

健康格差拡大の原因究明に向けた 高度な分析方法と理論の構築

(研究助成金 50万円)

代表研究者 岡山大学大学院医歯薬学総合研究科 疫学・衛生学分野

助教 鈴木 越 治

平成17年 岡山大学医学部卒業
平成21年 岡山大学大学院医歯薬学総合研究科修了

共同研究者 岡山大学 教授 津 田 敏 秀
岡山理科大学 特任教授 山 本 英 二
岡山大学 助教 三 橋 利 晴

研究目的

近年、健康格差の拡大への対応が急務となっている。私はこれまで国内外の関連文献を詳細に検討し、わが国の健康格差の全体像を捉えるためには、社会的健康格差および地理的健康格差の両面から検証するアプローチが重要と考えた。このため、平成22年にハーバード大学および広島大学の研究者らと国際共同研究グループを立ち上げた。このアプローチは国際的に先駆的なものであり、国内外から高い評価を得た^{1~4)}。例えば、1975年から2005年の自殺リスクの推移を評価した結果³⁾、男性では専門職や管理職などで自殺リスクが著しく増加するなど、リスクの社会的格差が広がっていることを明らかにした。また、地理的格差が1995年以降著しく拡大し、リスクが最も高い秋田県と最も低い奈良県で2.45倍の差であることを示した。本研究成果は、岡山大学のプレスリリースに取り上げられて、複数の新聞やテレビで報道された。

本研究はこれら一連の研究成果を踏まえ、様々なバイアスを制御するための高度な分析方法を構築して、健康格差の原因を究明することを目的とする。特に、交絡バイアスの同定・制御方法は健康格差の原因究明において大きな障壁になっているため、本研究の学術的意義は非常に高いと考える。この研究成果を展開させ、日本全国における健康格差の原因を究明して改善に結びつけることが、本研究の究極のゴールである。

参考文献

- 1) Suzuki E, et al. BMJ Open. 2012.
- 2) Suzuki E, et al. PLoS One. 2012.
- 3) Suzuki E, et al. PLoS One. 2013.
- 4) Suzuki E, et al. Eur J Public Health. 2014.

研究実施計画の概要

一連の研究で明らかにしてきた健康格差の推移を評価して原因を究明するため、様々なバイアスを制御する高度な分析方法を構築する。それに向けて専門家チームを形成し、以下の2点から研究を計画する。

1. 健康格差研究で生じる誤差の分類法及び制御法の構築：

誤差は系統的誤差とランダム誤差に大別されており、それらの制御法がこれまで提唱されているが、これらを包括的に扱った文献は皆無である。本研究では、これまで申請者が有用性を実証してきた因果ダイアグラムを用いて、誤差の新たな分類法を構築する。このアプローチにより、拡大する健康格差の原因を正確に同定する。

2. 適切なデータ分析のための変数選択法の構築：

健康格差研究では、ロジスティック回帰分析などの多変量解析が広く用いられている。しかし、調整すべき共変量の選択法についてはコンセンサスが得られていない。本研究では、重要な2種類の因果モデル（反事実モデルと十分原因モデル）の対応を明らかにし、共変量バランスの概念および分析法の構築を行う。これらの基となる手法はすでにこれまでの研究で確立しており、成果の実現性は高いと考える。

倫理面の配慮

本研究の一部はひとを対象にしたものであるが、ナショナルデータなどを活用するため、倫理委員会に諮る必要はない。

I 緒 言

健康格差については国内外で多くの研究が行われている。しかし、これらの研究の多くは、個人の特性を調整することなく地域の社会経済状態と健康リスクの関連を単純に評価しているため、健康格差の原因を正しく評価できていない。また、健康格差に関する研究のように無作為化比較試験の実施が困難な状況では、交絡バイアスの扱いが特に大きな問題となる¹⁾。しかし、長年にわたって交絡バイアスの様々な観念が十分に理解されていなかったため、医学研究を実施する際に大きな混乱が生じていた^{2,3)}。私たちの研究グループは交絡バイアスの概念を詳説する論文を発表し⁴⁾、交絡の四つの観念を区別すべきことを示した。また、交絡に関する深い理解が因果ダイアグラムの有用性を理解する上で必要不可欠であることも示した。

本研究の目的は、この研究成果をさらに発展させて、交絡バイアスを調整するための変数選択法を構築し、健康格差研究の進展に寄与することである。

II 研究方法

本研究では、反事実モデルと十分原因モデルという重要な2種類の因果モデルの対応を基に⁵⁻⁸⁾、交絡バイアスを調整するための変数選択法について考察した。反事実モデルは特定の原因について様々な結果を論じるモデルである一方で、十分原因モデルは特定の結果について様々な原因を論じるモデルである⁷⁾。本稿では、二値の曝露変数 E (1 : 曝露した, 0 : 曝露しなかった) と二値のアウトカム変数 Y (1 : アウトカムが生じた, 0 : アウトカムが生じなかった) を考慮する。

反事実モデルでは、各個人について、曝露した場合に生じるであろうアウトカムと、曝露しなかった場合に生じるであろうアウトカムを考慮する。すなわち、各個人において二つの潜在アウトカム (potential outcome) Y_e を考慮することになる。 Y_1 は個人が曝露した場合の潜在アウトカムであり、 Y_0 は個人が曝露しなかった場合の潜在アウトカムである。このように、曝露の有無とアウトカムの有無のパターンを考慮することにより、個人は表1に示すように四つの反応タイプ (response type) に分類される⁹⁾。

一方、十分原因モデルでは¹⁰⁾、二値のアウトカム変数 Y について、二値の背景因子 C_k を伴う三種類の十分原因 ($C_1, C_2E, C_3\bar{E}$) を考慮する。ここで \bar{E} は E の補事象を表す。本稿では、これらの背景因子が存在することを $C_k = 1$ と表し、背景因子が存在しないことを $C_k = 0$ と表す。特定の個人については、いずれの十分原因に関してもリスクがない状況に加え、いずれか一つの十分原因に関してリスクがある状況、または、複数の十分原因に関してリスクがある状況が考慮できる。そのため、個人は十分原因に関するリスク状況に基づいて、表1に示すように八つのリスク状況タイプ (risk status type) に分類される。

表1の対応関係にあるように、アウトカム変数 Y の潜在アウトカムは、背景因子 C_k を用いて $Y_1 = \max(C_1, C_2)$ 、および、 $Y_0 = \max(C_1, C_3)$ と表される。

表1 二値の曝露と二値のアウトカムを考慮する場合の反応タイプとリスク状況タイプの対応関係^{a)}

タイプ	反応タイプ					リスク状況タイプ						
	潜在アウトカム		タイプが占める割合 ^{b)}			タイプ	背景因子			タイプが占める割合 ^{b)}		
	Y_1	Y_0	曝露群	非曝露群	全集団		C_1	C_2	C_3	曝露群	非曝露群	全集団
1	1	1	p_1	q_1	r_1	1 ^{d)}	1	1	1	s_1	t_1	u_1
						2	1	1	0	s_2	t_2	u_2
						3 ^{d)}	1	0	1	s_3	t_3	u_3
						4	1	0	0	s_4	t_4	u_4
						5 ^{d)}	0	1	1	s_5	t_5	u_5
2	1	0	p_2	q_2	r_2	6	0	1	0	s_6	t_6	u_6
3 ^{c)}	0	1	p_3	q_3	r_3	7 ^{d)}	0	0	1	s_7	t_7	u_7
4	0	0	p_4	q_4	r_4	8	0	0	0	s_8	t_8	u_8

^{a)} 二値の曝露変数 E (1 : 曝露した, 0 : 曝露しなかった) と二値のアウトカム変数 Y (1 : アウトカムが生じた, 0 : アウトカムが生じなかった) を考慮する。この状況では、各個人において二つの潜在アウトカム Y_e がある。アウトカム変数 Y について、二値の背景因子を伴う三種類の十分原因 ($C_1, C_2E, C_3\bar{E}$) を考慮する。ここで \bar{E} は E の補事象を表す。

^{b)} r_j は $p_j \times P[E=1] + q_j \times P[E=0]$ により求められる。ここで $P[E=e]$ は全集団における $E=e$ の割合のことである。同様に、 u_j は $s_j \times P[E=1] + t_j \times P[E=0]$ により求められる。

^{c)} 反事実モデルにおける正の単調性の仮定 (すなわち、全個人において $Y_0 \leq Y_1$) の下では、この反応タイプは排除される^{8,11)}。

^{d)} 十分原因モデルにおける正の単調性の仮定 (すなわち、全個人において $C_3 = 0$) の下では^{11,12)}、これらのリスク状況タイプは排除される^{8,13)}。この仮定は、反事実モデルにおける正の単調性の仮定を示唆する。

Ⅲ 研究結果

交絡バイアスの概念は、反事実モデルの枠組みで広く議論されている^{4, 9, 14-22}。反事実モデルにおいて、交絡バイアスがない必要十分条件は「交換可能性 (exchangeability)」として知られている (表 2)。(以下では主に「分布における交絡」を念頭に置いている。「分布における交絡」の詳細は他を参照されたい^{3, 4, 23}。) 研究対象となる集団 (標的集団と呼ばれる) が曝露群である場合の交換可能性は以下のよう示される。

$$Y_0 \perp\!\!\!\perp E \\ (\Leftrightarrow (p_1 + p_3) = (q_1 + q_3))$$

また、標的集団が非曝露群である場合の交換可能性は以下のよう示される。

$$Y_1 \perp\!\!\!\perp E \\ (\Leftrightarrow (p_1 + p_2) = (q_1 + q_2))$$

これらの条件は「部分交換可能性 (partial exchangeability)」と呼ばれている^{9, 24}。一方で、全集団を標的集団とする場合に交絡バイアスがない必要十分条件は以下のよう示される。

$$Y_e \perp\!\!\!\perp E \quad (e=0,1) \\ (\Leftrightarrow \{(p_1 + p_2) = (q_1 + q_2)\} \wedge \{(p_1 + p_3) = (q_1 + q_3)\})$$

この条件は、「(狭義の) 交換可能性」と呼ばれている。最後に、アウトカム変数 Y の潜在アウトカムの同時分布と曝露変数 E の独立性 ($Y_0, Y_1 \perp\!\!\!\perp E \Leftrightarrow (p_1, p_2, p_3, p_4) = (q_1, q_2, q_3, q_4)$) は「完全交換可能性 (full exchangeability)」と呼ばれている²⁴。完全交換可能性は、部分交換可能性や狭義の交換可能性よりも強い条件である。言い換えるならば、標的集団にかかわらず、完全交換可能性は交絡がないことの十分条件ではあるが、必要条件ではない。

次いで、十分原因モデルを基に、交絡バイアスが存在しないための共変量バランスの条件を提唱した。本研究では、十分原因の各背景因子 C_1, C_2, C_3 を関心対象である一連の共変量として用いる。この概念化は、各因子がアウトカム変数 Y の予測因子であることを適切に表す。この議論を視覚的に示すため、以下では十分原因を因果ダイアグラム (非循環有向グラフ: directed acyclic graph) の枠組みに組み込んで説明する^{25, 26}。

理想的な無作為化比較試験では、曝露群と非曝露群は同等であり交絡バイアスがないと期待される。また、十分原因モデルの観点から考慮すると、曝露群と非曝露群の間で背景因子 C_1, C_2, C_3 がすべて同等に分布することが期待される。そのため、各背景因子と曝露の間には関連がなく、共変量バランスが成立することが期待される。図 1 は、このような理想的な無作為化比較試験の状況を表している。曝露 E からアウトカム Y への裏口パス (backdoor path) がないことは、曝露とアウトカムの間に交絡が生じていないことを示している。

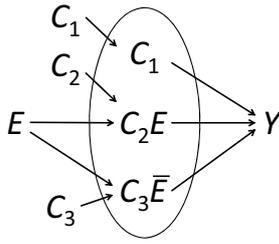


図1 十分原因を組み込んだ非循環有向グラフ

二値の曝露変数 E と二値のアウトカム変数 Y を考慮する。アウトカム変数 Y について、二値の背景因子を伴う三種類の十分原因 ($C_1, C_2E, C_3\bar{E}$) を考慮する。ここで \bar{E} は E の補事象を表す。

一方で、観察研究では一般に曝露群と非曝露群は同等ではなく、共変量バランスが成立していないことが示唆される。十分原因モデルの観点から考慮すると、これは曝露群と非曝露群の間で背景因子の分布が同等ではないことを意味する。図2は、各背景因子が曝露と正負いずれかに関連している状況を示している。曝露 E からアウトカム Y に遮断されていない裏口パスが3本存在することは、共変量バランスが成立していない状況を便宜的に表す。なお、ここでは背景因子が曝露からアウトカムの中間因子ではないと仮定している。十分原因モデルにおける仲介 (mediation) の扱いについては他を参照されたい²⁷⁻²⁹⁾。

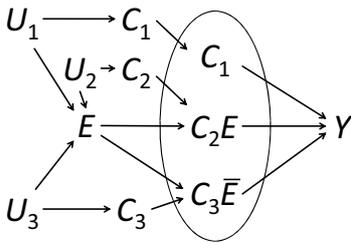


図2 十分原因モデルにおいて共変量バランスが成立しない状況を表す非循環有向グラフ

背景因子と曝露の間の (未知ないしは未測定) の共通原因を U_1, U_2, U_3 を用いて表す。

以上を踏まえて、関連のある背景因子の同時分布を曝露群と非曝露群の間で比較して共変量バランスの条件を考察した。その結果、交換可能性と同様、共変量バランスの条件は標的集団により異なることを明らかにした。標的集団が曝露群である場合の共変量バランスは以下のように示される。

$$(C_1, C_3) \parallel E \\ (\Leftrightarrow \{(s_1 + s_3) = (t_1 + t_3)\} \wedge \{(s_2 + s_4) = (t_2 + t_4)\} \wedge \{(s_5 + s_7) = (t_5 + t_7)\})$$

また、標的集団が非曝露群である場合の共変量バランスは以下のように示される。

$$(C_1, C_2) \parallel E \\ (\Leftrightarrow \{(s_1 + s_2) = (t_1 + t_2)\} \wedge \{(s_3 + s_4) = (t_3 + t_4)\} \wedge \{(s_5 + s_6) = (t_5 + t_6)\})$$

本研究では、これらの条件を「部分共変量バランス (partial covariate balance)」と呼ぶことを提唱する。これは部分交換可能性よりも強い条件である。そして、標的集団が全集団である場合の共変量バランスは以下のように示される。

$$(C_1, C_k) \parallel E \quad (k=2,3) \\ (\Leftrightarrow [\{(s_1 + s_2) = (t_1 + t_2)\} \wedge \{(s_3 + s_4) = (t_3 + t_4)\} \wedge \{(s_5 + s_6) = (t_5 + t_6)\}] \\ \wedge [\{(s_1 + s_3) = (t_1 + t_3)\} \wedge \{(s_5 + s_7) = (t_5 + t_7)\}])$$

この条件を「(狭義の) 共変量バランス」と呼ぶことを提唱する。(狭義の) 共変量バランスは、(狭義の) 交換可能性よりも強い条件である。最後に、三つの背景因子の同時分布と曝露変数 E の独立性 $(C_1, C_2, C_3) \perp\!\!\!\perp E \Leftrightarrow (s_1, s_2, s_3, s_4, s_5, s_6, s_7, s_8) = (t_1, t_2, t_3, t_4, t_5, t_6, t_7, t_8)$ を「完全共変量バランス (full covariate balance)」と呼ぶことを提唱する。完全共変量バランスは、部分共変量バランスや狭義の共変量バランスより強い条件である。留意すべき点として、完全共変量バランスは完全交換可能性よりも強い条件であるが、部分共変量バランスや狭義の共変量バランスは完全交換可能性より強くもなければ弱くもない条件である。表 2 に、交換可能性と共変量バランスの関係性をまとめる。

これらの議論は、標的集団にかかわらず、共変量バランスは交絡バイアスを制御するための十分条件であるものの、必要条件ではないことを示している。言い換えるならば、交絡バイアスの存在は共変量バランスが成立していないことを示唆するが、共変量バランスが成立していないことは必ずしも交絡バイアスの存在を示唆するわけではない。一般に共変量は、曝露と因果的相互作用 (causal coaction) を有しないもの (C_1) と、曝露と何らかの因果的相互作用を有するもの (C_2, C_3) に分けられるため、共変量バランスにおいて標的集団が重要な役割を果たすのである。

表 2 交換可能性と共変量バランスの関係性^a

標的集団	反事実モデル	十分原因モデル	
	交換可能性	背景因子の観点から見た交換可能性	共変量バランス
曝露群	$Y_0 \perp\!\!\!\perp E$	$\Leftrightarrow \max(C_1, C_3) \perp\!\!\!\perp E$	$\Leftarrow (C_1, C_3) \perp\!\!\!\perp E$
非曝露群	$Y_1 \perp\!\!\!\perp E$	$\Leftrightarrow \max(C_1, C_2) \perp\!\!\!\perp E$	$\Leftarrow (C_1, C_2) \perp\!\!\!\perp E$
全集団	$Y_e \perp\!\!\!\perp E \quad (e=0,1)$	$\Leftrightarrow \max(C_1, C_k) \perp\!\!\!\perp E \quad (k=2,3)$	$\Leftarrow (C_1, C_k) \perp\!\!\!\perp E \quad (k=2,3)$

	↑↑	↑↑	↑↑
	$(Y_0, Y_1) \perp\!\!\!\perp E$	$\Leftrightarrow (\max(C_1, C_3), \max(C_1, C_2)) \perp\!\!\!\perp E$	
		$\Leftarrow (C_1, C_2, C_3) \perp\!\!\!\perp E$	

^a 記号については表 1 を参照されたい。

IV 考 察

本研究では、十分原因モデルを基に、十分原因の各背景因子を関心対象である一連の共変量として用いた。その上で、関連のある背景因子の同時分布を曝露群と非曝露群間で比較して共変量バランスの条件を考察した。反事実モデルと十分原因モデルの対応関係を考察した結果、標的集団にかかわらず、共変量バランスが交絡バイアスを制御するための十分条件であるものの、必要条件ではないことを示した。この知見は、十分原因モデルが反事実モデルよりも“細かい”モデルであるという事実と合致する。加えて、十分原因を因果ダイアグラムの枠組みで示すことにより、共変量バランスの概念モデルを視覚的に表した。この結果、適切な分析において重要となる共変量バランスの概念を精緻化することに成功した。この概念化は、交絡の意味を明確にするうえで非常に有用なものである。

V 結 語

本研究成果は、無作為化比較試験が困難な状況でもバイアスを同定・制御するための新たな研究理論構築の基盤となるものである。特に、交絡バイアスの同定・制御方法は健康格差の原因究明において大きな障壁になっているため³⁰⁻³³⁾、共変量バランスの概念を精緻化した本研究の学術的意義は非常に高い。また、これらの理論的発展は、健康格差拡大の原因究明に向けた高度な分析方法と理論の再構築にもつながると期待できる。なお、本研究結果の詳細については国際学術誌 *Annals of Epidemiology* に発表した論文を参照されたい³⁴⁾。

最後に、本研究を実施するために公益財団法人総合健康推進財団より貴重な助成を受けたことに対して、深甚なる謝意を表する。

【参考文献】

- 1) Suzuki E. Toward a further understanding of causality in medicine: a contribution to "seizon and life sciences". *J Seizon and Life Sci.* 2016;27:97-106.
- 2) Suzuki E, Mitsuhashi T, Tsuda T, Yamamoto E. A simple example as a pedagogical device? *Ann Epidemiol.* 2014;24:560-1.
- 3) Suzuki E, Yamamoto E. Further refinements to the organizational schema for causal effects. *Epidemiology.* 2014;25:618-9.
- 4) Suzuki E, Mitsuhashi T, Tsuda T, Yamamoto E. A typology of four notions of confounding in epidemiology. *J Epidemiol.* 2017;27:49-55.
- 5) Greenland S, Poole C. Invariants and noninvariants in the concept of interdependent effects. *Scand J Work Environ Health.* 1988;14:125-9.
- 6) Flanders WD. On the relationship of sufficient component cause models with potential outcome (counterfactual) models. *Eur J Epidemiol.* 2006;21:847-53.
- 7) VanderWeele TJ, Hernán MA. From counterfactuals to sufficient component causes and vice versa. *Eur J Epidemiol.* 2006;21:855-8.
- 8) Suzuki E, Yamamoto E, Tsuda T. On the link between sufficient-cause model and potential-outcome model. *Epidemiology.* 2011;22:131-2.
- 9) Greenland S, Robins JM. Identifiability, exchangeability, and epidemiological confounding. *Int J Epidemiol.* 1986;15:413-9.
- 10) Rothman KJ. Causes. *Am J Epidemiol.* 1976;104:587-92.
- 11) Greenland S, Lash TL, Rothman KJ. Concepts of interaction. In: Rothman KJ, Greenland S, Lash TL, eds. *Modern Epidemiology*. 3rd ed. Philadelphia, PA: Lippincott Williams & Wilkins, 2008:71-83.
- 12) VanderWeele TJ. Attributable fractions for sufficient cause interactions. *Int J Biostat.* 2010;6:5. doi:10.2202/1557-4679.1202.
- 13) Suzuki E, Yamamoto E, Tsuda T. On the relations between excess fraction, attributable fraction, and etiologic fraction. *Am J Epidemiol.* 2012;175:567-75.
- 14) Greenland S, Robins JM, Pearl J. Confounding and collapsibility in causal inference. *Stat Sci.* 1999;14:29-46.
- 15) Greenland S, Morgenstern H. Confounding in health research. *Annu Rev Public Health.* 2001;22:189-212.
- 16) Greenland S, Rothman KJ, Lash TL. Measures of effect and measures of association. In: Rothman KJ, Greenland S, Lash TL, eds. *Modern Epidemiology*. 3rd ed. Philadelphia, PA: Lippincott Williams & Wilkins, 2008:51-70.
- 17) Greenland S. Confounding. In: Boslaugh S, ed. *Encyclopedia of Epidemiology*. Thousand Oaks, CA: Sage Publications, 2008:227-32.
- 18) Greenland S, Robins JM. Identifiability, exchangeability and confounding revisited. *Epidemiol Perspect Innov.* 2009;6:4. doi:10.1186/1742-5573-6-4.
- 19) Maldonado G, Greenland S. Estimating causal effects. *Int J Epidemiol.* 2002;31:422-9.
- 20) Flanders WD, Johnson CY, Howards PP, Greenland S. Dependence of confounding on the target population: a modification of causal graphs to account for co-action. *Ann Epidemiol.* 2011;21:698-705.

- 21) Suzuki E, Mitsuhashi T, Tsuda T, Yamamoto E. A counterfactual approach to bias and effect modification in terms of response types. *BMC Med Res Methodol.* 2013;13:101. doi:10.1186/1471-2288-13-101.
- 22) Suzuki E, Tsuda T, Mitsuhashi T, Mansournia MA, Yamamoto E. Errors in causal inference: an organizational schema for systematic error and random error. *Ann Epidemiol.* 2016;26:788-93.
- 23) VanderWeele TJ. Confounding and effect modification: distribution and measure. *Epidemiol Method.* 2012;1:55-82. doi:10.1515/2161-962X.1004.
- 24) Hernán MA, Robins JM. *Causal Inference.* Boca Raton, FL: Chapman & Hall/CRC, forthcoming; 2018.
- 25) VanderWeele TJ, Robins JM. Directed acyclic graphs, sufficient causes, and the properties of conditioning on a common effect. *Am J Epidemiol.* 2007;166:1096-104.
- 26) VanderWeele TJ, Robins JM. Minimal sufficient causation and directed acyclic graphs. *Ann Stat.* 2009;37:1437-65.
- 27) Hafeman DM. A sufficient cause based approach to the assessment of mediation. *Eur J Epidemiol.* 2008;23:711-21.
- 28) VanderWeele TJ. Mediation and mechanism. *Eur J Epidemiol.* 2009;24:217-24.
- 29) Suzuki E, Yamamoto E, Tsuda T. Identification of operating mediation and mechanism in the sufficient-component cause framework. *Eur J Epidemiol.* 2011;26:347-57.
- 30) Suzuki E, Kashima S, Kawachi I, Subramanian SV. Social and geographic inequalities in premature adult mortality in Japan: a multilevel observational study from 1970 to 2005. *BMJ Open.* 2012;2:e000425. doi:10.1136/bmjopen-2011-000425.
- 31) Suzuki E, Kashima S, Kawachi I, Subramanian SV. Geographic inequalities in all-cause mortality in Japan: compositional or contextual? *PLoS One.* 2012;7:e39876. doi:10.1371/journal.pone.0039876.
- 32) Suzuki E, Kashima S, Kawachi I, Subramanian SV. Social and geographical inequalities in suicide in Japan from 1975 through 2005: a census-based longitudinal analysis. *PLoS One.* 2013;8:e63443. doi:10.1371/journal.pone.0063443.
- 33) Suzuki E, Kashima S, Kawachi I, Subramanian SV. Prefecture-level economic conditions and risk of suicide in Japan: a repeated cross-sectional analysis 1975-2010. *Eur J Public Health.* 2014;24:949-54.
- 34) Suzuki E, Tsuda T, Yamamoto E. Covariate balance for no confounding in the sufficient-cause model. *Ann Epidemiol.* 2018;28:48-53.