

令和 5 年 6 月 1 日現在

機関番号：15301

研究種目：基盤研究(C)（一般）

研究期間：2020～2022

課題番号：20K10471

研究課題名（和文）大規模疫学研究における傾向スコアモデル開発：AIを組み込んだ因果推論手法の構築

研究課題名（英文）Development of propensity score models in large epidemiologic studies: causal inference and AI

研究代表者

鈴木 越治（SUZUKI, Etsuji）

岡山大学・医歯薬学域・研究准教授

研究者番号：10627764

交付決定額（研究期間全体）：（直接経費） 3,300,000円

研究成果の概要（和文）：新たな因果モデルとして提唱されているmarginal sufficient component cause modelの有用性について評価し、理論的問題や限界を考察した。また、因果メカニズムの観点からヒルの基準について考察し、因果律を評価するために、適切な因果モデルを相補的に用いる重要性を示した。さらに、因果推論の観点から介入効果を評価する指標を構築した。因果メカニズムの評価や交絡の問題について理論を構築し、AI時代における疫学の展望について考察した。

研究成果の学術的意義や社会的意義

本研究成果は、近年注目されている因果推論手法を活用したものであり、リアルワールドデータから因果関係を評価する上で有用な理論的知見である。これらの知見は、十分原因モデルと潜在アウトカムモデルという重要な2種類の因果モデルの対応を発展させており、傾向スコアモデルを開発する基盤となる。また、交絡バイアスなどを適切に制御し、因果メカニズムを評価するために重要な研究成果である。

研究成果の概要（英文）：We evaluated the usefulness of the marginal sufficient component cause model, a newly proposed causal model, by discussing its theoretical problems and limitations. Furthermore, we examined the Hill's viewpoints from the perspective of causal mechanisms and showed the importance of using appropriate causal models in a complementary manner to evaluate causality. Moreover, we developed measures for assessing intervention effects within the framework of causal inference. We developed theories regarding the assessment of causal mechanisms and confounding, offering a prospect for the field of epidemiology in the AI era.

研究分野：疫学

キーワード：因果推論 因果関係 AI 傾向スコア 疫学

1. 研究開始当初の背景

(1) 近年、医療ビッグデータやAIが大きな注目を集めている。しかし、単にビッグデータをAIで分析するだけでは、臨床や地域保健において誤った結論に至りかねない。なぜならば、これらは「相関関係」を表すに過ぎず、「因果関係」に迫ることができないからである。「因果関係」に迫るためには、高度な「因果推論」手法の構築とデータ分析により、交絡バイアスなどの誤差の問題に対処することが必要である。

(2) 交絡バイアスなどを制御するため、多くの疫学研究では「傾向スコア」モデルが用いられている。しかし、その理論的背景は十分に理解されておらず、専門家の間でもコンセンサスが得られていない。そのため、傾向スコアモデルや、それを発展させた高度な分析手法が正しく用いられず、大規模疫学研究でも因果関係に迫れていない懸念がある。

2. 研究の目的

本研究の目的は、大規模疫学研究で因果関係に迫る理論的問題に取り組み、交絡バイアスなどを制御するための新たな因果推論手法を構築することである。さらに、機械学習などのAI技術も応用することで、疫学研究の質の向上に資することをめざす。

3. 研究の方法

データに基づいて原因と結果の関係を見極め、その因果メカニズムを推論するためには、適切な因果モデルを構築して媒介や交互作用などの問題を扱う必要がある。本研究では、因果関係のメカニズムを「十分原因モデル」を用いて概念化した。「十分原因モデル」は、1976年に疫学分野で発展した因果モデルであり、多要因疾患の因果関係をモデル化する上で特に有用である。次いで、「十分原因モデル」を「潜在アウトカムモデル」という別の因果モデルと組み合わせることにより、データからメカニズムを識別した。「潜在アウトカムモデル」は因果関係を定量的に評価する重要なモデルであり、その起源は20世紀初頭に遡る。これら二つの異なる因果モデルの関係性を精緻に構築し、新たな因果推論手法を構築した。

4. 研究成果

(1) 近年、新たな因果モデルとして **marginal sufficient component cause model** (周辺十分構成原因モデル) が提唱され、幾つかの論文で用いられてきた。この新たな因果モデルを提唱した研究者らが強調している **agonism** の現象にフォーカスを当て、特に三つの観点から **marginal sufficient component cause model** の有用性を考察した。結論として、**marginal sufficient component cause model** には大きなメリットは見られないこと、また、従来の十分原因モデルと潜在アウトカムモデルの対応関係を理解することが因果律の深い理解につながることを明らかにした。研究成果を *Epidemiology* に論文発表し (①)、54th Annual Meeting of the Society for Epidemiologic Research で口頭発表を行った (図1)。

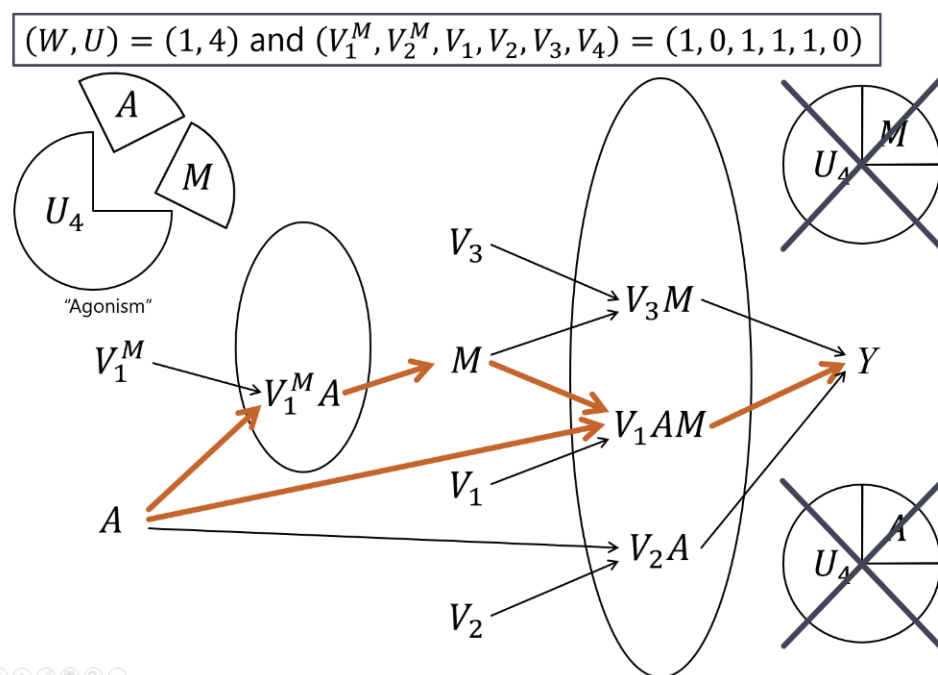


図1. 54th Annual Meeting of the Society for Epidemiologic Research の発表スライドより抜粋

(2) 医学分野では長年、Austin Bradford Hill 卿により提唱された「ヒルの基準 (Hill's viewpoints)」をもとに因果関係が論じられてきた。医学における因果関係の問題に関する考察を深めるため、メカニズムの観点から「ヒルの基準」の「強固性」、「関連の時間性」、「妥当性」に着目し、因果関係を評価する方法を考察した。さらに、哲学分野で因果関係を論じる際に用いられる「INUS条件」と十分原因モデルの関係性も考察した。「INUS条件」は、1965年に哲学者 Mackie により提唱されたものである。これらの考察は、本研究が扱った因果関係の問題が、自然科学の根幹であることを示している。結論として、因果律を評価するためには、適切な因果モデルを相補的に用いて吟味することが重要であることを示した。研究成果を *European Journal of Epidemiology* に論文発表した (2) (図2)。

Sufficient cause model		Potential-outcome model ^b				Proportions of risk status types in population with (A, C) = ^{c, d, e}					
Risk status type	Background factors of sufficient causes	Response type	Potential outcomes of Y				(1, 1)	(0, 1)	(1, 0)	(0, 0)	Total ^f
	U_1 U_2 U_3 U_4		Y^{11}	Y^{01}	Y^{10}	Y^{00}	(N = 100)	(N = 100)	(N = 100)	(N = 100)	(N = 400)
1	1 1 1 1	1	1 (S ₁ or S ₂ or S ₃ or S ₄)	1 (S ₁ or S ₂)	1 (S ₁ or S ₂)	1 (S ₁)	p_1	q_1	r_1	s_1	t_1
2	1 1 1 0	1	1 (S ₁ or S ₂ or S ₃)	1 (S ₁ or S ₂)	1 (S ₁ or S ₂)	1 (S ₁)	p_2	q_2	r_2	s_2	t_2
3	1 1 0 1	1	1 (S ₁ or S ₂ or S ₄)	1 (S ₁)	1 (S ₁ or S ₂)	1 (S ₁)	p_3	q_3	r_3	s_3	t_3
4	1 1 0 0	1	1 (S ₁ or S ₂)	1 (S ₁)	1 (S ₁ or S ₂)	1 (S ₁)	p_4	q_4	r_4	s_4	t_4
5	1 0 1 1	1	1 (S ₁ or S ₃ or S ₄)	1 (S ₁ or S ₃)	1 (S ₁)	1 (S ₁)	p_5	q_5	r_5	s_5	t_5
6	1 0 1 0	1	1 (S ₁ or S ₃)	1 (S ₁ or S ₃)	1 (S ₁)	1 (S ₁)	p_6	q_6	r_6	s_6	t_6
7	1 0 0 1	1	1 (S ₁ or S ₄)	1 (S ₁)	1 (S ₁)	1 (S ₁)	p_7	q_7	r_7	s_7	t_7
8	1 0 0 0	1	1 (S ₁)	1 (S ₁)	1 (S ₁)	1 (S ₁)	p_8	q_8	r_8	s_8	t_8
9	0 1 1 1	2	1 (S ₂ or S ₃ or S ₄)	1 (S ₂)	1 (S ₂)	0	p_9	q_9	r_9	s_9	t_9
10	0 1 1 0	2	1 (S ₂ or S ₃)	1 (S ₂)	1 (S ₂)	0	p_{10}	q_{10}	r_{10}	s_{10}	t_{10}
11	0 1 0 1	3	1 (S ₂ or S ₄)	0	1 (S ₂)	0	p_{11}	q_{11}	r_{11}	s_{11}	t_{11}
12	0 1 0 0	3	1 (S ₂)	0	1 (S ₂)	0	p_{12}	q_{12}	r_{12}	s_{12}	t_{12}
13	0 0 1 1	4	1 (S ₃ or S ₄)	1 (S ₃)	0	0	p_{13}	q_{13}	r_{13}	s_{13}	t_{13}
14	0 0 1 0	4	1 (S ₃)	1 (S ₃)	0	0	p_{14}	q_{14}	r_{14}	s_{14}	t_{14}
15	0 0 0 1	5	1 (S ₄)	0	0	0	p_{15}	q_{15}	r_{15}	s_{15}	t_{15}
16	0 0 0 0	6	0	0	0	0	p_{16}	q_{16}	r_{16}	s_{16}	t_{16}

図2. 十分原因の正の単調性仮定における十分原因と潜在アウトカムの対応 (文献②)

(3) 疫学研究では介入効果を評価するために、人口寄与分画などの指標がしばしば用いられる。本研究では、寄与分画 (attributable fraction) について、潜在アウトカムモデルの観点から新たな理論を明らかにした。疫学研究では、Levin formula や Miettinen formula を用いて人口寄与分画を算出することが多い。これらの式の関係性を示し、バイアスが生じない必要十分条件などを明らかにした。さらに、preventable fraction と prevented fraction について考察し、新たな指標として attributed fraction を提唱した。また、causal excess fraction、preventive excess fraction についても考察した (図3)。研究成果を *Journal of Causal Inference* に論文発表し (3)、55th Annual Meeting of the Society for Epidemiologic Research と The 2023 American Causal Inference Conference でポスター発表した。

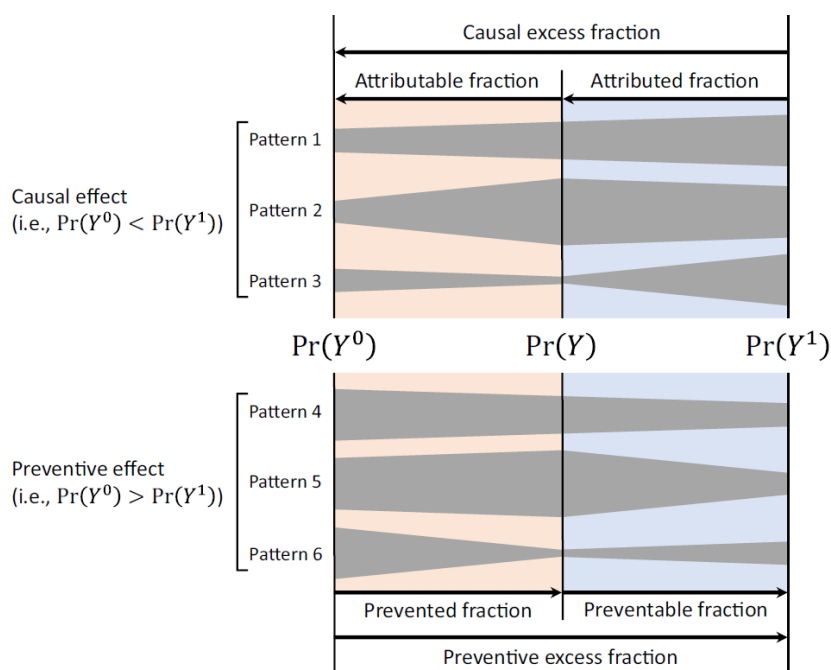


図3. 6つのパターンにおける6つの指標の関連性 (文献③)

(4) 対象集団で曝露と非曝露を入れ替えた場合に、関連指標の値が変わらないことがある。本研究では、曝露と非曝露を入れ替えた場合に関連指標の値が変わるか否かに基づくアプローチが、潜在アウトカムモデルにおける交絡とどのような関係にあるかを評価した。一般的に、曝露と非曝露を入れ替えた場合に関連指標の値が変わるときには、「分布交絡 (confounding in distribution)」が常に存在する。一方で、一般的に、曝露と非曝露を入れ替えた場合に関連指標の値が変わるときには、「指標交絡 (confounding in measure)」が存在するか否かは言えない。曝露と非曝露を入れ替えた場合に関連指標の値が変わらないとしても、分布交絡と指標交絡はいずれも存在し得る。結論として、本質的に、関連指標の値が変わるか否かに基づくアプローチは、交絡の定義を与えるものではないことを明らかにした。研究成果を *Journal of Epidemiology* と *Journal of Clinical Epidemiology* に論文発表し (④、⑤)、第 31 回日本疫学会学術総会で口頭発表した。また、第 94 回日本産業衛生学会で口頭発表し、交絡について考察した。

(5) 観察データに基づいて因果推論をするためには、調整すべき共変量を適切に選び、交絡バイアスを対処する必要がある。その際に用いられる傾向スコア法の理論的背景に着目し、変数選択法に関する理論を考察した。特に、標的集団が曝露群や非曝露群などの部分集団である場合に、部分交換可能性と部分共変量バランスの関係性を明らかにした。この知見は、十分原因モデルと潜在アウトカムモデルという重要な 2 種類の因果モデルの対応関係に基づいており、傾向スコアモデルを開発する基盤となる。研究成果を *Epidemiology* に論文発表した (⑥)。

(6) 周辺構造モデルに関する特別論文を *Journal of Epidemiology* に発表した (⑦)。本論文では、周辺構造モデルと逆確率重み付けは区別すべきこと、周辺構造モデルは推定対象とする因果母数への仮定を表す一方で、曝露確率モデルは観察分布に対する制約を与えること、周辺構造モデルと曝露確率モデルは用いる目的が異なっており、それぞれのモデルの誤特定によるバイアスの生じ方も異なることを詳説した。加えて、現実のデータ解析でモデルを特定する際に直面する課題は、周辺構造モデルと曝露確率モデルでは異なることや、g-公式は、因果効果の識別には逆確率重み付けと同じ条件を要するが、g-公式で周辺構造モデルが推定できるのは飽和した場合に限られることも論じた。

<引用文献>

- ① Suzuki E, Yamamoto E. Marginal sufficient component cause model: an emerging causal model with merits? *Epidemiology*. 2021;32(6):838–845. (doi: 10.1097/EDE.0000000000001411)
- ② Suzuki E, Yamamoto E. *Strength* in causality: discerning causal mechanisms in the sufficient cause model. *Eur J Epidemiol*. 2021;36(9):899–908. (doi: 10.1007/s10654-021-00798-6)
- ③ Suzuki E, Yamamoto E. Attributable fraction and related measures: Conceptual relations in the counterfactual framework. *J Causal Inference*. 2023;11(1):20210068. (doi: 10.1515/jci-2021-0068)
- ④ Suzuki E, Yamamoto M, Yamamoto E. Exchangeability of measures of association before and after exposure status is flipped: its relationship with confounding in the counterfactual model. *J Epidemiol*. (In press). (doi: 10.2188/jea.JE20210352)
- ⑤ Suzuki E, Yamamoto M, Yamamoto E. A general explanation of the counterfactual definition of confounding. *J Clin Epidemiol*. 2022;148:189–192. (doi: 10.1016/j.jclinepi.2022.02.002)
- ⑥ Suzuki E, Yamamoto E. Re: A graphical description of partial exchangeability. *Epidemiology*. 2021;32(2):e7–e9. (doi: 10.1097/EDE.0000000000001306)
- ⑦ Shinozaki T, Suzuki E. Understanding marginal structural models for time-varying exposures: pitfalls and tips. *J Epidemiol*. 2020;30(9):377–389. (doi: 10.2188/jea.JE20200226)

5. 主な発表論文等

〔雑誌論文〕 計23件（うち査読付論文 21件／うち国際共著 2件／うちオープンアクセス 13件）

1. 著者名 Suzuki Etsuji, Yamamoto Eiji	4. 巻 11
2. 論文標題 Attributable fraction and related measures: Conceptual relations in the counterfactual framework	5. 発行年 2023年
3. 雑誌名 Journal of Causal Inference	6. 最初と最後の頁 -
掲載論文のDOI（デジタルオブジェクト識別子） 10.1515/jci-2021-0068	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -
1. 著者名 Suzuki Etsuji, Yamamoto Michio, Yamamoto Eiji	4. 巻 -
2. 論文標題 Exchangeability of Measures of Association Before and After Exposure Status Is Flipped: Its Relationship With Confounding in the Counterfactual Model	5. 発行年 2023年
3. 雑誌名 Journal of Epidemiology	6. 最初と最後の頁 -
掲載論文のDOI（デジタルオブジェクト識別子） 10.2188/jea.JE20210352	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -
1. 著者名 Shrier Ian, Suzuki Etsuji	4. 巻 51
2. 論文標題 The primary importance of the research question: implications for understanding natural versus controlled direct effects	5. 発行年 2022年
3. 雑誌名 International Journal of Epidemiology	6. 最初と最後の頁 1041～1046
掲載論文のDOI（デジタルオブジェクト識別子） 10.1093/ije/dyac090	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 該当する
1. 著者名 Suzuki Etsuji, Yamamoto Michio, Yamamoto Eiji	4. 巻 148
2. 論文標題 A general explanation of the counterfactual definition of confounding	5. 発行年 2022年
3. 雑誌名 Journal of Clinical Epidemiology	6. 最初と最後の頁 189～192
掲載論文のDOI（デジタルオブジェクト識別子） 10.1016/j.jclinepi.2022.02.002	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

1. 著者名 Wang S, Puggioni G, Wu J, Meador KJ, Caffrey A, Wyss R, Slaughter JL, Suzuki E, Ward KE, Lewkowitz AK, Wen X	4. 巻 -
2. 論文標題 Prenatal exposure to opioids and neurodevelopmental disorders in children: a Bayesian mediation analysis	5. 発行年 2023年
3. 雑誌名 American Journal of Epidemiology	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 該当する

1. 著者名 Suzuki Etsuji, Yamamoto Eiji	4. 巻 32
2. 論文標題 Marginal Sufficient Component Cause Model: An Emerging Causal Model With Merits?	5. 発行年 2021年
3. 雑誌名 Epidemiology	6. 最初と最後の頁 838 ~ 845
掲載論文のDOI (デジタルオブジェクト識別子) 10.1097/EDE.0000000000001411	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Suzuki Etsuji, Yamamoto Eiji	4. 巻 36
2. 論文標題 Strength in causality: discerning causal mechanisms in the sufficient cause model	5. 発行年 2021年
3. 雑誌名 European Journal of Epidemiology	6. 最初と最後の頁 899 ~ 908
掲載論文のDOI (デジタルオブジェクト識別子) 10.1007/s10654-021-00798-6	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Suzuki Etsuji, Yamamoto Eiji	4. 巻 32
2. 論文標題 Re: A graphical description of partial exchangeability	5. 発行年 2020年
3. 雑誌名 Epidemiology	6. 最初と最後の頁 e7 ~ e9
掲載論文のDOI (デジタルオブジェクト識別子) 10.1097/EDE.0000000000001306	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Shinozaki Tomohiro, Suzuki Etsuji	4. 巻 30
2. 論文標題 Understanding Marginal Structural Models for Time-Varying Exposures: Pitfalls and Tips	5. 発行年 2020年
3. 雑誌名 Journal of Epidemiology	6. 最初と最後の頁 377 ~ 389
掲載論文のDOI (デジタルオブジェクト識別子) 10.2188/jea.JE20200226	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Yorifuji Takashi, Suzuki Etsuji	4. 巻 133
2. 論文標題 データサイエンスにおける人工知能 (AI) と疫学の位置付け 予測と因果推論の違い	5. 発行年 2021年
3. 雑誌名 岡山医学会雑誌	6. 最初と最後の頁 55 ~ 57
掲載論文のDOI (デジタルオブジェクト識別子) 10.4044/joma.133.55	査読の有無 無
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

[学会発表] 計17件 (うち招待講演 6件 / うち国際学会 8件)

1. 発表者名 Suzuki E, Yamamoto E
2. 発表標題 Attributable fraction and related measures: conceptual relations in the counterfactual framework
3. 学会等名 55th Annual Meeting of the Society for Epidemiologic Research (国際学会)
4. 発表年 2022年

1. 発表者名 Suzuki E
2. 発表標題 Mechanisms and sufficient causes: a theoretical perspective to understand the etiology of suicide
3. 学会等名 19th Congress of the WPA Epidemiology and Public Health Section (招待講演) (国際学会)
4. 発表年 2022年

1. 発表者名 Suzuki E, Yamamoto E
2. 発表標題 Attributable fraction and related measures: conceptual relations in the counterfactual framework
3. 学会等名 The 2023 American Causal Inference Conference (国際学会)
4. 発表年 2023年

1. 発表者名 Suzuki E, Yamamoto T, Yamamoto E
2. 発表標題 Understanding vaccine efficacy in the counterfactual framework
3. 学会等名 56th Annual Meeting of the Society for Epidemiologic Research (国際学会)
4. 発表年 2023年

1. 発表者名 鈴木越治
2. 発表標題 データから因果関係に迫る疫学理論の開発と実践
3. 学会等名 第33回日本疫学会学術総会 (招待講演)
4. 発表年 2023年

1. 発表者名 Suzuki E, Yamamoto E
2. 発表標題 Marginal sufficient component cause model: an emerging causal model with authenticity?
3. 学会等名 54th Annual Meeting of the Society for Epidemiologic Research (国際学会)
4. 発表年 2021年

1. 発表者名 鈴木越治, 篠崎智大, 三橋利晴, 津田敏秀, 山本英二
2. 発表標題 交絡はバイアスと区別すべきか AI時代における疫学の展望
3. 学会等名 第94回日本産業衛生学会
4. 発表年 2021年

1. 発表者名 Suzuki E, Shinozaki T, Yamamoto E
2. 発表標題 The meaning of arrows: reflections on the assumption of faithfulness in causal directed acyclic graphs
3. 学会等名 53rd Annual Meeting of the Society for Epidemiologic Research (国際学会)
4. 発表年 2020年

1. 発表者名 Shrier I, Suzuki E
2. 発表標題 A different perspective of “cross-world independence assumption” and the utility of natural effects compared to controlled effects
3. 学会等名 53rd Annual Meeting of the Society for Epidemiologic Research (国際学会)
4. 発表年 2020年

1. 発表者名 鈴木越治, 山本倫生, 山本英二
2. 発表標題 曝露と非曝露を入れ替えた場合に関連指標が変わるか否か 反事実モデルにおける交絡との関係
3. 学会等名 第31回日本疫学会学術総会
4. 発表年 2021年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

岡山大学大学院医歯薬学総合研究科疫学・衛生学分野ホームページ
<https://www.unit-gp.jp/eisei/wp/>

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究分担者	頼藤 貴志 (YORIFUJI Takashi) (00452566)	岡山大学・医歯薬学域・教授 (15301)	
研究分担者	篠崎 智大 (SHINOZAKI Tomohiro) (60644482)	東京理科大学・工学部情報工学科・講師 (32660)	

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関			
	米国	ハーバード大学	マサチューセッツ工科大学	ロードアイランド大学
カナダ	マギル大学			