

令和 2 年 6 月 8 日現在

機関番号：82606

研究種目：基盤研究(A) (一般)

研究期間：2017～2019

課題番号：17H01557

研究課題名(和文)大規模コホートの疾病横断的ゲノム解析に基づく個別化予防に資するエビデンスの構築

研究課題名(英文)Development of evidence for personalized prevention based on genetic analyses of various diseases in a large-scale cohort study.

研究代表者

津金 昌一郎 (Tsugane, Shoichiro)

国立研究開発法人国立がん研究センター・社会と健康研究センター・社会と健康研究センター長

研究者番号：40179982

交付決定額(研究期間全体)：(直接経費) 32,500,000円

研究成果の概要(和文)：多目的コホート研究で既に収集されている試料・情報を活用し、遺伝および環境情報を有する、がんと循環器疾患のケースコホートを構築した。構築したケースコホートを活用し、循環器疾患のゲノム網羅的関連解析、遺伝環境交互作用解析、メンデルランダム化解析を行うことにより、我が国の地域住民を対象に、遺伝要因と環境要因を組み合わせた疾患発症との関連がみられるかどうかを検証した。

研究成果の学術的意義や社会的意義

本成果は、欧米に比べてきわめて少ない、我が国での疾患発症・罹患に関する遺伝要因のエビデンスを補強するものである。

本研究において、構築された疾患横断的な罹患情報、遺伝および環境情報を有するケースコホートを、今後も活用して、個別化予防に資する更なるエビデンスの創出、検証を重ねることで、我が国の一般住民を対象とした疾患発症予測法の確立に繋がることが期待される。

研究成果の概要(英文)：This study established case cohorts of cancer and cardiovascular disease, having genome and environment information from the Japan Public Health Center-based Prospective Study (JPHC study). In these cohorts, genome-wide association analysis for cardiovascular disease, gene-environmental interaction analysis, and Mendelian randomization analysis investigated whether the combined impact of genomic and environmental factors was associated with these diseases.

研究分野：疫学

キーワード：コホート研究 循環器疾患 ゲノム解析 交互作用 発症予測法

1. 研究開始当初の背景

がん、循環器疾患などの多くの国民が罹患する疾患の発生には、喫煙や食事をはじめとする環境要因が重要な役割を果たしている。我々は、「多目的コホート研究 (JPHC study)」の調査情報に基づき、生活習慣などの環境要因とがんや循環器疾患等の発症・罹患との関連を検討し、日本人の生活習慣病予防のためのエビデンス構築に貢献してきた。このような既知のリスク要因を用いたリスク予測モデルから推計された個人の発症予測確率に基づき、最適な予防法の選択を可能とする個別化予防の開発が求められている。その取り組みの一環として、我々は、がんや循環器疾患について、生活習慣情報や健診データを用いたリスク予測モデルを構築し、web 上で各自の生活習慣等の情報を入力することにより予測確率が表示されるツールを開発し、公開している (<http://epi.ncc.go.jp/riskcheck/index.html>)。

がん、循環器疾患などの多因子疾患は環境要因の関与が大きいと言われているが、近年、主にケース・コントロール研究デザインによるゲノム網羅的関連解析 (Genome-wide-association studies: GWAS) により多数の疾患感受性遺伝子多型 (single nucleotide polymorphism: SNP) が同定され、遺伝要因の解明が顕著に進展した。しかしながら、これらの疾患感受性遺伝子多型について、わが国の前向きコホートでの検証はまだ行われていない。また、遺伝要因の違いにより、同じ環境要因に対する感受性が人によって異なるという、いわゆる遺伝環境交互作用はほとんど解明されていない。国際的には GWAS の知見を踏まえて、遺伝子多型を操作変数としたメンデルランダム化 (Mendelian Randomization: MR) 法による解析が実施されるようになってきたが、そのエビデンスはまだ僅かである。さらに、個別化予防法の実現においては、従来のリスク予測モデルにゲノム情報を加えることで予測精度が向上するか否かの検討が求められているものの、特に日本人のエビデンスはほとんどないのが現状である。

このように現時点ではゲノム網羅的 SNP 情報の利用は疾患感受性領域の同定を目的とした研究にとどまるものが多く、特に我が国においては、前向きコホート研究の利点を最大限に生かした疫学・予防研究への応用事例はほとんどない。そして、このような取り組みは、病因論に基づき疾患ごとに検討すべき課題であるが一方で、公衆衛生的な観点からは、将来どのような疾患に罹るかわからない健康な国民に対する予防対策においては、疾患横断的な解析により個々の疾患の発症確率の違いを提示することが重要である。遺伝要因および環境要因の、牽いては、疾患横断的な、前向きコホートでの疫学・予防研究を実施するため、遺伝および環境情報を有する研究基盤を構築し、個人にとって優先すべき予防対策を明確にするようなエビデンスが求められている。

2. 研究の目的

大規模前向きコホート研究において既に収集されている試料・情報を用いて、がん、循環器疾患などの多くの国民が罹患する疾患に関して、既知の疾患感受性遺伝子多型の検証、遺伝環境交互作用の同定、メンデルランダム化法を用いた解析による潜在的なリスク要因の因果関係の評価などを通して、遺伝要因、環境要因を合わせた疾患発症との関連を解明し、疾患発症予測法の確立を目的とした。

3. 研究の方法

疾患横断的なゲノム解析に基づく個別化予防に資するエビデンスを確立する研究を推進するための遺伝および環境情報を有する研究基盤となる(1)がんおよび循環器疾患を解析するためのケースコホートの構築を行い、続いて、構築したコホートを活用し、以下の(2)個別課題に対する研究を実施した。

既知の疾患感受性遺伝子多型の検証

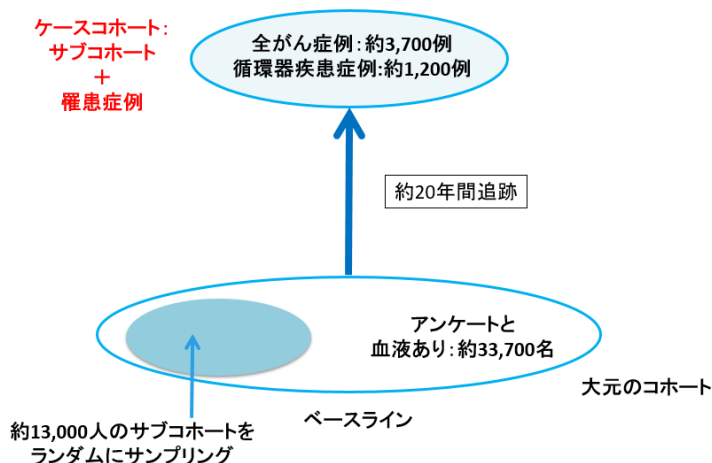
遺伝要因・環境要因の交互作用の解明

メンデルランダム化法を用いた潜在的なリスク要因の評価

(1)がんおよび循環器疾患を解析するためのケースコホートの構築

本研究は、全国 11 か所の保健所管内に在住する 40 歳から 69 歳の地域住民約 14 万人を対象とした「多目的コホート研究 (JPHC study)」(研究代表者 津金昌一郎)において収集・保管された以下の試料・情報を用いる。既存の分析値を最大限に活用しながら、サンプルサイズを極力大きくするという方針のもと対象者を定義した。ベースライン調査時点でアンケート調査に

図 1. 研究デザイン



回答し、血液を提供した約 3.4 万人から、性別、年齢 5 歳階級、保健所の条件により層別ランダムサンプリングした約 1.3 万人をサブコホート、追跡期間中(1990 年から 2009 年末までの約 20 年間に約 3.4 万人から発生したがん罹患約 3700 例および循環器疾患発症約 1200 例をケースとして、ケースコホート研究をデザインした(図 1 参照)。このケースコホートの全対象者のジェノタイピングおよびゲノムインピュテーションを行い、本ケースコホート

で、遺伝要因と環境要因を解析できる研究基盤を構築した。最終的に、ジェノタイピングの質、がんと循環器疾患既往者および近親者などの除外を行い、約 3700 人のがん罹患患者を含む約 1.2 万人のがんのケースコホートと、約 1200 人の循環器罹患患者を含む、約 1.1 万人の循環器疾患のケースコホートをそれぞれ構築した。

(2)個別課題

既知の疾患感受性遺伝子多型の検証

-心疾患、突然死、脳卒中ゲノム網羅的解析(GWAS)

構築した循環器疾患のケースコホートを用いて、心疾患、突然死、脳卒中をアウトカムにしたゲノム網羅的解析(GWAS)を行い、我々のコホートにおいても、既知の疾患感受性遺伝子多型との関連が認められるか検証を実施した。

循環器疾患のケースコホート約 1.1 万人において、ジェノタイピングデータを用いた主成分分析を行い、集団の特徴を示す 10 個の主成分を計算した。続いて、本集団において、マイナーアレル頻度が 1%以上であった約 50 万の遺伝子多型と心疾患、突然死、脳卒中の関連を、計算した 10 個の主成分、年齢、性別を共変量に加え、ロジスティック回帰モデルにより解析を行った。本課題では、インピュテーションされた遺伝子多型情報は用いず、ジェノタイピングされた遺伝子多型情報のみを利用し、遺伝子多型の効果は、アレルの数に比例するアディティブモデルを仮定した。また各 3 つの疾患の GWAS を行う際は、サブコホートと当該疾患罹患患者のみを対象として、以下のコントロール、ケースにおいて、解析を実施した。

表 1. GWAS 対象者数

アウトカム	コントロール数	ケース数
心疾患	10,772	257
突然死	10,813	119
脳卒中	10,370	1,237

既知の疾患感受性遺伝子多型の再現性を確認する基準は、通常の GWAS 基準に倣い、当該遺伝子多型の P 値が 5×10^{-8} よりも低い場合とする。また本課題では、ジェノタイピングされた遺伝子多型のみで解析を実施しているため、P 値が GWAS 基準に達した遺伝子多型の近傍の遺伝子多型についても考慮するために、PhenoScanner や LDtrait などの近傍の遺伝子多型の形質まで調べることのできるアノテーションツールを用いて、検証を行った。

Phenoscaner (<http://www.phenoscaner.medschl.cam.ac.uk/tools/>)

LDtrait(<https://ldlink.nci.nih.gov/?tab=home>)

遺伝要因・環境要因の交互作用の解明

-食塩感受性遺伝子などの候補遺伝子多型アプローチ解析

先行研究のレビューを行い、食塩感受性遺伝子など、高血圧との関連が報告されている既知の

遺伝子多型を選定し、高血圧との関連および遺伝子多型と環境要因との交互作用について分析した。

分担研究者の小久保らを中心にまとめた高血圧に関連する遺伝子多型についての国内外の先行研究のレビュー (Kokubo Y et al. Environ Health Prev Med. 2019;24(1):19.) に加えて、日本人を対象として実施された先行研究についての検索を行った。これらの情報をもとに、食塩感受性遺伝子や飲酒を介して高血圧と関連する可能性が考えられる下表の 21 の候補遺伝子多型を選択した。

本課題では、一般集団における解析が望まれるため、ケースコホートではなく、疾患症例を除いたランダムサンプリングしたサブコホートにおいて、解析を実施した。選出した候補遺伝子多型のうち、一部はインピュテーションによってアレルの確率を推定した。最終的に、近親者などの除外を行い、血圧の測定値を有する約 1 万人を対象とした。高血圧の定義は、ベースラインでの収縮期血圧 140mmHg 以上、拡張期血圧 90mmHg 以上、または降圧剤内服とした。Additive model を仮定して、高血圧と候補遺伝子多型との関連を分析した。さらに、食塩摂取、飲酒状況などの生活習慣による層別を行い、遺伝子多型と環境要因との交互作用の有無について検討を行った。

表 2. 候補遺伝子多型一覧

chr	position	rsID	候補遺伝子	chr	position	rsID	候補遺伝子
1	11907648	rs5063	NPPA	7	150696111	rs1799983	NOS3
1	221273158	rs7554672	LOC101929750	8	143999600	rs1799998	CYP11B2
1	230845794	rs699	AGT	11	78612721	rs10466739	TENM4
1	230845977	rs4762	AGT	12	6954875	rs5443	GNB3
4	2906707	rs4961	ADD1	12	88223685	rs11104632	LOC105
4	141323162	rs2567241	CLGN	12	112214459	rs2238152	ALDH2
5	148206440	rs1042713	ADRB2	12	112241766	rs671	ALDH2
5	148206473	rs1042714	ADRB2	15	75077367	rs1378942	CSK
6	149112190	rs13211840	UST	15	75082552	rs3784789	CSK-MIR4513
7	99270539	rs776746	CYP3A5	17	61565990	rs4341	ACE
7	131175494	rs1643270	MKLN1				

chr=chromosome number(染色体番号); position(染色体位置); rsID(遺伝子多型 ID).

メンデルアン・ランダムイゼーション (MR) 法を用いた潜在的なリスク要因の評価

-糖尿病とがん罹患リスクについての解析及び MR 法の方法論についての検討

構築したがんのケースコホートを用いて、糖尿病とがん罹患リスクについて、MR 法を適用して、がんにおける糖尿病の因果関係を検証した。またこの他に、さまざまな形質と疾患等の発症・罹患との関連と、幅広いテーマに応用できるように MR 法の方法論を検証した。

MR 法は、左下図のように、観察研究で問題となる未知・未観察の交絡要因の影響を回避するために、直接、糖尿病とがんの関係を解析するのではなく、糖尿病感受性遺伝子多型 (SNPs) を操作変数として、遺伝子多型と糖尿病の関連 (1)、遺伝子多型とがんの関連 (2) を推定し、この推定値を統合することで、糖尿病とがんの関連 (3 = 2 / 1) を求める手法である。糖尿病とがん罹患リスクについては、1.2 万人からなるがんのケースコホートにて、遺伝子多型と糖尿病の関連 (1)、遺伝子多型とがんの関連 (2) を、同じ 1 つの集団で推定する MR 法の中で one-sample 法と呼ばれる手法を用いて、解析を実施した。この操作変数となる遺伝子多型の選定方法は重要であるが、今回は、アジア人において、複数回報告されている 29 の糖尿病感受性遺伝子多型を操作変数に用いた。

また、MR 法の方法論として、1 と 2 を異なる 2 つの集団で推定する two-sample 法と操作変数となる遺伝子多型の選定方法について、検証を行った。

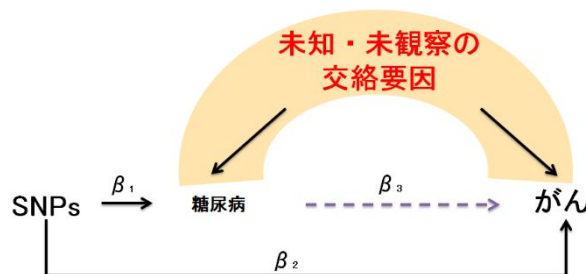


図 2. MR 法概念図

4. 研究成果

3. 研究方法において、掲げた(2)個別課題については、それぞれ以下の結果が得られた。

既知の疾患感受性遺伝子多型の検証

-心疾患、突然死、脳卒中ゲノム網羅的解析(GWAS)

循環器疾患の中で、心疾患、突然死、脳卒中について、GWAS 解析を行ったところ、GWAS 基準 (P 値 $< 5.0 \times 10^{-8}$) および準 GWAS 基準 (P 値 $< 1.0 \times 10^{-6}$) に達する遺伝子多型は、10 個にも満たず、ほぼ認められなかった。一般的な解析で使用される P 値 < 0.05 に達する遺伝子多型は、どの疾患においても大きくは変わらず、約 2.2 万個程度であった。

脳卒中においては、2 つの遺伝子多型が、GWAS 基準に達していたが、準 GWAS 基準に達した遺伝子多型と合わせて 8 つの遺伝子多型において、PhenoScanner と LDtrait で、近傍の遺伝子多型を含めて形質の確認を行ったところ、脳卒中に関連する遺伝子多型は認められなかった。心疾患についても同様の結果であった。我々のコホートでは、既存のこれらの疾患の感受性遺伝子多型の再現性は得られなかった。本研究で再現性が得られなかった 1 つの原因として、サンプルサイズの大きさが考えられる。遺伝子多型の効果は、あまり大きくないため、本サンプルサイズでは、偽陰性となっている可能性がある。今後、サンプルサイズの拡大を検討する。

一方で GWAS 基準に達した、脳卒中との関連が示唆された 2 つの遺伝子多型は、これまでに身長、白血球数との関連がそれぞれ報告されていた。GWAS は膨大な遺伝子多型を多重解析するため、勝者の呪いと呼ばれる、偽陽性の遺伝子多型を検出する可能性がある。脳卒中に関連した既報がないことから当該遺伝子多型も、その可能性があるため、新規の脳卒中感受性遺伝子であるか、今後慎重に検証を行う必要がある。

遺伝要因・環境要因の交互作用の解明

-食塩感受性遺伝子などの候補遺伝子多型アプローチ解析

3. 研究の方法に、掲げた 21 候補遺伝子多型について、解析を行ったところ、一部の遺伝子多型で高血圧との関連が認められた。さらに一部は飲酒によりその関連が増強される可能性が示された。今後、遺伝子多型と環境要因との交互作用について、さらに詳細な解析を行う。また、心血管疾患発症者を含めたケースコホートを用いて、心血管疾患の発症をアウトカムとした解析を行う。

メンデルアン・ランダムイゼーション (MR) 法を用いた潜在的なリスク要因の評価

-糖尿病とがん罹患リスクについての解析及び MR 法の方法論についての検討

がんのケースコホートにおいて、MR 法を適用したところ、2 型糖尿病が、がんに関連するという、強い遺伝的なエビデンスは得られなかった (Goto A et al. Int J Cancer. 2020;146(3))。一部の観察研究では、がんと糖尿病の関連が示唆されている。観察研究の結果は、糖尿病患者が、高インスリン血症、高血糖、慢性炎症などを有している交絡による、見かけ上の関連の可能性がある。また、MR 法の操作変数の遺伝子多型の選定が、不十分であった可能性もある。しかしながら、このように、異なる方法で、要因と疾患発症・罹患との関係を、検討することで、真のエビデンスに近づくことができる。

続いて MR 法の方法論について検証したところ、十分なサンプルサイズが得られる場合、two-sample 法が有用であること、また操作変数の選定する際は、メタアナリシスの文献を収集するように、小さい効果の遺伝子多型なども含め体系的に選定することで、安定した推計値が得られやすいことなどが、示唆された。得られた知見を活かし、順次、他の要因と疾患発症・罹患の関連の解析を進める。

総括

これらの個別課題の結果は、ネガティブな結果も含むが、精度の高い疾患発症予測法の確立するために、必要不可欠な情報である。また、薬剤の副作用反応で知られるように、遺伝子多型の分布は、人種間差が非常に大きいため、日本人に根差したエビデンスを確立していくことは重要であり、本成果は、欧米に比べて不足している遺伝要因を含めたエビデンスを補強するものである。本研究において、構築された、疾患横断的な罹患情報、遺伝および環境情報を有する研究基盤であるケースコホートを、今後も活用していき、個別化予防に資する更なるエビデンスを創出することを目指す。引き続き、これらの検証を重ねることで、日本人集団における疾患発症予測法の確立に繋がることが期待される。

5. 主な発表論文等

〔雑誌論文〕 計6件（うち査読付論文 6件/うち国際共著 1件/うちオープンアクセス 2件）

1. 著者名 Kokubo Y, Padmanabhan S, Iwashima Y, Yamagishi K, Goto A.	4. 巻 24
2. 論文標題 Gene and environmental interactions according to the components of lifestyle modifications in hypertension guidelines.	5. 発行年 2019年
3. 雑誌名 Environ Health Prev Med.	6. 最初と最後の頁 19
掲載論文のDOI（デジタルオブジェクト識別子） doi:10.1186/s12199-019-0771-2.	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 Goto A, Yamaji T, Sawada N, Momozawa Y, Kamatani Y, Kubo M, Shimazu T, Inoue M, Noda M, Tsugane S, Iwasaki M.	4. 巻 146
2. 論文標題 Diabetes and cancer risk: A Mendelian randomization study.	5. 発行年 2020年
3. 雑誌名 Int J Cancer.	6. 最初と最後の頁 712-719
掲載論文のDOI（デジタルオブジェクト識別子） doi: 10.1002/ijc.32310.	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 Iwasaki Motoki, Tanaka-Mizuno Sachiko, Kuchiba Aya, Yamaji Taiki, Sawada Norie, Goto Atsushi, Shimazu Taichi, Sasazuki Shizuka, Wang Hansong, Loic Le Marchand, Tsugane Shoichiro	4. 巻 10
2. 論文標題 Inclusion of a Genetic Risk Score into a Validated Risk Prediction Model for Colorectal Cancer in Japanese Men Improves Performance	5. 発行年 2017年
3. 雑誌名 Cancer Prevention Research	6. 最初と最後の頁 535-541
掲載論文のDOI（デジタルオブジェクト識別子） 10.1158/1940-6207.CAPR-17-0141	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 Akiyama M, Okada Y, Kanai M, Takahashi A, Momozawa Y, Ikeda M, Iwata N, Ikegawa S, Hirata M, Matsuda K, Iwasaki M, Yamaji T, Sawada N, Hachiya T, Tanno K, Shimizu A, Hozawa A, Minegishi N, Tsugane S, Yamamoto M, Kubo M, Kamatani Y.	4. 巻 49
2. 論文標題 Genome-wide association study identifies 112 new loci for body mass index in the Japanese population	5. 発行年 2017年
3. 雑誌名 Nature Genetics	6. 最初と最後の頁 1458-1467
掲載論文のDOI（デジタルオブジェクト識別子） 10.1038/ng.3951	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 -

1. 著者名 Sugasawa S, Noma H.	4. 巻 38
2. 論文標題 Estimating individual treatment effects by gradient boosting trees	5. 発行年 2019年
3. 雑誌名 Stat Med.	6. 最初と最後の頁 5146-5159
掲載論文のDOI (デジタルオブジェクト識別子) DOI: 10.1002/sim.8357.	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 -

1. 著者名 Sugasawa S, Noma H.	4. 巻 印刷中
2. 論文標題 Efficient screening of predictive biomarkers for individual treatment selection	5. 発行年 2020年
3. 雑誌名 Biometrics.	6. 最初と最後の頁 印刷中
掲載論文のDOI (デジタルオブジェクト識別子) DOI: 10.1111/biom.13279	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 -

〔学会発表〕 計3件 (うち招待講演 0件 / うち国際学会 0件)

1. 発表者名 後藤 温
2. 発表標題 メンデルアン・ランダムマイゼーションによる因果推論.
3. 学会等名 第29回日本疫学会学術総会
4. 発表年 2019年

1. 発表者名 野間久史
2. 発表標題 The most powerful overall test for GWAS to detect molecular predictive markers with heterogeneous subgroup effects.
3. 学会等名 医学統計研究会鹿児島秋季セミナー2019
4. 発表年 2019年

1. 発表者名 野間久史
2. 発表標題 疾患の発症に関わる稀な遺伝的変異の検出のための最強力検定.
3. 学会等名 大分統計談話会第61回大会
4. 発表年 2020年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

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	山地 太樹 (Yamaji Taiki) (10466203)	国立研究開発法人国立がん研究センター・社会と健康研究センター・室長 (82606)	
研究分担者	山岸 良匡 (Yamagishi Kazumasa) (20375504)	筑波大学・医学医療系・准教授 (12102)	
研究分担者	小久保 喜弘 (Kokubo Yoshihiro) (20393217)	国立研究開発法人国立循環器病研究センター・病院・医長 (84404)	
研究分担者	磯 博康 (Iso Hiroyasu) (50223053)	大阪大学・医学系研究科・教授 (14401)	
研究分担者	岩崎 基 (Iwasaki Motoki) (60392338)	国立研究開発法人国立がん研究センター・社会と健康研究センター・部長 (82606)	

6. 研究組織（つづき）

	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	野間 久史 (Noma Hisashi) (70633486)	統計数理研究所・データ科学研究系・准教授 (62603)	
研究分担者	後藤 温 (Goto Atsushi) (80644822)	国立研究開発法人国立がん研究センター・社会と健康研究センター・室長 (82606)	