

令和 4 年 6 月 30 日現在

機関番号：23803

研究種目：基盤研究(C) (一般)

研究期間：2019～2021

課題番号：19K11746

研究課題名(和文) 長期健診受診者の欠損値の多重代入データに基づく疾病罹患リスク予測と予防方法の確立

研究課題名(英文) Establishment of a method for predicting and preventing the risk of disease occurrence based on multiple imputation data of missing values of long-term health checkup examinees.

研究代表者

栗木 清典(Kuriki, Kiyonori)

静岡県立大学・食品栄養科学部・教授

研究者番号：20543705

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)：大規模コホート(J-MICC)の静岡・桜ヶ丘地区における約6,400人の10年間の健診データより、非継続受診者の欠損値をmultiple imputation by Markov Chain Monte Carlo methods (MCMC法)で補完して解析し、肥満や糖尿病などの罹患リスクの予測モデルを確立することを目指した。体重や血糖値の維持管理は、腸内細菌叢を解析して検討した。適正血圧値の維持管理では、尿試料中電解質を測定した。MCMC法による本研究の解析モデルとして、J-MICC Sakura Diet Studyにおいて、腸内細菌叢、季節が血糖値に影響を及ぼすことを明らかにした。

研究成果の学術的意義や社会的意義

本研究は、『継続受診者』と『非継続受診者』の人口学的特性、各健診項目、生活習慣要因、精神ストレスなどを比較することから、継続受診の重要性を示すだけでなく、メタボリックシンドロームや高血圧などの一次予防や血圧維持のための食事・保健指導の方法を確立するうえで重要な基礎資料となる。大規模な長期の時系列データで疾患の罹患を高精度に予測するモデルを確立するにあたり、従来では、(尿試料中電解質や)便試料の腸内細菌叢は検討されてこなかった。高精度の予測モデルに基づき、これらのバイオマーカーを用いて、食事・保健指導で生活習慣病を一次予防することの重要性を示すことに寄与した。

研究成果の概要(英文)：As part of the Japan Multicenter Cohort (J-MICC), data were collected on approximately 6,400 individuals in the Sakuragaoka area, Shizuoka, Japan, over a period of about 10 years. The Multiple Imputation by Markov Chain Monte Carlo (MCMC) method was used to complement missing values and to examine methods to predict and prevent the risk of developing obesity, hypertension, dyslipidemia, diabetes, and metabolic syndrome diseases. Electrolytes in urine samples were measured to maintain and control adequate blood pressure levels. As a model for analysis using the MCMC method in this study, we found that green tea, human gut microbiota, and season had an effect on blood glucose levels in the data analysis of the J-MICC Sakura Diet Survey (Four Seasons Diet Survey).

研究分野：栄養疫学

キーワード：生活習慣病 罹患リスク 予測モデル 時系列データ 欠損値 多重代入法

## 1. 研究開始当初の背景

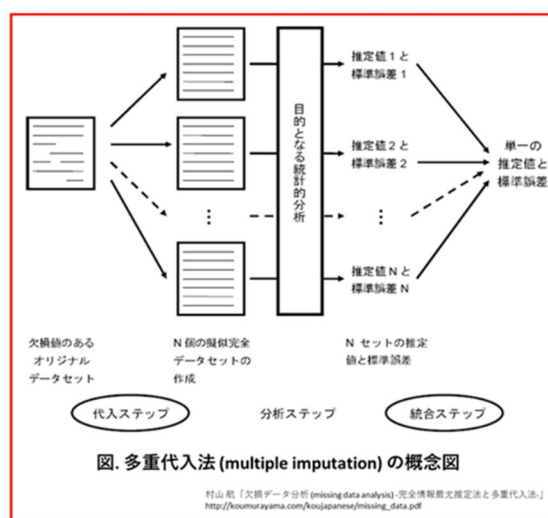
肥満、高血圧、脂質異常症、糖尿病、および、メタボリックシンドロームの罹患（発症）までの時間と生活習慣要因との関連を検討する時系列データの解析では、健診受診者を対象としたコーホート研究は良い研究デザインである。本申請者は、J-MICC 研究の静岡・桜ヶ丘地区の独自研究（研究代表者）として、5つの健診機関より約 6,400 人の研究参加の同意を得ている。既に、ベースライン時から過去 5 年間、および、現在までの 5 年間の計 10 年間の健診データを入力し、個人レベルでデータを突合している。

しかし、このような長期の時系列データの解析では欠損値が多いという問題点がある。従来の統計ソフトウェアは、リストワイズ法やペアワイズ法で欠損値を削除し、毎年のデータのある『継続受診者』に限定して解析している。平均値や代入法もあるが、『非継続受診者』の欠損値が完全にランダムでないと推定値にバイアスが生じるため、Healthy worker 's effects などにより、各種の疾患予測モデルは過小評価されるという大きな問題がある。

最近、SAS や R などの統計ソフトウェアは、欠損値の多重代入（multiple imputation）を推奨している。R の「欠損値代入法の比較」によると、Amelia、norm、mice などのパッケージ（計算アルゴリズム）は「欠損値なしの完全データ」と同等程度に統計学的に正しい結果を導き出すことができるとある。

多重代入では、マルコフ連鎖（1 個前の状態によって次の状態が決まる連鎖）とモンテカルロ法（乱数を発生させること）を組み合わせた“multiple imputation by Markov Chain Monte Carlo methods (MCMC 法)”により、確率分布からサンプリングした N 個の疑似完全データセットを作成し、目的とする統計解析を行った後、各々の推定値と標準誤差を統合することで「単一の推定値と標準誤差」を求める（図）。Multiple imputation における MCMC 法はベイズ統計学を中心に幅広い分野で利用され始めているが、実用例は少ない。

したがって、高精度の疾患予測モデルを確立するには、1) 『非継続受診者』の欠損値を多重代入（MCMC 法）して、『継続受診者』のデータと合わせた『疑似完全データ』を構築する必要がある。2) 横断調査ではなく、長期追跡調査による個々人の健診データが必要である。そして、生活習慣要因などの改善による疾患別の一次予防法を提案するには、3) 詳細な生活習慣要因に関するデータが必要である。さらに、詳細の大規模 GWAS 解析に向けて、4) ゲノム情報を考慮した検討も必要である。



## 2. 研究の目的

本申請の研究目的は、

- 1) 約 6,400 人の 10 年間の健診データより、『非継続受診者』の欠損値を MCMC 法による multiple imputation により『継続受診者と非継続受診者で構成される疑似完全データ』を構築して「単一の推定値と標準誤差」を求めることで、肥満、高血圧、脂質異常症、糖尿病、メタボリックシンドロームの罹患リスク予測モデルを高精度に確立する。
- 2) 1) では、各疾患の罹患リスクに対し、食事（食品・栄養素摂取量）、飲酒、喫煙、運動・身体活動などの生活習慣要因や精神ストレス等の関連を明らかにして、生活習慣要因などの改善による疾患別の一次予防法を提案する。
- 3) 2) より、『継続受診者』と『非継続受診者』の人口学的特性、各健診項目のデータを比較して、『非継続受診者』の特徴を明らかにし、継続受診の勧めと、食事・保健指導のアドバイス方法を確立する。

## 3. 研究の方法

J-MICC 研究では、全国研究の目的とともに、各研究機関の独自研究の目的を遂行している。当地区では、ベースライン時より過去 5 年分と毎年の健診データを収集し、長期健診受診者の追跡調査による各種の疾患罹患と生活習慣要因との関連を検討している。ベースライン時に収

集した尿試料中電解質より、食塩摂取量と血圧値との関連を検討する。

#### 4 . 研究成果

日本多施設共同コーホート (J-MICC) の静岡・桜ヶ丘地区における約 6,400 人の計 10 年間の健診データより、非継続受診者の欠損値を“multiple imputation by Markov Chain Monte Carlo methods (MCMC 法)”で補完して解析し、肥満、高血圧、脂質異常症、糖尿病、メタボリックシンドロームの罹患リスクの予測モデルを高精度に確立することを目指した。肥満予防、適正体重の維持管理については、腸内細菌叢を解析した高精度モデルの確立も検討した。適正血圧値の維持管理については、尿試料中電解質を測定した。

なお、MCMC 法による本研究の解析モデルとして、J-MICC Sakura Diet Study (四季食事調査) のデータ解析において、腸内細菌叢、季節が血糖値に影響を及ぼすことを明らかにした。また、季節の影響を考慮して、肥満 (体格指数 : body mass index) に対するアディポネクチンに影響を及ぼす遺伝子多型 (SNPs) と食事要因との関連について検討した。その結果、血中アディポネクチン濃度は、食事要因に関連なく、アディポネクチンに関連する遺伝子多型と季節による影響を受けることを明らかにした。

5. 主な発表論文等

〔雑誌論文〕 計2件（うち査読付論文 1件/うち国際共著 0件/うちオープンアクセス 1件）

1. 著者名 Akimoto S, Goto C, Kuriki K.	4. 巻 5(61)
2. 論文標題 Relationship between ethanol consumption and TBL2 rs17145738 on LDL-C concentration in Japanese adults: a four season 3-day weighed diet record study.	5. 発行年 2019年
3. 雑誌名 BMC Nutr.	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） 10.1186/s40795-019-0315-6	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 栗木清典	4. 巻 53(8)
2. 論文標題 疫学研究における遺伝的要因と食事要因	5. 発行年 2021年
3. 雑誌名 細胞	6. 最初と最後の頁 36-39
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 無
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 可児美月, 後藤千穂, 栗木清典
2. 発表標題 多重代入法を用いた血清アディポネクチン濃度と遺伝的要因、食事要因、季節の関連
3. 学会等名 第79回日本公衆衛生学会総会
4. 発表年 2020年

1. 発表者名 中野夏菜, 川田久美子, 可児美月, 秋元 柊, 後藤千穂, 栗木清典.
2. 発表標題 慢性腎臓病に対する推定糸球体濾過量と食事要因、遺伝的要因、季節の関連: the J-MICC Sakura Diet Study
3. 学会等名 第30回 日本疫学会学術総会
4. 発表年 2020年

1. 発表者名 古郡かすみ, 可児美月, 五十嵐麻希, 後藤千穂, 栗木清典.
2. 発表標題 肥満に対するアルコールの摂取量と代謝酵素遺伝子多型の関連, 季節の関連
3. 学会等名 第80回日本公衆衛生学会総会
4. 発表年 2021年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

静岡県立大学 食品栄養科学部 公衆衛生学研究室 <a href="https://dfns.u-shizuoka-ken.ac.jp/labs/phealth/">https://dfns.u-shizuoka-ken.ac.jp/labs/phealth/</a> 日本多施設共同コホート研究 (J-MICC Study静岡・桜ヶ丘地区) <a href="http://www.jmicc.com/feature/feature01-12/">http://www.jmicc.com/feature/feature01-12/</a> みらいぶっく：栄養素や緑茶が腸内細菌の働きで肥満予防に役立つことを明らかに～食と健康の大規模調査～ <a href="https://miraibook.jp/researcher/330">https://miraibook.jp/researcher/330</a>
---

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

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
---------	---------