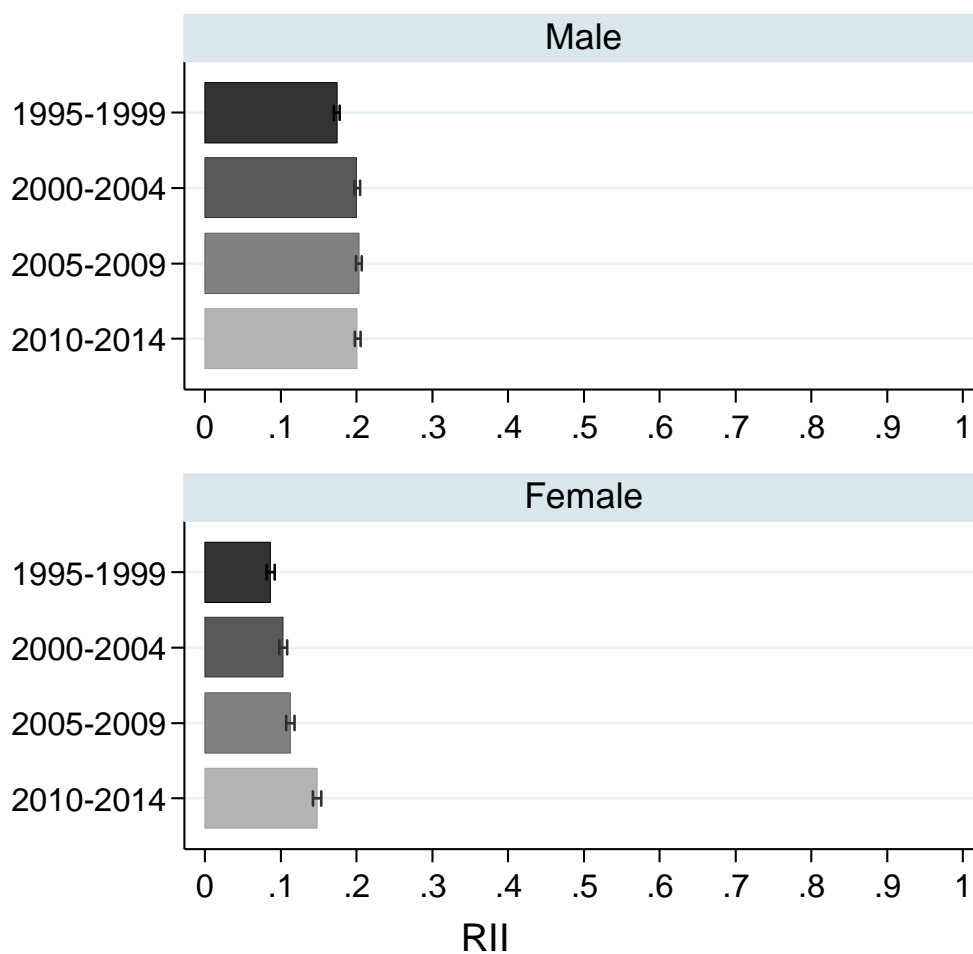


図 3-1. 全死亡における年齢調整死亡率の格差の相対指標 (Relative Index of Inequalities: RII): SII を Q1-Q5 全体の年齢調整死亡率で除したもの

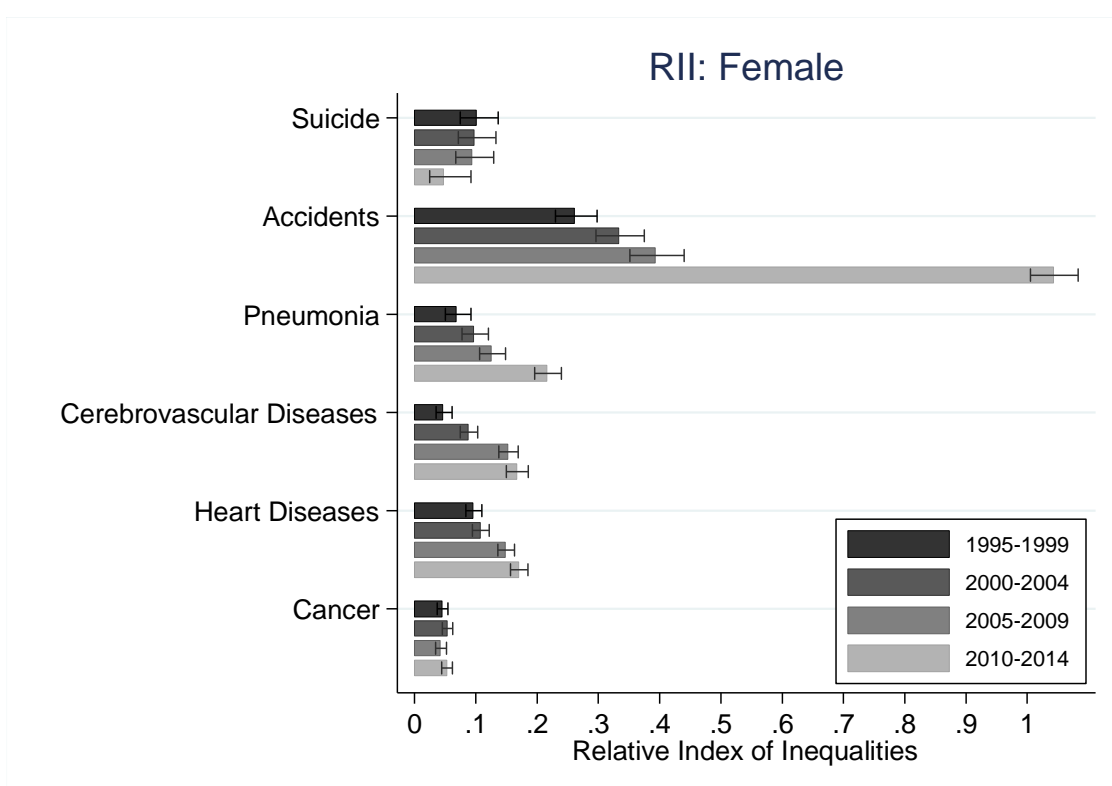
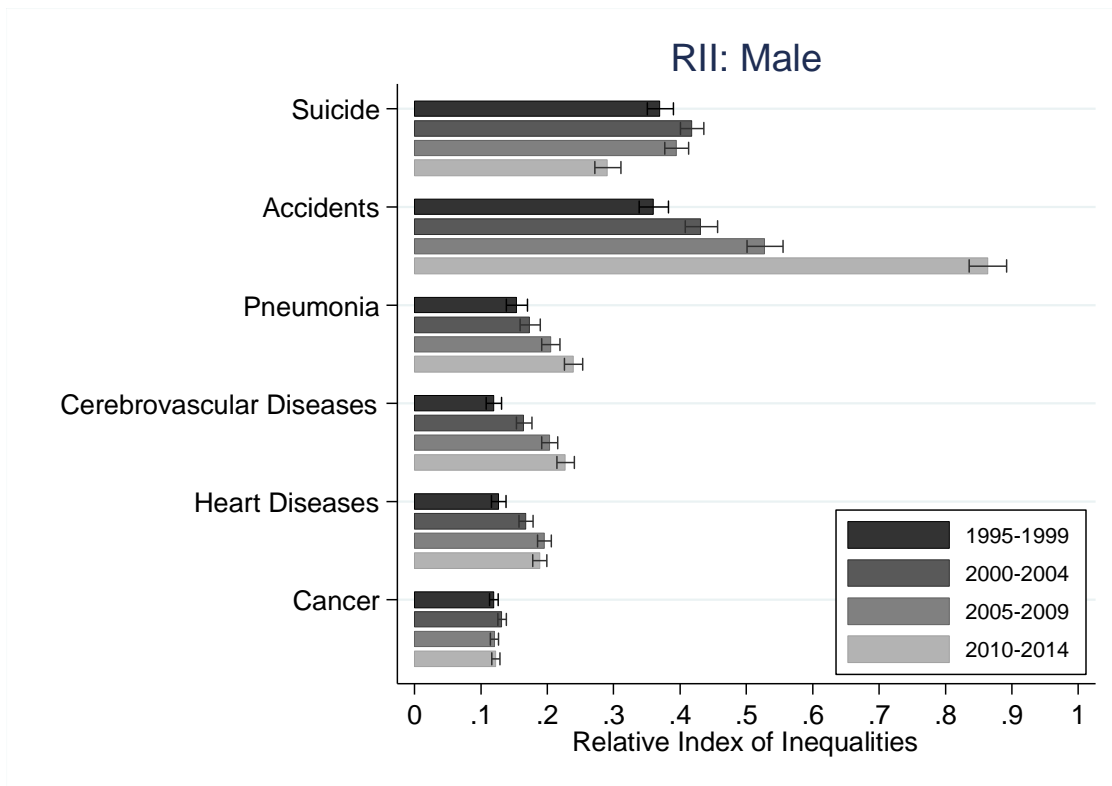


図 3-2. 主要死因別年齢調整死亡率の格差の相対指標 (Relative Index of Inequalities: RII): SII を Q1-Q5 全体の年齢調整死亡率で除したもの, 上・男性, 下・女性

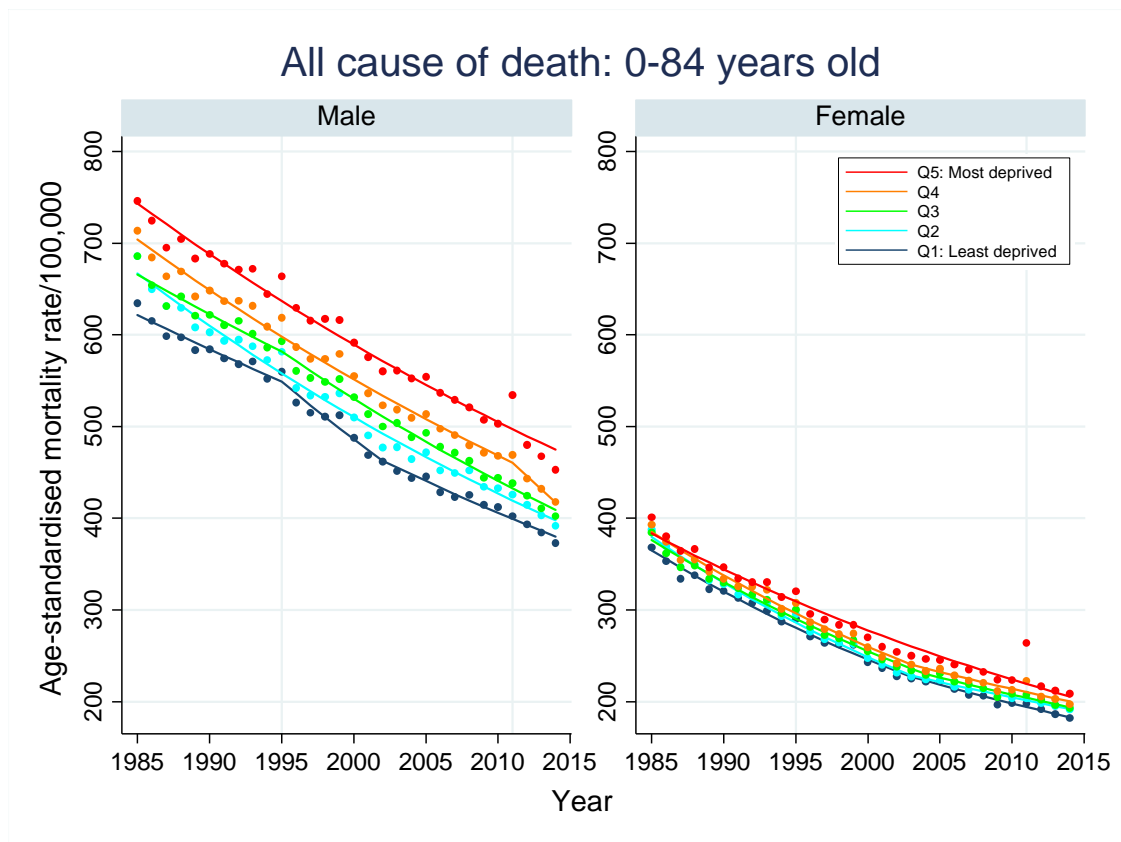


図 4-1 . 全死亡における ADI5 分位別年齢調整死亡率の推移：直線は Joinpoint regression model によるあてはめ

Cancer: 0-84 years old 1995-2014

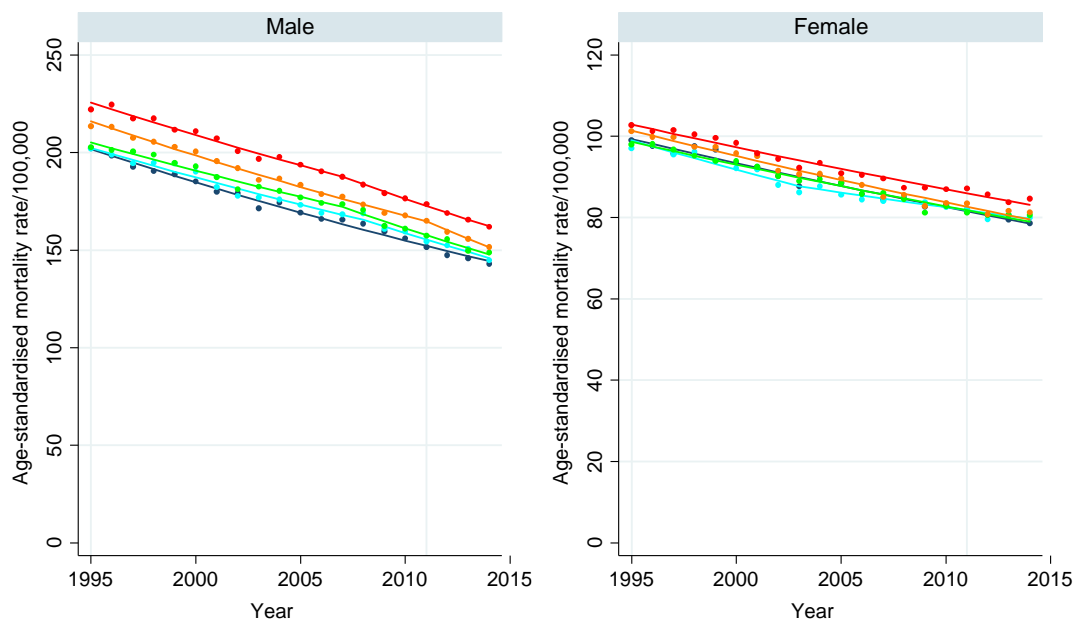


図 4-2 . ADI5 分位別年齢調整死亡率の推移：がん

Heart Diseases: 0-84 years old 1995-2014

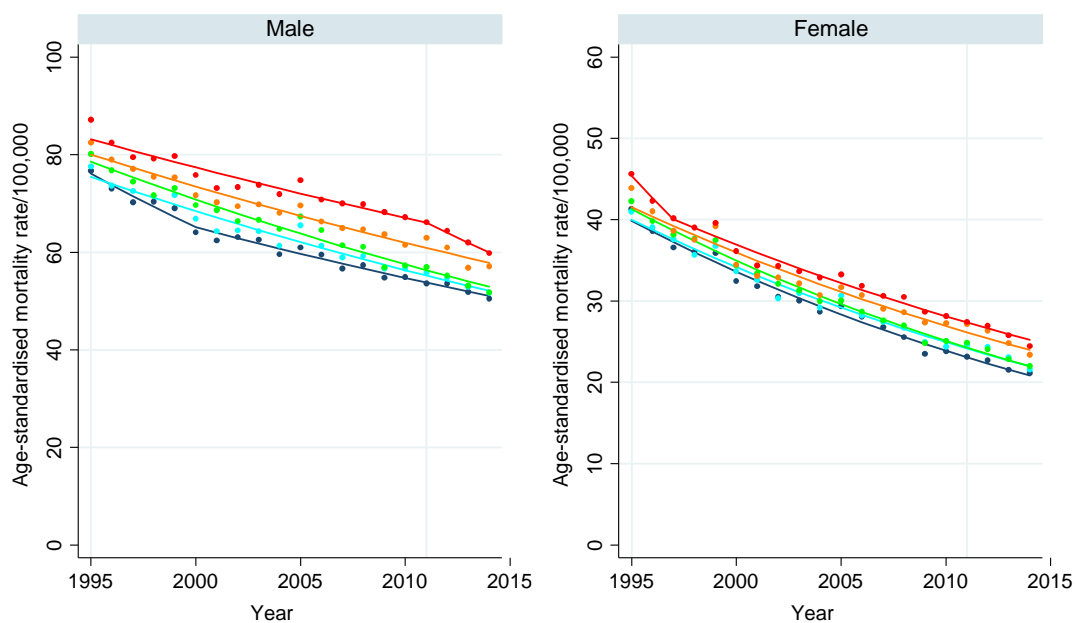


図 4-3 . ADI5 分位別年齢調整死亡率の推移：心疾患（高血圧性を除く）

Cerebrovascular Diseases: 0-84 years old 1995-2014

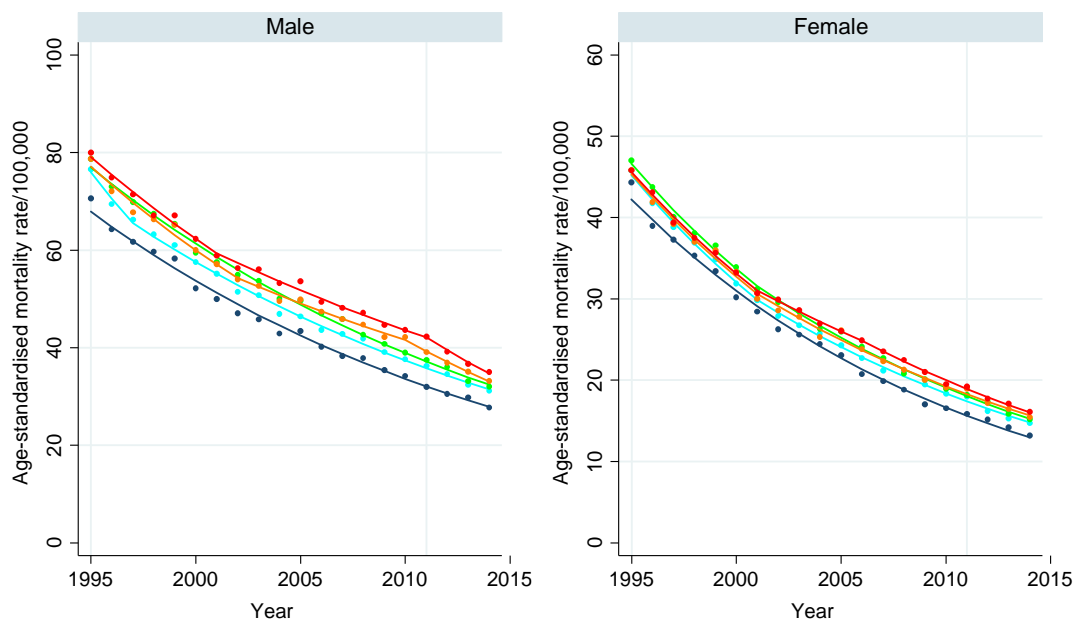


図 4-4 . ADI5 分位別年齢調整死亡率の推移：脳血管疾患

Pneumonia: 0-84 years old 1995-2014

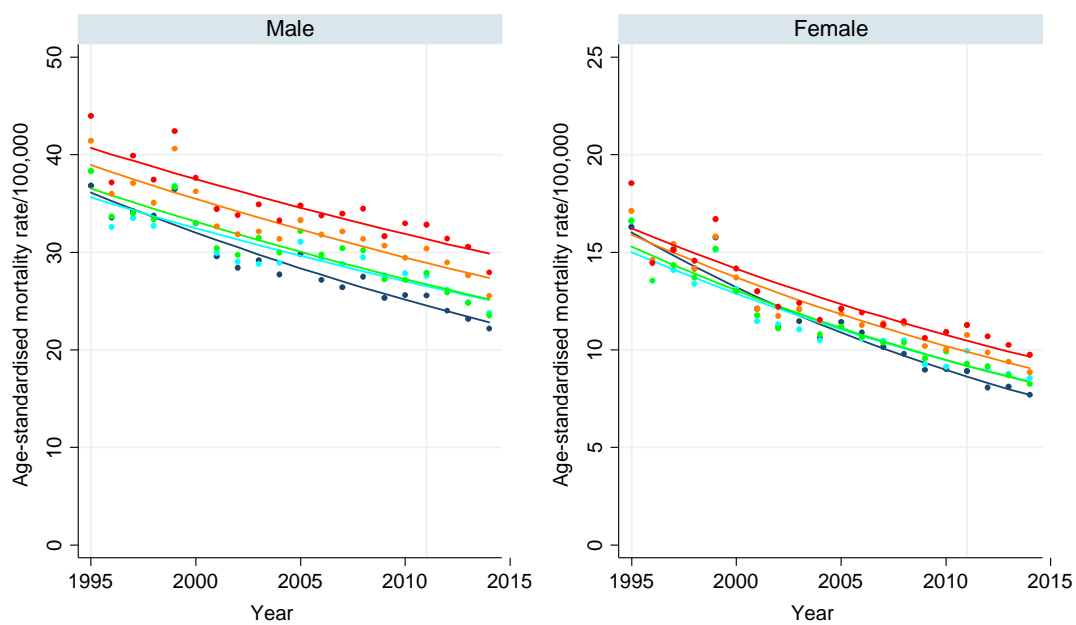


図 4-5 . ADI5 分位別年齢調整死亡率の推移：肺炎

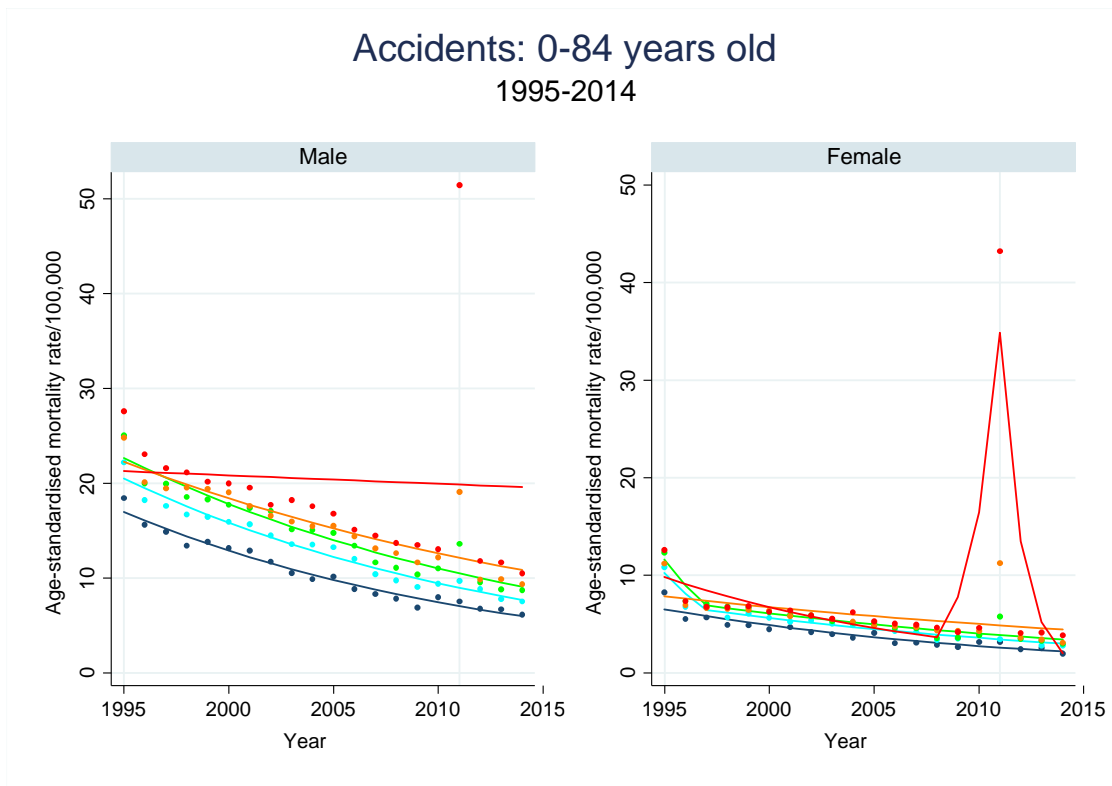


図 4-6 . ADI5 分位別年齢調整死亡率の推移：不慮の事故（大震災の年を含む）

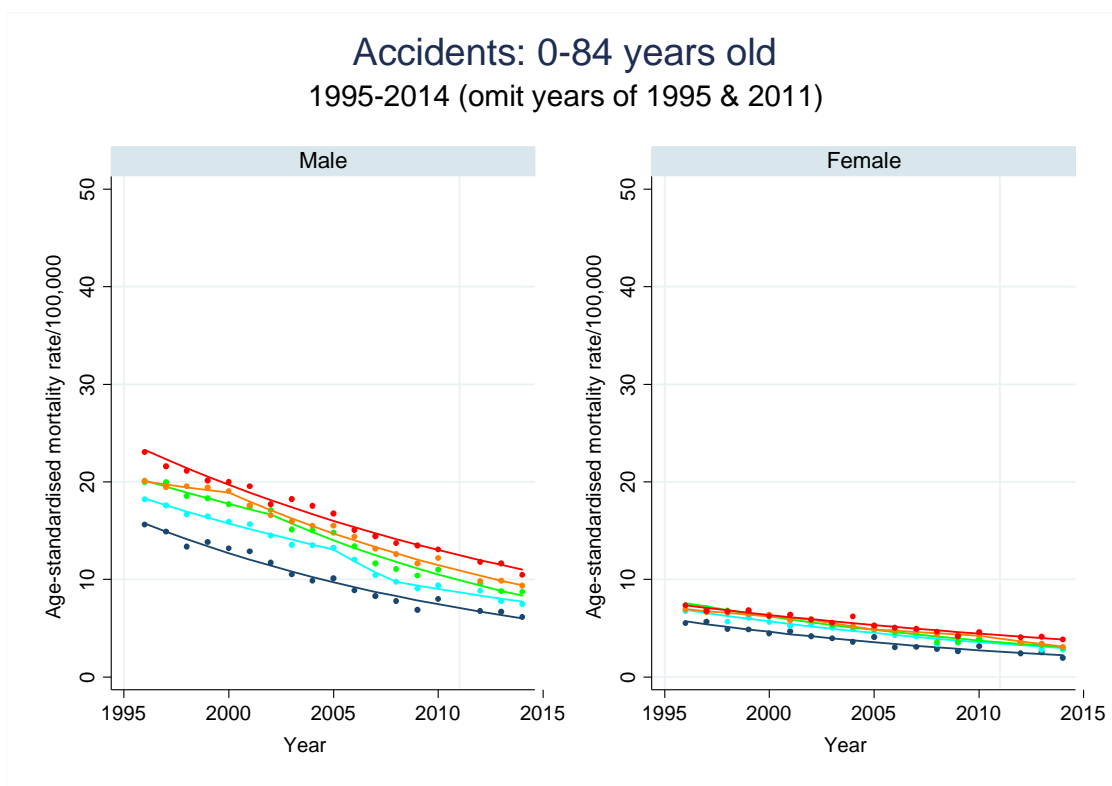


図 4-7 . ADI5 分位別年齢調整死亡率の推移：不慮の事故（大震災の年を含まない）

Suicide: 0-84 years old 1995-2014

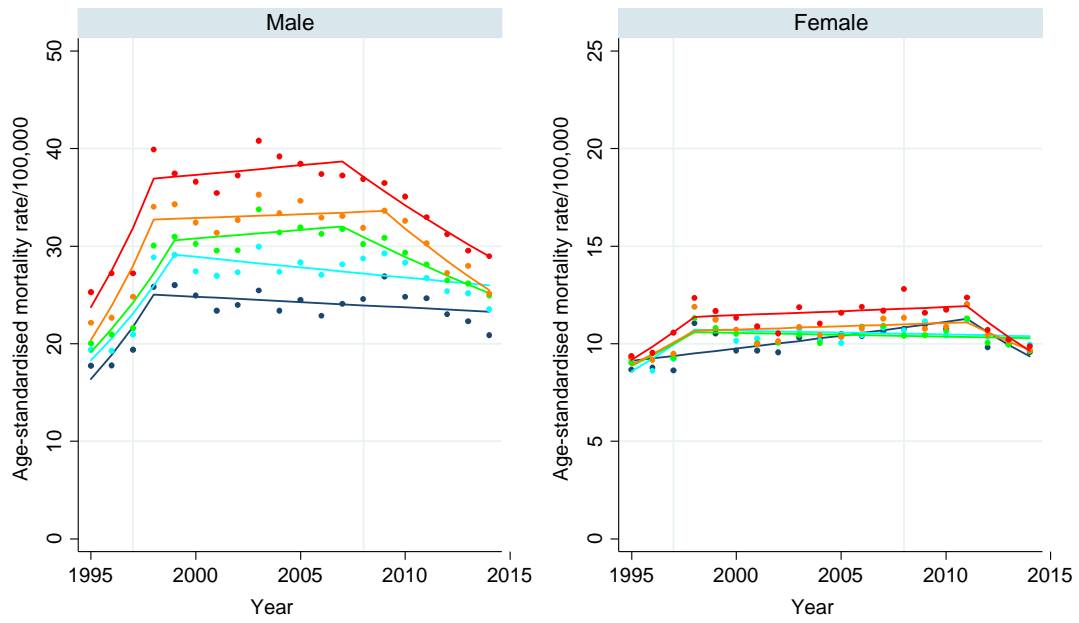


图 4-8 . ADI5 分位別年齢調整死亡率の推移：自殺

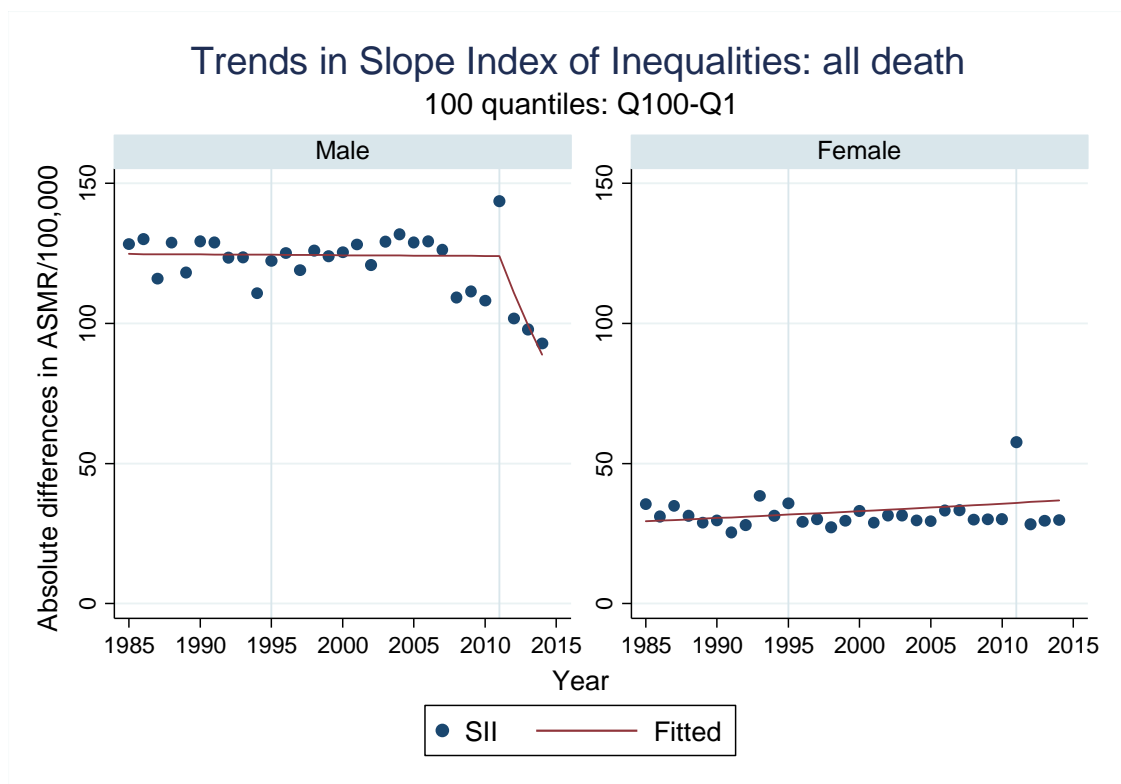


図 5-1-1A . ADI100 分位別年齢調整死亡率を用いた SII の推移：全死亡

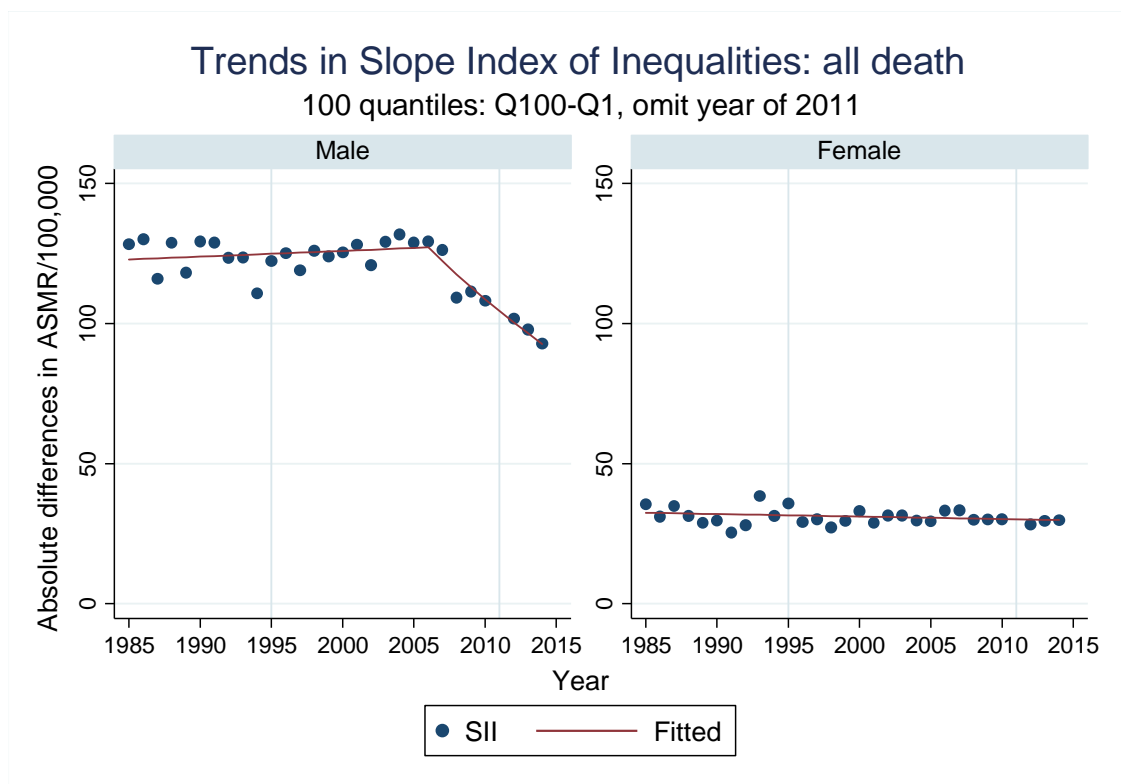


図 5-1-1B . ADI100 分位別年齢調整死亡率を用いた SII の推移：全死亡（大震災の年を含まない）

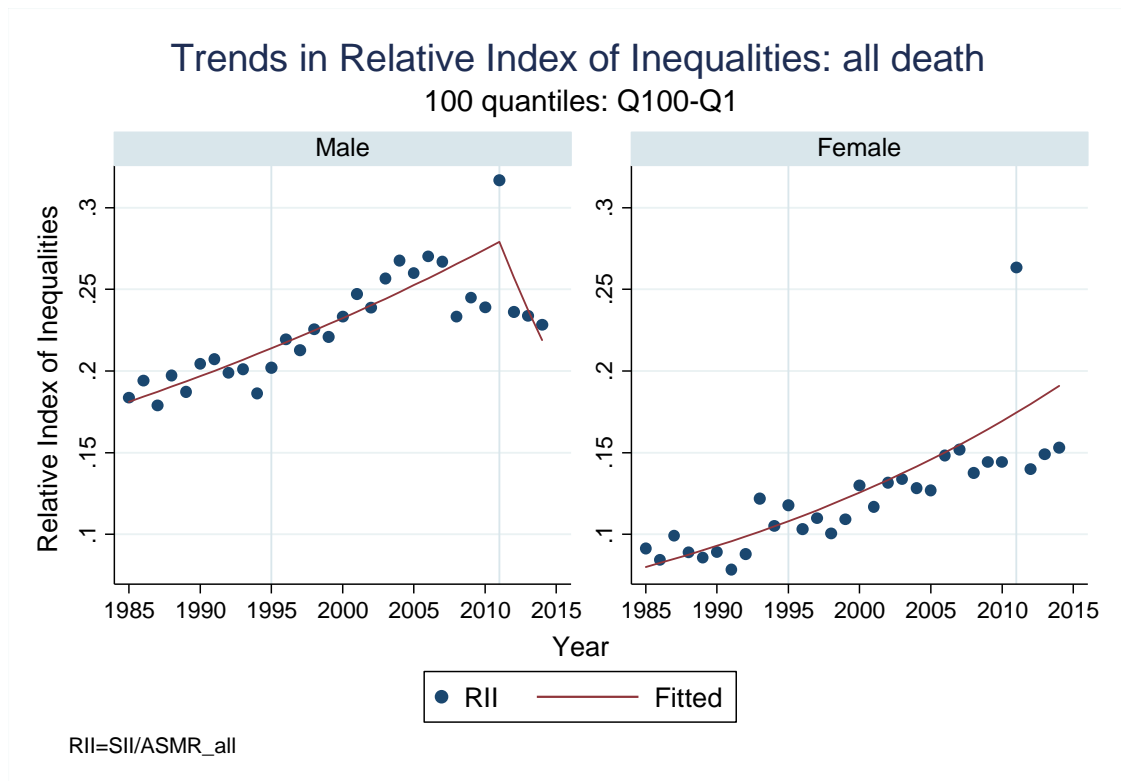


図 5-1-2A . ADI100 分位別年齢調整死亡率を用いた RII の推移：全死亡

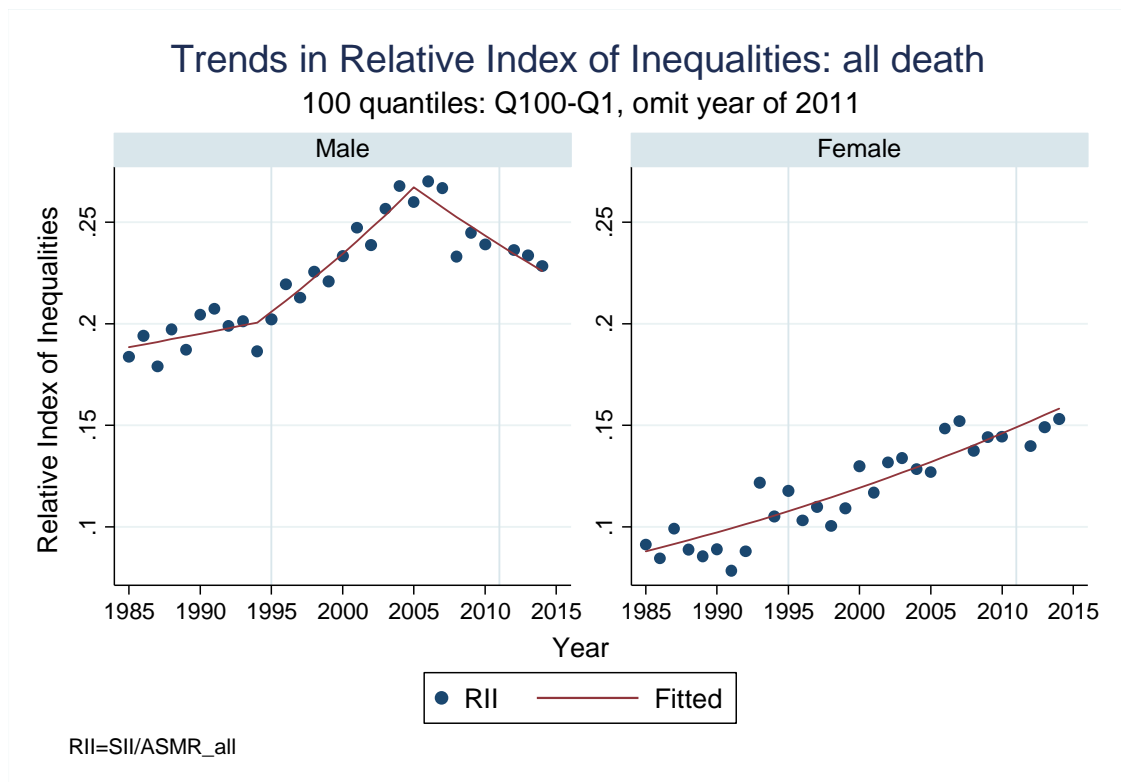


図 5-1-2B . ADI100 分位別年齢調整死亡率を用いた RII の推移：全死亡（大震災の年を含まない）

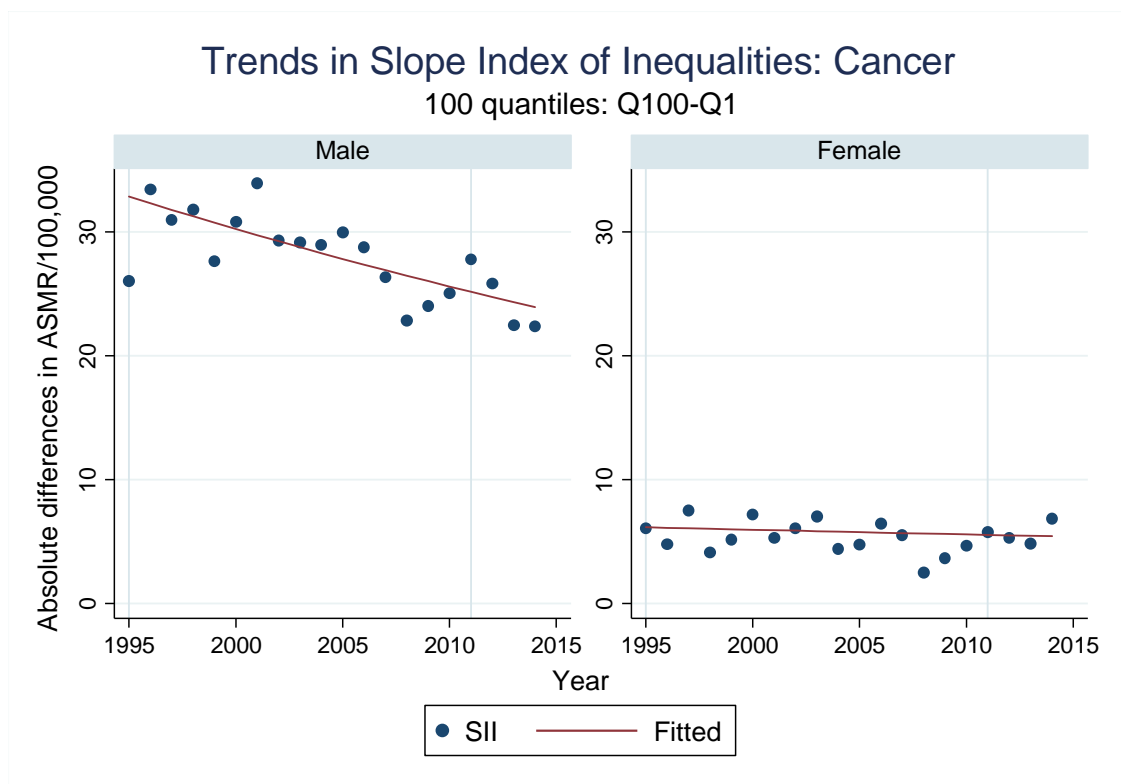


図 5-2-1 . ADI100 分位別年齢調整死亡率を用いた SII の推移 : がん

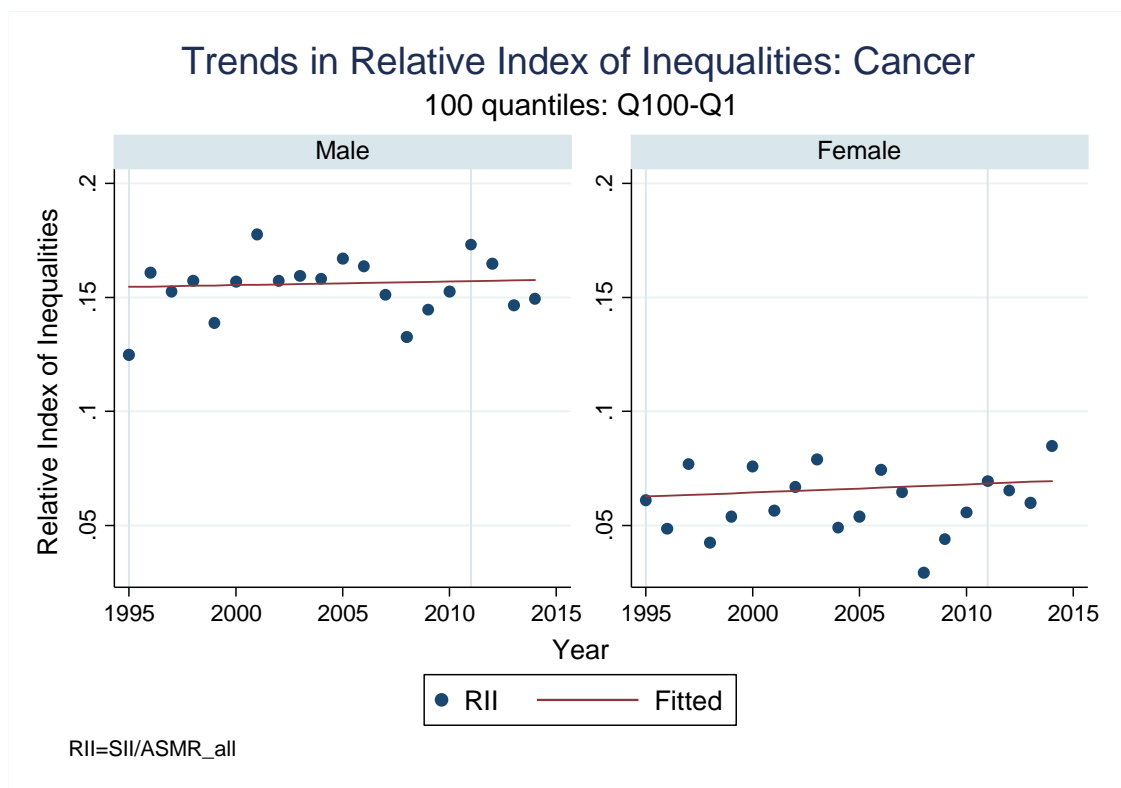


図 5-2-2 . ADI100 分位別年齢調整死亡率を用いた RII の推移 : がん

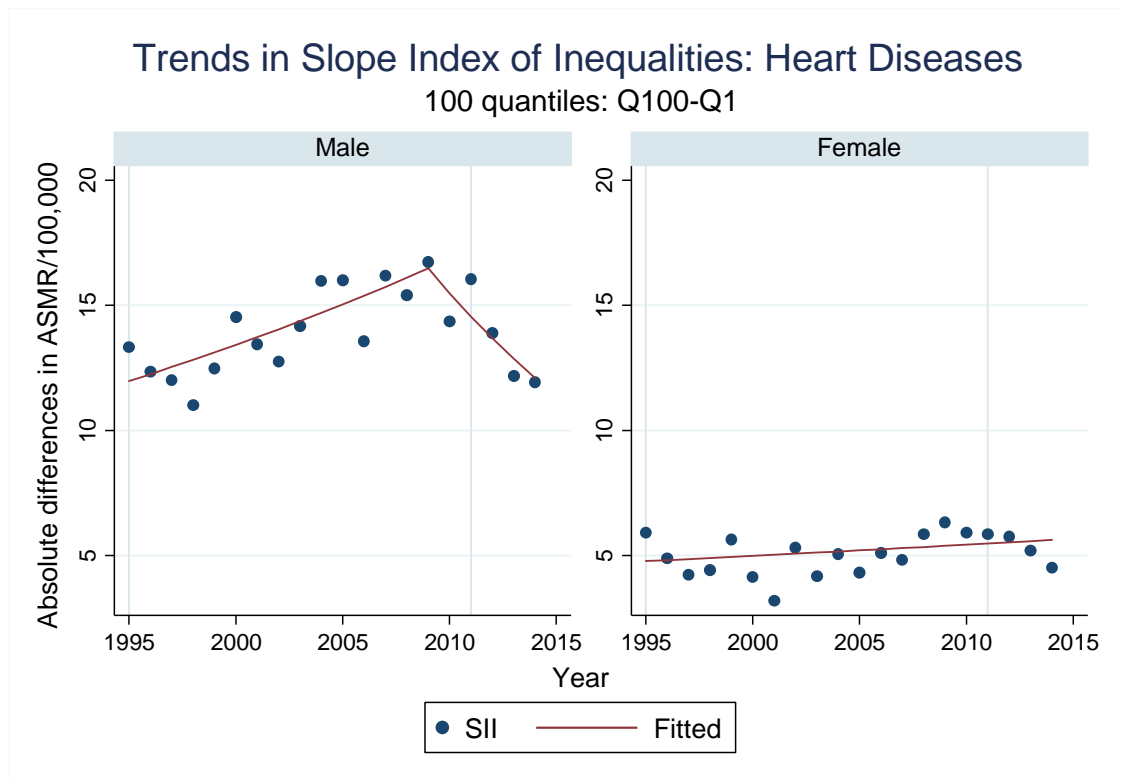


図 5-3-1 . ADI100 分位別年齢調整死亡率を用いた SII の推移：心疾患（高血圧性を除く）

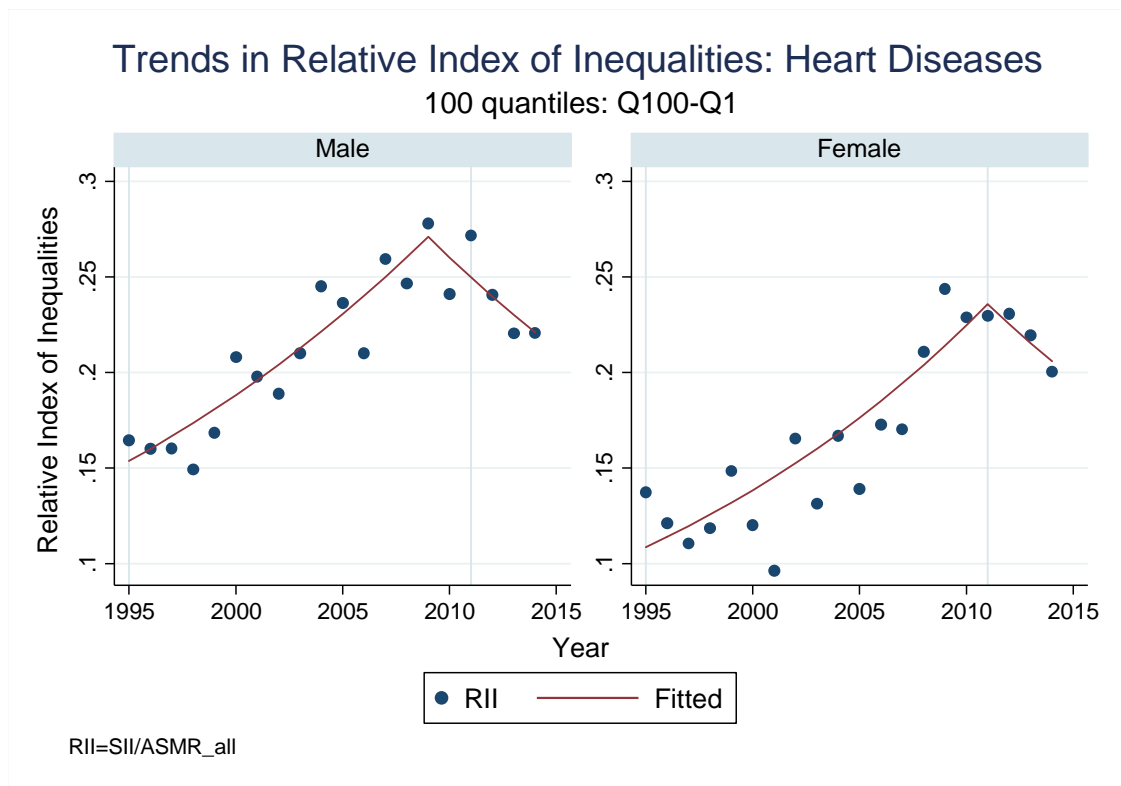


図 5-3-2 . ADI100 分位別年齢調整死亡率を用いた RII の推移：心疾患（高血圧性を除く）

Trends in Slope Index of Inequalities: Cerebrovascular Diseases

100 quantiles: Q100-Q1

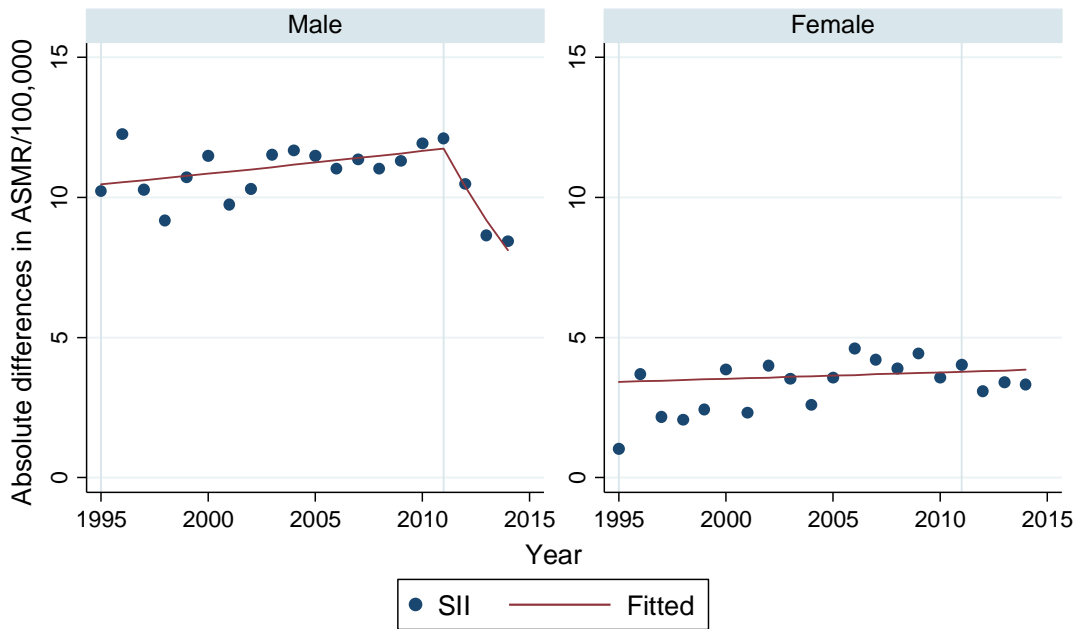
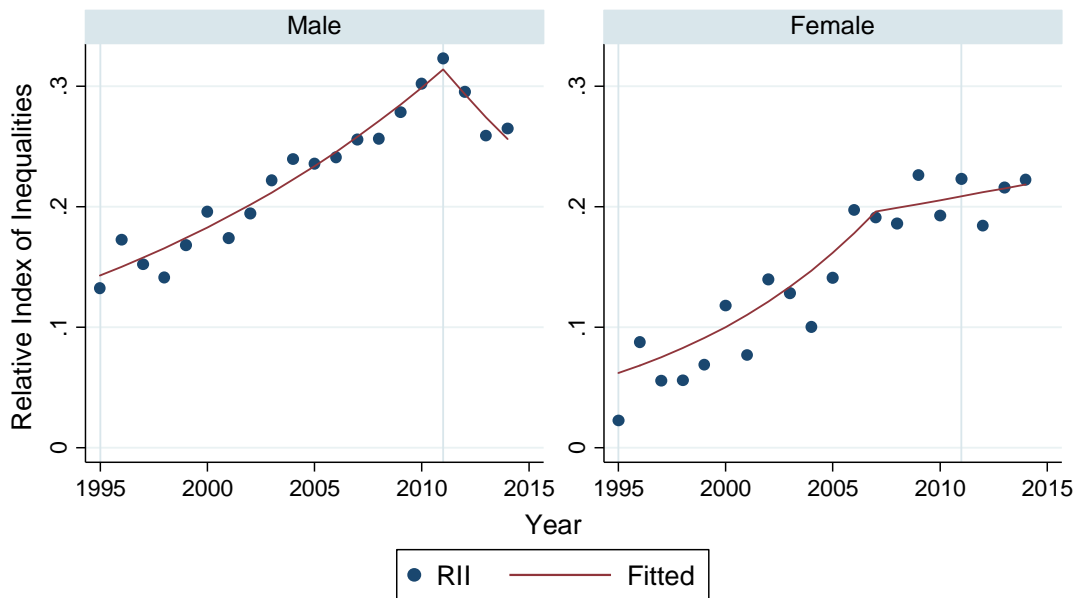


図 5-4-1 . ADI100 分位別年齢調整死亡率を用いた SII の推移：脳血管疾患

Trends in Relative Index of Inequalities: Cerebrovascular Diseases

100 quantiles: Q100-Q1



$RII = SII / ASMR_all$

図 5-4-2 . ADI100 分位別年齢調整死亡率を用いた RII の推移：脳血管疾患

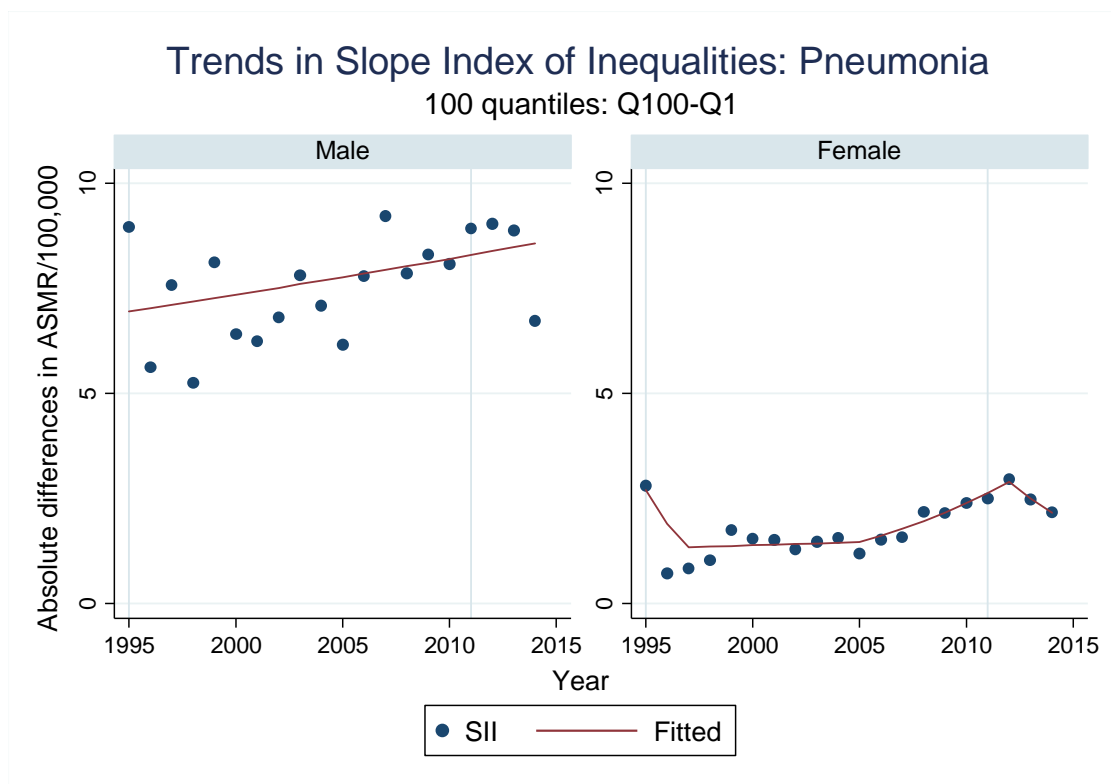


図 5-5-1 . ADI100 分位別年齢調整死亡率を用いた SII の推移 : 肺炎

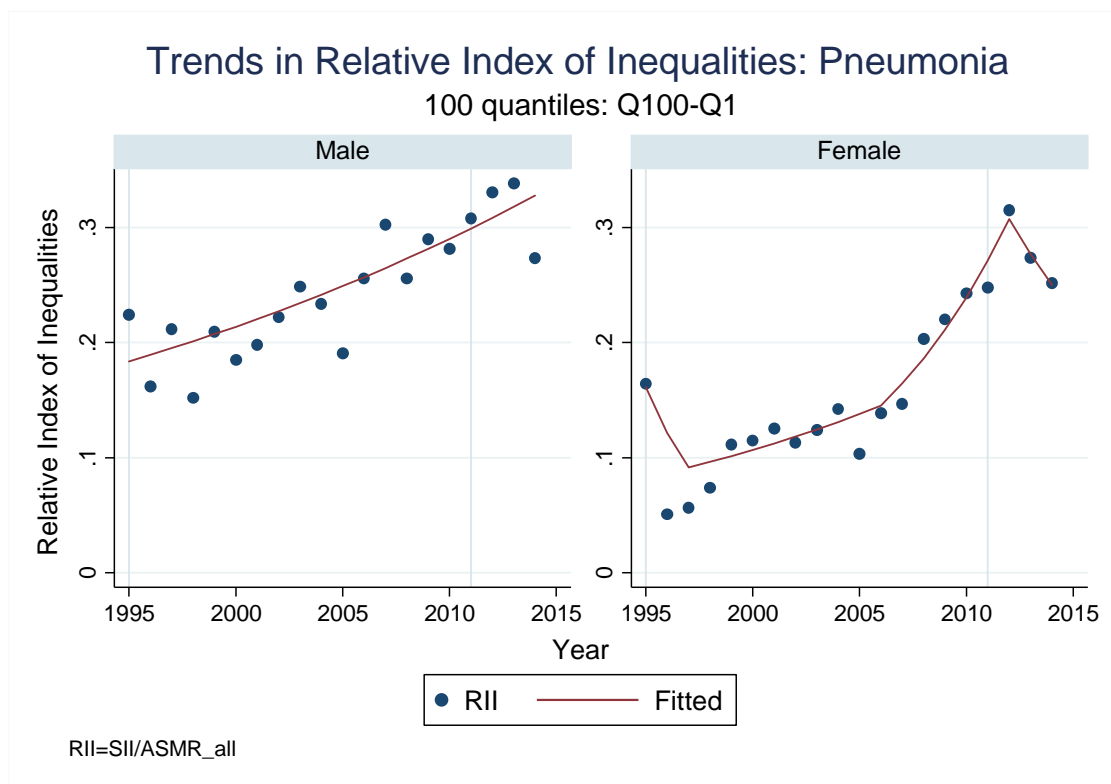


図 5-5-2 . ADI100 分位別年齢調整死亡率を用いた RII の推移 : 肺炎

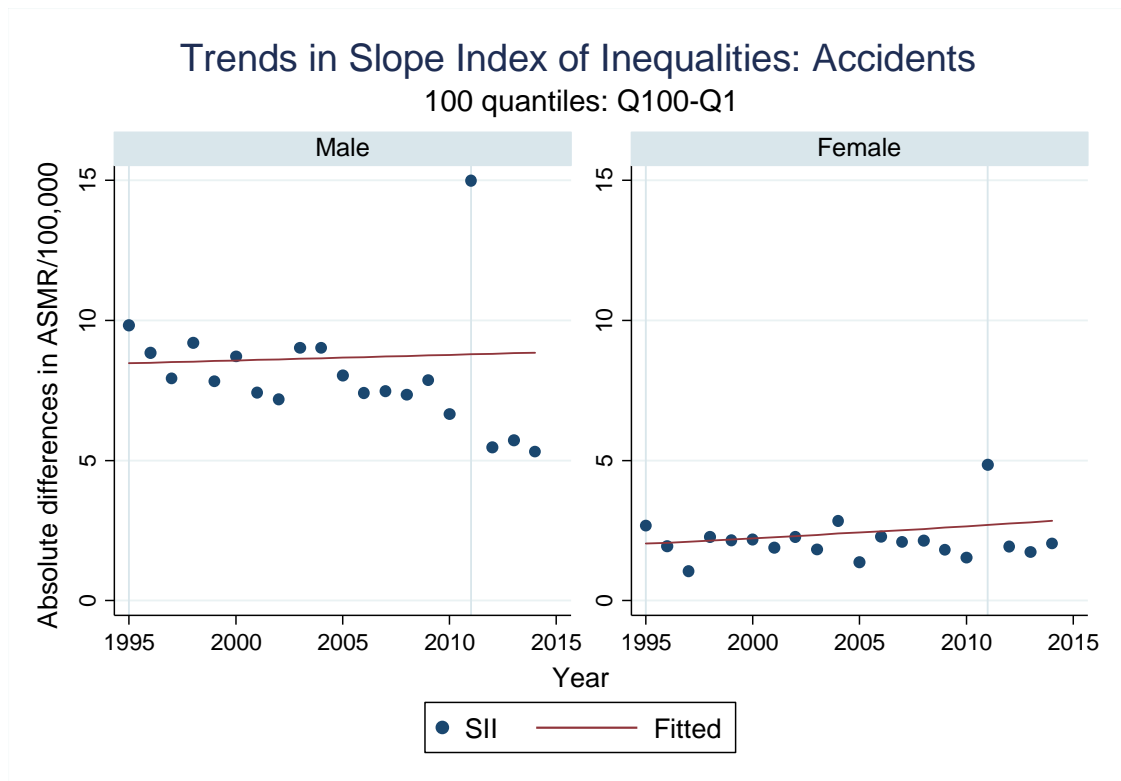


図 5-6-1A . ADI100 分位別年齢調整死亡率を用いた SII の推移：不慮の事故

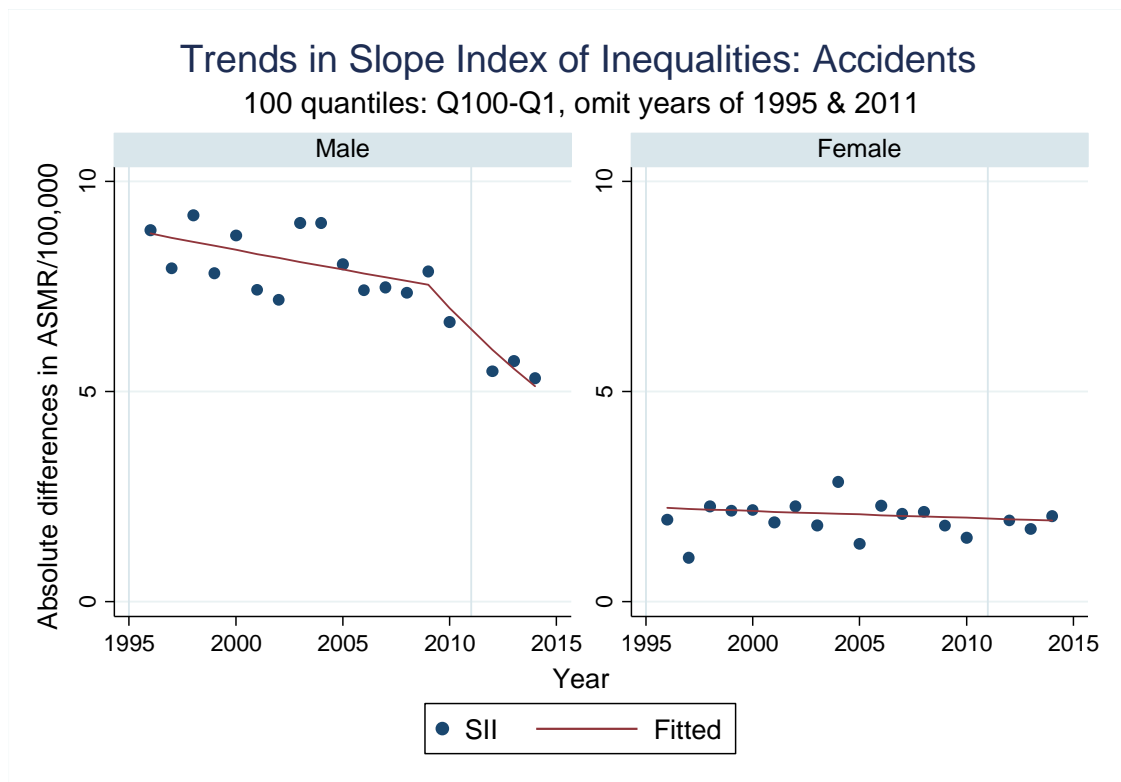


図 5-6-1B . ADI100 分位別年齢調整死亡率を用いた SII の推移：不慮の事故（大震災の年を除外）

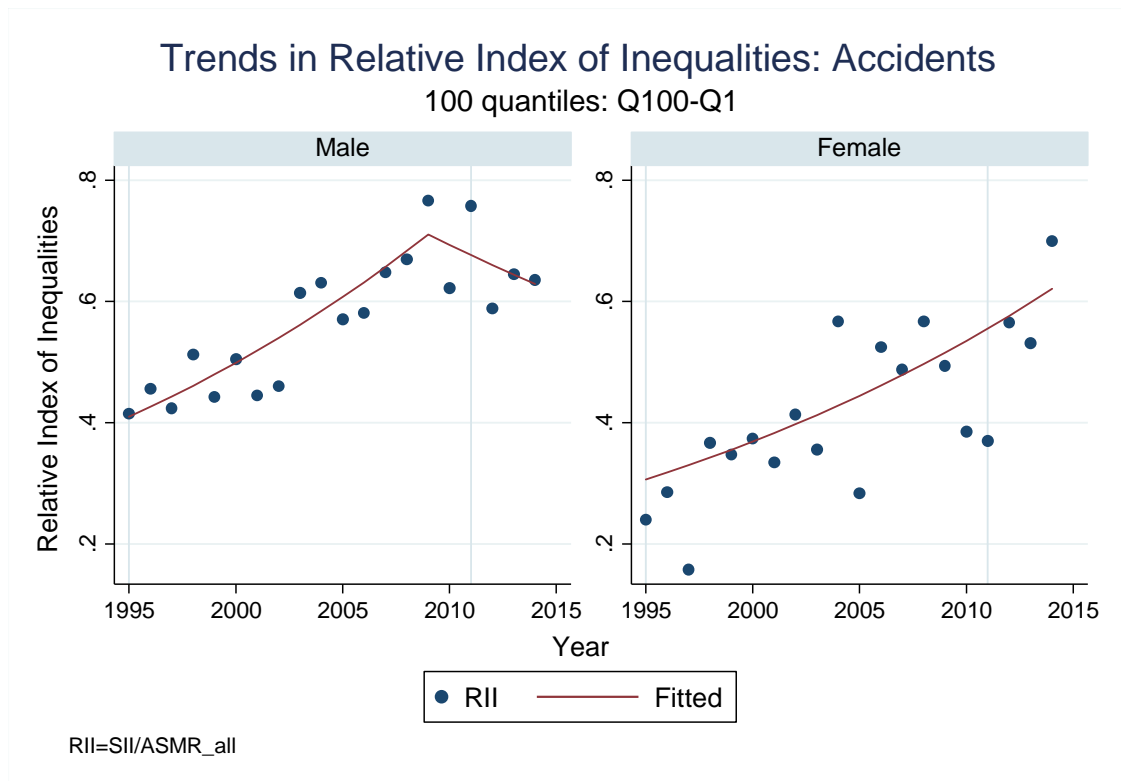


図 5-6-2A . ADI100 分位別年齢調整死亡率を用いた RII の推移：不慮の事故

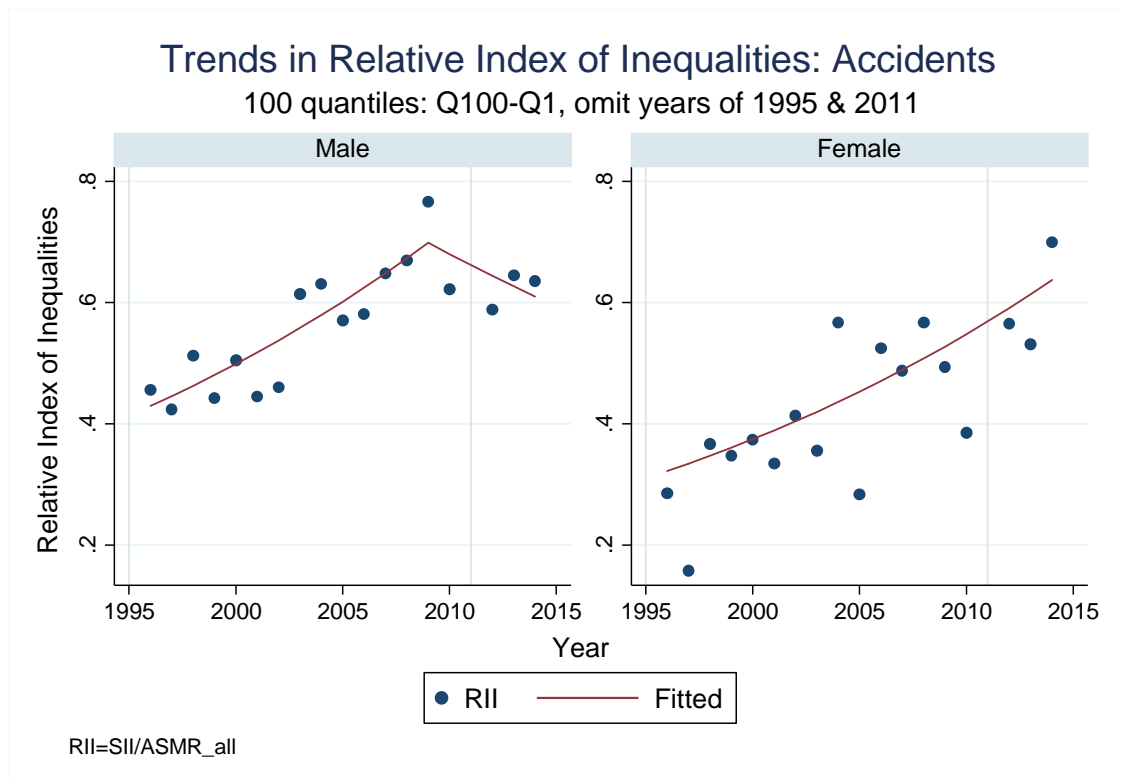


図 5-6-2B . ADI100 分位別年齢調整死亡率を用いた RII の推移：不慮の事故（大震災の年を除外）

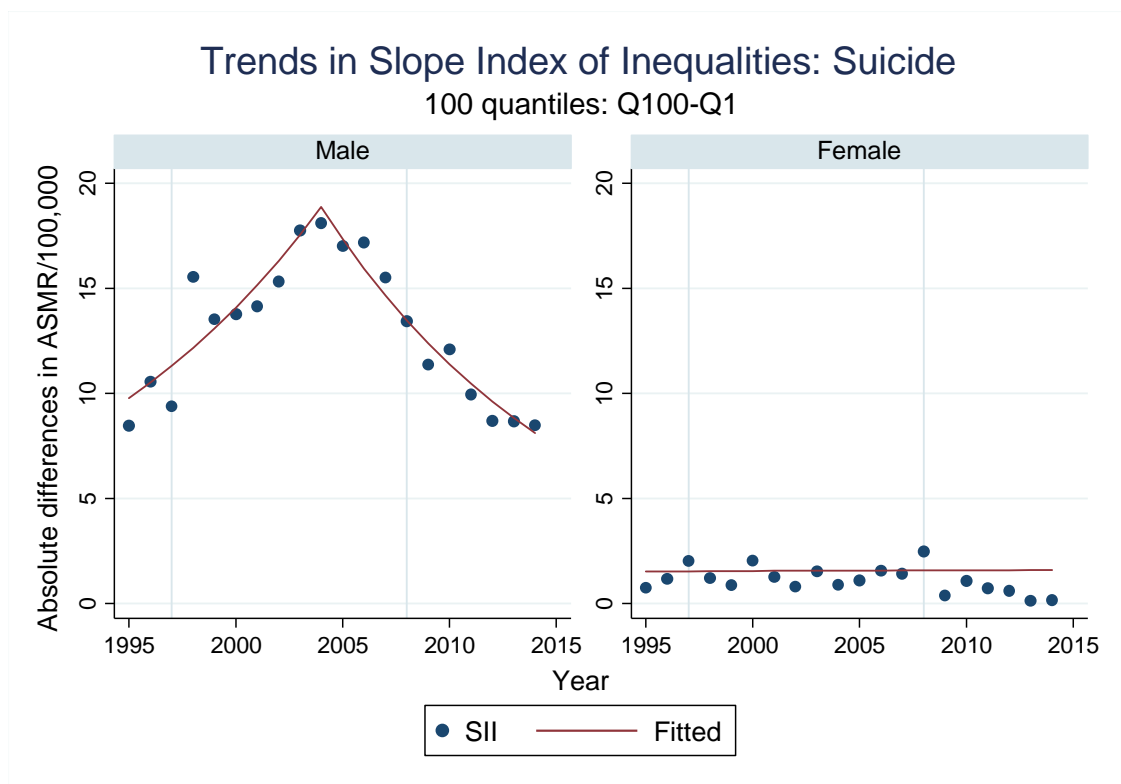


図 5-7-1 . ADI100 分位別年齢調整死亡率を用いた SII の推移：自殺

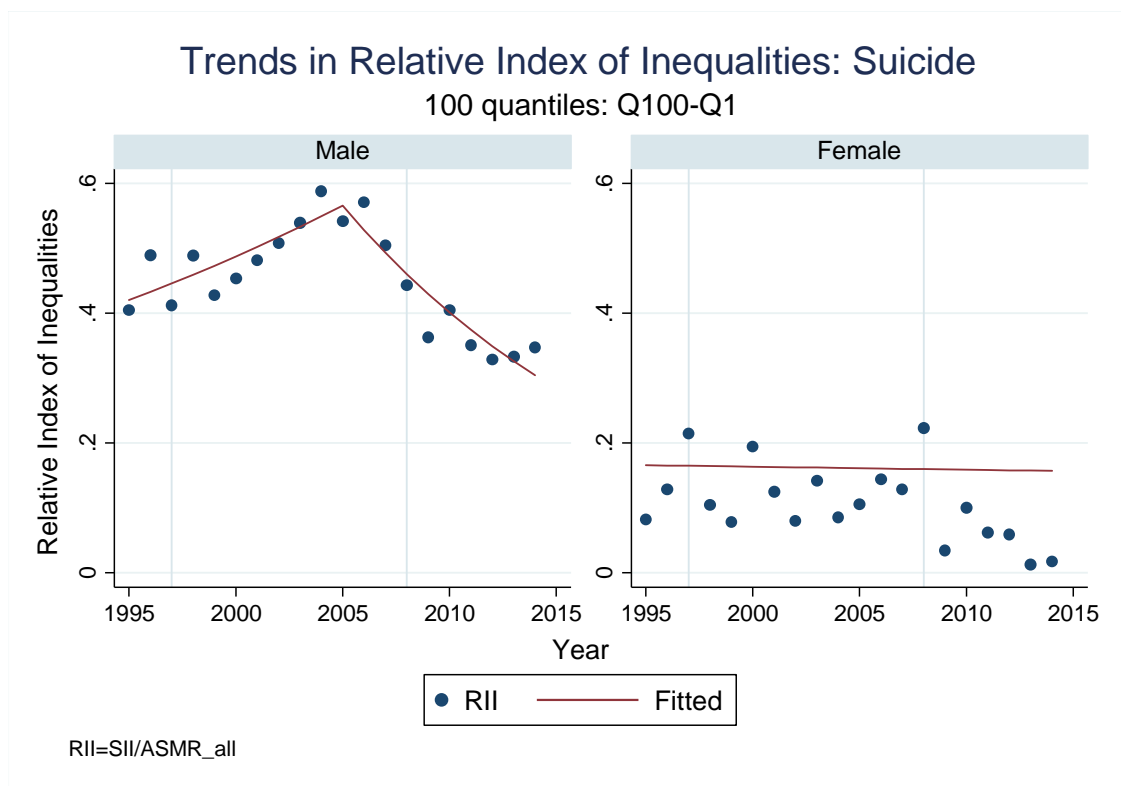


図 5-7-2 . ADI100 分位別年齢調整死亡率を用いた RII の推移：自殺

地理的剥奪指標データセットの構築

研究分担者 中谷友樹 立命館大学文学部(立命館大学歴史都市防災研究所 兼任)教授
研究協力者 米島万有子 立命館大学衣笠総合研究機構 専門研究員

研究要旨

本研究では、過去に提案されたマイクロデータによる貧困世帯分析の結果と一定整合的な日本版の地理的剥奪指標(ADI)について、1990年から2010年にかけて5年間隔の全国市区町村データセットを構築した。このADIは、国勢調査によって得られる単身高齢者世帯割合や母性世帯割合、失業率など8つの地域指標を利用する重み付線形和の式として定義される。異なる空間スケールや市町村合併に関する地理的単位の統一についても、世帯数を重みづけした平均値によって容易に再計算が可能である。今後は健康リスクの社会格差を計測するための社会経済的位置に関する地域指標としてこのADIを活用し、健康の社会格差に関する推移の把握に供することが期待される。

A. 研究目的

地理的単位を利用した健康リスクの地域差と、社会経済的な地域較差との相関関係は、健康リスクの社会格差を計測する1つの手段をもたらす。例えば、人口動態統計を利用した標準化死亡比など地域の健康リスク指標を、地域の社会経済的位置(socio economic position: SEP)の指標に関連づけ、SEPに応じた健康リスクの変化から社会格差の大きさを計測するのである。このような地域レベルのSEP指標として、居住地の貧困の度合いの尺度である地理的剥奪指標(ADI: areal deprivation index)がしばしば利用されてきた(中谷・矢野, 2014; 中谷・埴淵, 2015)。本研究では、先行する研究をふまえ、日本で利用可能なADI指標の

多時点・市区町村別データセットを構築する。

ADIは、何らかの定義に基づいて貧困と定義される人々の多寡をもって地区の貧困度を定義し、通常は値が大きい地区ほどより貧困度の高い地区であることを意味するように尺度構成される。現代的な意味での貧困を基礎づける相対的剥奪概念においては、社会通念的に「あるべき生活」に必要な、消費や社会参加などの活動が遂行できない状況が重なった状態を貧困と想定するが、この相対的剥奪をもたらす所得や資産、教育、労働、家族等の様々な剥奪の次元が存在する。そのため、ADIは複数の相対的な剥奪と関係する地域指標の合成によって操作化されることが多い。

この ADI として、多次元の地域指標が定期的に得られ、小地域での情報も公開される国勢調査データを利用した合成指標が、タウンゼンドの物的剥奪指標 (Townsend et al., 1988)をはじめ、数多く提案されてきた (Senior, 1988)。タウンゼンドは剥奪の次元に関する先験的な知識に従って 4 つの指標を選定し、それぞれを基準化した上で合計しているが、多変量解析を利用して、地理的な貧困度を示すと解釈しうる合成指標を経験的に抽出する手段もよく用いられる。それらに対して、貧困を規定するミクロレベル (世帯レベル) の研究成果 (例えば、阿部, 2006; 橋木・浦川, 2006) と概ね整合的な地域指標として ADI を得るための合成推定の方法も提案されてきた (Gordon, 1988)。日本のデータを利用した指標として、この合成推定の方法に従い、国勢調査指標を合成した ADI がある (中谷, 2011b; Nakaya et al., 2014)。本研究では、その先行して提案された方法に従い、市区町村別で、かつ異なる時期における ADI のデータセットを構築する。

B . 研究方法

提案された合成推定の方法では、貧困世帯を客観的基準 (世帯の年間等価所得が社会全体の中位数の半分以下であること) と主観的基準 (世帯の代表者の持つ主観的な社会的地位が中位より低いこと) を同時に満たす世帯とみなし、これを予測するマイクロデータの分析 (ロジスティック回帰分析) を通して単身高齢者や母子世帯、失業などといった指標と貧困世帯を関連づける重みを求める。その上で、

地域指標の重み付きの合計によって単独の地域指標を合成する。ただし、地域指標の重みは、ロジスティック回帰分析によって得られる各指標のオッズ比である。また、客観的基準での所得水準については、地域による物価水準による格差を考慮するため、4 人モデル世帯での最低生活費 (地域等級および冬季加算を考慮) について 1 級地の 1 の額を 1.0 とし、これに対する都道府県 × 市町村規模 3 区分 (政令指定市, 市部, 郡部) 別の額の比率を居住地域調整項 (0.70 ~ 1.00) として、所得水準を調整した (中谷, 2011)。

マイクロデータの貧困世帯の分析には、日本版総合社会調査 (JGSS) 累積データ 2000-2003 を、ADI の合成には、各年次の国勢調査統計を利用する。なお、この方法では、マイクロデータの分析と地域レベルでの ADI の合成では、共通する指標が必要である。そのため、学歴等、年次によって国勢調査の項目に含まれない指標は除外した上で、分析に利用する指標および ADI 合成にあたっての変数の選択を行った。

最終的に得られる ADI 計算式 (Nakaya et al., 2014) は、各地域の指標の重み付けの合計によって次のように定義される。

$$\begin{aligned}
 \text{ADI} = & k (2.99 \times \text{高齢夫婦世帯割合} \\
 & + 7.57 \times \text{高齢単身者世帯割合} \\
 & + 17.37 \times \text{母子世帯割合} \\
 & + 2.22 \times \text{賃貸住居居住世帯割合} \\
 & + 4.03 \times \text{グレーカラー就業者割合} \\
 & + 6.05 \times \text{農業漁業就業者割合} \\
 & + 5.38 \times \text{ブルーカラー就業者割合} \\
 & + 18.29 \times \text{完全失業率})
 \end{aligned}$$

ここで k は任意の正の定数であり、ADI を

利用して社会経済的にみた相対的な位置を調べるにあたっては(例えば、人口5分位グループを作成する場合)いずれの数値でも結果は一致する(そのため、 $k = 1$ で便宜的に計算する)。ただし、Gordon (1995)では、各地区における重み付けの指標の線形和が、各地区の貧困世帯割合に比例すると仮定し、全国の集計値(ADIの世帯重みづけ平均値)が全国の貧困世帯割合に一致するように k を定める方法も提案している。これに従い、JGSS 累積データ 2000-2003での貧困世帯割合 8.62%と2000年におけるADIの(世帯数を重みづけした)平均値が一致するように k を定め、便宜的に $k = 1$ の場合の合成指標をADIで表記すると素朴な貧困世帯割合の推定値 DepPoorRは、以下の関係式で得られる。

$$\text{DepPoorR} = \text{ADI} * 0.015641204039864$$

これらを1990年から2010年まで各国勢調査年次の市区町村別統計にあてはめて、各年次のADIのデータセットを作成した。

なお、ここで示した Nakaya et al. (2014)による重み係数は、同じデータを用いて当初に提案された中谷(2011)のものと若干異なっているが、これは職業分類の見直し(保安職業をグレーカラーに追加)および所得欠損値を補間していたものを分析上除外して再分析したことによる。両者の違いはそれほど大きくはないが、現時点では Nakaya et al. (2014)による重み係数の利用を推奨している。

(倫理面への配慮)

本研究はすべて公開資料に基づいており、倫理面での問題は存在しない。

C. 研究結果

市区町村別ADIを、1990年から2010年にわたる期間を対象に作成した。利用した統計資料は5年間隔で実施された国勢調査の地域統計資料による。ただし、空間データ解析や地理的視覚化には、地理情報システム用のデータセットとして、ESRI Japan社が販売する「データコレクション・スタンダードパック(2015年版)」(DCS2015)を利用すると便利である。DCS2015には、2010年の国勢調査による市区町村別統計データが、市区町村の境界に関する幾何的データとあわせて用意されており、2010年のADIについては、これを利用して計算した。それ以外の年次については、これに該当する国勢調査の各年次の統計表を参照してADIを作成した。このADIの合成に利用した国勢調査統計の地域指標に関する具体的な定義と、DCS2015を利用した場合の変数名について、表1に整理した。

計算された結果の例として、2010年の市区町村別国勢調査資料による、DepPoorRの分布図を図1に、ヒストグラムと箱ひげ図、正規確率プロットを図2に示す。図1より、概して貧困度の高い地区は、大都市圏よりも地方圏に広く分布するが、大都市圏内を拡大した図をみると、とくに京阪神大都市圏内にも貧困度の著しく高い地区が存在する。図2によれば、DepPoorR(ADI)の統計的な分布は平均値付近では正規分布に近いが、著しく貧困度の高い外れ値的な地区が一定存在することも分かる。

算出されたADIおよび関連する指標群の平均および標準偏差(市区町村規模による重みづけなし)を表2に示す。

D．考察

ここで整理した ADI および DepPoorR は、国勢調査指標の重み付合計によって容易に計算可能である。市区町村別 ADI に限らず、町丁字別など、他の空間スケールの地理的単位についても、国勢調査指標がその単位で得られている限り、利用できる。なお、国勢調査の小地域統計（町丁字等集計）では、2005 年以降、母子世帯比率が標章されなくなったが、小地域で利用可能な世帯類型の情報と市区町村別統計で得られる母子世帯割合の情報から、おおよその数値を推計して ADI を合成することも可能である（詳細はここでは略す）。また、貧困世帯の割合の推定値という考えに基づけば、世帯数を重みづけすることで、異なる単位での ADI を世帯の重み付け ADI 平均値によって求めることもできる。すなわち、ここで提案している ADI は、単純な重み付線形和として定義されているため、計算およびその取扱いが容易である点に利点がある。

なお、表 2 をみると、高齢化に関連する指標や賃貸住宅での居住、母子世帯・失業に関する指標値の増加を反映して、全体として ADI の平均値は年次とともに上昇傾向にある。ただし、これは指標に関する重みの絶対値に変動がないことを前提とした議論であり、年次を違えた ADI の絶対値の比較に意味があるかどうかは、現時点では判断が難しい。そのため、今回整備した ADI については、原則的に各時期における地域の相対的な貧困度を示すものと考えておくことにしたい。より厳密には、ADI を合成する重みや利用する国勢調査指標についても、一定の期間ごとに見直すことが将来的には望ましいが（cf. Dorling et al.,

2007）、統一的な基準で貧困世帯のマイクロデータ分析が可能な社会調査資料を長期間にわたって揃えることが難しい。

また、各年次の ADI は、それぞれの時期に実施された国勢調査資料に基づいているため、地理的単位数（市区町村数）が異なる点にも留意する必要がある。原則として、近年になるほど、市町村合併により市町村数は減少し、地理的な分解能は低下する。そのため、過去では捉えられた社会経済的な居住分離が近年の市区町村単位では補足しづらくなった可能性がある。

最後に、市区町村を識別する JIS コードが定義されているため、国勢調査の時点（各 国勢調査年の 10 月 1 日）の市区町村境界に関する GIS ファイルを用意すれば、ADI に関する分布図の作成や各種の空間分析が容易である。ただし、人口動態統計は各年末の市区町村を基準として作成されており、年度途中での市区町村変更による影響も受けるため、ここで作成した ADI データセットを、人口動態統計等の資料と結合する場合には、分析単位と統計作成時期の関係に注意する必要がある。任意の年月日に対応した市区町村別境界の GIS ファイルと、異なる 2 時点間の市区町村の対応関係については、桐村ほか(2011)を参照されたい。

E．結論

本研究では、国勢調査指標を重み付合成する地理的剥奪指標（ADI）をとりあげ、1990 年から 2010 年まで 5 年間隔の市区町村別 ADI データを整備した。当該 ADI データは、地理的単位の変更に伴う再計算等も容易である。今後は健康リスクの社会格差を計測するための社会経済的位置に関する地域指標としてこの ADI を活用し、健康

の社会格差に関する推移の把握に供することが期待される。

F . 健康危険情報

記載すべき情報なし。

G . 研究発表

1 . 論文発表

中谷友樹・埴淵知哉 (2016) 健康の社会格差と地域格差. 地理 61-1, 51-57.

2 . 学会発表

Nakaya T, Ito Y, Ioka A and Nakayama T (September 2015): Modelling geographic disparities of relative cancer survival at the neighbourhood level in Osaka prefecture, Japan, GEOMED 2015, 11 September 2015, at Università di Firenze, Florence, Italy
埴淵知哉・中谷友樹・本庄かおり (2015年11月4日): JGSS 累積データからみた健康格差の推移, 第74回日本公衆衛生学会総会, 長崎新聞文化ホール, 長崎市.

伊藤ゆり・中谷友樹・近藤尚己・福井敬祐・中田佳世・井岡亜希子・宮代 勲・中山富雄 (2015年11月4日): 大阪府におけるがん進行度別罹患率の社会経済格差: 1993-2004年における格差の変化, 第74回日本公衆衛生学会総会, 長崎新聞文化ホール, 長崎市.

富岡慎一・藤野善久・中谷友樹・久保達彦・松田晋哉 (2015年11月4日): 患者居住地域の SES 及び精神疾患の有無が入院中に患者が受療する医療行為に与える影響についての解析, 第74回日本公衆衛生学会

総会, 長崎新聞文化ホール, 長崎市.

中谷友樹 (2015年12月12日): 健康リスクの社会格差と地域格差. 貧困研究会第8回研究大会シンポジウム「地域が抱える健康・貧困リスク問題への学術的視点」, 九州大学箱崎キャンパス.

中谷友樹 (2015年10月3日)

GISを利用した健康づくり支援環境研究の推進. シンポジウム: 病院を中心とする街づくり まちなか集積医療の提言, 東京医科歯科大学.

H . 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

引用文献

阿部彩 (2006): 相対的剥奪の実態と分析 - 日本のマイクロデータを用いた実証研究 -, 社会政策学会誌 16, 251-275.

橘木俊詔・浦川邦夫 (2006): 『日本の貧困研究』東京大学出版会.

桐村喬・中谷友樹・矢野桂司(2011): 市区町村の区域に関する時空間的な地理情報データベースの開発 Municipality Map Maker for Web, GIS 理論と応用 19(2), 83-92.

中谷友樹(2011): 地理統計に基づくがん死亡の社会経済的格差の評価 - 市区町村別がん死亡と地理的剥奪指標との関連性 -.統計数理 59, 239-265.

- 中谷友樹・埴淵知哉 (2015): 健康の社会格差と地域格差, 地理 61-1, 51-57.
- 中谷友樹・矢野桂司(2014:)社会格差を視る小地域のセンサス指標:地理的剥奪指標とジオデモグラフィクス. 地域開発 59, 35-40.
- Dorling, D., Rigby, J., Wheeler, B., Ballas, D., Thomas, B., Fahmy, E., Gordon, D. and Lupton, R. (2007): Poverty and wealth across Britain 1968 to 2005, The Policy Press, Bristol.
- Gordon, D. (1995): Census based deprivation indices: their weighting and validation, Journal of Epidemiology and Community Health, 49 (Suppl 2), pp. S39-S44.
- Nakaya, T. (2010): 'Geo-morphology' of population health in Japan: Looking through the cartogram lens, Environment and Planning A, 42, pp. 2807-2808.
- Senior, M. (2002): Deprivation indicators. in Rees, P., Martin, D. and Williamson, P. eds., The Census Data System, Chichester: Wiley.
- Townsend, P., Phillimore, P. and Beattie, A. (1988): Health and deprivation: inequality and the North, London: Croom Helm.

表1 市区町村別地理的剥奪指標に関する国勢調査指標一覧

剥奪指標を構成する指標項目	原データ (2010年版の場合、それより過去の場合は国勢調査の市区町村別集計を利用)	指標の計算定義	備考
一般世帯総数	ArcGISデータコレクション統計バック2014> 国勢調査 2010 町丁・字等別集計 全指標 >全国 >CITY_POP_H22.shp	GENHH_ALLPER_H22 (一般世帯数_総数_H22)	
高齢単独世帯		GENHHFA_WT65UPONL_SOL_H22(一般世帯数(65歳以上世帯員のみ在住)単独世帯_H22)	高齢単独世帯割合=高齢単独世帯/一般世帯総数
高齢夫婦世帯	平成22年国勢調査 [平成22年] 都道府県・市区町村別統計表(一覧表) https://www.e-stat.go.jp/SG1/estat/GL02020101.do?method=extendTClass&refTarget=toukeihyo&listFormat=hierarchy&statCode=00200521&tstatCode=000001049104&tclass1=&tclass2=&tclass3=&tclass4=&tclass5=	(再掲) 高齢夫婦世帯(夫65歳以上妻60歳以上の夫婦1組のみの一般世帯)	高齢夫婦世帯割合=高齢夫婦世帯/一般世帯総数
母子世帯	平成22年国勢調査人口等基本集計(男女・年齢・配偶関係、世帯の構成、住居の状態など)(総務省統計局) 第10表 世帯の家族類型(22区分)別一般世帯数及び一般世帯人員(6歳未満・18歳未満世帯員のいる一般世帯及び三世帯世帯並びに母子世帯及び父子世帯 - 特掲) - 全国、都道府県、市町村 http://www.e-stat.go.jp/SG1/estat/NewList.do?tid=000001039448 (再掲) 母子世帯数 一般世帯数	(再掲) 母子世帯数 一般世帯数	母子世帯割合 = 母子世帯 / 一般世帯総数 「-」は0に変換
賃貸住宅	ArcGISデータコレクション統計バック2014> 国勢調査 2010 町丁・字等別集計 全指標 >全国 >CITY_POP_H22.shp	GENHHHO_PUBRENT_H22(一般世帯数_公営・都市再生機構・公社の借家_H22) GENHHHO_RNTPRV_H22 (一般世帯数_民営の借家_H22) GENHHHO_LOD_H22 (一般世帯数_間借り_H22)	賃貸住宅 = (公営・都市再生機構・公社の借家 + 民営の借家 + 間借り) 賃貸住宅居住世帯割合=賃貸住宅/住宅に住む一般世帯数
労働力人口	ArcGISデータコレクション統計バック2014 >国勢調査 2010 町丁・字等別集計 全指標 >全国 >CITY_IND_H22.shp	LAB_LABPOP_H22 (人口(15歳以上)労働力人口_H22)	
農林漁業従事者		JOB_AFS_H22 (就業者数(15歳以上)職業_農林漁業従事者_H22)	農林漁業従事者割合=農林漁業従事者/労働力人口
グレーカラー-就業者	ArcGISデータコレクション統計バック2014 >国勢調査 2010 町丁・字等別集計 全指標 >全国 >CITY_JOB_H22.shp	JOB_SAL_H22 (就業者数(15歳以上)職業_販売従事者_H22) JOB_SER_H22 (就業者数(15歳以上)職業_サービス職従事者_H22) JOB_PRE_H22 (就業者数(15歳以上)職業_保安職従事者_H22)	グレーカラー従事者 = 販売従事者 + サービス職従事者 + 保安職従事者 グレーカラー従事者割合=グレーカラー従事者/労働力人口
ブルーカラー-就業者		JOB_IND_H22 (就業者数(15歳以上)職業_生産工程従事者_H22) JOB_PROD_H22 (就業者数(15歳以上)職業_輸送・機械運転従事者_H22) JOB_BUI_H22 (就業者数(15歳以上)職業_建設・採掘従事者_H22) CAR_H22 (就業者数(15歳以上)職業_運搬・清掃・包装等従事者_H22)	ブルーカラー従事者 = 生産工程従事者 + 輸送・機械運転従事者 + 建設・採掘従事者 + 運搬・清掃・包装等従事者 ブルーカラー-就業者割合=ブルーカラー-就業者/労働力人口
完全失業者	平成22年国勢調査 産業等基本集計(労働力状態、就業者の産業など) 労働力状態・産業・従業上の地位 1-2 労働力状態(8区分)、年齢(5歳階級)、男女別15歳以上人口 - 都道府県、市町村 http://www.e-stat.go.jp/SG1/estat/NewList.do?tid=000001039448	総数(15歳以上年齢) 完全失業者	完全失業者割合 = 完全失業者 / 労働力人口

表2 地理的剥奪指標と関連した地域指標の年次別集計値

指標	要約統計量	1990年	1995年	2000年	2005年	2010年
DepPoorR	平均	8.40%	8.63%	8.92%	9.50%	9.67%
	標準偏差	1.01%	1.02%	1.02%	1.13%	1.16%
ADI	平均	5.37	5.52	5.70	6.07	6.19
	標準偏差	0.64	0.65	0.65	0.72	0.74
高齢夫婦世帯割合	平均	7.34%	8.72%	10.60%	11.35%	12.18%
	標準偏差	3.75%	4.26%	4.64%	4.13%	3.81%
高齢単独世帯割合	平均	5.26%	6.58%	8.10%	9.17%	10.48%
	標準偏差	3.33%	3.84%	4.27%	4.23%	4.19%
母子世帯割合	平均	1.04%	0.94%	1.05%	1.34%	1.41%
	標準偏差	0.55%	0.50%	0.54%	0.58%	0.55%
賃貸住宅居住世帯割合	平均	16.64%	17.78%	18.28%	21.57%	23.89%
	標準偏差	12.85%	12.99%	12.64%	12.60%	12.66%
グレーカラー就業者割合	平均	16.97%	18.08%	19.75%	21.65%	23.37%
	標準偏差	5.83%	5.62%	5.29%	4.91%	4.40%
農林漁業就業者割合	平均	18.04%	15.66%	13.45%	11.81%	9.64%
	標準偏差	13.13%	11.76%	10.71%	10.37%	9.64%
ブルーカラー就業者割合	平均	36.78%	35.70%	34.75%	32.10%	29.21%
	標準偏差	8.46%	7.66%	7.35%	7.11%	6.86%
完全失業者割合	平均	2.47%	3.31%	3.86%	5.56%	6.33%
	標準偏差	1.45%	1.64%	1.64%	2.13%	2.14%
市区町村数		3,376	3,370	3,368	2,366	1,901

注：人口（世帯数）0の自治体は含めていない。

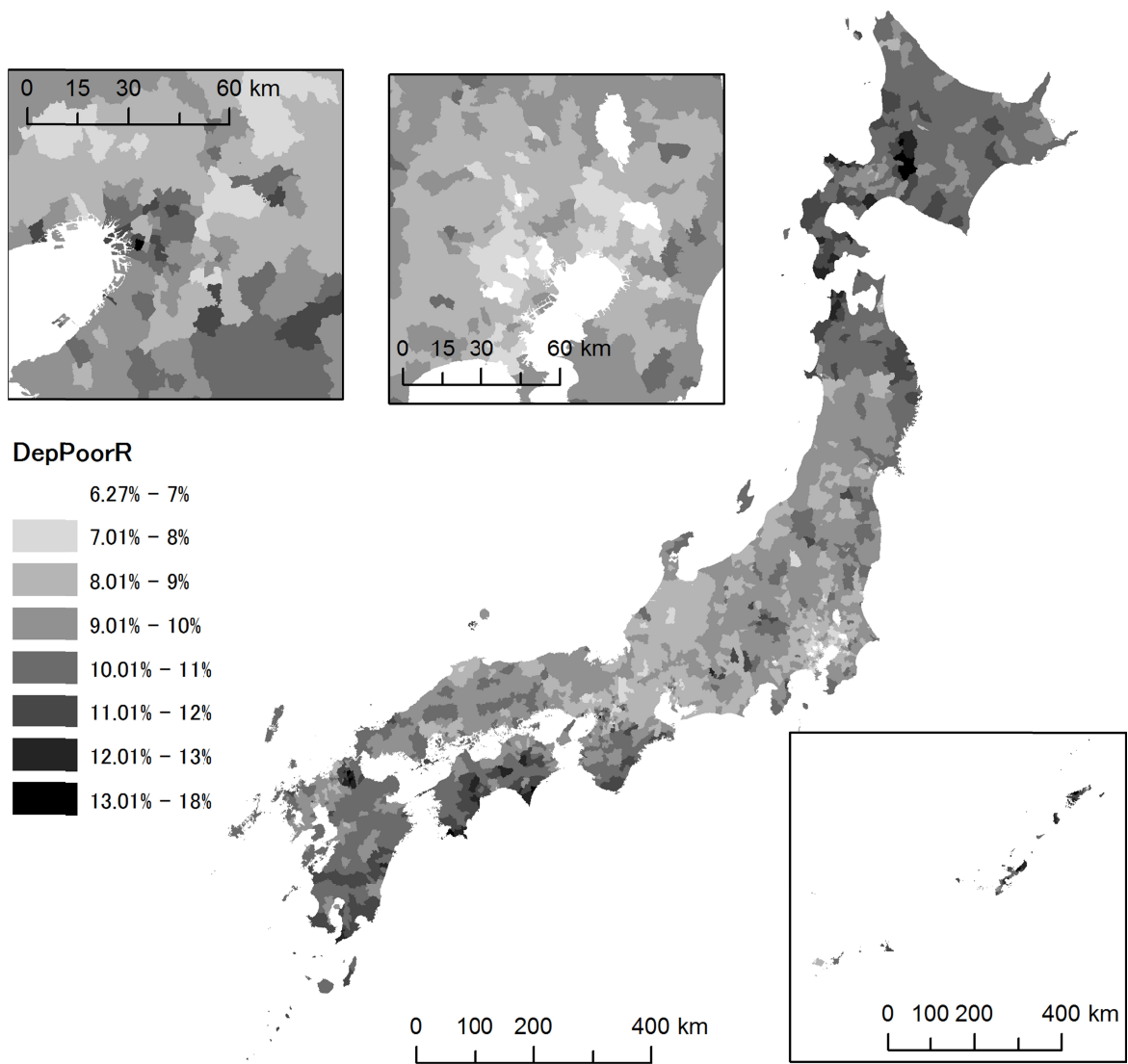


図1 2010年地理的剥奪指標の地理的分布

ここでは DepPoorR をプロットしている。値が大きいほど、貧困な地区とみなされる。

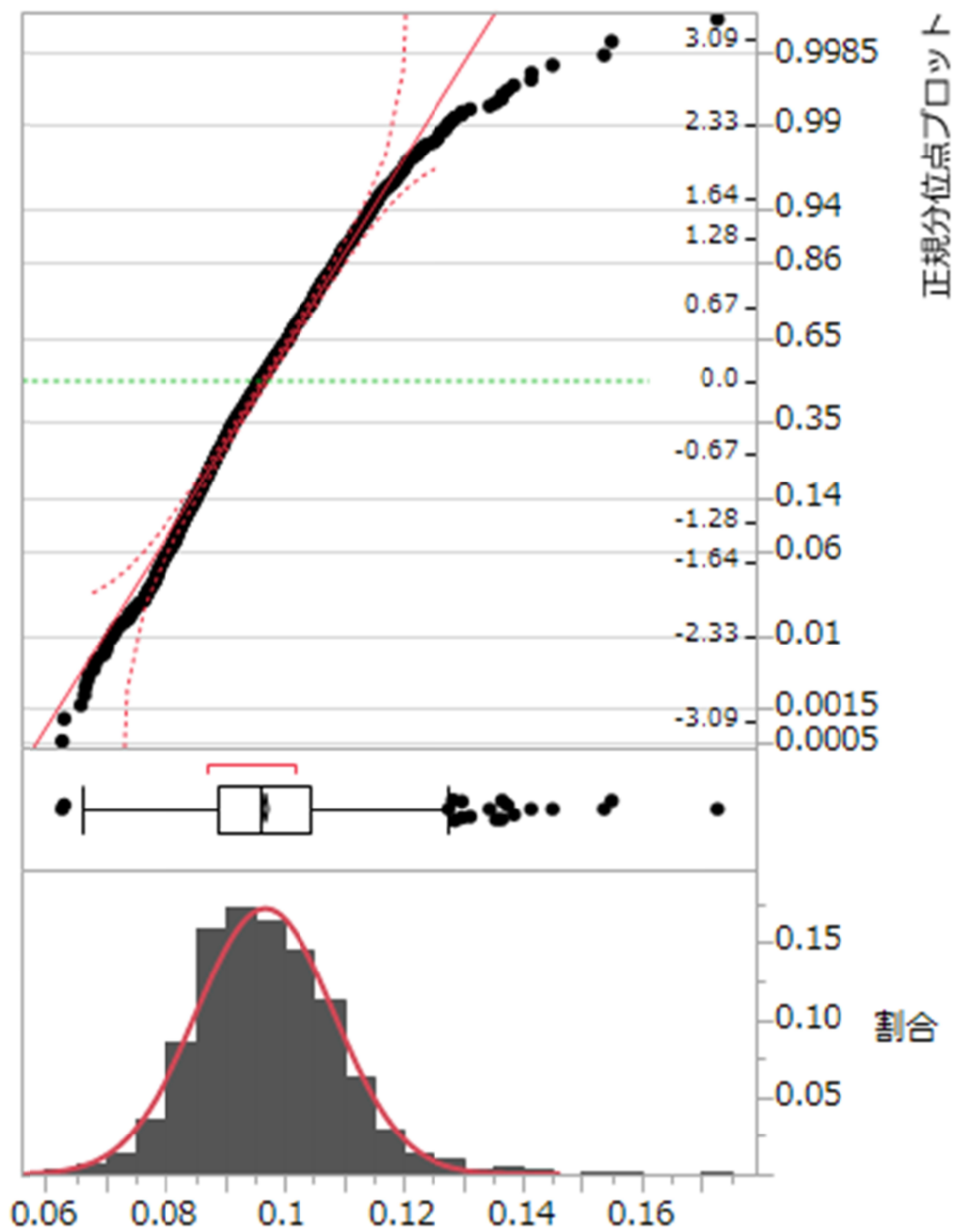


図2 地理的剥奪指標の分布

ここでは DepPoorR をプロットしている。値が大きいほど、貧困な地区とみなされる。

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))
分担研究報告書

市区町村別性年齢階級別人口の線形補間について

研究協力者 福井敬祐 大阪府立成人病センターがん予防情報センター 研究員

研究要旨

市区町村別地理的剥奪指標を用いた全死亡・主要死因別の年齢調整死亡率を算出するためには、市区町村別の性別・年齢 5 歳階級別の人口が必要となる。しかしながらこれらの人口データは国勢調査実施年のものしか提供されていない。そこで、本研究における人口動態統計の分析対象期間である 1985～2014 年における国勢調査実施年以外の年における市区町村別性年齢階級別人口の線形補間の実施手順および課題についてまとめた。対象期間の間、市区町村は合併・分割・分割後合併が行われている。Web にて桐村らが提供する「Municipality Map Maker ウェブ版 市区町村区域の GIS データ生成ツール」を用いて、市区町村構成の変遷に対応した。市区町村構成の変遷パターンに応じ、2 時点の市区町村における性別・年齢 5 歳階級別人口の線形補間を行った。本手法による補間において以下の課題が残った。東京都三宅村の人口が 0 になる、外挿した場合の人口が負の値になるところがある、直近の国勢調査実施年の間で線形補間したが、対象期間共通の市区町村区分で統一し、線形補間した方がよいか、基準日の詳細設定、面積による重み付けの問題などである。今後、上記課題を解決し、より精緻な人口データセットを作成する必要がある。

A . 研究目的

国勢調査における市区町村別の人口データは 1985 年, 1990 年, 1995 年, 2000 年, 2005 年, 2010 年までの 5 年毎のものとなっている。より安定した分析を行うためには各 5 年の間のデータを補間し利用することが考えられる。補間法として一般的な線形補間は隣接する 2 時点の対応するデータを用いて行われるが、市区町村別に着目すれば場合には合併や分割など影響を考慮する必要がある。本報告は国勢調査から得られた人口データにおいて、合併を考慮した上で

線形補完を行う方法についてまとめたものである。

B . 研究方法

1. 具体例

図 1 は熊本県球磨郡上村, 免田町, 岡原村, 須恵村, 深田村の 5 町村の合併の 2000 年から 2005 年の変遷を表している。5 町村が 2003 年 10 月 1 日を以てあさぎり町として合併していることがわかる。国勢調査は 5 年毎に行われるため、今, 2000 年と 2005 年の人口データのみが得られている。図内

の5町村は2005年時点では合併により存在しないため、2000年と2005年の間でそのまま線形補間を行うことができない。このような場合においては2000年を合併後の市区町村に、もしくは2005年を合併前の市区町村に作り変え、仮想的に対応した2時点を作り出すことで線形補間を行う。図2は2000年時点の5町村を2000年時点で仮想的にあさぎり町に合併することで、2005年時点と対応可能にし、線形補間を行うイメージ図を表している。仮に、2000年時点の上村、免田町、岡原村、須恵村、深田村の人口がそれぞれ、45千人、100千人、30千人、20千人、15千人であったとすれば、2000年時点で仮想的に作成されたあさぎり町の人口は5町村の人口を足し合わせた210千人である。さらに2005年時点のあさぎり町の人口が135千人であったと仮定すると、仮想的に作成された2000年時点の人口210千人との線形補間により2001年、2002年、2003年、2004年の人口はそれぞれ195千人、180千人、165千人、150千人と計算される。

2. 使用したデータ

人口データ

国勢調査より入手した1985年、1990年、1995年、2000年、2005年、2010年の市区町村・性・年齢階級別人口データ

市区町村変遷対応表

桐村らが提供する「Municipality Map Maker ウェブ版 市区町村区域のGISデータ生成ツール」[1]より以下の5つのcsvファイルを作成・入手した。

1985年10月1日時点から1990年10月1日時点への市区町村対応表

1990年10月1日時点から1995年10月1日時点への市区町村対応表

1995年10月1日時点から2000年10月1日時点への市区町村対応表

2000年10月1日時点から2005年10月1日時点への市区町村対応表

2005年10月1日時点から2010年10月1日時点への市区町村対応表

図3は入手したの市区町村対応表の一例である。市区町対応表を用いて、2000年時の住所区分けを表す住所コード(JISCODE1)に2005年時の住所区分けを表す住所コード(JISCODE2)を対応させる(紐付ける)ことができる。

3. 補間方法

線形補間は市区町村変遷対応表のJISCODE2をJISCODE1に紐付ける(最新年の市区町村分けに対応させる)ことで線形補間に用いる1対1の対応を作成した後に行う。紐付けの方法は市区町村の変遷パターン(エラー! 参照元が見つかりません。)に大きく依存するため、線形補間の方法についてもこのパターンに沿って説明する。なお2時点 t_1, t_2 年($t_1 < t_2$)に対応する人口を y_1, y_2 としたとき、区間 $[t_1, t_2]$ 内の任意の時点 t の人口 y は

$$y = y_1 + \frac{y_2 - y_1}{t_2 - t_1} \times (t - t_1),$$

によって計算される。

(1) 変化しない

時間が経過しても市区町村が変化しない場合にはJISCODE2をJISCODE1に紐付け、2時点の人口を使用して線形補間を行う。

(2) 合併

合併が起きた場合には、対応表を元に JISCODE2(合併後住所)を JISCODE1(合併前住所)に紐付けしたあと、JISCODE2 が同じ市区町村の人口を合算することで仮想的に最新年と同じ市区町村分けを作成し線形補間を行う(図 5)。

(3) 分割

分割の場合には、JISCODE2 を JISCODE1 に紐付けたあと、市区町村変遷対応表内にある WEIGHT を用いて最新年の市区町村分けに重み付けで人口を分割する。その後、対応した 2 時点間でそれぞれ線形補間を行う(図 6)。

(4) 分割後合併

分割後合併の場合には JISCODE2 を JISCODE1 に紐付けたあと、市区町村変遷対応表内にある WEIGHT を用いて最新年の市区町村分けに重み付けで人口を分割する。その後、紐付けられた JISCODE2 が同じ市区町村の人口を合算し仮想的な人口を作成し線形補間を行う(図 7)。

(倫理面への配慮)

本研究に用いた資料は全て公開データに基づいているため、倫理面において問題になることはない。

C・D. 研究結果および考察

本手法による補間では以下の課題が残っている。

(ア)三宅村の人口

平成 12 年国勢調査の際、三宅島噴火により、全島民が島外へ避難したことにより、東京都三宅村の人口は 0 となっている。こ

の場合どのように取り扱うのか。

(イ)外挿した場合の取り扱い

外挿において地区・年齢階級によっては人口が負の値になる地区も存在する。単純に 0 と置き換えてよいか。

(ウ)線形補間の区間(基準点をどうするか)

線形補間は国勢調査が行われた 5 年間で直近分の市区町村区分けに紐づけして行っているが実際には全期間で統一した方がよいのではないか。

(エ)使用する市区町村変遷対応表基準日

市区町村変遷対応表は現在取得を各国政調査が行われた年の 10 月 1 日を基準日としているが、より細かく基準日を設定する必要がある可能性がある。

(オ)WEIGHT の使用について

分割の際に利用する市区町村変遷対応表の重み(WEIGHT)は土地面積比により作成されているが、面積が大きい人口は少ないという土地に分割を行った場合には、実際の人口から大きく乖離し、線形補間が不安定になる場合がある。

静岡県浜松市の例

静岡県浜松市は 2009 年 9 月 1 日に中区、東区、西区、南区、北区、浜北区、天竜区の 7 区に分割された。2010 年における各区の人口割合は、浜松市人口割合

図 9. . 浜松市人口割合

図 9 の通りであ

るが、図 9 にある通り本来人口が最も少ない天竜区に対する WEIGHT が最大となっている。2005 年時点の浜松市の人口を 2010

年の区分けに対応させるための WEIGHT を用いた分割時に天竜区の人口は 2005 年時点の人口 399 千人に 0.626657 をかけたおよそ 250 千人と算出される。しかし、2010 年時点の人口は 16 千人とその差が大きい、そのため補間も実際の人口推移とは乖離するものになってしまう。

E . 結論

本報告書に記載した方法で作成した補間人口データは 1985 年、1990 年、1995 年、2000 年、2005 年、2010 年の国勢調査のデータの線形補間法について記述した。作成したデータは 1986 年～1989 年、1991 年～1994 年、1996 年～1999 年、2001 年～2004 年、2006 年～2009 年の単年データである。これらのデータはそれぞれ直近の国勢調査に対応する市区町村区分けに変換しているのみ(例えば、1986 年～1989 年であれば 1990 年に変換)であり、全期間を通して同じ市町村区分けを利用している訳でないことに注意されたい。今後は 2000 年などのようにある特定の時点の市区町村区分けに変換し、全期間で比較・対応可能なデータの作成を行いたい。

F . 健康危険情報

なし

G . 研究発表

1 . 論文発表

なし

2 . 学会発表

なし

H . 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

引用文献

[1] 桐村 喬, 「Municipality Map Maker ウェブ版 市区町村区域の GIS データ生成ツール」,
<<http://www.tkimura.com/mmm/>> (参照 2016 年 4 月 2 日).

2000年	2001年	2002年	2003年		2004年	2005年
			～10/1	10/1～		
上村	上村	上村	上村	あさぎり町	あさぎり町	あさぎり町
免田町	免田町	免田町	免田町			
岡原村	岡原村	岡原村	岡原村			
須恵村	須恵村	須恵村	須恵村			
深田村	深田村	深田村	深田村			

図 1. 2000 年から 2005 年間で行われた熊本県球磨郡あさぎり町合併の変遷



図 2. 2000 年から 2005 年における合併を考慮した線形補間のイメージ

NO	DATE1	JISCODE1	PNAME1	GNAME1	CNAME1	WEIGHT	DATE2	JISCODE2	PNAME2	GNAME2	CNAME2
469	20001001	43502	熊本県	球磨郡	上村	1	20051001	43514	熊本県	球磨郡	あさぎり町
470	20001001	43503	熊本県	球磨郡	免田町	1	20051001	43514	熊本県	球磨郡	あさぎり町
471	20001001	43504	熊本県	球磨郡	岡原村	1	20051001	43514	熊本県	球磨郡	あさぎり町
472	20001001	43508	熊本県	球磨郡	須恵村	1	20051001	43514	熊本県	球磨郡	あさぎり町
473	20001001	43509	熊本県	球磨郡	深田村	1	20051001	43514	熊本県	球磨郡	あさぎり町

図 3. 市区町村変遷対応表の例

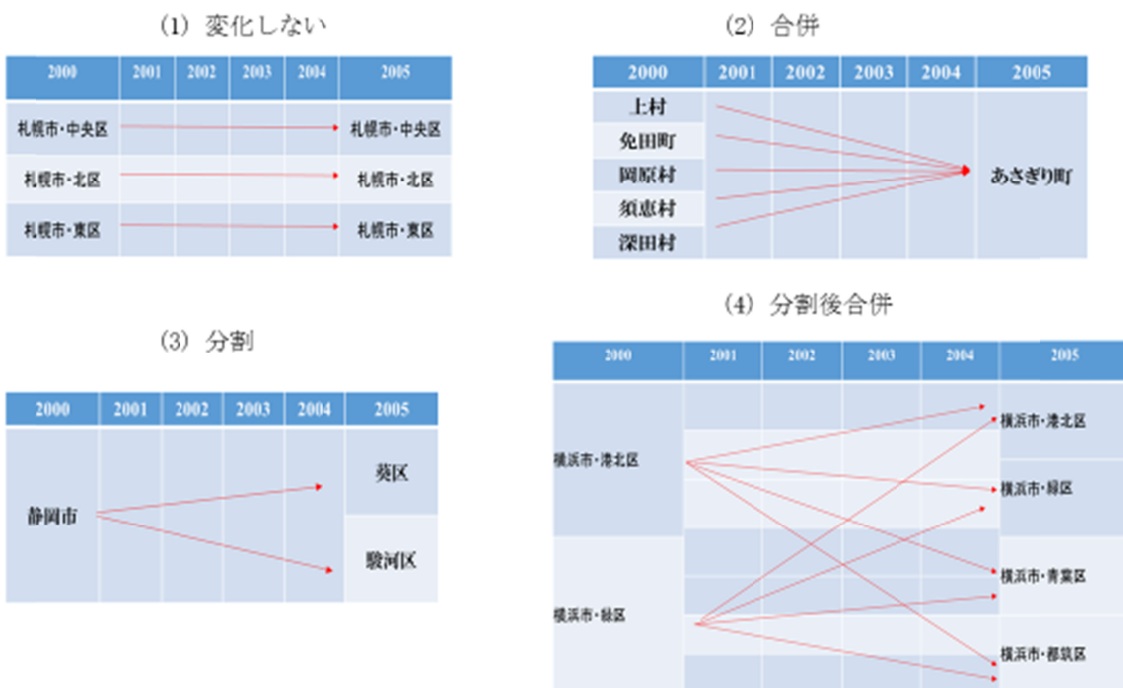


図 4. 市区町村変遷パターン



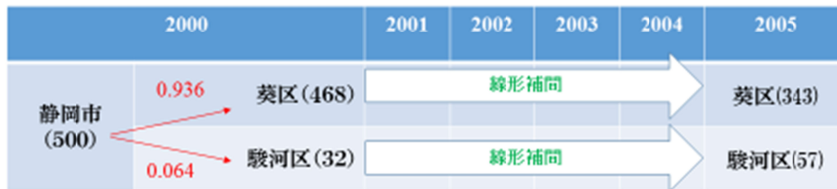
① 市区町村変遷対応表を元にあさぎり町を2000年時点の上村・免田町・岡原村・須恵村・深田村に紐付け、人口を合算することで2000年時点の仮想的なあさぎり町人口を作成。

② ①にて作成した2000年のあさぎり町人口と2005年のあさぎり町人口で線形補間。つまり、

$$x\text{年の人口} = 210 + \frac{135 - 210}{2005 - 2000} \times (x - 2000), \quad x = 2001, \dots, 2004$$

図 5. 市区町村変遷・合併の例。カッコ内は仮想的な人口(千人)を表す

NO	DATE1	JISCODE1	PNAME1	GNAME1	CNAME1	WEIGHT	DATE2	JISCODE2	PNAME2	GNAME2	CNAME2
448	20001001	22201	静岡県	静岡市		0.936	20051001	22101	静岡県	静岡市	葵区
448	20001001	22201	静岡県	静岡市		0.064	20051001	22102	静岡県	静岡市	駿河区

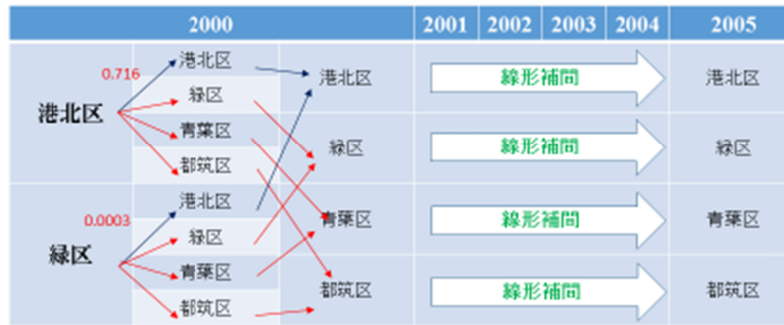


① 市区町村変遷対応表を元に葵区・駿河区を2000年時点の静岡市と紐付け、対応表内のWEIGHTを用いて人口を分割

② 分割させた2000年時点の葵区・駿河区人口と2005年時点の葵区・駿河区人口でそれぞれ線形補間

図 6. 分割時に関する市区町村変遷対応表と市区町村変遷・分割の例

NO	DATE1	JISCODE1	PNAME1	GNAME1	CNAME1	WEIGHT	DATE2	JISCODE2	PNAME2	GNAME2	CNAME2
360	19901001	14109	神奈川県	横浜市	港北区	0.71641	19951001	14109	神奈川県	横浜市	港北区
360	19901001	14109	神奈川県	横浜市	港北区	0.000658	19951001	14113	神奈川県	横浜市	緑区
360	19901001	14109	神奈川県	横浜市	港北区	0.000889	19951001	14117	神奈川県	横浜市	青葉区
360	19901001	14109	神奈川県	横浜市	港北区	0.282043	19951001	14118	神奈川県	横浜市	都筑区
361	19901001	14113	神奈川県	横浜市	緑区	0.000319	19951001	14109	神奈川県	横浜市	港北区
361	19901001	14113	神奈川県	横浜市	緑区	0.332717	19951001	14113	神奈川県	横浜市	緑区
361	19901001	14113	神奈川県	横浜市	緑区	0.46172	19951001	14117	神奈川県	横浜市	青葉区
361	19901001	14113	神奈川県	横浜市	緑区	0.205244	19951001	14118	神奈川県	横浜市	都筑区



- ① 市区町村変遷対応表を元に2000年時点の区を, 対応表内のWEIGHTを用いて人口を分割.
- ② 分割させた2000年時点の区を2005年時点の区分けを用いて合算
- ③ 合算させた人口を2005年人口と対応させ線形補間

図 7. 分割後合併時に関する市区町村変遷対応表と市区町村変遷・分割の例

市	区名	2010年総人口	DATE1	JISCODE1	GNAME1	CNAME1	WEIGHT	DATE2	JISCODE2	GNAME2	CNAME2
浜松市	中区	119044	20051001	22202		浜松市	0.029425	20101001	22131	浜松市	中区
	東区	63053	20051001	22202		浜松市	0.03097	20101001	22132	浜松市	東区
	西区	56162	20051001	22202		浜松市	0.056619	20101001	22133	浜松市	西区
	南区	51420	20051001	22202		浜松市	0.031523	20101001	22134	浜松市	南区
	北区	46260	20051001	22202		浜松市	0.180607	20101001	22135	浜松市	北区
	浜北区	44915	20051001	22202		浜松市	0.044198	20101001	22136	浜松市	浜北区
	天竜区	16292	20051001	22202		浜松市	0.626657	20101001	22137	浜松市	天竜区

図 8. 浜松市人口割合

図 9. 市町変遷対応表(一部抜粋)

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))
分担研究報告書

Probabilistic linkage を用いた大規模公的統計データベースの活用に関する研究

研究協力者 福井敬祐 大阪府立成人病センターがん予防情報センター 研究員
研究分担者 近藤尚己 東京大学大学院医学系研究科 准教授

研究要旨

我が国の人口動態特殊報告を用いた職業・産業別死亡情報は、死亡時職業による分類であるため、実際に長く従事した職業や産業による影響を正確に計ることができないことが懸念されてきた。国勢調査における職業・産業情報と人口動態統計を突合することができれば、その問題点の解決につながる。個人識別番号によるリンケージが不可能な大規模統計データ同士を、個人単位でリンケージする手法として、複数の変数同士の一一致確率を統計的に算出して行う probabilistic record linkage の理論と実際について整理した。Probabilistic record linkage においては、リンケージを行うデータベース間での共通のマッチング変数が重要となる。しかしながら、本研究で想定している国勢調査および人口動態統計の二次利用データで入手可能な変数では、あまり正確にリンケージできない可能性がある。また、住所等の情報の入り方の問題もあるため、かなりの作業量を要することが予想された。将来的には Probabilistic record linkage に頼る必要がないよう個人識別番号による突合が可能となるような基盤整備が必要であることが示唆された。

A . 研究目的

我が国の人口動態特殊報告を用いた職業・産業別死亡情報は、死亡時職業による分類であるため、実際に長く従事した職業や産業による影響を正確に計ることができないことが懸念されてきた。国勢調査における職業・産業情報と人口動態統計を突合することができれば、その問題点の解決につながるが、現状では氏名・生年月日・住所等、個人をつなぐ情報を使うことができないため、完全なリンケージデータの作成は困難である。

二次利用申請により提供可能である限ら

れた個人属性を用いて、国勢調査データと人口動態調査データをリンケージすることが可能であるかどうか、諸外国で活用されている Probabilistic Record Linkage 法について紹介し、我が国における活用可能性を検討する。

B . 研究方法

Probabilistic Record Linkage を含む一般的な Record Linkage 手法について、その実行方法を調査し、整理した。特に、Probabilistic Record Linkage については理論的な背景を中心にまとめ、実行可能なソフ

トウェア等についても紹介した。

(倫理面の配慮)

本報告では実際のデータを用いていないため、倫理面での問題は生じない。

C. 研究結果

1. Record Linkage の概要

今、2つのデータセットA、Bを考える。データセットを構成する各データはレコードと呼ばれ、各データセットの中からペアとなるレコードを作成することがRecord Linkageの目的である。Linkageに使用する各データセットに共通した変数をMatching Variables (MV)と呼び、MVに基づいてLinkageを行う。例えば、Record Linkage法として最も単純なDeterministic Record LinkageにおいてはMVの値が同じレコードをLinkageする。図1はデータセットAとデータセットBをRecord Linkageを行う場合の例である。ここではIDと名前をMVとして、レコード*a*とレコード*b*をLinkageしている。

一般的なLinkageの手順は研究毎に様々であるが(例えば、[1, 2, 3]など)それらの方法は簡単に以下の手順に包含することができる(図2)

Data cleaning and standardization

データの整形にあたる部分である。Linkageを行うデータの誤った入力の修正や、データ間で異なった入力値を共通なものへと変換し、標準化する。

Blocking

各データの間でのLinkage作業における比較数を削減するため、データをBlockと呼ばれるいくつかのグループに

分割する。Linkageは対応するBlock間のみで行われる。例えば、性別という変数を共通に持つデータAとデータBをリンケージする場合には、あらかじめ男性のBlock、女性のBlockというようにBlockingを行う。このようにすることで、Linkageの際には男性のBlock同士、女性のBlock同士のみを比較すればよく、計算量を削減することができる。

Linkage

Blockingによって作成されたBlock間でMVを基にしてLinkage作業を行う。Linkageの方法としてはDeterministicな方法とProbabilisticな方法がある。

Deterministic法はMVが一致するレコード同士をペアとしてLinkageする方法であり、Probabilistic法はMVを基にして作ったWeightを使ってLinkageを行う方法である。Deterministic法は非常に単純な方法であり、一般的にMVを種別化・順序化し、データセット同士で結合作業をすればよい。しかし、例えばMVがユニークでない値を含む場合や欠損値が存在する場合などにはLinkageの精度は大きく低下する。一般的にRecord Linkageを行うデータ量は巨大になりがちであり、データの質の担保が困難な場合が多い。そのため、Deterministic法を用いたRecord Linkageでは期待した通りの結果が得られないことが多い。

Probabilistic法はMVを基に作成された重み(Weight)を利用してLinkageを行う。WeightはLinkさせたレコードが真のペアであるか否かの確率を反映しており、MVの値をExactに用いないため、データの質が低い場合でも使用できる可能性がある。

Clerical review

Linkage 作業で Linked もしくは Non-Linked に判別されないようなデータを人為的に判別することを指す。また、Linkage 作業に用いた閾値などのパラメータ設定やソフトウェアの実行が正しいものであったのかを判別することも含む。

Evaluating data quality

省略

2. Probabilistic Record Linkage について
Probabilistic Record Linkage の基本的な考え方は[4]により提案された。今、 A 、 B を Linkage 対象のデータ、 $a \in A$ 、 $b \in B$ を任意のそれぞれのレコードとする。このとき、直積集合 $A \times B = \{(a, b) | a \in A, b \in B\}$ の 2 つの部分集合を

$$M = \{(a, b) | a = b, a \in A, b \in B\},$$
$$U = \{(a, b) | a \neq b, a \in A, b \in B\},$$

とすれば、 M は真に Link 関係にあるレコードの集合、 U は真には Link 関係にないレコードの集合を表す。また、 $t(a, b)$ をレコード a, b の一致度を測る一致度ベクトルとし、レコード a, b が真に Link 関係にあるとき、一致度ベクトルが t となる確率を $m(t)$ と定義する。すなわち、

$$m(t) = P(t(a, b) = t | (a, b) \in M).$$

同様に、レコード a, b が真に Link 関係にないときに、一致度ベクトルが t となる確率を $u(t)$ で定義する。すなわち、

$$u(t) = P(t(a, b) = t | (a, b) \in U).$$

このとき、[4]は次の Weight を用いて Link が否かを決定する方法を提案している。

$$w(t) = \log \frac{m(t)}{u(t)}.$$

実際には、一致度ベクトルが取り得る値 t_1, \dots, t_k の全てに対して $w(t_1), \dots, w(t_k)$ を計算し、あらかじめ設定した閾値と比較するという方法をとる。上記の Weight の計算には、レコード内の角変数に独立性を仮定し、EM アルゴリズムを利用する方法が提案されている(参照[5])。

3. 応用ソフトウェア等

Probabilistic Record Linkage を行うことができるソフトウェアは有償・無償のものも含めて様々開発がされている。また一般的な統計ソフトウェアのパッケージとして提供されているなど、導入しやすい。例えば、National Program of Cancer Registries (NPCR) によって開発・提供されている無償の Record Linkage ソフトウェア Link Plus (図 3)は GUI ユーザーインターフェイスで直感的な操作が可能であり、比較的簡単に Record Linkage が可能である。他にも Record Linkage を行える GUI ベースのソフトとしては、無償のものでは D-Dupe、DuDe、Merge Tool Box などがある。その他のソフトについては[6]を参照されたい。また、統計解析ソフトウェアの R 言語における RecordLinkage パッケージは多少のプログラミング知識を有するが、多量なデータを自動で Linkage したい場合や Linkage したデータを直接分析する必要がある場合などに有用である。

これらのソフトウェアを用いる上での注意点としては、無償版のソフトウェアのほとんどが日本語に完全に対応しているわけではないということである。そのため、入力値や MV として日本語が含まれるデータの Linkage の際にはその精度に対して十分な

注意が必要である。

4. 実適用について

Probabilistic Record Linkage の基礎理論は上述したとおり、[4]により提案されており、その歴史は長く、海外を中心に活用されている（例えば、[7,8,9]など）。一方で、日本における適用例はまだ少なく、[10]においては、Record Linkage を必要とする分野がそれぞれ領域固有の知識を必要とするために、学術的な一般化が難しかったこと、我が国では戸籍制度が整備されており、国勢調査等で人物同定の必要性がほとんどなかったという社会背景をその理由として挙げている。

D . 考察

個人識別番号によるリンケージが不可能な大規模統計データベースを、個人単位でリンケージする手法として、複数の変数同士の一致確率を統計的に算出して行うProbabilistic Record Linkageの理論と実際について調べた。

Record Linkage法は上記に上げた手順～の作業を必要とするが、応用ソフトウェアではこれらの作業を支援することができるものがほとんどである。しかし、Linkageに有効なMVとして用いられがちな氏名や住所等の情報は基本的に日本語での入力が行われるため、ソフトウェアでMVとして正確に作用しない可能性がある。

本研究でLinkageを考える人口動態統計と国勢調査データを使用する場合、現行でLinkageに使用可能な変数は以下の通りである。

人口動態統計

- 事件簿番号

- 性別
- 生年月日
- 死亡年月日
- 死亡した人の住所
死亡票：市区町村コード
オンライン報告分：詳細住所

国勢調査

- 性別
- 生年月
- 調査区番号
- 市区町村（町丁目）情報

氏名や個人識別番号などの利用が困難である我が国の現状においては、利用可能な変数は有用なMVとなりやすく、ユニークでない組み合わせがかなり存在するリンケージデータとなる可能性が高い。住所情報に関しても共通のコード化などの工夫が必要となり、かなりの作業量を要することが想定される。

相当な作業量を要する上に、そのリンケージデータの精度が低いことが想定されるため、将来的には、北欧諸国や英国、米国のように、個人識別番号の整備を経て、各種公的統計のリンケージを公的機関が行い、個人識別可能な情報を削除した匿名化データを利用者に提供する仕組みが必要であると考えられる。

E . 結論

Probabilistic Record Linkageの実行においては、現状の日本の国勢調査と人口動態統計データでは有用な共通のマッチング変数が利用可能でないため、精度の低いリンケージデータとなる可能性がある。Probabilistic Record Linkageにおける理論やソフトウ

エアの整備が進む一方で、得られる結果の整合性を考慮すれば、将来的には各種統計データベース間での共通個人識別番号の整備およびその活用について、検討していく必要がある。

F . 健康危険情報

G . 研究発表

1 . 論文発表

2 . 学会発表

H . 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

引用文献

[1] Gu, L., Baxter, R., Vickers, D. and Rainsford, C. (2003). Record linkage: Current practice and future directions. *CSIRO Mathematical and Information Sciences Technical Report*, **3**, 83.

[2] Elfeky, M. G., Verykios, V. S. and Elmagarmid, A. K. (2002). TAILOR: A record linkage toolbox. In *Data Engineering, 2002. Proceedings. 18th International Conference on* (pp. 17-28). IEEE.

[3] Lee, M. L., Ling, T. W. and Low, W. L. (2000). IntelliClean: a knowledge-based intelligent data cleaner. In *Proceedings of the sixth ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 290-294). ACM.

[4] Fellegi, I. P. and Sunter, A. B. (1969). A theory for record linkage. *Journal of the American Statistical Association*, **64**(328), 1183-1210.

[5] Bauman, G. J. (2006). Computation of weights for probabilistic record linkage using the EM algorithm.

[6] Christen, P., (2012). *Data Matching: Concepts and Techniques for Record Linkage, Entity Resolution, and Duplicate Detection*, Springer Science and Business Media.

[7] Whop, L. J., Diaz, A., Baade, P., Garvey, G., Cunningham, J., Brotherton, J. M., ... & Moore, S. P. (2016). Using probabilistic record linkage methods to identify Australian Indigenous women on the Queensland Pap Smear Register: the National Indigenous Cervical Screening Project. *BMJ open*, **6**(2).

[8] Kesinger, M. R., Kumar, R. G., Ritter, A. C., Sperry, J. L., & Wagner, A. K. (2016). Probabilistic Matching Approach to Link Deidentified Data from a Trauma Registry and a Traumatic Brain Injury Model System Center. *American Journal of Physical Medicine & Rehabilitation*.

[9] Adam, M., Kuehni, C. E., Spoerri, A., Schmidlin, K., Gumy-Pause, F., Brazzola, P., ... & Zwahlen, M. (2015). Socioeconomic Status and Childhood Leukemia Incidence in Switzerland. *Frontiers in oncology*, **5**.

[10] 相澤彰子, 高須淳宏, 大山敬三, & 安達淳. (2004). 異種データベース間でのレコード照合に関する研究動向. *NII journal*, (8), 43-51.

データセットA				データセットB			
ID	名前	変数1	変数2	ID	住所	名前	変数3
12	B			1	MV	E	
13	A		レコードa	3		F	
14	C			13		A	レコードb
15	D			26		G	

図 1 Record Linkage の例

データセット A 内のレコード a とデータセット B 内のレコード b を ID と名前を基に Linkage



図 2 Record Linkageのフロー

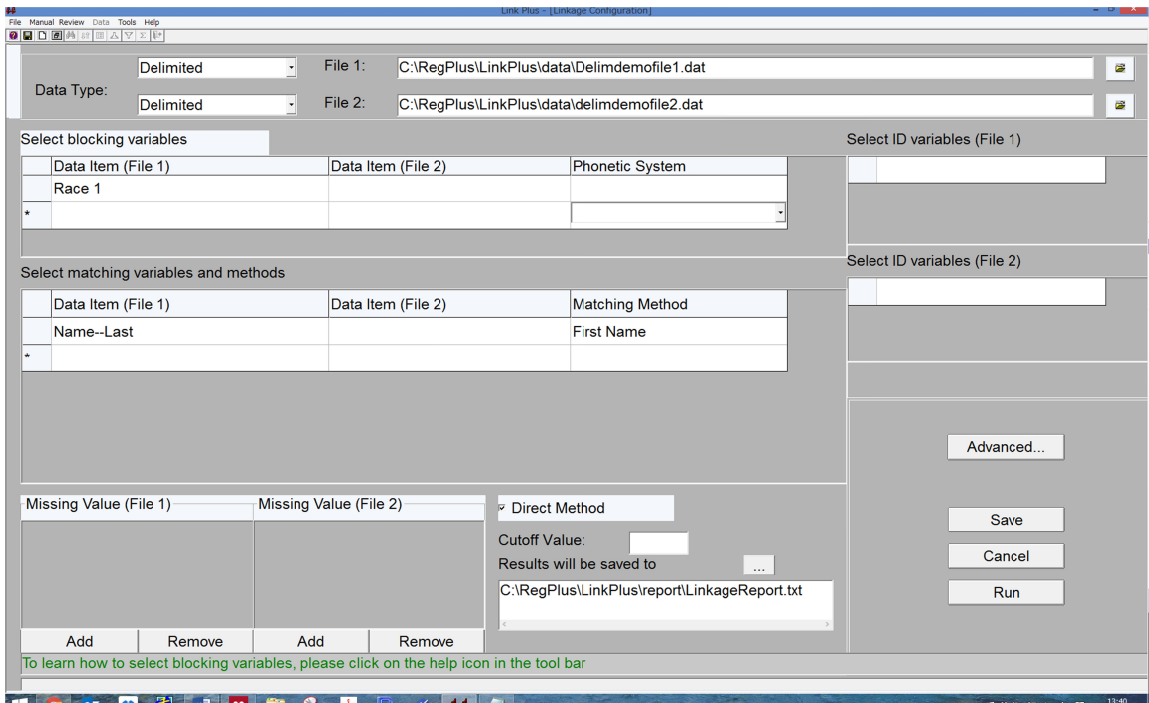


図 3 Link Plus のインターフェイス画面

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))
分担研究報告書

がん進行度別罹患率における社会経済格差

研究代表者	伊藤ゆり	大阪府立成人病センター	がん予防情報センター	主任研究員
研究分担者	近藤尚己	東京大学大学院医学系研究科		准教授
研究分担者	中谷友樹	立命館大学文学部(立命館大学歴史都市防災研究所 兼任)		教授
研究協力者	米島万有子	立命館大学衣笠総合研究機構		専門研究員
研究協力者	中山富雄	大阪府立成人病センター	がん予防情報センター	疫学予防課長
研究協力者	福井敬祐	大阪府立成人病センター	がん予防情報センター	研究員

研究要旨

既存統計資料のひとつであるがん登録資料を用いて、がん罹患率における社会経済格差のトレンドを評価することを目的とする。大阪府がん登録資料より、大阪府において1993-2004年に診断された胃、大腸、肺、乳房、子宮頸部、前立腺のがん患者の情報を入手し、居住地(町字単位)を基に地理的剥奪指標(Areal Deprivation Index: ADI)を付与した。ADIは数値が大きいほど地域の剥奪度が高い、つまり社会的に不利な経済状況にある人々の割合が高いという指標で、各地域の人口で重み付けし5分位に分けた。がん罹患率は検診受診等の予防行動に影響を受けるため(検診受診率の高い地域の罹患率が高い)、診断時の進行度別(早期がん: 上皮内+限局/進行がん: 領域+遠隔転移)の年齢調整罹患率をADIごとに算出した。部位・性別・診断時期別に、分散重み付き最小二乗法によりADIと進行度別罹患率の関連を分析した。前立腺がん以外の全ての部位で、進行がんの罹患率はADIの高い地域ほど高かった。ADI第5分位と第1分位間の進行がんの罹患率差は、男女とも1993-1998年診断の肺がんが最大で、その差は1999-2004年には統計的に有意に減少した。他の部位の進行がん罹患率においても、統計的に有意には拡大しなかった。一方、早期がんの罹患率は、男性で前立腺、胃、大腸において、ADIの高い地域ほど罹患率が低かった。その傾向は前立腺で顕著であった。女性では子宮頸がんにおいて、ADIの高い地域の罹患率が高かった。格差の縮小が最も大きかったのは、男性の胃・大腸の進行がんであったが、この傾向は女性ではみられなかった。本研究により明らかとなった進行度別がん罹患率の社会経済格差には、喫煙やハイリスクな性行動など、がん発症のリスクとなりうる行動の違いやがん検診の受診率の違いなどが影響していると考えられる。

A . 研究目的

平成 28 年 1 月 1 日より、がん登録等推進法に基づき、全国がん登録制度が開始した。地域がん登録資料は 1960 年代から一部の府県において収集されている既存統計資料であり、今後ますます健康政策への活用が求められている。健康格差の縮小が健康日本 21 の目標に掲げられ、健康格差の評価が求められているが、地域がん登録資料を用いた健康格差のモニタリング体制は十分ではない。本研究は既存統計資料である大阪府がん登録資料を用いて、がん罹患率における社会経済格差のトレンドを評価することを目的とする。

B . 研究方法

大阪府において 1993-2004 年に診断された胃、大腸、肺、乳房、子宮頸部、前立腺のがん患者の情報に対し、居住地（町字単位）を基に地理的剥奪指標（Areal Deprivation Index: ADI）を付与した。ADI は数値が大きいほど地域の剥奪度が高い、つまり社会的に不利な経済状況にある人々の割合が高いという指標である。これを各地域の人口で重み付けし、5 分位に分けた。第 1 分位が最も剥奪度が低く裕福な地域（高 SES 群）、第 5 分位が最も剥奪度が高い地域（低 SES 群）。がん罹患率は検診受診等の予防行動に影響を受ける。例えば、検診受診率の高い地域では、より多くの患者ががんと診断されるため、罹患率が高くなる。そのため、診断時の進行度を早期がん：上皮内 + 限局、進行がん：領域 + 遠隔転移に区分し、進行度別の年齢調整罹患率を ADI5 分位グループごとに算出した。部位・性別・診断時期別に、分散重み付き最小二乗法により ADI と進行度別罹患率の関

連を分析し、ADI 第 5 分位地域と第 1 分位地域の罹患率差を推定した（Model 1）。

$$\text{Model 1: } ASIR_i = \beta_{ADI} a_i + e_i$$

また、診断時期により格差が拡大したか、縮小したかについて検討するために、診断時期と ADI との交互作用項で評価した（Model 2）。

Model 2:

$$ASIR_{ij} = \beta_{ADI} a_i + \beta_{per} p_j + \beta_{ADIPer} a_i p_j + e_{ij}$$

（倫理面への配慮）

大阪府がん登録より入手した情報は匿名化された情報であるが、居住地住所等の情報を含むデータを扱う際は、スタンドアロン環境の PC において作業を行う。また社会経済指標等と突合後には個人同定が不可能な状態に変換する。

C . 研究結果

性別・部位別・診断時期別・進行度別に ADI 第 5 分位（低 SES 群）と第 1 分位（高 SES 群）の年齢調整罹患率の差を分散重み付き最小二乗法における推定値を元に算出し図 1、2 に示した。進行度別年齢調整罹患率の分位ごとの傾向は性別・部位別に図 3 ~ 11 および表 1 に示した。前立腺がん以外の全ての部位で、進行がんの罹患率は ADI の高い地域（低 SES 群）ほど高かった。ADI 第 5 分位（低 SES 群）と第 1 分位（高 SES 群）間の進行がんの罹患率差は、男女とも 1993-1998 年診断の肺がんが最大で、それぞれ人口 10 万人対 12.0（95%信頼区間：9.4-14.5）5.7（4.4-7.0）であった。しかし、その差は 1999-2004 年には統計的に有意に減少し、他の部位の進行がん罹患率においても、統計的に有意には拡大しなかった。一方、

早期がんの罹患率は、男性で前立腺、胃、大腸において、ADIの高い地域(低SES群)ほど罹患率が低かった。その傾向は前立腺で顕著であり、1999-2004年ではADI第5分位と第1分位間の罹患率差は-7.3と拡大した。女性では子宮頸がんにおいて、ADIの高い地域(低SES群)の罹患率が高かった。格差の縮小が最も大きかったのは、男性の胃・大腸の進行がんであったが、この傾向は女性ではみられなかった。

D. 考察

大阪府がん登録資料より、前立腺以外の主要部位の進行がんの罹患率で社会経済格差が生じていることが明らかとなった。進行がんにおいては低SES群の罹患率は高SES群よりも高く、男性の早期がんでは、その逆の関連が見られた。

進行がんにおける社会経済格差の要因は、喫煙率やハイリスクな性行動などがん発症のリスクとなりうる行動の違いや、がん検診受診やがんの自覚症状への気づき(awareness)から医療機関へのアクセスの違いなど様々な要因が融合して影響していると考えられる。例えば、喫煙率に関しては平成26年度国民健康栄養調査において、男女とも世帯年間収入が600万円以上の群に比べ、200万円以上600万円未満、200万円未満の群の喫煙率が統計的に有意に高かった。¹喫煙をリスクとする肺がん罹患率の社会経済格差は喫煙率の違いにより部分的に説明可能であるといえる。

また、がん検診受診率についても、国民生活基礎調査をもとに、医療保険別にみた場合、共済組合加入者は市町村国民健康保険加入者よりも25~40ポイント受診率が高い傾向にあった。²男性において、高SES

群の方が低SES群より早期がん罹患率が高い理由としては、裕福な地域に住む住民の方が職場をはじめ、がん検診を受診する機会が多く、早期に診断された可能性が示唆された。しかし、この逆の相関関係は、女性では観測されなかった。女性においては、常勤として従事している人の割合が低いいため、職場における検診受診体制の差の影響が男性ほど大きく出なかったと推察される。一方、進行がんの罹患率においては、女性ではどのがんにおいても低SES群の方が高SES群より高く、肺がん以外ではその格差は縮小していなかった。

米国の8つのがん登録データを用いた地域別進行乳がん罹患率は、マンモグラフィ施設集中度や高学歴者割合、英語識字率吐逆相関し、黒人割合と相関していた³。また、米国9つのがん登録データより、County-levelの剥奪指標4分位を用いて、詳細のステージ別罹患率および死亡率を分析した研究においては、高SES群の上皮内がん罹患率が低SES群よりも高く、その差は拡大傾向にあった。それ以外の進行がん罹患率や死亡率に関しては、格差が顕著ではなかった。⁴米国における子宮頸がんに関しては、死亡率・進行がん罹患率ともに格差が拡大傾向にある報告もある。⁵

がん登録資料より、診断されたがん患者内における進行度分布における進行がんをアウトカムにした研究においても、SESが進行がんでの診断に影響を与える結果が各種報告されている⁶⁻⁸。しかし、この研究デザインの場合、過剰診断による影響を大きく受けるため、人口を分母とした本研究のような進行度別罹患率による検討が望ましい。

がんにおける総合的なアウトカム指標は

死亡率である。がんにおける社会経済格差を検討する際には、がん進行度別罹患率だけでなく死亡率を最終アウトカムとし、他に生存率および喫煙率、検診受診率、医療アクセスなど各種関連指標と総合的にその関係性を分析しメカニズムを解明する必要がある。その上で、がんによる死亡の社会経済指標による格差を縮小するための効果的な手立てを検討できるといえよう。

現時点（H28年3月）で、利用可能なデータとして、本研究のように町字レベルまで使用可能であるのは、がん登録資料（生存率および医療アクセス）および近年の人口動態統計オンライン届出分のデータに限られる。国民生活基礎調査から得られる喫煙率およびがん検診受診率は都道府県レベル、自治体で行われるがん検診受診率は地域保健・健康増進事業報告の市町村レベルのデータに限定される。これらの指標に関しては、自治体におけるデータ提供体制の整備が急務であるが、サンプル調査に関しては空間的マイクロシミュレーションを用いた推定法を適用し、小地域における推定結果を用いる必要がある。

E . 結論

大阪府がん登録資料より、進行度別がん罹患率の社会経済格差を分析した。高SES群に比べ低SES群では進行がんの罹患率が高かった。男性において、高SES群において早期がん罹患率が高い傾向も観測された。進行度別罹患率の社会経済格差は喫煙やハイリスクな性行動など、がん発症のリスクとなりうる行動の違いやがん検診の受診率の違いなどが影響していると考えられ、さらなる検討が必要であることが示唆された。

F . 健康危険情報

なし

G . 研究発表

1 . 論文発表

Ito Y, Nakaya T, Ioka A, Nakayama T, Tsukuma H, Uehara S, Sato KK, Endo G, Hayashi T: Investigation of Spatial Clustering of Biliary Tract Cancer Incidence in Osaka, Japan: Neighborhood Effect of a Printing Factory. J Epidemiol 2016, [in press].

Kinoshita F, Ito Y, Nakayama T: Trends in lung cancer incidence rates by histological type in 1975-2008: a population-based study in Osaka, Japan. J Epidemiol 2016;in press].

伊藤ゆり, 中山富雄: 肺がん生存率の国際比較. 肺癌 2015, 55:266-272.

2 . 学会発表

伊藤ゆり, 中谷友樹, 近藤尚己, 福井敬祐, 中田佳世, 井岡亜希子, 宮代勲, 中山富雄. 大阪府におけるがん進行度別罹患率の社会経済格差: 1993-2004年における格差の変化. 第74回日本公衆衛生学会総会. 2015:402 (P-0802-10). 長崎

Ito Y, Nakaya T, Kondo N, Fukui K, Nakaya K, Ioka A, Miyashiro I, Nakayama T, Rachet B. SOCIO-ECONOMIC DIFFERENCES IN STAGE-SPECIFIC CANCER INCIDENCE IN OSAKA, JAPAN: 1993-2004. 37th International Association of Cancer Registries, Annual

Scientific Conference 2015. 8-10. Oct. 2015 [9th Oct. 2015, Oral Presentation, CANCER CONTROL: Data and Studies (Track 1)] 2015:054 [O179]. Mumbai, India

H . 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

引用文献

1. 平成 26 年国民健康・栄養調査報告 [homepage on the Internet]; c2016 [cited 2016 28 Apr]. Available from: <http://www.mhlw.go.jp/stf/houdou/2r98520000020qbb.html>
2. 田淵貴大, 中山富雄, 津熊秀明. 日本におけるがん検診受診率格差 医療保険のインパクト. 日本医事新報. 2012:84-8. (in 日本語).
3. Tatalovich Z, Zhu L, Rolin A, Lewis DR, Harlan LC, Winn DM. Geographic disparities in late stage breast cancer incidence: results from eight states in the United States. Int J Health Geogr. 2015;14:31. (in eng).
4. Schootman M, Lian M, Deshpande AD, Baker EA, Pruitt SL, Aft R, et al. Temporal trends in area socioeconomic disparities in breast-cancer incidence and mortality, 1988-2005. Breast Cancer Res Treat. 2010;122:533-43. (in eng).
5. Simard EP, Fedewa S, Ma J, Siegel R, Jemal A. Widening socioeconomic disparities in cervical cancer mortality among women in 26 states, 1993-2007. Cancer. 2012;118:5110-6. (in eng).
6. Chatterjee NA, He Y, Keating NL. Racial differences in breast cancer stage at diagnosis in the mammography era. Am J Public Health. 2013;103:170-6. (in eng).
7. Islami F, Kahn AR, Bickell NA, Schymura MJ, Boffetta P. Disentangling the effects of race/ethnicity and socioeconomic status of neighborhood in cancer stage distribution in New York City. Cancer Causes Control. 2013;24:1069-78. (in eng).
8. Chor JS, Lam HC, Chan A, Lee HM, Fok E, Griffiths S, et al. Socioeconomic disparity in breast cancer detection in Hong Kong--a high income city: retrospective epidemiological study using the Breast Cancer Registry. PLoS One. 2014;9:e107630. (in eng).

表1. 大阪府における進行度別罹患率の社会経済指標による格差: 1993-2004年(時代変化)

	1993-1998													1999-2004																
	年齢調整罹患率(実測)					年齢調整罹患率(あてはめ)					格差(Q5-Q1)			年齢調整罹患率(実測)					年齢調整罹患率(あてはめ)					格差(Q5-Q1)						
	Q1	Q2	Q3	Q4	Q5	Q1	Q2	Q3	Q4	Q5	95% CI	Q1	Q2	Q3	Q4	Q5	Q1	Q2	Q3	Q4	Q5	95% CI								
男性 早期がん																														
胃	24.4	26.2	24.3	24.8	22.7	25.5	25.0	24.4	23.9	23.4	-2.0	-3.6	-0.5	27.4	27.5	25.8	26.6	23.7	27.9	27.0	26.2	25.3	24.5	-3.4	-5	-1.8				
大腸	19.5	19.5	19.8	19.8	18.0	19.9	19.6	19.3	19.0	18.7	-1.2	-2.6	0.2	25.0	22.0	21.9	23.0	20.3	24.0	23.2	22.4	21.5	20.7	-3.3	-4.8	-1.8 *				
肺	5.9	6.8	6.6	6.8	7.1	6.2	6.4	6.6	6.9	7.1	0.9	0.1	1.7	10.1	10.4	9.4	9.3	9.5	10.2	10.0	9.7	9.5	9.3	-0.9	-1.9	0.1 *				
前立腺	4.7	4.2	3.4	3.4	2.9	4.5	4.1	3.7	3.3	2.9	-1.6	-2.3	-1	12.3	9.4	9.1	9.0	6.5	11.6	10.4	9.2	8.0	6.8	-4.8	-5.8	-3.9 *				
進行がん																														
胃	22.6	25.8	28.3	28.8	29.3	23.6	25.3	26.9	28.6	30.2	6.6	5	8.2	26.6	27.2	29.7	28.7	29.4	26.9	27.6	28.3	29.0	29.7	2.9	1.2	4.6 *				
大腸	13.5	15.3	17.8	18.3	18.0	14.1	15.3	16.5	17.7	18.9	4.8	3.5	6	18.3	20.4	20.4	22.1	20.7	19.1	19.7	20.3	21.0	21.6	2.5	1.1	3.9 *				
肺	24.5	27.3	28.5	30.9	32.6	24.8	26.8	28.8	30.8	32.8	8.0	6.3	9.7	32.2	32.8	34.3	35.5	37.6	31.8	33.1	34.5	35.9	37.2	5.4	3.6	7.3 *				
前立腺	4.1	4.4	4.0	3.8	3.9	4.3	4.2	4.0	3.9	3.8	-0.5	-1.1	0.2	6.3	5.5	5.1	5.4	5.0	5.9	5.7	5.4	5.2	4.9	-1.0	-1.8	-0.3				
女性 早期がん																														
胃	8.8	9.5	10.1	10.0	10.2	9.0	9.4	9.7	10.1	10.4	1.4	0.5	2.4	9.6	9.0	9.0	9.2	10.1	9.1	9.2	9.4	9.5	9.6	0.5	-0.4	1.4				
大腸	10.0	9.4	10.9	11.0	9.6	10.0	10.1	10.1	10.2	10.2	0.2	-0.7	1.1	12.8	11.6	11.3	12.3	11.2	12.3	12.0	11.8	11.6	11.4	-0.9	-1.9	0.1				
肺	2.4	1.9	2.5	2.6	2.7	2.1	2.3	2.4	2.6	2.7	0.6	0.1	1.1	4.4	4.0	4.3	4.3	4.2	4.3	4.2	4.2	4.2	4.2	-0.1	-0.6	0.5				
乳房	19.4	19.0	19.6	21.2	18.6	19.4	19.5	19.5	19.6	19.6	0.2	-1.3	1.6	24.5	22.8	23.3	24.4	24.0	23.6	23.7	23.8	23.9	23.9	0.3	-1.3	1.9				
子宮頸	7.3	7.9	8.7	8.7	9.2	7.4	7.9	8.4	8.8	9.3	1.9	0.9	2.9	6.8	7.8	6.9	8.0	7.9	6.9	7.2	7.4	7.7	8.0	1.0	0.1	2				
進行がん																														
胃	9.9	11.8	12.3	11.9	12.7	10.5	11.1	11.7	12.3	12.9	2.3	1.3	3.4	9.1	11.9	10.8	11.9	11.8	9.9	10.5	11.1	11.6	12.2	2.3	1.3	3.2				
大腸	9.2	11.1	11.5	12.0	12.1	9.8	10.4	11.1	11.8	12.5	2.7	1.8	3.7	11.5	12.2	13.2	13.7	13.4	11.8	12.3	12.8	13.3	13.8	2.1	1	3.1				
肺	6.7	7.9	8.9	9.2	10.9	6.8	7.7	8.7	9.7	10.6	3.8	2.9	4.7	8.8	9.7	9.8	10.1	11.2	8.9	9.4	9.9	10.4	10.9	2.0	1.1	2.9 *				
乳房	13.8	14.5	14.6	15.2	16.0	13.8	14.3	14.8	15.3	15.8	2.1	0.8	3.3	13.9	15.0	15.9	17.6	16.4	14.2	15.0	15.7	16.5	17.3	3.1	1.8	4.3				
子宮頸	2.5	2.8	3.7	3.4	4.2	2.5	2.9	3.3	3.7	4.1	1.6	1	2.2	2.1	2.9	3.0	2.9	3.5	2.3	2.6	2.9	3.2	3.4	1.1	0.6	1.7				

年齢調整罹患率は人口10万人対

*は診断期間で格差の大きさに統計的に有意に変化があったもの(p<0.05)

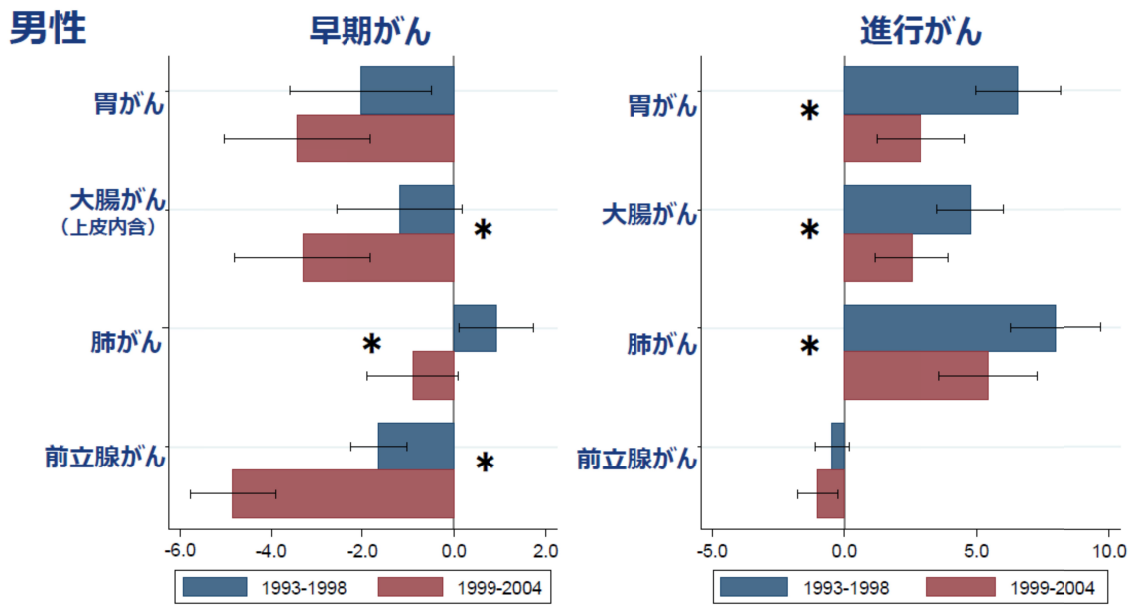


図 1 . 大阪府男性における部位別進行度別がん罹患率の格差 (Q5 の年齢調整罹患率と Q1 の年齢調整罹患率): 1993 - 1998 年 / 1999 - 2004 年

* は、診断時期間で格差が統計的有意に変化した部位 (p<0.05)

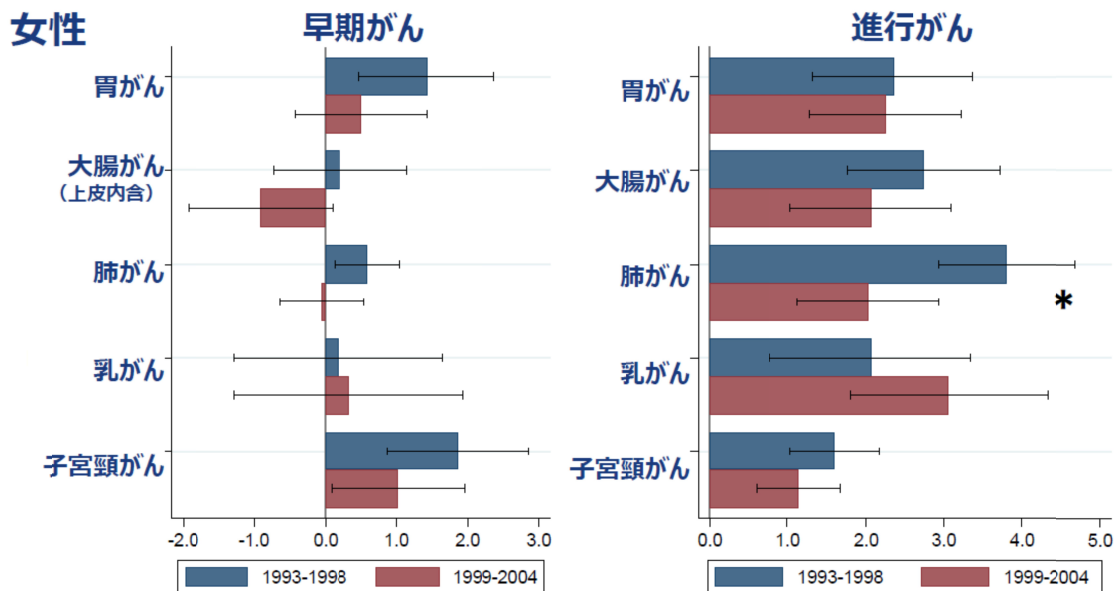


図 2 . 大阪府女性における部位別進行度別がん罹患率の格差 (Q5 の年齢調整罹患率と Q1 の年齢調整罹患率): 1993 - 1998 年 / 1999 - 2004 年

* は、診断時期間で格差が統計的有意に変化した (p<0.05)



図 3 . 大阪府胃がん男性の ADI5 分位別進行度別年齢調整罹患率：1993～1998 年 / 1999～2004 年（直線は分散重み付き最小二乗法によるあてはめ）

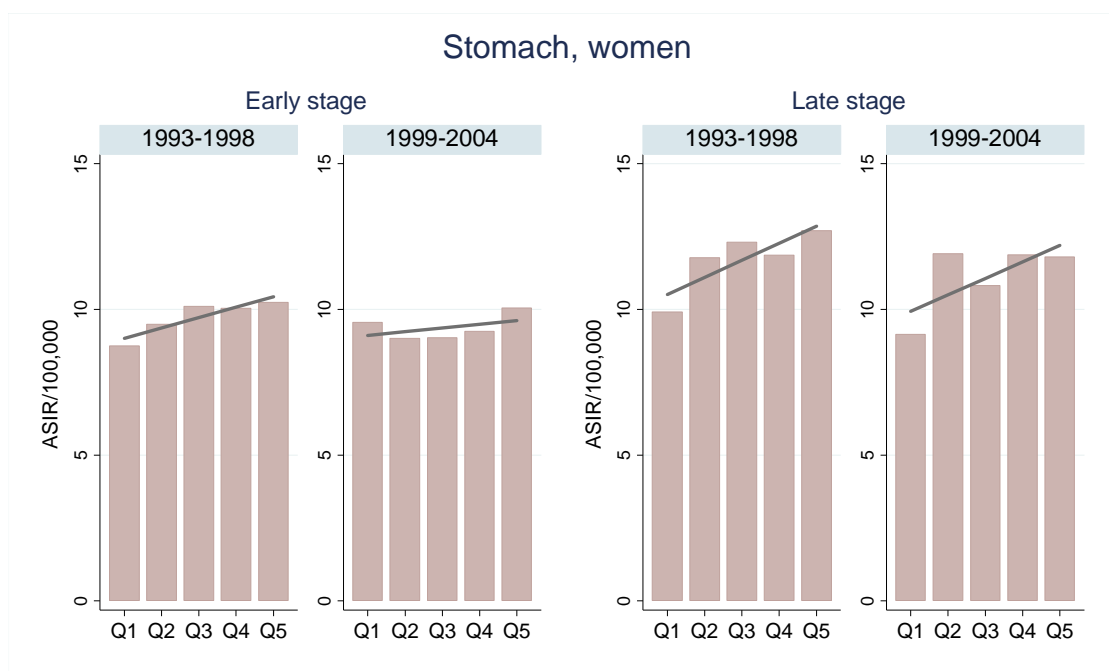


図 4 . 大阪府胃がん女性の ADI5 分位別進行度別年齢調整罹患率：1993～1998 年 / 1999～2004 年（直線は分散重み付き最小二乗法によるあてはめ）

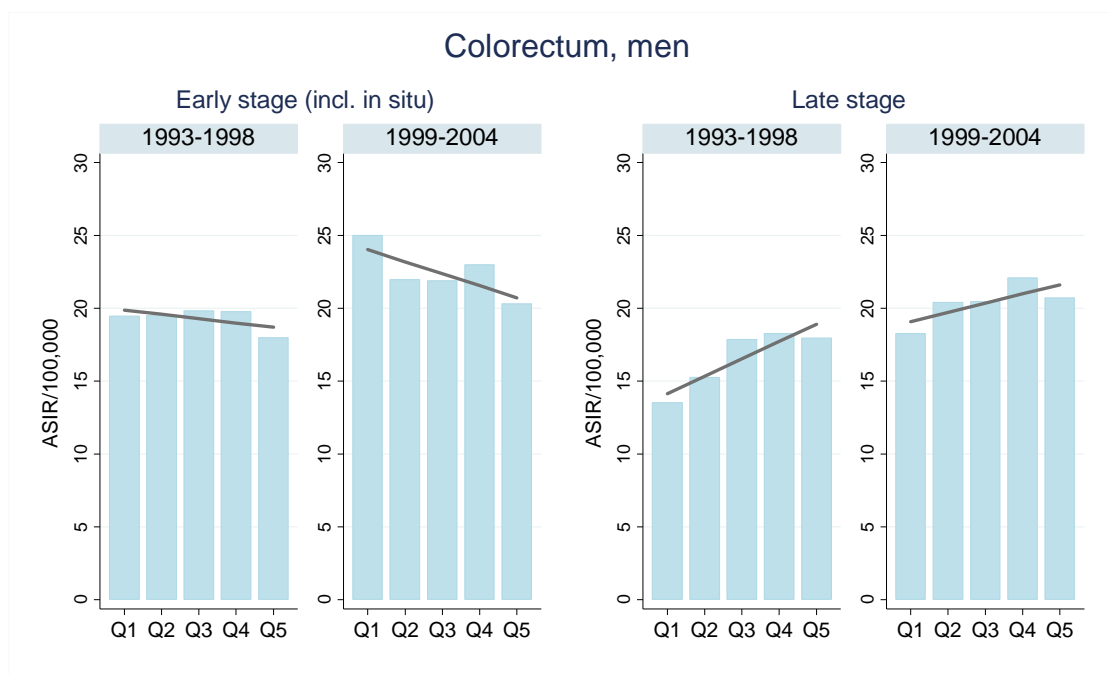


図5 .大阪府大腸がん男性のADI5分位別進行度別年齢調整罹患率：1993～1998年 / 1999～2004年（直線は分散重み付き最小二乗法によるあてはめ）

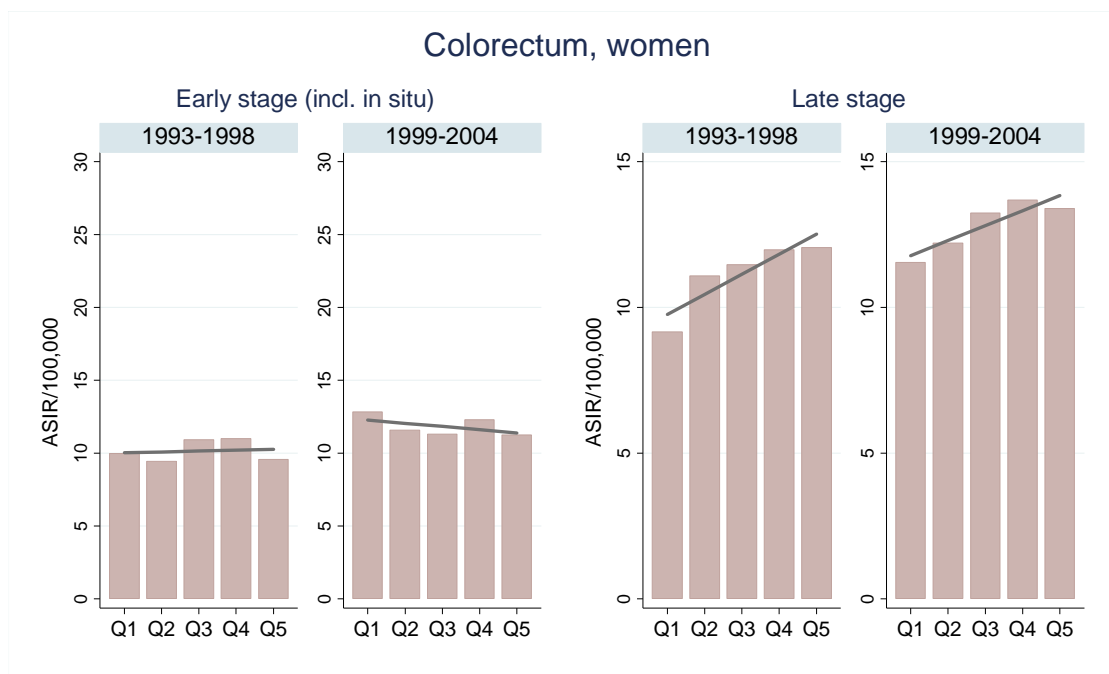


図6 .大阪府大腸がん女性のADI5分位別進行度別年齢調整罹患率：1993～1998年 / 1999～2004年（直線は分散重み付き最小二乗法によるあてはめ）

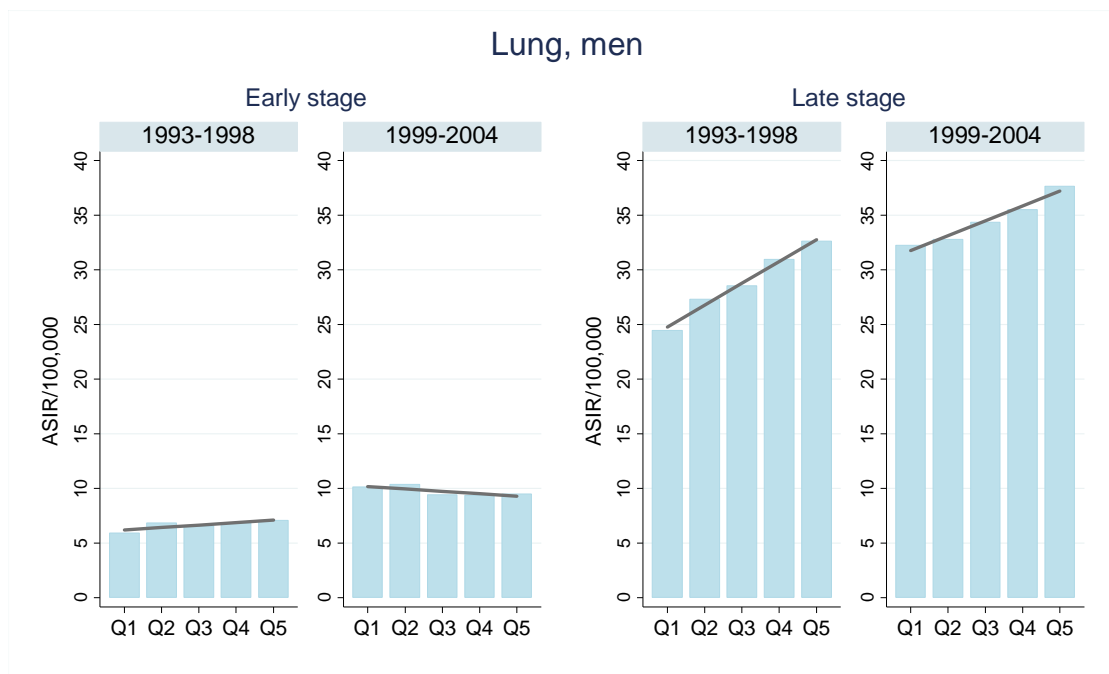


図 7 . 大阪府肺がん男性の ADI5 分位別進行度別年齢調整罹患率：1993～1998 年 / 1999～2004 年（直線は分散重み付き最小二乗法によるあてはめ）

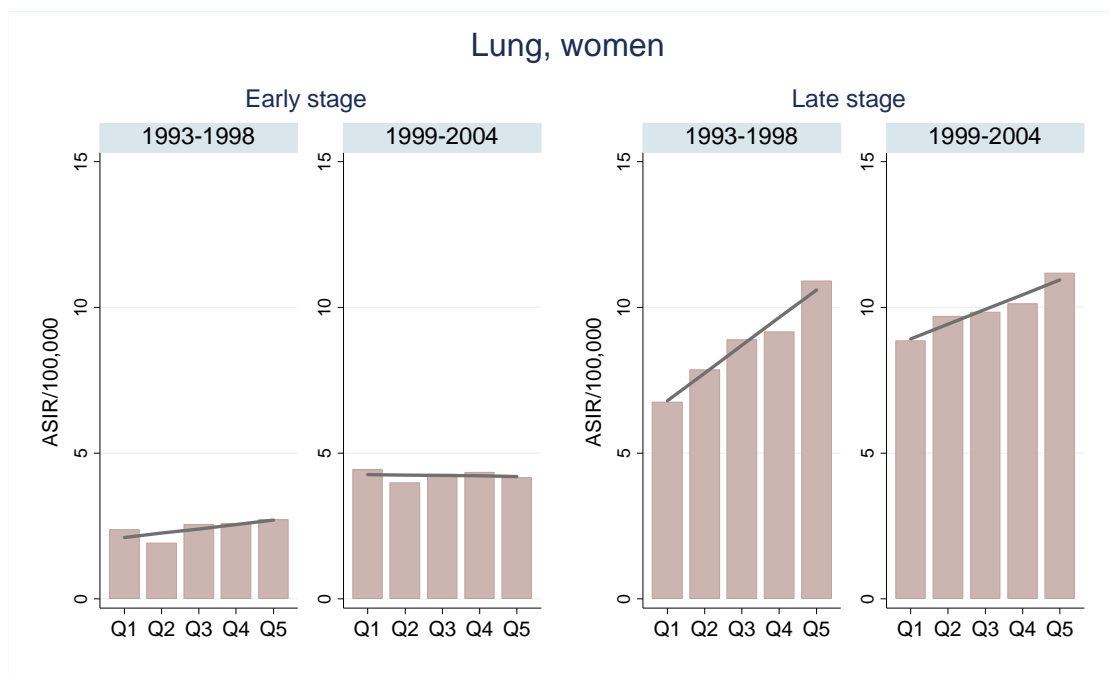


図 8 . 大阪府肺がん女性の ADI5 分位別進行度別年齢調整罹患率：1993～1998 年 / 1999～2004 年（直線は分散重み付き最小二乗法によるあてはめ）

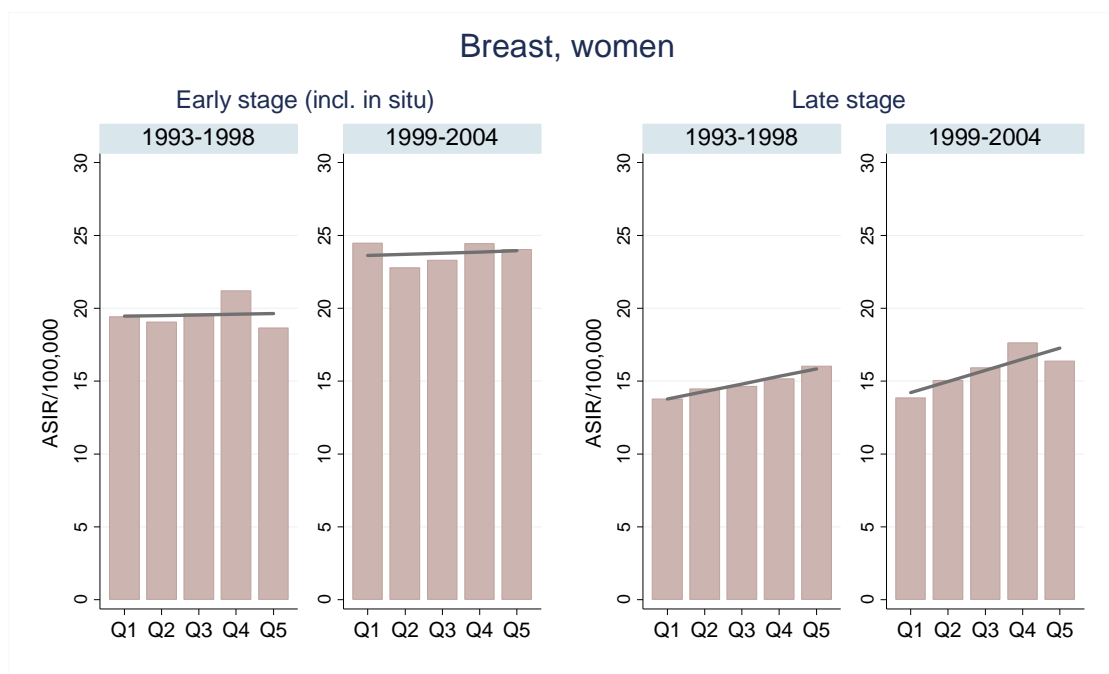


図 9 . 大阪府乳がん女性の ADI5 分位別進行度別年齢調整罹患率：1993～1998 年 / 1999～2004 年（直線は分散重み付き最小二乗法によるあてはめ）

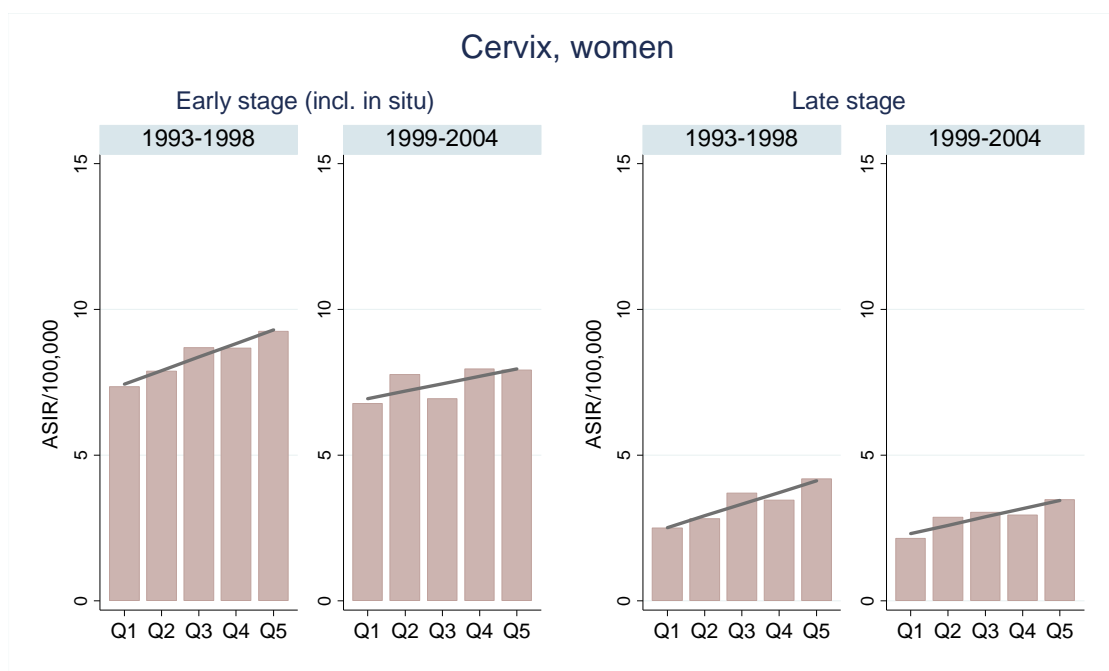


図 10 . 大阪府子宮頸がん女性の ADI5 分位別進行度別年齢調整罹患率：1993～1998 年 / 1999～2004 年（直線は分散重み付き最小二乗法によるあてはめ）

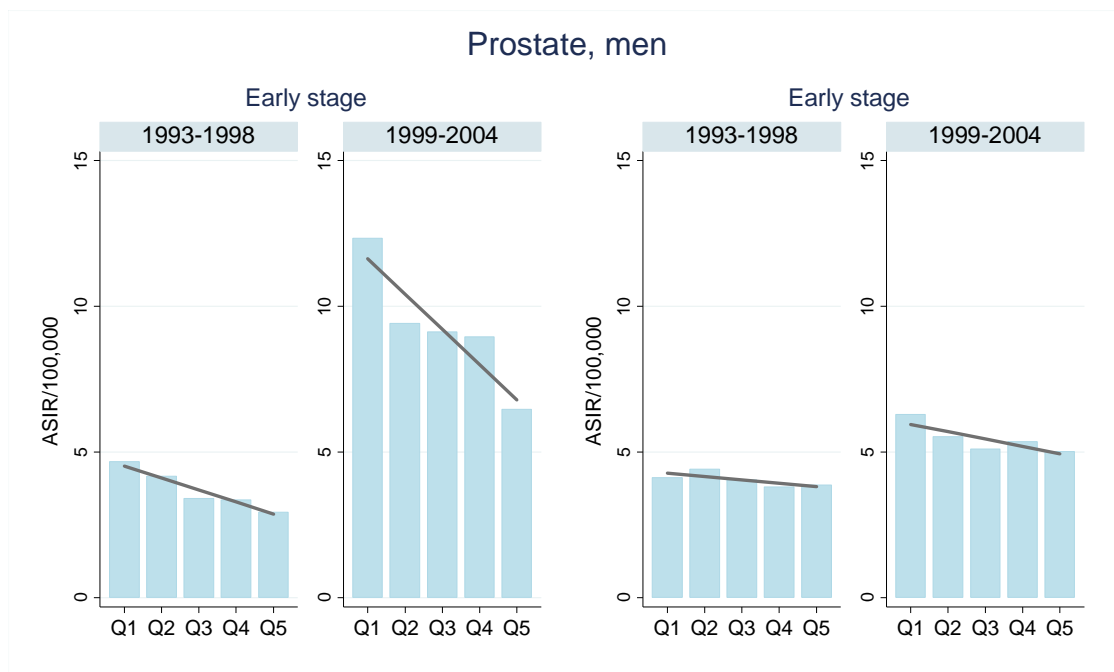


図 11 . 大阪府前立腺がんの ADI5 分位別進行度別年齢調整罹患率：1993～1998 年 / 1999～2004 年（直線は分散重み付き最小二乗法によるあてはめ）

研究成果の刊行に関する一覧表

書籍

著者氏名	論文タイトル名	書籍全体の編集者名	書籍名	出版社名	出版地	出版年	ページ
なし							

雑誌

発表者氏名	論文タイトル名	発表誌名	巻号	ページ	出版年
Ito Y, Nakaya T, Ioka A, Nakayama T, Tsukuma H, Uehara S, Sato KK, Endo G, Hayashi T	Investigation of Spatial Clustering of Biliary Tract Cancer Incidence in Osaka, Japan: Neighborhood Effect of a Printing Factory.	J Epidemiol			2016 In press
Kinoshita F, Ito Y, Nakayama T	Trends in lung cancer incidence rates by histological type in 1975-2008: a population-based study in Osaka, Japan.	J Epidemiol			2016 In press
伊藤ゆり, 中山富雄	肺がん生存率の国際比較	肺癌	55	266-272	2015
中谷友樹・埴淵知哉	健康の社会格差と地域格差	地理	61(1)	51-57	2016