

大規模疫学データを活用した口腔の疫学研究

松山祐輔

Oral epidemiologic research utilizing big data

Yusuke Matsuyama, DDS, PhD

抄 録

ランダム化比較試験（RCT）は単一の研究としてエビデンスレベルが最も高いとされる。しかし、RCT は大規模かつ長期的な実施が困難であり、研究参加者の代表性にも限界がある。一方、近年の疫学理論の発展と大規模データの蓄積により、観察研究に基づくエビデンスの活用が注目されている。RCT と観察研究はそれぞれ利点と限界があり、両者の特性を理解し研究を計画・実施することが重要である。本稿では、大規模データを用いた歯科領域の観察研究を紹介しながら、観察研究の質を高めるポイントを述べる。

キーワード

観察研究，エビデンス，口腔と全身

ABSTRACT

Randomized controlled trials (RCTs) are considered as the study design with the highest level of evidence. However, it is often challenging to conduct large-scale and long-term RCTs, and the generalizability of the findings from RCTs is limited. In contrast, recent advances in epidemiologic theory and the accumulation of large-scale data have increased the value of evidence derived from observational studies. RCTs and observational studies have different advantages and disadvantages. Therefore, it is important to understand their characteristics. This article introduces observational studies in dentistry and explains key points for increasing the value of evidence derived from observational studies.

Key words:

Observational study, Evidence, Oral and systemic health

I. はじめに

Evidence-Based Dentistry の概念が広く普及し、エビデンスに基づく歯科医療の実践が求められている。一般に用いられているエビデンスレベルのピラミッド（図 1）において、ランダム化比較試験（RCT）は単一の研究として最も高いエビデンスレベルを有するとされている。しかし、RCT は大規模かつ長期間にわたる実施が困難であり、研究参加者の代表性にも

限界がある¹⁾。

一方、近年の疫学理論の発展と大規模データの蓄積により、観察研究に基づく質の高いエビデンスの活用が注目されている。RCT と観察研究はそれぞれに利点と限界があり、質の高いエビデンスを生み出すためには、それぞれの特性を正しく理解し研究を計画・実施する必要がある。たとえば、口腔と全身の健康との関連を長期にわたり追跡する研究や、口腔疾患が社会に及ぼす影響を明らかにする研究は、観察研究が適している領域である。ただし、観察研究では、逆の因果

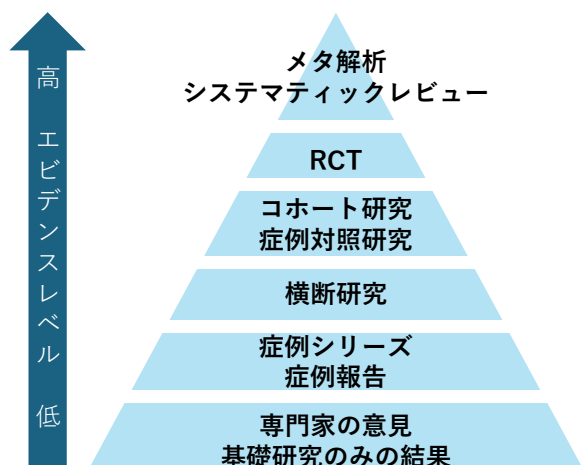


図1 エビデンスレベルのピラミッド

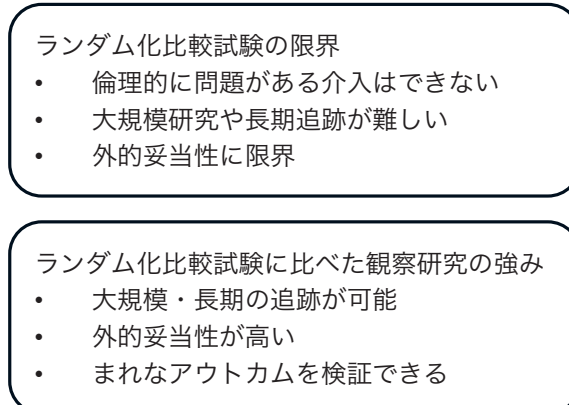


図2 ランダム化比較試験と観察研究の比較

関係や交絡といった問題への対処が依然として大きな課題である。

本稿では、口腔と全身の健康との因果関係や、実社会における口腔疾患の影響に焦点を当て、大規模データを活用した歯科領域の疫学研究を紹介する。あわせて、観察研究に基づく質の高いエビデンスを生み出すために重要となるポイントを述べる。RCTと観察研究は検証できる研究仮説が異なるため、どちらが優れているかを問うのではなく、両者の強みを活かし、互いの限界を補い合いながら補完的に活用していくことが望ましい。本稿が、エビデンスに基づく歯科医療の推進に資する一助となれば幸いである。

II. RCTと観察研究の特徴

よくデザインされたRCTは、研究参加者の背景要因（交絡因子）が介入群と対照群でバランスされており、集団の差をとることで介入効果を推定できる。しかし、RCTにはいくつかの限界がある。第一に、倫理的に問題がある介入はできない。たとえば、有効性が確立している治療を無作為に割り付けることは、対照群の参加者に不利益をもたらすため実施が困難である。また、残存歯数の影響を検証するために抜歯を無作為に割り付けることは、倫理的に明らかに不適切である。第二に、RCTは大規模研究や長期追跡が困難である。これには研究資金や人員といったリソースの制限が関係している。たとえば、Fotiらは、食塩摂取制限の心血管疾患による死亡への影響を検証するRCTに必要なコストを試算した²⁾。その結果、必要なサンプルサイズは約1.7万人、人件費などを含む費

用は4億ドルに達すると報告している。このことから、口腔と全身疾患の関連のように長期の追跡を必要とする研究テーマでは、RCTの実施が現実的でない場合が多く、観察研究が適していると考えられる。第三に、RCTの参加者は、一般集団と比べて健康状態などの特性に偏りがある場合が多く、外的妥当性に限界がある。Kennedy-Martinらの報告によれば、心疾患、精神疾患、がんを対象としたRCTでは、実際の臨床患者に比べて、高齢者・女性・併存疾患を有する者が除外されやすく、RCTの適格基準を適用すると50%以上の患者が対象外となることも少なくないと報告した³⁾。このことから、RCTの結果をそのまま実際の臨床患者や実社会の人々へ外挿できるとは限らないといえる。

一方、観察研究は介入を伴わず、実社会での診療や調査データを用いるため、大規模かつ長期の追跡が比較的容易であり、まれなアウトカムの検討や、高い集団の代表性といった強みを有している（図2）。このように、観察研究と介入研究はそれぞれ異なる利点と限界を持つ。よって、研究デザインの種類にのみ基づいてエビデンスレベルの優劣を一律に判断することには限界がある。近年では、研究デザインで分けられたエビデンスの階級を再考し、個々の研究の質を重視しつつ、同時に系統的レビューをエビデンスの俯瞰のために活用するという考え方も提唱されている⁴⁾。

適切に計画された観察研究が、どの程度RCTに一致した結果を示すか検討した研究も存在する。Wangらは、レセプトデータベースを用いて、複数のRCTを厳密な研究計画に基づいた観察研究で模倣した⁵⁾。その結果、推定値はRCTとデータベース研究の間で、

全体として高い相関 (Pearson 相関係数 0.82) が示された。さらに、質の高い模倣ができたものに限定した場合、相関はさらに高く (Pearson 相関係数 0.93)、模倣が不十分な場合は相関が低下する (Pearson 相関係数 0.53) ことが報告された。つまり、条件が整えば、綿密な研究計画のもと、観察研究で RCT に近い効果を推定できることが示唆される。また、Dagan らはイスラエルにおける 120 万人分の COVID-19 ワクチン接種記録を用いて⁶⁾、先に実施された RCT⁷⁾ に近い効果の推定値を得た。これは理想的な条件下での介入効果 (efficacy) を RCT で示し、実社会における介入の有効性 (effectiveness) を観察研究で示した好例である。このように RCT と観察研究が互いに補完し合うことで、より強固なエビデンスが形成されるといえる。本稿では、こうした背景を踏まえつつ、観察研究の弱点を補い、強みを活かすための研究デザインや統計的アプローチを、歯科疫学の例を交えながら紹介する。

Ⅲ. 観察研究の弱みを補う

交絡因子によるバイアスは観察研究では避けて通れない問題である。観察研究の多くは、測定された交絡因子の影響を多変量解析や傾向スコア法などの手法で考慮し、曝露とアウトカムの関連を推定する。しかし、未測定の要因が存在する可能性や、測定指標の妥当性の問題もある。測定指標の妥当性とは、たとえば、タバコを交絡因子として考え、喫煙の有無を質問紙調査で測定したとする。この変数を統計的に調整したとしても、喫煙の期間や量の違いは考慮できておらず、交絡が残る可能性がある。

このセクションでは、こうした未測定交絡に対処できる分析手法を紹介する。しかしながら、これらの手法も完璧ではなく、強い仮定を必要とする場合も多い。仮定が妥当かどうかはデータからは判断できないことも多く、理論に基づき仮定が妥当と考えられるかを検証することが重要である。

1. 固定効果分析

固定効果分析は、複数時点のパネルデータを用いることで、性別や遺伝子など、時間で変化しない要因の影響を除外できる分析手法である⁸⁾。パネルデータで個人間の違いを除外し、個人内の変化のみを推定に活用することで、時間で変化しない要因の影響を除外する。

固定効果分析は歯科疫学領域でもしばしば用いられ

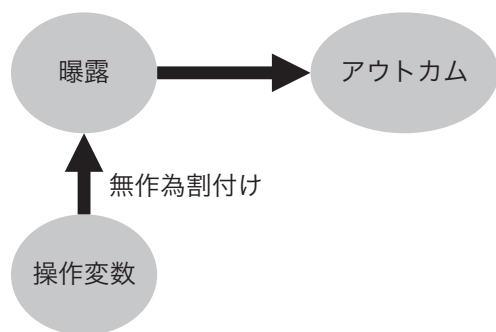
る分析手法である。Kiuchi らは、日本人高齢者のパネルデータを用いて、嚥下機能などの口腔機能が悪化した高齢者は主観的認知機能が低下すると報告した⁹⁾。また、米国の研究では、延べ測定人数 11 万人の大規模データを用いて、無歯顎になった人で心疾患既往が増加したことを報告した¹⁰⁾。さらに、英国のコホート研究からは、無歯顎になった人で孤独感が高まることも報告されている¹¹⁾。

固定効果分析は比較的シンプルな方法だが、口腔と全身の健康の関連の研究において重要な意義を持つ。その理由の一つとして、口腔疾患と全身疾患との関連が、幼少期の認知機能や社会経済的背景といった、個人に固有の交絡因子によって影響を受けている可能性が指摘されていることが挙げられる。例えば、歯の喪失と認知機能低下の関連を論じた報告¹²⁾では、幼少期に認知機能が低い人は高齢期においても認知機能が低い傾向があり、歯の喪失はライフコース全体を通じた認知機能の低さの指標にすぎない可能性があるとして述べられている。このような指摘を科学的に考慮するために、時間で変化しない個人特性を統制できる固定効果分析は有効であろう。今後は、こうした分析手法を活用し、ライフコースにわたる交絡因子の影響を統制したうえで、口腔と全身の健康との関連を明らかにすることが求められる。しかしながら、時間とともに変化する要因は、未測定の場合、固定効果分析では考慮できないことに注意が必要である。異なる仮定に基づく複数のアプローチで、結果の頑健性を確認していく必要がある。

2. 操作変数法

操作変数法は、研究者の意図によらない出来事（自然実験とも呼ばれる）を活用し、曝露の影響を検証する分析手法である¹³⁾。妥当な操作変数はいくつかの条件を満たす必要がある（図 3）。操作変数を用いた二段階推定法により、あたかも RCT が実施されたかのような状況を作り、曝露の影響を検証できる。

操作変数法は活用できる場面は限られるが、適切な操作変数を見つければ、因果関係の推定において強力な手法になりうる。Matsuyama らは、米国や英国の水道水フロリデーションの普及率の地域的・経時的な差を残存歯数に対する操作変数として、歯数と全身疾患の因果関係を検証した。その結果、子ども期に水道水フロリデーションに曝露したと考えられる参加者は、成人期に歯が 0.5-0.7 本程度多く残っており、このことが心血管疾患やうつ症状などのリスク低下に関連すると報告した¹⁴⁻¹⁶⁾。



妥当な操作変数の条件

- 曝露とアウトカムの交絡因子から独立である
- 曝露を強く予測する
- 曝露以外の経路でアウトカムに影響しない

図3 操作変数法

操作変数法を用いる場合、「どの集団における効果を推定しているか」に注意が必要である。操作変数の値と曝露の有無により、集団は4パターンに分けることができる(図4)。操作変数法は、「操作変数によりもたらされた曝露のばらつきの影響」を推定しているため、Defier(操作変数の影響を受けたことで“曝露なし”になる群)がいないという単調性の仮定のもと、Complier(操作変数の影響を受けたことで“曝露あり”になる群)を効果推定の対象としている。誰がcomplierに属するかはデータからはわからず、その属性が集団全体とかけ離れている場合、操作変数法から得られた結果は、集団全体における効果とは異なる可能性がある。

3. E-value による定量化

固定効果分析や操作変数法のように、未測定交絡に対処できる統計手法は強力だが、実施できる状況は限られる。そこで、より適用範囲の広いものとして、未測定交絡があることを受け入れたうえで、その影響が結果を覆すほど大きなものなのか、定量的に評価する方法を紹介する。E-valueはその指標のひとつで、「観察された結果を打ち消すだけの未測定交絡の強さ」を表す。より正確には、「得られた推定値を null(リスク比 = 1)にするために必要な、未測定交絡が曝露とアウトカムに関連する大きさの閾値」を示し、以下の式で求められる¹⁷⁾。求められたE-valueを既存の文献やデータと比較し、E-valueよりも大きな影響をもつ未測定交絡因子が存在するかを評価する。

Always taker 操作変数にかかわらず “曝露あり”になる群	Defier 操作変数の影響を受けた ことで“曝露なし”になる群
Complier 操作変数の影響を受けた ことで“曝露あり”になる群	Never taker 操作変数にかかわらず “曝露なし”になる群

図4 操作変数と曝露の状態による集団の分類。

$$E-value = \text{リスク比} + \sqrt{\text{リスク比} * (\text{リスク比} - 1)}$$

ここで、歯科疫学研究におけるE-valueの活用事例を紹介する。Kinoらは、残存歯数と様々な全身の健康アウトカムとの関連を検証するアウトカムワイド研究を実施し、残存歯数9本未満で歯科補綴物を使っていない場合、死亡リスクが1.33倍に高まると報告した¹⁸⁾。そして、この死亡リスク増加に対するE-valueは1.99であった。これは、観察された死亡リスク増加を未測定交絡が完全に説明するには、曝露(残存歯数・歯科補綴物使用)とアウトカム(死亡)のそれぞれについて、リスク比で1.99以上の大きさの関連が必要であることを示す。この研究において重要な交絡因子はすでに測定され統計的に考慮されていることを踏まえると、リスク比で1.99以上の関連をもつ未測定交絡が存在するとは考えにくい。よって、観察された死亡リスクの増加は因果関係である可能性が高いと解釈できる。観察研究である以上、未測定交絡がないとは言えないため、その影響の大きさを定量的に評価することは現実的で有効なアプローチであり、今後さらに普及すると考えられる。

IV. 観察研究の強みを活かす研究

このセクションでは、観察研究だからこそできた研究事例を紹介する。特に、介入研究に比べ、データの代表性が高いこと、測定時点が多いことおよび、多様な背景要因が測定されていることは観察研究の大きな強みである。それぞれについて歯科疫学の研究事例を

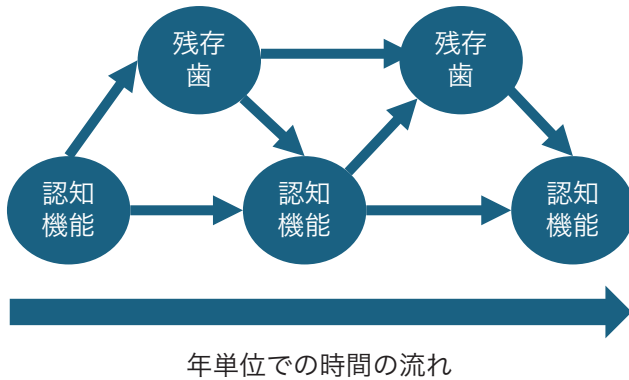


図5 時間依存性曝露の例

紹介する。

1. データの代表性の高さ

観察研究は一般に介入研究に比べ集団の代表性が高い。このため、口腔疾患の社会への影響を定量化する研究に適している。Nakazawaらは、日本人高齢者5万人を6年間追跡したデータから、さまざまな要因の死亡への集団レベルの寄与（人口寄与割合；population attributable fraction, PAF）を算出した¹⁹⁾。その結果、男性では、年齢に次いで現在歯数の寄与が2番目に大きかった。女性では、残存歯数の寄与は喫煙よりも大きかった。修正可能なリスク要因（modifiable risk factor）に限れば、残存歯数は男性では最も死亡への集団レベルの寄与が大きく、歯科疾患の社会への影響の大きさを示している。

2. 測定時点の多さ

追跡期間中に変化する口腔状態の影響を検証するためには多時点データが有用である。このため、長期追跡が比較的实施しやすい観察研究の意義は大きい。たとえば、残存歯数と認知機能の関連が指摘されているが、どちらも互いに影響し合って長期的な予後をもたらすと考えられる。図5に示すように、認知機能と残存歯数が相互に影響し合う場合、追跡期間中の認知機能を統計的に調整すると、残存歯数から最終アウトカムとしての認知機能へのパスがブロックされ、関連を過小評価する可能性がある。一方、調整しなかった場合、残存歯数と認知機能の間に交絡が残り、バイアスが生じる。このような状況では、通常の変量解析ではなく、適切な統計手法を用いることで、時間依存性曝露としての残存歯数の影響を推定する必要がある²⁰⁾。

Matsuyamaは、米国のデータを用いて、残存歯の長期的な維持が認知機能障害リスクを低下するか検証した²¹⁾。G-methodと呼ばれる分析の結果、追跡開始時から無歯だった場合に比べて、歯を維持した場合、認知機能障害リスクが3%ポイント低下すると報告した。また、Kusamaらは、口腔機能が悪いと認知症リスクが増加することを、追跡期間中の認知機能の変化とその口腔への影響を考慮し明らかにした²²⁾。また、さらに、日本人高齢者を対象とした研究から、歯科補綴物の継続使用が死亡リスクの低下に関連するとの報告もある²³⁾。時間依存性曝露の分析は歯科疫学研究ではまだ応用例が少ないが、今後増加していくと思われる。

3. 背景要因の多彩さ

大規模観察研究で測定される背景要因の多彩さを活かした研究も、観察研究の強みを活かした研究といえよう。これまでの研究では主に、背景要因はいわばノイズとして、交絡として調整に用いる研究が多かった。しかし近年、これらの背景要因に着目し、どのような背景要因を持つ人で介入や曝露の影響が大きいか調べる手法が開発され、疫学分野でも用いられてきている^{24,25)}。Matsuyamaらは、日本人高齢者の追跡データから、残存歯数が身体認知機能に及ぼす影響が、背景要因でどのように異なるか検証した²⁶⁾。分析の結果、残存歯数が少ないと身体認知機能が悪化しやすいことおよび、その効果量の大きさには個人の特性でばらつきがあり、高齢・男性・未婚・健康状態が悪い・社会経済状況が低い人で影響が大きいと報告した。ある治療が特に有効な患者の特性を明らかにする研究や、臨床や政策レベルでのより手厚い歯科介入が求められる集団を明らかにする研究に活用できる手法である。

V. おわりに

本稿では、観察研究を活用し質の高いエビデンスを生み出すための研究手法を、歯科疫学の事例を交えながら紹介した。近年の情報技術の発展により、大規模データの蓄積が進んでいる。エビデンスに基づく歯科医療の推進に資する観察研究を実施するにあたっては、観察研究のデータ構造や研究デザイン上の特性を正しく理解し、利点を活かすとともに、限界を補完できるような研究デザインを組むことが重要である。

文 献

- 1) Frieden TR. Evidence for health decision making – beyond randomized, controlled trials. *N Engl J Med* 2017; 377: 465-75.
- 2) Foti K, Foraker RE, Martyn-Nemeth P, Anderson CAM, Cook NR, Lichtenstein AH et al. Evidence-based policy making for public health interventions in cardiovascular diseases: formally assessing the feasibility of clinical trials. *Circ Cardiovascular Quality and Outcomes* 2020; 13: e006378.
- 3) Kennedy-Martin T, Curtis S, Faries D, Robinson S, Johnston J. A literature review on the representativeness of randomized controlled trial samples and implications for the external validity of trial results. *Trials* 2015; 16: 495.
- 4) Murad MH, Asi N, Alsawas M, Alahdab F. New evidence pyramid. *Evid Based Med* 2016; 21: 125-7.
- 5) Wang SV, Schneeweiss S, RCT-DUPLICATE Initiative, Franklin JM, Desai RJ, Feldman W et al. Emulation of randomized clinical trials with nonrandomized database analyses: Results of 32 clinical trials. *JAMA* 2023; 329: 1376.
- 6) Dagan N, Barda N, Kepten E, Miron O, Perchik S, Katz MA et al. BNT162b2 mRNA Covid-19 vaccine in a nationwide mass vaccination setting. *N Engl J Med* 2021; 384: 1412-23.
- 7) Polack FP, Thomas SJ, Kitchin N, Absalon J, Gurtman A, Lockhart S et al. Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine. *N Engl J Med* 2020; 383(27): 2603-15.
- 8) Gunasekara FI, Richardson K, Carter K, Blakely T. Fixed effects analysis of repeated measures data. *Int J Epidemiol* 2014; 43: 264-9.
- 9) Kiuchi S, Kusama T, Sugiyama K, Yamamoto T, Cooray U, Yamamoto T et al. Longitudinal association between oral status and cognitive decline using fixed-effects analysis. *J Epidemiol* 2022; 5: 32: 330-6.
- 10) Meyerhoefer CD, Pepper JV, Manski RJ, Moeller JF. Dental care use, edentulism, and systemic health among older adults. *J Dent Res* 2021; 100: 1468-74.
- 11) Matsuyama Y. Complete loss of natural teeth and loneliness: a fixed-effect analysis. *J Dent Res* 2024; 103: 973-9.
- 12) Thomson WM, Barak Y. Tooth loss and dementia: A critical examination. *J Dent Res* 2021; 100: 226-31.
- 13) Greenland S. An introduction to instrumental variables for epidemiologists. *Int J Epidemiol* 2000; 29: 722-9.
- 14) Matsuyama Y, Jürges H, Listl S. Causal effect of tooth loss on cardiovascular diseases. *J Dent Res* 2023; 102: 37-44.
- 15) Matsuyama Y, Jürges H, Dewey M, Listl S. Causal effect of tooth loss on depression: evidence from a population-wide natural experiment in the United States. *Epidemiology and Psychiatric Sciences* 2021; 5(30): e38.
- 16) Matsuyama Y, Listl S, Jürges H, Watt RG, Aida J, Tsakos G. Causal effect of tooth loss on functional capacity among older adults in England. *J Am Geriatr Soc* 2021; 69: 1319-27.
- 17) VanderWeele TJ, Ding P. Sensitivity analysis in observational research: Introducing the E-Value. *Ann Intern Med* 2017; 167: 268-74.
- 18) Kino S, Tamada Y, Takeuchi K, Nakagomi A, Shiba K, Kusama T et al. Exploring the relationship between oral health and multiple health conditions: An outcome-wide approach. *J Prosthodont Res* 2023; 68: 415-24.
- 19) Nakazawa N, Kusama T, Cooray U, Yamamoto T, Kiuchi S, Abbas H et al. Large contribution of oral status for death among modifiable risk factors in older adults: the JAGES prospective cohort study. *The Journals of Gerontology: Series A* 2023; 78: 167-73.
- 20) Hernán M, Robins J. *What If*. Boca Raton: Chapman & Hall/CRC; 2020.
- 21) Matsuyama Y. Time-varying exposure analysis of the relationship between sustained natural dentition and cognitive decline. *J Clin Periodontol* 2023; 50: 727-35.
- 22) Kusama T, Takeuchi K, Kiuchi S, Aida J, Osaka K. Poor oral health and dementia risk under time-varying confounding: A cohort study based on marginal structural models. *J Am Geriatr Soc* 2024; 72: 729-41.
- 23) Matsuyama Y, Aida J. Dental prosthesis use and mortality: A time-varying exposure analysis with machine learning. *J Prosthet Dent* 2025; 134: 730-8.
- 24) Athey S, Tibshirani J, Wager S. Generalized random forests. *Ann Statist* 2019; 47: 1148-78.
- 25) Shiba K, Inoue K. Harnessing causal forests for epidemiologic research: key considerations. *Am J Epidemiol* 2024; 193: 813-8.
- 26) Matsuyama Y, Aida J, Kondo K, Shiba K. Heterogeneous Association of Tooth Loss with Functional Limitations. *J Dent Res* 2024; 103: 369-77.

著者連絡先：松山 祐輔

〒113-8549 東京都文京区湯島 1-5-45

Tel: 03-5803-5476

Fax: 03-5803-0194

E-mail: matsuyama.y.db99@m.isct.ac.jp