

ミクロ健康データ利用の2つの課題

著者	熊谷 成将
雑誌名	経済学論究
巻	76
号	3
ページ	21-39
発行年	2022-12
URL	http://hdl.handle.net/10236/00030532

マイクロ健康データ利用の2つの課題

The Two Challenges of the Use of Micro-Health Data

熊谷成将*

The invisible social pressure to engage in informal caregiving is a burden for the family members with care recipients in Japan. This study examined the determinants of high-intensity caregiving using an aggregated data as the proxy variable of the social pressure on the family members. Results showed that cohabitating with care recipients who will enter a long-term care facility was the main determinant of high-intensity caregiving. However, due to the data limitations, no possible intervention on the family members with care recipients could be found. The environment of both the collection and the use of micro data in Japan should be improved rapidly.

Narimasa Kumagai

JEL : C31, D13, I12

キーワード : うつ病、高負荷介護、社会的圧力、マッチング、ライフスタイル

Keywords : depression, high-intensity caregiving, social pressure, matching, lifestyle

1. はじめに

個人の健康と生活の質を決定する要因の60%がライフスタイルと相関する(Campagne, 2021)ので、健康状態や生活満足度に影響するライフスタイル要因を探求する研究者は、健康状態などの結果変数と関連する生活習慣のみならず、人々の嗜好や価値観が反映される生活様式を含む広義のライフスタイルをカバーするデータセットを用いるのが望ましい。

だが、実際に日本国内のマイクロデータを用いてライフスタイルと健康感の

* 西南学院大学経済学部教授

関係を分析しようとする、熊谷 [2022] が指摘するように、研究者がひとつのデータセットで、勤労世代や退職世代のライフスタイルと健康感の関係を包括的にとらえるのは容易ではない。運動習慣のような「自助」、居住地域の住民と助け合う「互助」や社会保険に代表される「公助」を網羅的にカバーするデータセットを入手できないからである。

加えて、政策的な分析を行ううえで不可欠な回答者の属性や居住地域の情報に欠けるデータが珍しくない。近年の内閣府の調査（内閣府, 2021）は、市場労働に従事している女性の生活満足度が、介護時間が長いほど低下する傾向があることを明らかにしたが、居住地域の情報との接合が十分ではない。そのためか、人々の嗜好や価値観が反映される地域特有の要因、例えば、伝統的な家族価値観が支配的な地域において「義務感や社会的圧力から介護を続けており、家族介護者の負担が解消されていない」（デンリ & ベッカー, 2014）ことを考慮した分析は行われていない。

可視化しづらい特性が家族介護に従事することや長時間介護の決定要因になり得るので、研究者は家族介護時間や介護者の健康状態の地域差を分析するとき、社会的圧力のような地域の特性に注意を払う必要がある。この点は、ミクロ健康データ利用の課題のひとつであると思われる。本研究は、社会的圧力の代理変数として「主たる家族介護者の地域別介護時間」を用い、家族介護に係る地域の異質性を考慮した回帰分析を行う。

一方、被調査者が母集団を代表しないビッグデータ研究における研究者の関心は、結果変数に対する因果効果の探求にある。マイクロデータを用いて患者の行動選択が結果変数の健康状態に与えた因果効果を計測するとき、研究者は観察研究における 2 群間の不均衡な患者の行動選択に留意し、2 群間のランニングを行って患者の行動選択に係る共変量の影響を最小限にしようとして、因果効果を評価する必要がある。

ビッグデータを用いて、寛解期のうつ病患者の活動記録を分析した Kumagai et al. [2021] はゴロゴロもしくはボーっと過ごした時間の長さが、うつ病再発回数の過剰分散を減らして、うつ病再発回数の正確な計測に寄与することを見出した。しかし、この研究は、ゴロゴロ・ボーっと過ごしたことの決定要因

を探究しておらず、患者の行動選択が自身の精神的健康状態に影響したメカニズムは明らかになっていない。

本研究は、因果効果計測の手順に従って、患者の行動選択が自身の精神的健康状態に与えた影響を推定し、寛解期のうつ病患者がゴロゴロ・ボーッと過ごすことを選択する際の決定要因を探究する。その分析結果に基づいて、もうひとつのマイクロ健康データ利用の課題であるマッチング法の使用を念頭に置いたデータ収集を議論する。

本稿の構成は次の通りである。第2節で、家族介護のマイクロデータ分析において地域差が大きい集計データをマッチングする意義を説明する。第3節では、患者の行動選択が自身の精神的健康状態に与えた因果効果を計測する手順を表す。第4節で、2種類のマイクロデータの記述統計と、家族介護者の精神的健康状態悪化や寛解期のうつ病患者の再発を防ぐ方策を検討するための回帰分析の結果を要約する。前半の分析で、総務省「平成28年 社会生活基本調査」と内閣府「満足度・生活の質に関する調査」（2019-2020年）のデータが用いられる。後半の分析では Kumagai et al. [2019, 2021] のビッグデータが再利用される。第5節で、観察研究におけるデータの利活用を考察し、因果効果の推定を念頭に置いたデータ収集の課題を議論する。

2. 可視化しづらい特性を考慮した家族介護時間の分析

家族への介護提供と就労の間にトレードオフが存在することを多くの研究 (Johnson & Lo Sasso 2000; Spiess & Schneider 2003; Johnson & Lo Sasso 2006; Lilly, Laporte & Coyte 2007; Bolin, Lindgren & Lundborg 2008) が支持しており、長時間介護の研究 (Heitmueller 2007; Lilly, Laporte & Coyte 2010; Van Houtven, Coe & Skira 2013) の結果は、長時間介護が勤労世代の就労に負の影響を与えることで一致している。Kumagai [2017] は、日本において高負荷の長時間介護の継続により、無業（介護離職者を含む）の家族介護者が精神的健康状態を悪化させることを明らかにした。

一方、いくつかの研究 (Arber & Ginn 1997; Vlachantoni 2010) は、家族介護提供と就労の時間配分決定は、介護者の人格や家族の状況によるとしてい

る。それゆえ、家族介護時間の差異を分析するとき、可視化しづらい個人や地域の特性に注意を払う必要がある。

他方、日本では、菊池 [2012] が同居介護者の家族介護が施設サービス利用に負に影響したことを、菊池 [2014] が居住地域の介護施設密度が回答者の居宅介護選好と相関していることを、さらに、Kumagai [2018] が施設介護サービスを利用するのが容易でないことを都市部の一人暮らし高齢者が認識していることを明らかにしている。

介護サービス提供体制の地域差を表す介護施設密度を考慮した上述の分析により、介護施設密度と居宅介護選好の地域差が小さくないことが知られている。しかしながら、可視化が難しい地域固有の特性によって介護保険サービス利用の潜在需要が顕在化していないとき、家族介護時間の差を介護サービス提供体制の差で十分に説明できない。娘が父母、もしくは配偶者の嫁が義父母と同居して親の介護を担うのが当然とされている地域では、義務感や社会的圧力から主介護者となり、主介護者の家族介護時間が長くなる傾向があるかもしれない。そのような側面に注意を払って、本研究では、都道府県別に集計データ「50代女性の地域別平均介護時間（介護時間の全国平均比）」を作成し、ミクロデータにマッチングした。この変数の特徴は第4節で説明される。

3. 患者の行動選択とその結果変数の因果効果推定

ライフログデータを用いて研究者が患者の再発を予測するとき、①メンタルヘルスを悪化させるライフスタイル要因と再発の“双方向の関係”を念頭に置いて再発の予測モデルを構築することと、②真の状態が再発であるにも関わらず再発していないと診断する偽陰性の予測因子を見極めて再発回数の決定要因を分析することが重要である [熊谷ほか, 2021]。②に着目した Kumagai et al. [2021] はゴロゴロ、ボーッと過ごした時間が増えると再発確率が高まった半面、長時間睡眠の増加がゼロ（非再発）の分散を大きくしたことを見出した。このファインディングに基づいて、ゴロゴロ、ボーッと過ごす時間と睡眠の正確な記録が分析結果の解釈を容易にすると考えられる。

そこで、患者の行動選択が自身の精神的健康状態に与えた因果効果を分析す

るために、本研究ではマッチング法を用いる。マッチング法は、政策や環境変化の処置を受けた個人の結果と、その個人属性や環境等の外的条件が非常に似ており、これらの処置を受けなかった個人の結果を比較する分析手法である。

以下では、患者固有の個人差を除去した形で、標本期間内におけるゴロゴロ・ボーの有無が、患者の再発／非再発に与えた因果効果をバイアスなく推定するために、マッチング法を用いる。

共変量 X の情報を「共変量 X を有する患者が処置群に割り当てられる確率」(傾向スコア: propensity score) に一元化することによって因果効果を推定したい。その推定に際し、この割り当てが潜在的結果変数に依存することなく、傾向スコアの推定に用いた共変量 X のみに依存するとき、「強く無視できる割り当て条件 (strongly ignorable treatment assignment)」が満たされる。

Panel VAR を用いた Kumagai et al. [2019] は、潜在的結果変数 (再発／非再発) を定義する K6 のラグ項が今期のゴロゴロ・ボーに対して統計的に有意に影響を与えないことを示した。このことから、患者の処置群への割り当てが潜在的結果変数のラグ項に依存せず、「強く無視できる割り当て条件」が満たされると考えることができる。

傾向スコアマッチングの手順に沿って、研究者は対照群と処置群の参加者のセットを作り、マッチされた2群のセットに対して、次の手順で処置群に対する平均的な効果 (the average treatment effect on the treated: ATE_T) を求めることができる。(1) 式は、対照群と処置群の間における平均的な効果の差 (ΔATE_T) を表している。

$$\Delta ATE_T = \left(\sum y_{t,i} - y_{c,i} \right) / n_e = \sum \Delta_e / n_e \quad (1)$$

ここで、 $y_{t,i}$ は処置群の個人 i のアウトカム、 $y_{c,i}$ はマッチングした対照群の個人 i のアウトカム、 n_e はマッチングしたペアのサンプルサイズ、 Δ_e はマッチングしたペアのアウトカムの差である。

次の手順に従って、処置群と対照群をバランスさせて ATE_T を求めた。第一に、2値変数 (ゴロゴロ・ボーかつしなかった=1) を被説明変数とする Generalized linear model (リンク関数: ロジットモデル) を推定した。第二

に、処置群と対照群の共変量の分布をバランスさせるために傾向スコアの逆数を重みとする逆確率重み付け (inverse probability weighting: IPW) 推定を行った。IPW 推定は、処置群に割り当てられた確率の逆数 (傾向スコアの逆数) で処置群を、処置群に割り当てられなかった確率の逆数 (1-傾向スコアの逆数) で対照群をそれぞれ重みづけて、平均値を求めるという方法である。IPW 推定により、交絡項の影響を取り除いて偏りが無い処置効果を計測できる (Cattaneo, 2010)。

第三に、傾向スコアによる調整後の群間差が小さくなったか、2 群間のバラシグがうまくいったかを確認するために、(2) 式を用いて処置群と対照群の独立変数の標準化差 (standardized differences) (D'Agostino 1998; Rosenbaum & Rubin 1985) が 0.1 以下であるかを調べた。標準化差 (d) が 0.1 を超える共変量の使用は推奨されない (D'Agostino 1998)。(2) 式の分母を構成する 2 変数はそれぞれが処置群 ($trea$) と対照群 ($cont$) の共変量の分散を表しており、(2) 式の分子は 2 群の共変量の平均値である。

$$d = (\bar{x}_{trea} - \bar{x}_{cont}) / \sqrt{\frac{S_{trea}^2 + S_{cont}^2}{2}} \quad (2)$$

4. 高負荷介護と患者の行動選択の決定要因

4-1. 主介護・高負荷介護の分析

Kumagai [2017] は、週あたり 20 時間以上の介護を高負荷の介護 (high intensity caregiving) と定義しており、この定義は家族介護の研究において標準的である。その定義に準じて本研究では、1 日あたり 3 時間以上を家族介護に費やす主介護者を「高負荷の家族介護者」と定義した。

内閣府「満足度・生活の質に関する調査」の 2020 年調査に回答した者のうち 17.6% の 1,814 名が、介護を必要とする家族がいる者であった。この年の調査は、平日・休日のそれぞれにおいて介護や看護に費やした平均的な時間を尋ねており、本研究は、自身が中心となって介護が必要な家族を介護した者を主介護者と定義した。主介護者のうち介護時間を回答した者の比率は 82% (187/229) であった。女性主介護者の 46.5% (59/127)、男性介護者の 33% (20/60) が

高負荷の主介護者であった（表1）。

在宅主介護者229名のうち65%は女性であり、女性の主介護者の平均年齢は53歳であった。介護を必要とする家族の68%は施設サービスもしくは通所サービスを利用していた（表2）。不等分散を仮定した平均値の差の検定により、女性の主介護者の主観的健康感が非主介護者に比べて10%水準で、同様に生活満足度が1%水準で低いことが分かった。女性と異なり、男性の主介護者の主観的健康感是非主介護者と統計的有意な差がなかった（付表）。

表1 主介護者の家族介護時間（平日、2020年）

	人数	3時間以上	平均	標準偏差	最小値	最大値
女性	127	59	3.66	3.91	0.5	24
男性	60	20	2.66	3.03	0.5	13

（出所）内閣府「満足度・生活の質に関する調査」

一方、総務省「平成28年社会生活基本調査」によると、世代別人口に占める家族介護者比率が最も高いのは50代女性である。同年の女性介護者421.1万人のうち、50-60代が57.61%を占めている。世代別人口比では60代の13.3%（123.3万人）を勤労世代の50代の15.6%（119.3万人）が上回っている。この点に着目して、都道府県別「50代女性の地域別平均介護時間（平日、介護時間の全国平均比）」を作成した。その最小値は0.46、最大値は1.36であり、岐阜県と徳島県が1.3超の値を記録した（表2）。

Stata cmp module (Roodman [2011]) を用いて、平日・休日の別に、私には simulated Maximum Likelihood により (3) 式の変数のプロビットモデルを推定した。

$$\begin{aligned}
 y_{1it}^* &= \mathbf{x}'_{1it}\beta_1 + \varepsilon_{1it}, \\
 y_{2it}^* &= \mathbf{x}'_{2it}\beta_2 + \varepsilon_{2it},
 \end{aligned}
 \tag{3}$$

ここで、 y_{jit} , $j = 1, 2$ は高負荷介護と主介護者である。 y_{jit}^* , $j = 1, 2$ はこれらの潜在変数であり、 \mathbf{x}_{jit} , $j = 1, 2$ は説明変数、 β_j , $j = 1, 2$ は係数ベクトル

ルである。(3) 式の 2 本の関数に共通の説明変数は性別と年齢であり、高負荷介護関数の説明変数として、主介護者、等価世帯所得、婚姻状態、要介護者の通所サービス利用と施設に入所予定の要介護者が用いられた。

多変量プロビットモデルの推定結果は、50 代女性の地域別平均介護時間が長い地域の居住が主介護者の決定要因であることを示した (表 3)。一方、平日・休日の 2 本とも連立方程式の誤差項は統計的に有意に相関しなかった。この結果から、主介護者となることに影響を与える観察されない地域の特性が高負荷介護の確率を高めるとは考えられない。

平日・休日の双方とも高負荷介護になりやすい要因は、施設に入所予定の要介護者がいること、低い等価世帯所得、より高齢であることと女性であった。

表 2 家族介護者とその世帯の状況 (2020 年)

変数	N	平均	標準偏差	最小値	最大値
介護を必要とする家族がいる	10,293	0.18	0.38	0	1
50代女性の地域別家族介護時間 (全国平均比、都道府県別)	10,293	0.91	0.23	0.46	1.36
在宅で主介護	1,814	0.13	0.33	0	1
在宅で他の家族が主介護	1,814	0.14	0.34	0	1
主介護者の女性の年齢	149	53.21	14.36	15	83
主介護者の男性の年齢	80	49.21	17.17	19	81
施設サービス利用	1,814	0.35	0.48	0	1
施設に入所予定 (未入所)	1,814	0.07	0.26	0	1
通所サービス利用	1,814	0.33	0.47	0	1
性 (男性=1)	10,293	0.50	0.50	0	1
年齢	10,293	43.85	16.40	15	86
世帯人員数	10,293	2.91	1.38	1	12
等価世帯所得	10,293	399.19	847.14	18.9	20000
未婚	10,293	0.22	0.41	0	1
世帯に父母(義父母)がいる	8,650	0.33	0.47	0	1
主観的健康感	10,293	3.33	1.03	1	5
生活満足度	10,293	6.78	2.34	1	11

(出所) 内閣府「満足度・生活の質に関する調査」(2020 年)、総務省「平成 28 年 社会生活基本調査」

表3 高負荷介護の決定要因（2020年、N=1,808）

	平日		休日	
	高負荷介護	主介護者	高負荷介護	主介護者
主介護者	1.726*** (0.287)		1.742*** (0.296)	
1+等価世帯所得（自然対数値）	-0.240*** (0.0620)		-0.127** (0.0558)	
性（男性=1）	-0.281*** (0.106)	-0.297*** (0.0864)	-0.185** (0.0931)	-0.297*** (0.0865)
年齢	0.0243*** (0.00405)	0.0142*** (0.00289)	0.0178*** (0.00348)	0.0142*** (0.00289)
未婚	0.336** (0.140)		0.184 (0.126)	
施設に入所予定（未入所）	0.539*** (0.165)		0.492*** (0.150)	
通所サービス利用	0.387*** (0.102)	-0.330*** (0.0931)	0.346*** (0.0913)	-0.332*** (0.0931)
50代女性の地域別介護時間（全国平均比）		0.364** (0.185)		0.387** (0.184)
世帯に父母(義父母)がいる		0.544*** (0.0915)		0.547*** (0.0915)
在宅で他の家族が主介護		-0.284** (0.127)		-0.269** (0.128)
他の家族が介護に従事（本人は非従事）		-1.464*** (0.262)		-1.466*** (0.262)
定数項	-1.747*** (0.412)	-2.009*** (0.248)	-1.809*** (0.372)	-2.032*** (0.247)

（注）高負荷介護関数の説明変数に離別と死別が含まれている。平日、休日の双方とも、主介護関数と高負荷介護関数の残差は相関していない。

括弧内は標準誤差である。*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

通所サービス利用に比べて施設入所予定の係数推定値が大きい。このことから、特別養護老人ホームなどに入所可能な要介護度 3 以上の要介護者を介護することが高負荷介護の主たる要因であると考えられる。従って、高負荷介護の負のアウトカムである家族介護者の精神的健康状態の悪化（うつ病の発症など）を防ぐためには、施設入所予定者に対して長期間の介護を提供しないことが肝要である。

4-2. 患者の行動選択の分析

2016 年 10 月から 2017 年 3 月にかけて患者がエントリーした Kumagai et al. [2019, 2021] の患者データの特性を説明する。調査の詳細は、熊谷 [2022] を参照されたい。調査開始時の患者の平均年齢は 44.3 歳、患者の約 54% の 48 名が正規労働従事者である（表 4）。市場労働に従事していない 14 名とパートタイム労働者 12 名が続いた。婚姻状態別では、既婚者が 54%、未婚者が 33%、離婚者が 12% である。高負荷介護による家族介護者の精神的健康状態悪化を分析した Kumagai [2017] と異なり、このサンプルに家族介護者は含まれていない。

うつ病の最初のエピソードからの平均年数は 9.4 年、うつ病のエピソード平均回数は 2.5、うつ病の最初のエピソードからの年数とエピソード回数の相関係数は、女性が 0.38、男性が 0.33 であり、ともに 1%水準で有意であった。

2 値変数（ゴロゴロ・ボーつとしなかった=1）を被説明変数とする Generalized linear model の推定結果が表 5 に、2 週間後の再発（再発=1）を被説明変数とする IPW 推定の結果が表 6 に表されている。オッズ比の推定値が統計的に有意に 1 を下回っているため、昼食の欠食回数が多いこと、高校卒と正規労働に従事していることが、「ゴロゴロ・ボーつとした」ことと関係していたと考えられる（表 5）。

表 6 の結果は、「ゴロゴロ・ボーつとしなかった」ことによつて 2 週間後の再発回数が 1 週あたり 0.024 回減ることを示しており、これは 1 年あたり約 1.25 回の再発回数の減少を意味する。従つて、寛解期の患者がゴロゴロ・ボーつとしなくなるような介入を実施できれば、再発を防ぐことが可能と考えられる。

表4 患者の特性

		患者数 89
変数	人数 (%) / 平均 (標準偏差)	
性 (男性=1)	49 (55.1)	
年齢	44.3 (10.8)	
就業形態		
正規労働者	48 (53.9)	
パートタイム労働者	12 (13.5)	
休業中	3 (3.4)	
専業主婦	6 (6.7)	
退職者	3 (3.4)	
非就労(失業者を含む)	14 (15.7)	
その他	3 (3.4)	
婚姻状態		
未婚	29 (32.6)	
離別	11 (12.4)	
死別	1 (1.1)	
既婚	48 (53.9)	

(出所) 熊谷 [2022] から一部を抜粋して転載。

表5 Generalized linear model

N=4,679

ゴロゴロ・ボーっとしない	オッズ比	標準誤差	Z	P>z	[95% 信頼区間]	
1/(1+紫外線の摂取量)	1.34	0.16	2.50	0.01	1.06	1.69
昼食の欠食回数 (標準化変数)	0.87	0.03	-4.07	0.00	0.82	0.93
長時間睡眠のダミー変数	1.47	0.25	2.25	0.02	1.05	2.06
正規労働に従事している	0.84	0.07	-2.21	0.03	0.72	0.98
休業中	10.88	2.56	10.13	0.00	6.86	17.27
失業中	1.35	0.14	2.94	0.00	1.10	1.65
中学卒	1.36	0.24	1.72	0.09	0.96	1.93
高校卒	0.52	0.04	-7.82	0.00	0.44	0.61
性(男性=1)	2.81	0.19	15.04	0.00	2.46	3.22
年齢	1.03	0.00	9.42	0.00	1.02	1.04
定数項	0.18	0.03	-9.69	0.00	0.13	0.26

(注) 大学卒 (66名) を reference として、中学卒 (3名) や高校卒 (20名) と比較した。

「ゴロゴロ・ボーっとしなかった」を説明する変数として、weighted の系列のうち standardized differences が 0.1 を下回った、正規労働に従事していることと婚姻していることが用いられた (表 7)。IPW 推定の共変量候補変数のうち、正規労働従事者と既婚者の variance ratio は順に 0.998 と 1.003 であり、1 に近い値でマッチングは良好であった。しかし、長時間睡眠のダミー変

表 6 Treatment-effects estimation

			N=4,527
2週間後の再発	ATET	潜在的なアウトカムの平均値	ゴロゴロ・ボーっとしない
ゴロゴロ・ボーっとしないことの再発回数に対する効果	-0.0241*** (0.00416)	0.0308*** (0.00383)	
正規労働に従事している			-0.163** (0.0638)
婚姻している			0.195*** (0.0635)
定数項			0.227*** (0.0503)

括弧内は頑健な標準誤差である。*** p<0.01, ** p<0.05, * p<0.1

表 7 Standardized differences

	Raw	Weighted
1/(1+紫外線の摂取量)	0.107	0.112
昼食の欠食回数 (標準化変数)	-0.176	-0.183
長時間睡眠のダミー変数	0.057	0.056
性(男性=1)	0.447	0.441
年齢	0.226	0.203
正規労働に従事している	-0.053	0.017
失業中	0.106	0.100
婚姻している	0.074	-0.014

(注) 太字の変数を共変量として用いることが推奨されている。

数のそれは1を大きく上回っていたため、この変数を表6の説明変数から除いた。

表6の推定結果から、正規労働に従事することがゴロゴロ・ボーっとしない確率を引き下げ、対照的に、婚姻していることがゴロゴロ・ボーっとしない確率を高めることが分かった。これらの係数推定値の合計が正值であることから、婚姻して正規労働に従事することが、うつ病の再発抑制に寄与すると考えられる。しかし、婚姻もしくは正規労働が患者のゴロゴロ・ボーっとしない選択に与える影響のメカニズムは明らかではない。

正規労働従事者の長い勤務時間が患者の行動選択に与えた影響を正規労働に従事した効果と見誤っている可能性が疑われるが、残念ながら就業時間の長さや就業パターンの情報を利用できないため、ゴロゴロ・ボーっとしない確率を高めるために、どのような介入が望ましいかを論じることができない。

5. 議論：観察研究におけるデータの利活用

本研究は、高負荷介護に起因する家族介護者の精神的健康状態の悪化を防ぐために、主介護者が長期間に渡って施設入所予定者に高負荷介護を提供しないことが重要であると指摘した。だが、本研究で用いたデータからは、施設入所予定者に対して家族が高負荷介護を提供せざるを得ない理由が明らかでない。もし、高負荷介護の理由が義務感もしくは社会的圧力ではなく、介護サービスの供給側にあれば、保険者の介入によって主介護者の介護時間を減らすことが可能かもしれない。

上述のファインディングを見出すために集計データを個票データにマッチングせざるを得なかった。このことは個票データの利用環境が「先進国最低」(重岡, 2022)であることと無関係ではない。「日本が一步踏み出している間に主要先進国は地球を2~3周している」(同)と喩えられるほど、日本の個票データ利用環境は先進国の中で格段に悪い。

政府統計を用いて政策効果を検証するために家計や企業が回答した調査票情報を利用しようとすると、政府統計の目的外利用に該当するため、研究者は統計法33条に基づく利用申請を行わねばならない。この申請手続きに「数カ

月からひどい場合 1 年近くかかる」(渡辺, 2022)。長い待ち時間が生じる理由は「審査者は、分析にとって必要最小限な情報のみが利用可能であると考えて申請内容を確認する」(重岡 2022, 渡辺 2022) ことに起因しており、申請書類の修正作業と書類提出後の審査に長い時間を要する。

厚生労働省「中高年者縦断調査」の調査票情報を用いた論文 (Kumagai & Ogura, 2014) を iHEA の Sydney 2013 Congress で発表した私に、英国の研究者が「外国人研究者は、そのデータを利用できないか」「なぜ、英語でデータの質問項目が表示されていないか」と尋ねたことがある。科研費採択者のみが利用可能であることと日英併記のデータが日本には少ないことを説明したところ、「英国の政府統計の個票データは誰でもクリックして簡単にダウンロードできる」「(外国人研究者にとって) 日本のデータ利用の障壁は高い」との回答であった。

このようなデータ利用環境の大きな差を縮めることは日本にとって喫緊の課題である。効果的な政策を議論するために個票データの収集を進めるだけでなく、その利活用を容易に行えるようデータ利用に係る手続きを簡素化する必要がある。

他方、国全体を代表する調査に基づいていない convenience sampling によって収集されたデータは、その分析結果についての議論の一般化は難しいが、因果効果推定を念頭に置いたデータ収集方法を議論することは意義があると思われる。本研究では、寛解期のうつ病患者が再発しないようゴロゴロ・ボーっとする行動の決定要因を探索したが、正規労働従事者の働き方が患者の行動選択に与えた影響を正確に見定めることができなかった。

ワークライフバランスを重視した就労は再発のリスクが低いといったことが観察できれば、うつ病患者に対して介入しやすくなると思われる (田近・熊谷・古川, 2022)。実際、配偶者が家庭の掃除に高い頻度で従事すれば、等価世帯所得の 9.5%相当額だけ女性の健康損失額を減らすことができることを Kumagai [2021] が明らかにしている。働き方の変化が健康に与える影響を詳しく分析できるよう、就業時間の長さや就業パターンに留意してデータを収集する必要がある。

付表 主介護の有無による主観的健康感と生活満足度の有意差

男性の主観的健康感 (N=820)

	N	平均	標準誤差	標準偏差	95%信頼区間	
非主介護者	740	3.258	0.039	1.060	3.182	3.335
主介護者	80	3.250	0.116	1.037	3.019	3.481
差		0.008	0.122		-0.235	0.251
t = 0.066						

女性の主観的健康感 (N=994)

	N	平均	標準誤差	標準偏差	95%信頼区間	
非主介護者	845	3.212	0.036	1.059	3.140	3.283
主介護者	149	3.054	0.082	1.005	2.891	3.216
差		0.158	0.090		-0.019	0.336
t = 1.756						

男性の生活満足度 (N=820)

	N	平均	標準誤差	標準偏差	95%信頼区間	
非主介護者	740	6.745	0.089	2.423	6.570	6.919
主介護者	80	6.088	0.291	2.606	5.508	6.667
差		0.657	0.305		0.052	1.262
t = 2.157						

女性の生活満足度 (N=994)

	N	平均	標準誤差	標準偏差	95%信頼区間	
非主介護者	845	6.840	0.079	2.300	6.685	6.996
主介護者	149	6.148	0.199	2.428	5.755	6.541
差		0.693	0.214		0.270	1.115
t = 3.235						

謝辞

本研究は科研費基盤研究 (C) 20K01739 の助成を受けた。

東京大学社会科学研究所附属社会調査・データアーカイブ研究センター SSJ データアーカイブから、内閣府「満足度・生活の質に関する調査」(2019-2020 年) の個票データの提供を受けた。

追想

広島大学経済学部の高林研究室で初めて、広井良典 [1994] 『医療の経済学』日本経済新聞社の表紙を見て、「医療を経済学で分析するのですか？」と高林先生に尋ねてから 7 年後の 2002 年、見えざる手に導かれて、私は医療経済学の領域に参入しました。

当時、プロスポーツのデータを用いる分析のアイデアを練っていた高林先生から共同研究の御誘いを頂きましたが、「私は阪神ファンではないので」とお断り申し上げました。東京大学で開催された日本財政学会のランチタイムの会話であったように記憶しています。いくつかの私のアイデアが反映された高林喜久生 [2004] 『「今年も阪神優勝!」の経済学』光文社新書の出版後に再び、「本格的な単行本を作らないか、という話が届いているけど、熊谷君、一緒にどう？」とお声をかけて頂いた時も、応用ミクロ計量経済分析の力量が不足していた私は、スポーツ分析に参入せず、医療経済学の研究に力を注ぎました。

それから長い時間を経て、スポーツ分析参入の余裕ができた私は、高林先生の退職記念号には是非、高林 [2004] を引用してプロ野球のデータを用いた論文を、と考えて単著論文「日本プロ野球におけるレギュラーシーズン勝率の決定要因」を作りました。しかし、出版スケジュールの関係から、完成稿を 1 年余り寝かさざるを得ないことが分かり、私はこの論文を査読誌『スポーツ産業学研究』に投稿し、本年 1 月に拙稿が出版されました。希望通りの出版とならなかったこともあり、私は、この続編を考えています。

阪神の戦績に一喜一憂できる間、高林先生は現役です。プロ野球の分析の続編を作ることを中核とする研究を高林先生と共同で行う。これを阪神のセ・リーグ優勝よりも先に実現したいと願っています。

参考文献

- Arber, S., Ginn, J. [1997] Informal care-givers for the elderly people. In Ungerson, C., Kember, M. eds. *Women and social policy: A reader*. London: Macmillan. Arber & Ginn
- Bolin, K., Lindgren, B., Lundborg, P. [2008] Your next of kin or your own career? Caring and working among the 50+ of Europe. *Journal of Health Economics* 27(3), 718-738.
- Campagne, DM. [2021] Accountability for an unhealthy lifestyle. *The European Journal of Health Economics* 22, 351-355.
- Cattaneo, MD. [2010] Efficient semiparametric estimation of multi-valued treatment effects under ignorability. *Journal of Econometrics* 155, 138-154.
- デーシリ, ジェイソン & ベッカー, カール [2014] 高齢者介護の意味づけと家族介護者の支え方. 『こころの未来』13, 38-39.
- D'Agostino, RB., Jr [1998] Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. *Statistics in Medicine* 17(19), 2265-2281.
- Heitmueller, A. [2007] The chicken or the egg? Endogeneity in the labor market participation of informal carers. *Journal of Health Economics* 26(3), 536-559.
- Johnson, R., Lo Sasso, A. [2000] *The tradeoff between hours of paid employment and time assistance to elderly parents at midlife*. Washington DC: The Urban Institute.
- Johnson, R., Lo Sasso, A. [2006] The impact of elder care on women's labor supply. *Inquiry: A Journal of Medical Care Organization, Provision and Financing* 43(3), 195-210.
- 菊池潤 [2012] 介護サービスは家族による介護を代替するか. 井堀利宏・金子能宏・野口晴子 編 『新たなリスクと社会保障 生涯を通じた支援策の構築』東大出版会, 211-230.
- 菊池潤 [2014] 高齢者の希望介護場所と社会的ネットワーク. 『季刊 社会保障研究』49(4): 396-407.
- Kumagai, N., Ogura, S. [2014] Persistence of physical activity in middle age: a nonlinear dynamic panel approach. *The European Journal of Health Economics* 15, 717-735.
- Kumagai, N. [2017] Distinct Impacts of High Intensity Caregiving on Caregivers' Mental Health and Continuation of Caregiving. *Health Economics Review* 7(15), 1-14.

- Kumagai, N. [2018] Care Preferences of Elderly People Living Alone in Japan. *Health Education and Public Health* 1(2), 101-109.
- Kumagai, N., Tajika, A. et al. [2019] Predicting recurrence of depression using lifelog data: an explanatory feasibility study with a panel VAR approach. *BMC Psychiatry* 19, 391-402.
- 熊谷成将・田近亜蘭・後藤励・古川壽亮. [2021] うつ病患者の活動記録を利用した社会医学研究者と医療経済学者の共同研究. 『医療経済研究』 32(2): 78-89.
- Kumagai, N., Tajika, A. et al. [2021] Assessing Recurrence of Depression using a Zero-Inflated Negative Binomial Model: A Secondary Analysis of Lifelog Data. *Psychiatry Research* 300, 113919-113925
- Kumagai, N. [2021] Valuation of health losses of women with multiple roles using a well-being valuation approach: Evidence from Japan *PLOS ONE*. 16(5): e0251468.
- 熊谷成将 [2022] 『ライフスタイルと健康感の経済分析』 晃洋書房
- Lilly, MB., Laporte, A., Coyte, PC. [2007] Labor market work and home care's unpaid caregivers: A systematic review of labor force participation rates, predictors of labor market withdrawal, and hours of work. *The Milbank Quarterly* 85(4), 641-690.
- Lilly, MB., Laporte, A., Coyte, PC. [2010] Do they care too much? The influence of caregiving intensity on the labor force participation of unpaid caregivers in Canada. *Journal of Health Economics* 29(6), 895-903.
- 内閣府 [2021] 『満足度・生活の質に関する調査報告書 2021』 (https://www5.cao.go.jp/keizai2/wellbeing/manzoku/pdf/report05_1.pdf) (2021 年 9 月 1 日公表)
- Roodman, D. [2011] Fitting fully observed recursive mixed-process models with cmp. *The Stata J.* 11(2): 159-206.
- Rosenbaum, PR. and Rubin, DB. [1985] The bias due to incomplete matching. *Biometrics* 41(1), 103-116.
- 重岡仁 [2022] 花粉症の経済学. 日本経済学会春季大会招待講演 (オンライン開催) 2022 年 5 月 28 日
- Spieß, CK., Schneider, AU. [2003] Interactions between care-giving and paid work hours among European midlife women, 1994 to 1996. *Ageing and Society* 23(1), 41-68.
- 田近亜蘭・熊谷成将・古川壽亮. [2022] スマートフォンアプリとウェアラブルデバイスを用いた、寛解期のうつ病患者の再発予測. 『予防精神医学』 7(1), 印刷中 (2022 年 12 月刊行予定).
- Van Houtven, CH., Coe, NB., Skira, MM. [2013] The effect of informal care on work and wages. *Journal of Health Economics* 32(1), 240-252.

Vlachantoni, A. [2010] The demographic characteristics and economic activity patterns of carers over 50: Evidence from the English Longitudinal Study of Ageing. *Population Trends* 141(1), 51-73.

渡辺安虎 [2022] 「データによる政策」阻む法の壁. 日本経済新聞 2022 年 7 月 8 日付け朝刊.