

博士論文

日本人の体質を判定する簡便な体質質問票の作成および
日本語版中医体質調査票(CCMQ-J)における体質との関連の研究

佐藤 真梨子

2021年 9月 6日

奈良先端科学技術大学院大学
先端科学技術研究科 情報科学領域

本論文は奈良先端科学技術大学院大学先端科学技術研究科に
博士(工学)授与の要件として提出した博士論文である.

佐藤 真梨子

審査委員:

金谷 重彦 教授	(主指導教官)
松本 健一 教授	(副指導教官)
小野 直亮 准教授	(副指導教官)
MD. ALTAF-UL-AMIN 准教授	(副指導教官)
黄 銘 助教	(副指導教官)

日本人の体質を判定する簡便な体質質問票の作成および日本語版中医 体質調査票 (CCMQ-J) における体質との関連の研究

佐藤 真梨子

内容概要

中国医学では個人の遺伝的要素、社会・生活環境・地理・気候的要素、ライフスタイルにより体質が特徴づけられると考えられている。この体質を評価するための質問票 Constitution in Chinese Medicine Questionnaire (CCMQ) の日本語版 (CCMQ-J) が開発された。CCMQ-J は 60 の質問項目から構成され、個人の体質は 9 種の体質に分類される。これまでに、CCMQ-J で分類される体質が気分や QOL と関連することや、質問の回答から年齢や BMI が推定できることが報告された。このことから、質問票による体質判定によってほかの健康状態の把握もできるのではないかと考えた。質問票により体質を判定することができれば、個人の体質にあった健康施策を選択することが可能になる。しかし、CCMQ-J は 60 問で構成されており、一人当たりの体質を判定するのに時間を要する。また、中国医学に基づく質問票であるため、現代の日本人特有の体質を反映しているかは定かではない。

そこで、本研究において日本人の体質を分類することができる、簡便な最適な新体質質問票を作成することとした。まず、クラウドソーシングを用いて 851 名の日本人から CCMQ-J 60 問と既存の健康に関するアンケート 272 問、合計 332 問に対する回答を取得した。851 名の CCMQ-J に対する回答をクラスタリング解析することにより、日本人の最適体質数は 5 であることが確認された。さらに、851 名の 332 問に対する回答について、体質を 5 つに分類する際に特徴量が大きい質問を、ランダムフォレストを用いて抽出し、質問票を 18 問まで圧縮することができた。また、圧縮された 18 の質問を用いて決定木を作成し、1 人当たり 4-5 問で体質を判定できる質問票を開発した。本研究の結果から日本人の体質数が 5 であることや、5 体質に分類するための質問票を作成することができた。この質問票を疫学調査などで活用していくことで、より日本人の体質の特徴を明らかにしていくことができると考えられる。

キーワード

日本語版中医体質調査票(CCMQ-J), 体質、ランダムフォレスト、クラスタリング、決定木

*奈良先端科学技術大学院大学 先端科学技術研究科 情報科学領域 博士論文, 2021年9月6日

Evaluation of Convenient Questionnaire about Japanese Composition and Relationship between the Constitution in Chinese Medicine Questionnaire (CCMQ-J) and Japanese Composition Based on Multivariate Analysis

Mariko Sato

Abstract

In Chinese medicine, it is considered that the constitution is characterized by individual genetic factors, social, living environment, geography, climatic factors, and lifestyle. Chinese Medicine Questionnaire (CCMQ-J) consists of 60 independent questions and can determine nine physical constitutions called subscales. So far, it has been reported that the constitution classified by CCMQ-J correlates with mood and QOL, and that age and BMI can be estimated from the answers to the questions. So, it might be possible to predict other health conditions by judging the constitution. However, CCMQ-J consists of 60 questions, and it takes time to judge the constitution. Also, since CCMQ-J is a questionnaire based on Chinese, it is uncertain whether it reflects the constitution of modern Japanese people. Therefore, in this study, we tried to determine optimal number of constitutions and develop simple and convenient questionnaire that can classify the constitution of Japanese people. First, we used crowdsourcing, and obtained answers to 60 CCMQ-J questions and 272 existing health questionnaires, total 332 questions, from 851 Japanese people. By using clustering analysis of the responses to the answer CCMQ-J from 851 people, it was suggested CCMQ-J can classify Japanese individuals into five different groups. Next, we extracted 18 important questions when classifying the constitution into 5 groups and verified reproducibility. Furthermore, we created the decision tree using 18 questions, and develop the questionnaire that can judge the constitution with 4-5 questions per person. By utilizing this questionnaire in epidemiological surveys, it is thought that the characteristics of the constitution of Japanese people can be clarified.

Keywords

CCMQ-J, composition, constitution, random forest, clustering, decision tree

*Doctor's Thesis, Department of Information Science, Graduate School of Information Science, Nara Institute of Science and Technology, September 6, 2021

目次	
図目次	ix
表目次	ix
第1章 はじめに	1
1.1 はじめに	1
1.2 中医体質質問票について	1
1.3 日本における体質診断について	5
1.4 本研究の目的	5
1.5 本論文の構成	6
第2章 対象者	6
2.1 クラウドソーシングについて	6
2.2 本研究の対象者	7
第3章 データサイエンスの活用	7
3.1 多変量解析の活用	7
3.2 回帰	8
3.3 分類	9
3.4 機械学習	9
3.4.1 教師あり学習	9
3.4.2 教師なし学習	10
3.4.3 強化学習	12
第4章 クラウドソーシングを活用した日本全体の CCMQ-J 問診票調査における最適体質数とその体質特性	13
4.1 目的	13
4.2 方法	13
4.2.1 非階層的クラスタリング	14
4.2.2 階層的クラスタリング	15
4.3 結果	15
4.3.1 非階層的クラスタリングの結果	15
4.3.2 階層的クラスタリングの結果	16
4.3.3 CCMQ-J の質問項目におけるスコアの k 平均法と階層的クラスタリングの共通性	17
4.3.4 CCMQ-J の体質におけるスコアの k 平均法と階層的クラスタリングの共通性	19
4.3.5 主成分分析	22
4.4 考察	24
4.4.1 日本人の最適な体質数とその妥当性	24
4.4.2 日本人の5体質の特徴	25
4.5 課題	26
第5章 日本人特異的な体質を判定する質問票の開発	27
5.1 目的	27
5.2 方法	28
5.2.1 質問票	28
5.2.2 階層的クラスタリング	42

5.2.3	ランダムフォレスト.....	42
5.2.4	質問票のバリデーション.....	43
5.3	結果:.....	43
5.3.1	クラスタリング.....	43
5.3.2	質問数の削減.....	44
5.3.3	バリデーション.....	48
5.3.4	中医体質の比較.....	48
5.3.5	質問票の作成.....	51
5.4	考察.....	53
5.4.1	ランダムフォレストを用いた質問数の削減.....	53
5.4.2	39問によるバリデーション.....	53
5.4.3	決定木による簡便な質問票の作成.....	54
5.5	課題.....	54
第6章	総括.....	56
6.1	第4章の日本人の体質分類数に関するまとめ.....	56
6.2	日本人の体質を簡便に判定する体質質問表開発に関するまとめ.....	56
6.3	今後の展望.....	56
	謝辞.....	58
	研究業績.....	59
	参考文献.....	60

図目次

図 1. 日本語版中医体質質問票 (CCMQ-J)	3
図 2. 9 種類の体質の特徴と 60 の質問における特典からの計算法	4
図 3. デンドログラムの例	11
図 4. ギャップ関数の推移	16
図 5. グループ KA-KE、HA-HE に属する被験者数の体質 9 分類の相対数 (%)	21
図 6. 60 項目の質問による主成分プロット	23
図 7. 851 名を階層的クラスタリングで 5 分類にした際のデンドログラム	44
図 8. 332 問の質問から、特徴量が低い質問から 1 問ずつ削減していったランダムフォレストの精度	45
図 9. 162 名を階層的クラスタリングで 5 分類にした際のデンドログラム	48
図 10. 5 グループにおける中医体質の相対比	49
図 11. 18 問を用いて作成した決定木	52

表目次

表 1. 多変量データ	8
表 2. データ区分	8
表 3. クラスタ間の距離	11
表 4. 5 グループにおける平均スコアについて下位 5%と上位 5%の質問項目	19
表 5. 表 4 から読み取れるグループ A~E の特徴	19
表 6. CCMQ-J の体質分類における KA-HA, KB-HB, KC-HC, KD-HD, KE-HE の 相関係数	21
表 7. 9 つの質問票に含まれる 272 問	41
表 8. 5 体質の分類に寄与が大きい上位 40 問	46
表 9. 質問票の文言変更	47
表 10. 851 名の CCMQ-J への回答と 162 名の 18 問に対する回答から得られる 中医体質の相関係数	50

第1章はじめに

1.1 はじめに

現在、日本において高齢者が総人口に占める割合は 28.7%であり、2040 年には 35.3%まで上昇すると報告されている¹。

このように、我々の寿命は延び続け、2019年の日本人の平均寿命は男性が81.41歳、女性が87.45歳と過去最高を記録している²。一方で、健康寿命、すなわち、「日常生活に制限のない期間の平均」は、男性は72.14歳、女性は74.79歳と算出されている³。平均寿命と健康寿命には10年以上の乖離があり、これはすなわち、晩年の10年に寝たきりになってしまう状態であることを示す。その場合、寿命が長くても医療費、介護費の負担が増えることに加えて、老々介護などの社会問題につながる。そのため、年を重ねても健康な状態を維持することが重要になっている。約2000年前に中国の医学書「黄帝内経」に「未病」という言葉が出てくるが、未病は健康と病気の状態のことを示す。肥満や高血圧などのメタボリックシンドロームや骨や関節、筋肉などの運動器の衰えが原因で立つことや歩くことが困難になるロコモティブシンドロームも未病状態といえる。この未病をいかに予防するかが重要になってきている。

1.2 中医体質質問票について

近年、ヒトゲノムの解読や次世代シーケンサーによる大量の遺伝情報を収集、解析できる技術が発展してきたことから個別化医療が可能になってきた。ヒトは高齢になるほど生活してきた環境や食生活が異なるため、個人差が大きくなる。そのため、超高齢化社会である日本においては、今後、個々人において効果のある対策を選択していく必要があると考える。しかしこれまでの医療の考え方は西洋医学を基盤としており、疾患を対象とし、疾患の治療法を確立することを目的としていた。その一方で、疾患の症状は人それぞれであり、同じ病気であっても症状が異なることや、治療薬の効果が異なることが知られてきた。

そこで我々は東洋医学の「体質」の考えを導入することを考えた。体質とは、遺伝的素因と環境要因の相互作用によって形成される、個々人の総合的な形質のことである。

東洋医学は、疾患ではなくヒトに対して問診、脈診などを行うことで、「証」とよばれる診断を下し、またそれに合った対応策を選択する。中国では 60 問と 5 つの選択肢（1. まったくない; 2. まれにある; 3. ときどきある; 4. よくある; 5. いつもある）から構成される中医体質調査票（Constitution in Chinese Medicine Questionnaire; CCMQ）という、被験者の体質を 9 つに分類するシステムを活用している⁴。9 体質とは「気虚質」「陽虚質」「陰虚質」「淡湿質」「湿熱質」「血瘀質」「気鬱質」「特稟質」「平和質」を差し、このうち、「平和質」は健康体質を示し、ほかの 8 体質は未病体質を示す。この CCMQ は中国の国家基準として採用されており、国民の健康維持増進に活用されている。2008 年には日本語版中医体質調査票（CCMQ-J）が開発されて以来^{5,6}、地域の未病体質の調査^{6,7}および食品の機能性の研究⁸⁻¹⁰など様々な研究に利用されているが、日本人の独自の体質について研究は進められていない。

CCMQ-J 60 問を図 1、また、CCMQ-J 60 問における得点から 9 体質の計算方法および各体質の特徴を図 2 に示す。

ID	質問項目
Q ₁	元気がありますか
Q ₂	疲れやすいですか
Q ₃	息切れしますか
Q ₄	動悸（脈が速い）がしますか
Q ₅	めまいや立ちくらみがしますか
Q ₆	静けさが好きで、人と話をするのもおっくうですか
Q ₇	話し声は弱々しいですか
Q ₈	物忘れをしますか
Q ₉	憂鬱になったり、気が沈む感じがしますか
Q ₁₀	精神が緊張したり、イライラしますか
Q ₁₁	感傷的になったり、涙もろいですか
Q ₁₂	怖がりやすく、不安感がありますか
Q ₁₃	脇の下方もしくは乳房が張って痛みますか
Q ₁₄	胸が重苦しかったり、腹が張ったりしますか
Q ₁₅	無意識によくため息をつきますか
Q ₁₆	体や手足が重だるく感じますか
Q ₁₇	手のひらや足の裏がほてりますか
Q ₁₈	夏でも手足が冷えますか
Q ₁₉	背中・腹部・腰・膝などが冷えますか
Q ₂₀	寒がりです、人より厚着を好みますか
Q ₂₁	顔や体のほてりを感じますか
Q ₂₂	寒さに弱いですか（冬の寒さ・夏のエアコンなど）
Q ₂₃	ひとり風邪をひきやすいですか
Q ₂₄	風邪をひいていないのに、くしゃみが出ますか
Q ₂₅	風邪をひいていないのに、鼻水がつまったりしますか
Q ₂₆	季節の変わり目・温度の変化・いやなおいなどによって息苦しくなることがありますか
Q ₂₇	何もしないのに汗をかきますか
Q ₂₈	おでこが潮っぱいですか
Q ₂₉	皮膚やくちびるが乾きやすいですか
Q ₃₀	過敏症ですか（薬、食べ物、花粉、季節、気候など）
Q ₃₁	蕁麻疹がでやすいですか
Q ₃₂	皮膚に赤紫色の点状・斑状のうっ血斑が出ますか
Q ₃₃	気がつかないうちに、皮膚に青あざができますか
Q ₃₄	皮膚を掻くと赤くなったり、跡がつきますか
Q ₃₅	くちびるの色が、ひとりより赤いですか
Q ₃₆	頬に毛細血管がみえますか
Q ₃₇	どこか体に痛みがありますか
Q ₃₈	頬が赤く、ほてりを感じますか
Q ₃₉	鼻が脂っこく、テカリがありますか
Q ₄₀	顔色が暗く、しみができやすいですか
Q ₄₁	にきびや吹き出物ができやすいですか
Q ₄₂	上瞼がむくみやすいですか
Q ₄₃	目の下にくまができてやすいですか
Q ₄₄	目の乾燥感がありますか
Q ₄₅	くちびるが黒ずんでいますか
Q ₄₆	口や咽が渇きますか
Q ₄₇	咽になにかつまっているという違和感がありますか
Q ₄₈	口臭や口が苦い感じがしますか
Q ₄₉	口の中がねばつきますか
Q ₅₀	舌苔が厚く、べっとりしていますか
Q ₅₁	ふだん、痰が多いですか
Q ₅₂	冷たいものをとると、体の調子が悪くなりますか
Q ₅₃	自然環境や社会環境の変化に順応できますか
Q ₅₄	不眠がちですか
Q ₅₅	冷たいものをとると、下痢をしますか
Q ₅₆	便がねばついて、排便後もすっきりしない感じがしますか
Q ₅₇	便が硬く、便秘しやすいですか
Q ₅₈	お腹が太ってプヨプヨしていますか
Q ₅₉	排尿時、尿道に熱さを感じ、色の濃い尿が出ますか
Q ₆₀	おりもの色は黄色いですか（女性のみ回答）、陰囊が湿っぽい感じがしますか（男性のみ回答）

(回答は、以下の5段階評価で行う。1.まったくない； 2.まれにある； 3.ときどきある；4.よくある； 5.いつもある)

図 1. 日本語版中医体質質問票(CCMQ-J)

日本補完代替医療学会誌 13(2), 43-56, 2016 より引用

体質	特徴
平和質	心身ともにバランスがとれた理想的な「健康タイプ」
Gentleness type	$\frac{100}{32}(Q_1 + Q_{53} - Q_2 - Q_7 - Q_8 - Q_9 - Q_{22} - Q_{54} - 28)$
気虚質	身体の機能状態が低下し、気が虚（不足）した「気不足タイプ」
Qi-deficiency type	$\frac{100}{32}(Q_2 + Q_3 + Q_4 + Q_5 + Q_6 + Q_7 + Q_{23} + Q_{27} - 8)$
陽虚質	腸（エネルギー）が虚（不足）して、冷えやすく、寒さに弱い「腸不足タイプ」
Yang-deficiency type	$\frac{100}{28}(Q_{18} + Q_{19} + Q_{20} + Q_{22} + Q_{23} + Q_{52} + Q_{55} - 7)$
陰虚質	口やのどが渇く、のぼせる、イライラしやすい「水不足タイプ」
Yin-deficiency type	$\frac{100}{32}(Q_{17} + Q_{21} + Q_{29} + Q_{35} + Q_{38} + Q_{44} + Q_{46} + Q_{57} - 8)$
痰湿質	代謝産物がたまりやすい「水停滞タイプ」
Phlegm-wetness type	$\frac{100}{32}(Q_{14} + Q_{16} + Q_{28} + Q_{42} + Q_{49} + Q_{50} + Q_{51} + Q_{58} - 8)$
濕熱質	濕邪と熱邪が体内にこもり、排出機能が不十分な「湿・熱停滞タイプ」
Wet-heat type	$\frac{100}{24}(Q_{39} + Q_{41} + Q_{48} + Q_{56} + Q_{59} + Q_{60} - 6)$
血瘀質	血の流れが悪く、シミができやすい「血停滞タイプ」
Blood-stasis type	$\frac{100}{28}(Q_8 + Q_{33} + Q_{36} + Q_{37} + Q_{40} + Q_{43} + Q_{45} - 7)$
気鬱質	気分が沈みがちな「気停滞タイプ」
Qi-depression type	$\frac{100}{28}(Q_9 + Q_{10} + Q_{11} + Q_{12} + Q_{13} + Q_{15} + Q_{47} - 7)$
特稟質	環境要素など外部刺激に反応しやすい「過敏タイプ」
Special diathesis type	$\frac{100}{28}(Q_{24} + Q_{25} + Q_{26} + Q_{30} + Q_{31} + Q_{32} + Q_{34} - 7)$

図 2.9 種類の体質の特徴と 60 の質問における特典からの計算法

日本補完代替医療学会誌 13(2), 43-56, 2016

1.3 日本における体質診断について

富山県黒部市は2016年10月より、未病対策事業に取り組んでいる。金沢大学の許らは、2017年に富山県黒部市在住の20歳以上の住民を対象とした体質研究が許らによって実施された⁷。この時、黒部市以外に在住する20歳以上の人を対象群として設定された。結果は、黒部市在住の人は平和質の割合が高かったことから、環境が体質に影響を与える可能性が示されている。また、同筆者は、明日葉青汁を1日2回、4週間摂取させた場合の体質変化についても調査を行った¹⁰。その結果、明日葉青汁摂取によって、気虚質や陰虚質を改善する作用があることが示された。このような研究報告から、体質は環境や食事によって変化するものであるため、介入によって未病体質を改善できる可能性がある。

また、CCMQ-Jで分類される体質が気分やQOL(SF-36、SF-12)と相関すること¹¹や、質問の回答から年齢やBMIが推定できること¹²が報告された。このことから、質問票による体質判定によってほかの健康状態の把握もできるのではないかと考えられる。さらに、体質と健康診断値等の生体データを紐づけることで、体質に特徴的な健康状態を把握することができ、生活習慣病のリスク因子や防御因子を見出すことが可能になると考えられる。特に、人によって効果や体感が得にくい、健康食品市場において、体質判定に基づいたソリューション提供が可能になれば、奏効率や体感の向上が期待できる。

1.4 本研究の目的

中医体質質問票は体質を判定するのに有用であると考えられるが、質問が60問あり、回答に時間がかかることから、健康診断や臨床での活用にはまだハードルが高い。また、中国医学に基づく質問票であるため、現代の日本人特有の体質を反映しているかは定かではない。そこで、簡便かつ正確に判定できる新体質質問票を作成することとした。また、CCMQは中国の300か所以上の治未病センターにおいて活用されているが、対象者はすべて中国人であるため、CCMQによって得られた体質が日本人の体質を反映しているかはわからない。

そこで、本研究の目的は、日本人の体質数の推定、ならびに、簡便なかつ日本人の体質を判定することのできる新体質質問票を作成することとした。また、簡便な質問

票を作成するにあたり、CCMQ-J とヘルスケアに関する質問、332 問を健康な日本人男女から回答を収集し、体質に関する質問以外からも体質の推定が可能かどうか、質問を圧縮できるかどうかを検討した。

1.5 本論文の構成

本論文は第 1～6 章で構成されており、第 2 章では対象者について、第 3 章ではデータサイエンスの活用について説明を行っている。そして、第 4 章では日本人の体質数を把握し、第 5 章では既存の質問票からデータサイエンスの利用によって質問を圧縮し、簡便な質問票の開発を試みた。第 6 章では本研究の総括と今後の展望について述べる。

第2章 対象者

2.1 クラウドソーシングについて

クラウドソーシング (crowdsourcing) とは crowd (大衆) とアウトソーシング (outsourcing) を掛け合わせた造語であり、インターネットを通じて、不特定多数の人に仕事や課題を外注できるサービスのことを示す。クラウドソーシングは 2010 年ごろから普及され始め、業務の発注者が依頼する仕事・課題をワーカー (worker) が受注し、こなすことによって成立する。学術的なクラウドソーシングの利用について、定量的な指標はないが、近年、質問紙への回答やアンケートを用いた研究ではしばしば利用される機会が増えてきている。このクラウドソーシングを用いることによって、被験者の募集、研究同意、データ収集、報酬の支払いがすべてオンライン上で実施できる。そのため、今まで行ってきた、人を 1 か所に集めてアンケート調査を実施するなど、試験実施にかかる費用や労力をも削減できることから、アンケート調査などはクラウドソーシングの利用に向いていると考えられる。一方で、クラウドソーシングの利用者は主婦やフリーランス、退職者が多いことも知られている。回答者の基準を厳しくす

ると、対象選抜に該当する人が少なくなってしまうなど、集団のばらつきは課題として残る。

2.2 本研究の対象者

本研究の対象は2017年にクラウドソーシングにより収集した。試験前にインターネットを通し、本研究に対する説明を行ったのち、参加の同意を得た20-85歳の851名を対象者とした。CCMQ-J60問および、その他の健康に関する質問票についての回答を収集し、解析した。本研究では同じ回答となるダミー質問を複数配置することにより、回答の信頼性を確認した。本研究は奈良先端科学技術大学院大学の倫理審査ならびにサントリーグローバルイノベーションセンターの期間内倫理承認を得て実施されたものである。

第3章 データサイエンスの活用

3.1 多変量解析の活用

データサイエンスは1974年にデンマークのコンピューター科学者であるPeter Naurにより使用されたことが知られているが¹³、2010年以降AIやビッグデータが登場したことによってより着目されるようになった。データサイエンスとは、データを収集し、情報科学、統計、数学的なアプローチからデータを解析することに新たなパターンや法則を予測する、さらには画像や動画などをまたは新たな価値や規則を見出すことに利用され、今では数学、統計学、経営学など様々な学問に応用されている。

表1のように n 個のサンプルのそれぞれに対して p 個の変数 x_1, x_2, \dots, x_p の値が観測された(サンプル) \times (変数)の形式のデータを多変量データと呼び、データサイエンスの分野では、しばしば多変量データが解析に用いられる。変数は、表のように質的変数と量的変数に分類され、さらに名義尺度、順序尺度、間隔尺度、比率尺度に分類される。

サンプル No.	変数(変量)			
	x_1	x_2	...	x_p
1	x_{11}	x_{12}	...	x_{1p}
2	x_{21}	x_{22}	...	x_{2p}
⋮	⋮	⋮	⋮	⋮
i	x_{i1}	x_{i2}	...	x_{ip}
⋮	⋮	⋮	⋮	⋮
n	x_{n1}	x_{n2}	...	x_{np}

表 1. 多変量データ

データ	名称	特徴
質的データ	名義尺度	性別や職業などのカテゴリーの違いを表す
	順序尺度	満足度や順位などを表す
量的データ	間隔尺度	温度のように順序も間隔も意味があるが原点の位置はどこでもよい
	比率尺度	長さや重さのように間隔尺度であり、原点が決まっている

表 2. データ区分

3.2 回帰

回帰とは、連続値を予測することであり、1つの説明変数で予測することを単回帰分析、複数の説明変数で予測することを重回帰分析という。単回帰分析や重回帰分析は、回帰式と各データの誤差が最も小さくなるように最小 2 乗法を用いて予測を行う。また、回帰には直線で予測できる線形回帰と曲線で予測ができるロジスティック回帰などの非線形回帰がある。重回帰分析を実施する際に、説明変数間の相関関係が高い組み合わせがある場合には多重共線性生じ、回帰係数が不安定になることがある。その場合には、部分最小二乗回帰 (Partial Least Squares Regression; PLS) がしばしば用いられる。PLS は説明変数がサンプル数より多い場合に用いられることが多く、主成分と目的変数を用いて重回帰分析を行うことで、回帰係数を安定させることができる。

3.3 分類

分類とは、データを整理し、グループに割り当てていくことであり、具体的なモデルとしてロジスティック回帰やk近傍法、決定木などがある。ロジスティック回帰は目的変数が0から1の間であり、疾患の発生率や迷惑メールの判別、商品の購入確率などに用いることができる。k近傍法は周りのデータの分類を学習し、未知データの分類を算出することができる。

3.4 機械学習

機械学習とは、多変量データなどを読み込ませ学習し(インプット)、データに含まれるパターンを見つけることで分類や予測する(アウトプット)ことができるアルゴリズムを構築することのできる技術であり、パソコンのスパムメールの検出や、医療診断、チャットボット、需要予測や自動発注に活用されている。機械学習には「教師あり学習」、「教師なし学習」、「強化学習」の3種類に分類される。

3.4.1 教師あり学習

教師あり学習とは、正解がわかっているデータをインプットし、ルールやパターンを学習させることで、インプットに対し正しいアウトプットを出せるモデルを構築することができる。データが気温や売り上げなどの量的変数を予測する場合には回帰を行い、迷惑メールかどうかなど、質的データを判断する場合には分類を行う。アルゴリズムの例としては線形回帰、ロジスティック回帰、サポートベクタマシン、決定木、ランダムフォレストなどがある。

決定木とは、決定理論の分野において決定を行うための手法であり、意思決定を予測するモデルである。決定木は木構造をとり、回帰および分類の両方の目的で使用することが可能である。回帰木は被説明変数が量的変数をとる場合、分類木は質的変数の場合に利用される。木構造のターミナルノードについて条件分岐を作成できているかどうかは不純度で判断をしている。不純度としてはエントロピーや誤り率、ジニ係数などが用いられる。決定木はデータが変化すると木構造や判別ルールが変わってしまうため、不安定な学習器であるといわれている。そこで、複数の学習器を作成し、

多数決で判定する方法であるバギングを用いることで、学習器を安定させることができる。バギングを用いた手法の1つにランダムフォレストがある。

3.4.2 教師なし学習

教師なし学習は正解データを与えずにデータをインプットし学習させ、データの構造やパターンを抽出し、似た特徴を有するデータ同士を同じグループに分類することができる。教師なし学習には主成分分析やクラスタリングが含まれる。

主成分分析とは、変数が多いデータを扱う際に、変数を低次元の合成変数に変換し、データを解釈しやすくすることを目的に用いる手法である。

クラスタリングとは、何かしらの基準に従ってデータを分類していき、分類の結果や分類されたデータを観察するために使用される。クラスタリングは階層的な手法と非階層的な手法に分類できる。階層的な手法としては、階層的クラスタリングがあり、多変量データに含まれるサンプル同士の距離を計算することにより、サンプル間の距離が小さいものから統合し、クラスターを作成していく方法である。できたクラスターと残りのサンプルの距離が近いもの同士を統合し、すべてのクラスターが統合されるまで繰り返す手法である。最終的には図3のようなデンドログラムが作成され、適当な距離で切断することにより、いくつかのグループに分類することができる。また各グループに含まれる対象を調べることにより、グループの特徴を把握することが可能となる。階層的クラスタリングは近いものから順番にクラスターを作成していくため、あらかじめクラスター数を決定しなくていいことがメリットである。また、クラスター間の距離は表3のように5種類ほど知られている。

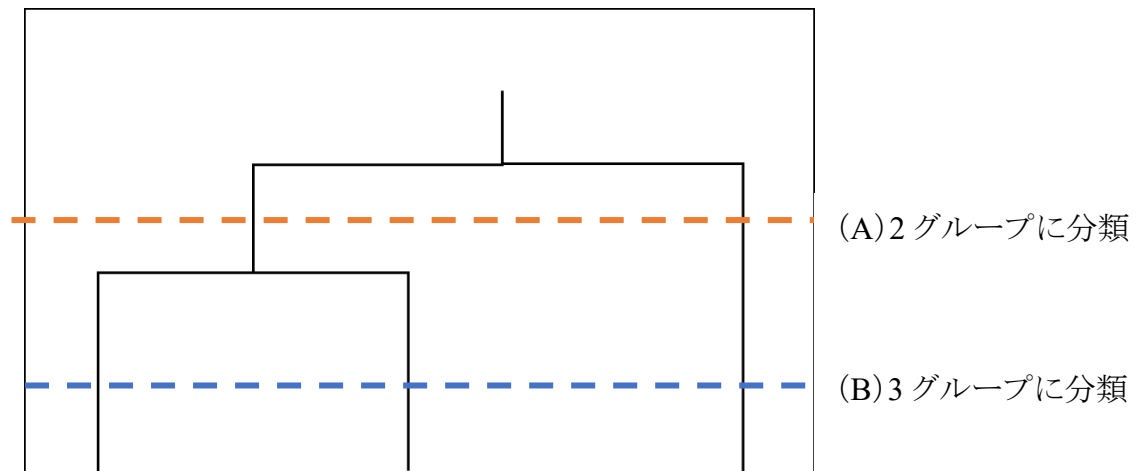


図 3. デンドログラムの例

(A)で区切るとデータは 2 グループに、(B)で区切るとデータは 3 グループに分割される。

	特徴
最近隣法	2つのクラスター内の個体の距離を測定し、最も近いものを群間距離として採用する
最遠隣法	2つのクラスター内の個体の距離を測定し、最も遠いものを群間距離として採用する
群平均法	2つのクラスター内の個体の距離を測定し、その平均距離を群間距離として採用する
重心法	クラスターの重心を計算し、重心間の距離を群間距離として採用する
Ward法	2つのクラスターA,Bを1つのクラスターにまとめたときの個体の分散 $\text{Var}(A \cup B)$ とそれぞれのクラスターの個体の分散 $\text{Var}(A)$ と $\text{Var}(B)$ を計算し、 $\text{Var}(A \cup B) - \text{Var}(A) - \text{Var}(B)$ を群間距離として採用する。分類感度が高く、最も使用されている。

表 3. クラスター間の距離

非階層的クラスタリングとは、階層を作らずにデータをクラスターにしていく手法のことであり、代表的な手法として k -平均法がある。 k -平均法は事前にクラスター数を k 個とし、 k 個の重心をランダムに設定する。それぞれのデータが重心に最も近くなるようにクラスターを作成し、各クラスターの重心を再度算出する。新しく設定された重心をもとに新たにクラスターを形成していく手順を何回か繰り返し、クラスターが変化しなくなるまでこの作業を繰り返すことで、最終的なクラスターを決定する。 k -平均法は計算量が少ないため、膨大なデータをクラスター分析する際に使用することが可能である一方で、初期値の取り方の違いで結果が変わる可能性がある。また、 k -平均法においてはクラスター数をあらかじめ決めておく必要がある。本研究において、クラスター数の推定にはギャップ統計量¹⁴を用いた。

3.4.3 強化学習

強化学習とは教師なし学習と同様、正解データは与えられないが、目的として設定された報酬を最大化するように試行錯誤を繰り返す手法であり、将棋AI、囲碁AIや自動運転の判断に用いられている。

第4章 クラウドソーシングを活用した日本全体の CCMQ-J 問診

票調査における最適体質数とその体質特性

4.1 目的

本章では、クラウドソーシングを利用して収集した、健康な日本人男女 851 名の CCMQ-J の質問 60 問に対する回答についてクラスタリング解析を実施し、最適な日本人の体質分類数を決定することを目的とした。また、分類された体質と中医体質を比較し、日本人の体質の特徴を評価した。

4.2 方法

第 2 章で収集した健康な日本人男女 851 名を対象とし、CCMQ-J への回答データに対して階層的小よび非階層的クラスタ解析を行うことにより、最適体質数の決定を行った。

CCMQ-J は 60 問と 5 段階の回答で構成される。いま、 i 番目の個人の、 j 番目の質問の回答を x_{ij} とすると、60 問における i 番目の個人の回答は、式(1)のように $M=60$ 次元のベクトルで示すことができる。

$$\boldsymbol{x}_i = (x_{i1}, x_{i2}, \dots, x_{i60}) \quad \dots (1)$$

さらに、 N 人の回答が得られると、 N 人の回答は、式(2)で示す行列で表すことができる。

$$X = \begin{pmatrix} x_{11} & \dots & x_{1j} & \dots & x_{1M} \\ \dots & \dots & \dots & \dots & \dots \\ x_{i1} & \dots & x_{ij} & \dots & x_{iM} \\ \dots & \dots & \dots & \dots & \dots \\ x_{N1} & \dots & x_{Nj} & \dots & x_{NM} \end{pmatrix} \quad \dots (2)$$

この行列データについて、非階層的、あるいは階層的クラスタリングによって1から K 個のクラスターを順次形成することで、最適な体質数を推定する。

4.2.1 非階層的クラスタリング

非階層的クラスタリングには k 平均法を用いた。 k 平均法において、 $k = 1$ は全体の行列を示す。また、 $k = 2$ はランダムに全体の行列を2つに分類し、それぞれのグループにおける平均値を計算する(式(3))。

$$\bar{x}^{(k)} = (\bar{x}_1^{(k)}, \bar{x}_2^{(k)}, \dots, \bar{x}_M^{(k)}) \quad \dots(3)$$

次に、サンプル $i = 1, \dots, N$ と、それぞれの平均ベクトルの距離を比較し、最小距離を有するグループに分類し直し、新しく作成されたグループの平均値、すなわち代表点を求めることを、グループへ分類されるサンプルが変動しなくなるまで行う。この時、各クラスターの代表点を c_l ($l = 1, 2, \dots, k$) とすると評価関数 L は式(4)のように表すことができ、評価関数 L が最小になる代表点を更新しながら探索する。

$$L = \sum_{i=1}^n \min_{l=1, \dots, k} \|x_i - c_l\| \quad \dots(4)$$

本研究では、 $k = 2, 3, \dots, 9$ における分類を検討した。

k 平均法は、全体の被験サンプルの分布から、分類をランダムに行うことを特徴とし、多変量データのクラスタリングなどで分類数を推定するのに適している。一方、最初に選択されたサンプルによって構築されるグループが決定するため、再現性を構築することが難しい。そこで、 k 平均法によって推定された分類数の妥当性を評価するために、ギャップ統計量を用いることとした。真のクラスター数が K 個である時、 $k < K$ であればクラスター数を $k + 1$ に増やし、真のクラスター数に近づくのであれば、 L が十分に減少する、すなわち $L_k \gg L_{k+1}$ になることが予測される。一方で、 $k > K$ であればクラスター数を $k + 1$ に増やしても分割がよりよくなるとは考えづらく、 $L_k - L_{k+1}$ は大きくなりませんと考えられる。クラスター数が1であり、式(1)と最小値や最大値が同様の範囲であり、次元が同じデータ X' とし、 X' に対する評価関数を L'_1, L'_2, \dots, L'_k としたときに、ギャップ関数を式(5)のように示すことができる。

$$G_k = \log L'_k - \log L_k \quad \dots(5)$$

式(5)を満たす最小の k を最適体質数とする。

4.2.2 階層的クラスタリング

階層的クラスタリングは類似度の高い順にクラスターを結合し、1 つのデンドログラムを形成する。決定論的にグループを定義するため、形成されるデンドログラムの再現性は高い。本解析では、ward 法 (Ward, 1963) を用いた。ward 法ではグループ内の要素の分散の合計を最小になるようにデンドログラムを形成していくので、鎖効果が置きづらく、凝集性の高いグループ形成が可能であることが知られている。また、ward 法におけるクラスター間の距離は下記のように評価される。

たとえばクラスター A と B を統合したクラスターと、クラスター C の距離には式 (6) の関係が成り立つことがわかっている。このとき n は各クラスターに含まれるサンプル数を示す。

$$D(A \cup B, C) = \frac{n_a + n_c}{n_a + n_b + n_c} D(A, C) + \frac{n_b + n_c}{n_a + n_b + n_c} D(B, C) - \frac{n_c}{n_a + n_b + n_c} D(A, B) \quad \dots(6)$$

ward 法を用いた方法では、決定論的にグループを定義するため、形成されるデンドログラムの再現性は高い。そのため、様々なデータ解析に使われている。山森らが過去に実施した CCMQ-J の 60 項目からなる問診スコアと体質の関連解析も ward 法を用いている¹¹。

4.3 結果

4.3.1 非階層的クラスタリングの結果

最適体質数の決定には全体の空間を把握している k 平均法を使用し、再現性を階層的クラスタリングで確認した。グループ数が 1-9 の時のギャップ関数の推移を図 4 に示す。図 4 からわかるように、 $k = 5$ 以上においてギャップ関数が増加しないことから、最適な体質数は 5 であると考えられた。この時、作成された 5 つのグループを KA, KB, KC, KD, KE とすると、それぞれの人数は 120 人、250 人、160 人、177 人、144 人となった。

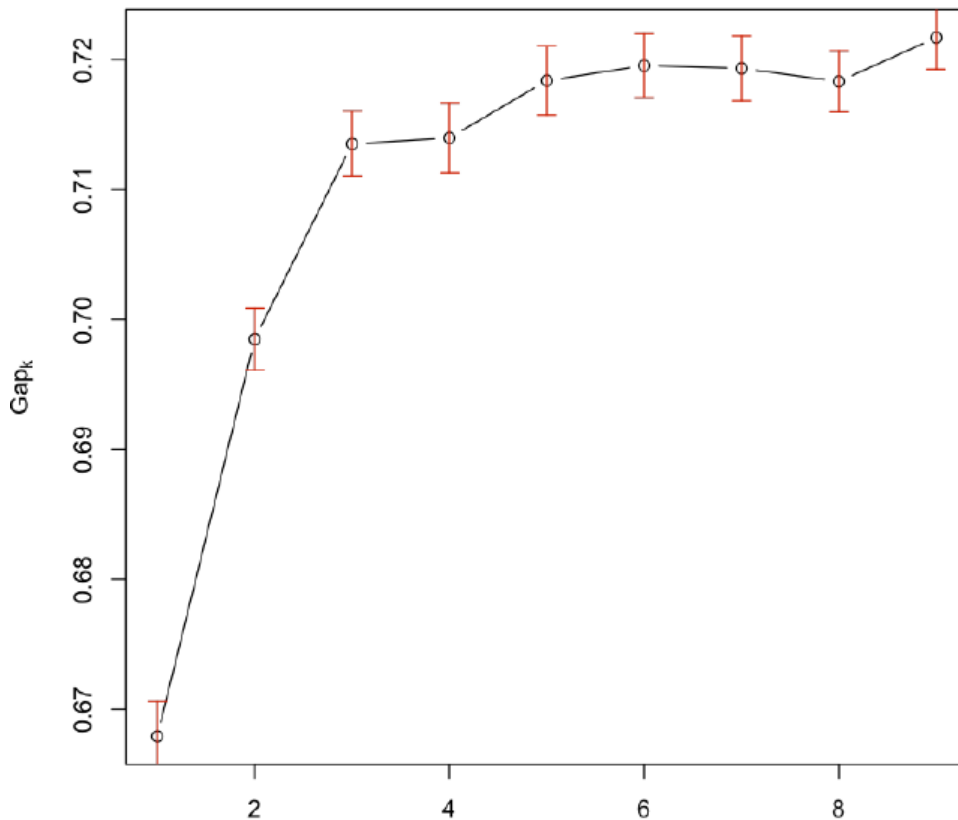


図 4. ギャップ関数の推移

4.3.2 階層的クラスタリングの結果

階層的クラスタリングにおけるクラスター間の距離測定法には **ward** 法を用いた。**ward** 法はクラスターを結合する際に、クラスター内のサンプルからその中心までの距離を最小化するため、凝集性の高いグループを形成することができる。**k**平均法では最適な体質数は 5 と算出されたため、階層的クラスタリングにおいてもグループ数を 5 とした場合の再現性を確認した。その際、作成された 5 つのグループを HA, HB, HC, HD, HE とおいた。**k**平均法と階層的クラスタリングで作成されたグループに直接関係はないが、KA-HA, KB-HB, KC-HC, KD-HD, KE-HE に含まれる共通のサンプルの人数が最大となるように、**k**平均法と階層的クラスタリングのグループの対応付けを行った。その結果、HA, HB, HC, HD, HE の各グループの人数は 176 人、308 人、109 人、119 人、139 人となった。

4.3.3 CCMQ-J の質問項目におけるスコアの k 平均法と階層的クラスタリングの共通性

k平均法により分類された 5 グループ KA, KB, KC, KD, KE と階層的クラスタリングによって分類された 5 グループ HA, HB, HC, HD, HE のグループそれぞれについて CCMQ-J の 60 項目の平均スコアを算出し、中医体質 9 分類を判定した。なお、体質の判定方法は山森らの論文を参照した¹¹⁾。さらに対応するグループごとの相関関数を算出した。KA-HA, KB-HB, KC-HC, KD-HD, KE-HE それぞれについて、相関関数を算出した。その結果、0.97, 0.93, 0.94, 0.97 および 0.90 となり平均値は 0.95 となった。このことから k 平均法と階層的クラスタリングによって得られるグループ内の体質 9 分類の頻度は相関関係があるといえる。

また、非階層クラスタリングと階層クラスタリングの 2 つの方法による CCMQ-J の 60 項目の平均スコアの上位 5% (高得点) および下位 5% (低得点) を選択し、両方の方法で共通に選択された項目を表 4 に示す。さらに、上位および下位スコアから読み取れるグループ A から E の特徴を表 5 に示す。グループ A, B, C, D, E はそれぞれ KA-HA, KB-HB, KC-HC, KD-HD, KE-HE の共通項目グループである。

A

グループ	下位 5%	上位 5%
A		Q2(疲れ)
		Q6(対話がおっくう)
		Q9(憂鬱さ)
		Q10(イライラ度)
		Q11(感傷的, 涙もろい)
		Q12(不安)
		Q15(ため息)
		Q16(手足の重だるさ)
		Q22(寒さへの弱さ)
		Q29(皮膚, 唇の渴きやすさ)
		Q34(皮膚の跡のつきやすさ)
		Q38(ほおのほてり)
		Q44(目の乾燥感)
		Q46(口の渴きやすさ)
	Q54(不眠がち)	
B	Q32(皮膚のうっ血斑の出やすさ)	
C		Q29(皮膚, 唇の渴きやすさ)
D	Q3(息切れ)	
	Q4(動悸)	
	Q13(脇の下方の痛み)	
	Q14(胸の重苦しさ)	
	Q17(体, 手足のほてり)	Q1(元気さ)
	Q21(顔や体のほてり)	
	Q27(汗)	
	Q31(蕁麻疹の出やすさ)	
	Q32(皮膚のうっ血の出やすさ)	
Q33(皮膚の青あざの出やすさ)		

	Q36(頬の毛細血管の見えやすさ)
	Q38(頬のほてり感)
	Q45(唇の黒ずみ)
	Q47(口や咽の渇き)
	Q50(舌苔の厚さとべっとり感)
	Q59(排尿の色の濃さ)
	Q60(おりものの色, 陰囊の湿っぽさ)
E	Q32(皮膚のうっ血の出やすさ)
	Q36(頬の毛細血管の見えやすさ)

表 4. 5 グループにおける平均スコアについて下位 5%と上位 5%の質問項目

グループ	特徴
A	疲労、精神的ダメージ、皮膚、目、口の健康不漁に関するスコアが高い
B	皮膚に関する質問項目 (Q32) について、健康状態が良好である。
C	皮膚と唇の渇きやすさについての質問項目 (Q29) のスコアが高い。
D	元気であること(Q1) のスコアが高い。
E	皮膚のうっ血の出やすさ(Q32)、頬の血管の見えやすさ(Q36) のスコアが高い

表 5. 表 4 から読み取れるグループ A~E の特徴

4.3.4 CCMQ-J の体質におけるスコアの k 平均法と階層的クラスタリングの共通性

k 平均法と階層的クラスタリングによる各グループの被験者数について、相関係数を算出した結果、KA-HA, KB-HB, KC-HC, KD-HD, KE-HE についてそれぞれ、0.90, 0.45, .97, 0.98 および 0.82 となり、平均値は 0.82 となり、CCMQ-J 9 体質の被験者の頻度においても強い相関があることが明らかになった(表 6)。

KA-HA, KB-HB, KC-HC, KD-HD, KE-HE について判定された体質について、被験者の体質 9 分類の相対数を棒グラフにより比較した (図 2)。この図から、グループ C, D, E はそれぞれ陽虚質、平和質、気鬱質の割合が高いことが確認された。一方グループ B は平和質とほかの 8 つの未病体質が同等に含まれ、様々な体質が複雑に存

在する境界型といえる。また、グループ A は平和質を含まず、主に陽虚質と気鬱質が混在する体質であることが明らかになった。

	k60A-A	k60A-B	k60A-C	k60A-D	k60A-E
S60A-A	0.906	0.145	0.478	-0.416	0.698
S60A-B	-0.430	0.453	-0.240	0.580	-0.210
S60A-C	0.744	0.036	0.975	-0.185	-0.216
S60A-D	-0.205	0.863	-0.135	0.988	-0.370
S60A-E	0.738	0.302	0.172	-0.296	0.823

表 6. CCMQ-J の体質分類における KA-HA, KB-HB, KC-HC, KD-HD, KE-HE の相関係数

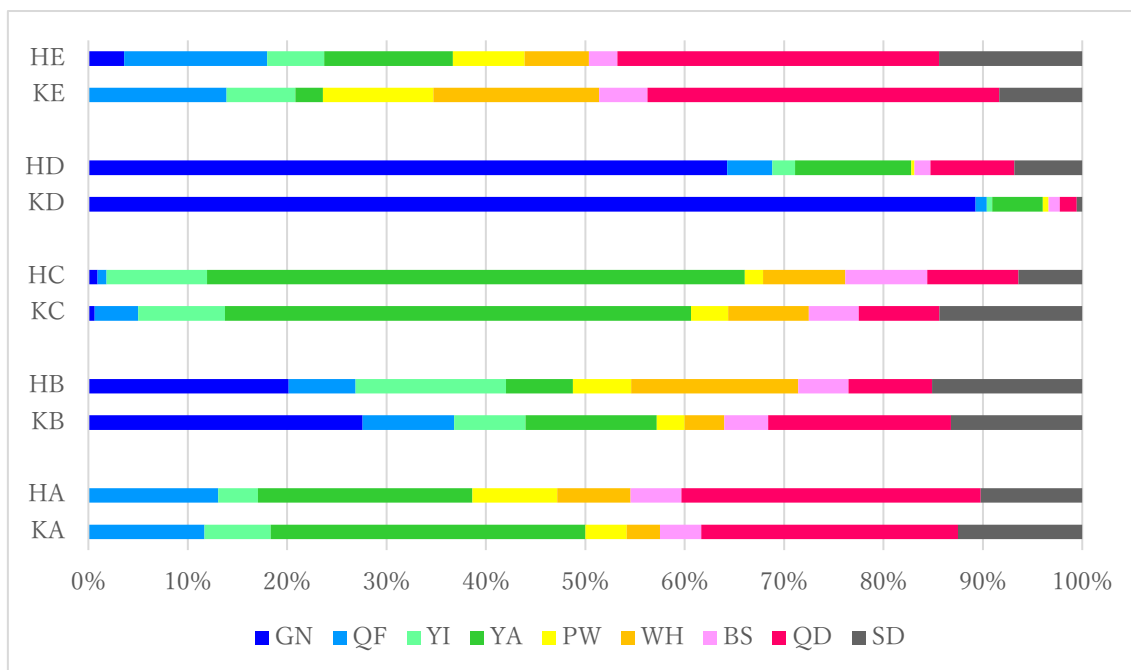


図 5. グループ KA-KE、HA-HE に属する被験者数の体質 9 分類の相対数(%)
 9 体質の略号をカッコ内に示す. 平和質[GN], 気虚質[QF], 陽虚質[YA], 陰虚質[YI],
 痰湿質[PW], 湿熱質[WH], 血瘀質[BS], 気鬱質[QD], 特稟質[SD]。

4.3.5 主成分分析

主成分分析は多変量データに対し、次元を削減し、新たな座標軸を設定しなおすことでデータを解釈させることが可能になる。本研究では、851名の60項目の質問について主成分プロットを行い、CCMQ-Jによる9体質とk平均法による5グループによってラベリングを行った。その結果、k平均法による分類の方が、被験者分布が明確であることが確認できた。

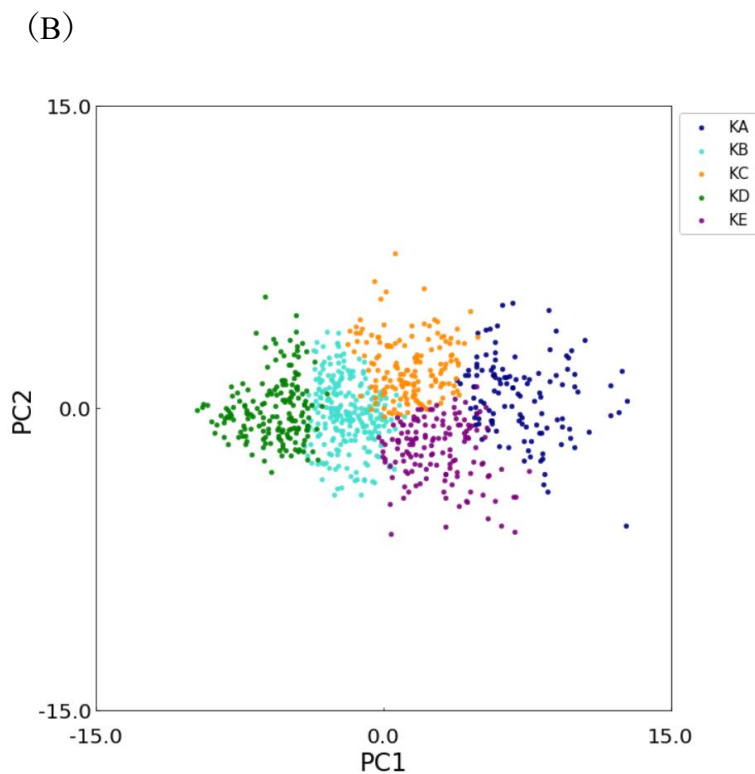
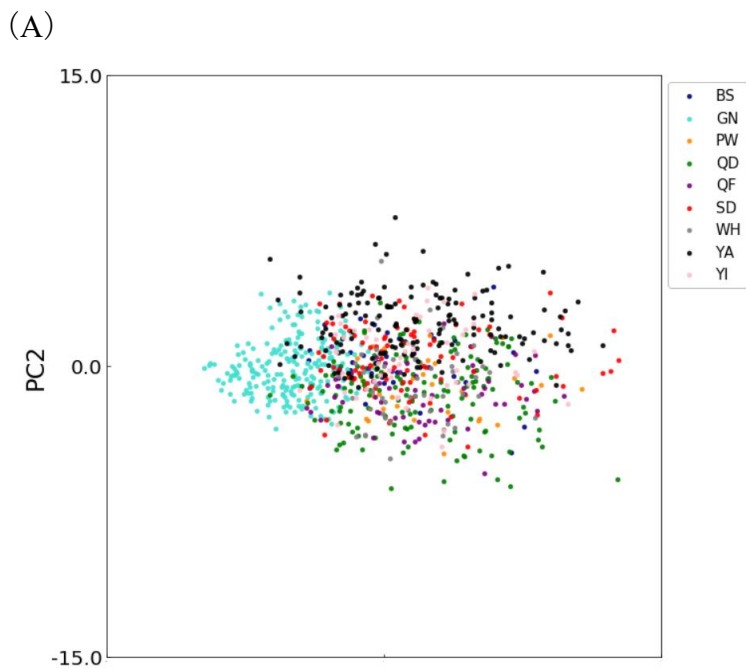


図 6. 60 項目の質問による主成分プロット

(A)CCMQ-J による体質 9 分類 (B)k 平均法による 5 分類

9 分離体質の略号をカッコ内に示す. 平和質[GN]、気虚質[QF]、陽虚質[YA]、陰虚質[YI]、痰湿質[PW]、湿熱質[WH]、血瘀質[BS]、気鬱質[QD]、特稟質[SD]。

4.4 考察

4.4.1 日本人の最適な体質数とその妥当性

中国では古くから体は「気・血・水」で構成されていると考えられており、「気」は生命活動の基本で活力のエネルギーの元、「血」は全身への栄養供給や精神活動の基本物質、「水」は体内の水分であるといわれている。その中でも「気」には遺伝的に定められている先天的な「気」と、生活習慣や環境に影響を受ける後天的な「気」があると定められており、体質は先天的および後天的要素の両方によって定められることがわかる。一方で西洋医学においては、ドイツの精神学者のクレッチマーが、人は leptosome (やせ型)、athletic (闘士型)、pyknic (肥満型) の 3 つの体系で分類できると報告している¹⁵。そのため、西洋医学の考え方からすると人の体質には先天的要素が影響しないことになる。しかし、遺伝的要因が人々の健康に大きく影響を与えることは明らかであり、体質は遺伝的要因と後天的要因の両者によって成り立つものだと考えられるため、日本人の体質を研究するにあたり、先天性の因子についても考慮されている中医学を参考にするのは妥当であると考えられた。また、第 1 章で述べたように、今まで日本においても、体質に関する研究が行われてきた。しかし、中医体質との関連を調査した研究が主体であり、本研究のような日本人独自の体質について調査した研究は初めてである。したがって、本研究は今後の日本人の健康維持増進を図る目的からも意義があると考えられた。

本研究では、クラウドソーシングによって、日本全国 851 名から CCMQ-J 60 問についての回答を得た。この多変量データを階層的クラスタリングおよび非階層的クラスタリングの 2 種類の方法を用いることで日本人の体質数は 5 である可能性が示された。またこの 2 種類の方法から得られたグループについて CCMQ-J の質問に対するスコアから予測できる特徴や中医体質との共通性を見出すことができた。

非階層的クラスタリングは全体のサンプル分布からグループを構築するため、多変量データの分析に適している。非階層的クラスタリングの代表的な手法である k 平均法においてはクラスター数の決定をする必要があるため、本研究においてはギャップ統計量を用いてクラスター数を算出した。また、 k 平均法は、最初のクラスターをランダム

に決定することから、結果が初期値に依存するため、再現性が得られないというデメリットがある。

階層的クラスタリングは被験サンプルの類似度によってクラスターを形成し、デンドログラムを作成して行く手法である。そのため、階層的クラスタリングで作成されるデンドログラムには再現性が保証されるが、一方、クラスタリングの結果は、多変量データ中に偏って存在する被験サンプルの影響を受けやすい可能性がある。

上記に述べたように異なる特性を持つ非階層的と階層的の2種類のクラスタリング方法を用いることで、妥当性の高いグループを設定することが可能になる。また、2種類の方法により設定されたグループ数が異なることや、グループ間に共通性を見出せないときには、設定されたグループの妥当性は低いと判断できる。本研究において、2種類の方法により見出された日本人の体質数が5であることの妥当性を、CCMQ-Jの平均スコアと中医体質の分布の両面において検討した結果、いずれも高い相関を示すことが確認できた。また、図6では、60項目の質問による主成分プロットを示している。この図から、CCMQ-Jによって判定される9体質による被験者の分布と比較して、 k 平均法による被験者の分布は明確に特徴づけられていることがわかる。この結果より、本研究の被験者は5体質に分類できることの妥当性が高いことが示された。

4.4.2 日本人の5体質の特徴

表4において、階層的および非階層的クラスタリングの2種類の方法で共通にみられる高得点ならびに低得点を示した質問項目について、また表5では表4から読み取れる5グループの特徴を示されている。さらに図5は5グループに分類された被験者が中医体質9分類のどの体質に判定されたかを相対数で示したグラフである。図5において、グループDは平和質に判定された人が多いことが特徴づけられていることから、グループDを「平和質型」と名付けることとする。表4において、「平和質型」の人が高得点を有する質問項目はQ1（元気さ）であり、また、頬のほてり感や口やのどの渇きなどの17項目が低得点である。このことから、Q1のスコアが高いほど、17項目は低くなる傾向を示しており、「平和質型」は17項目の兆候と相反する体質であることがわかる。グループBは、グループD（平和質型）に加えて8つの未病体質が同様に含まれている状態であることから、固有の体質を特徴づけられない。そのため、いわゆる「境界型」と考えられる。グループAは平和質が全く含まれず、「陽虚質」、

「気鬱質」を特徴とする「陽虚質・気鬱質混合型」であると考えられる。グループ C は「陽虚質」の分布が多いことを特徴とすることから「陽虚質型」、E は「気鬱質」の分布が多いことを特徴とすることから「気鬱質型」と定義することができる。山森らの研究では「気鬱質」は「陰虚質」、「陽虚質」は「特稟質」と強く関連しており、「気鬱質」と「陽虚質」は未病8体質のなかで最も遠い関係にあることも先行研究によりわかっている¹²。本研究において、日本人の5体質が平和質、陽虚質、気鬱質の3体質のみによって特徴づけられた理由として、中国との文化の違いや近年の食生活の多様化が考えられる。そのため、現代の中国においても9体質の分布には偏りがある可能性もあり、国ごとの体質の違いを明らかにすることで、今後の体質研究自体を発展させることができると考える。

朱ら⁶によると、陽虚質は冷え性、気鬱質は情緒不安定なタイプであると報告されている。さらに厚生労働省により神経症性障害や気分障害などのこころの病気の人数が年々増加していることが報告されている¹⁶。また、食品企業の江崎グリコ株式会社は20代～60代の女性500人を対象に冷え性に関する調査を行い、日本人女性の7割は冷え性であるという結果を取得している¹⁷。このように、本研究で得られた日本人の体質は現代社会の特徴と一致しており、これらの未病体質を改善することは健康維持に重要であると考えられる。

4.5 課題

本研究の被験者はクラウドソーシングという手法を用いていることから、日本各地から多数のデータを低コストで取得できるというメリットがある。一方で、年齢層や性別には30～50代の主婦層が多いことなど母集団に偏りが出してしまうことが懸念される。本研究では、日本人の体質を5つに分けることができたが、ほかの母集団での検討はまだ実施されていない。今後はより年齢や性別を偏りなく集めたうえで、同様の研究を行った際に、本研究と同様に5グループに分類できることが明らかになれば、より5体質の妥当性を保証でき、日本人の体質分類の確度を上昇させることができると考える。

第5章 日本人特異的な体質を判定する質問票の開発

5.1 目的

第4章では、日本人の体質数が5であると推定され、そのグループが(1)陽虚質・気鬱質混合型、(2)境界型、(3)陽虚質型、(4)平和質型、(5)気鬱質型と特徴づけられることを明らかにした。また、時々は、病気や社会的因子、精神的因子、食習慣など、既存のヘルスケアに関する質問票から体質スコアを推定することが可能であり、さらに、健康体質の推定には精神的因子、睡眠状態、食事状態など様々な因子が関与することを報告している¹⁸。

また、体質判定は健康診断や健康食品の開発に利用されることで、日本人の健康維持増進に貢献できると考えられるが、そのためには、判定するための質問が少ないほうが利用しやすい。実際に、アンケートは10分以内で回答可能であることが回答者にとって望ましいことも論じられている¹⁹。また、観察研究において被験者に体質質問を回答してもらう際に、10分の時間を確保することが難しい場合もあるため、1人当たり4、5問で回答可能な質問票を作成することができれば、今後の観察研究への導入は容易になると考えられる。

現代の日本において、疾患の判定や気分、ストレス状態の把握、健康状態の把握を目的として、健康診断や臨床の現場で多くの質問票が利用されているが、それぞれの質問票は別々で調査され、同時に解析されることはほぼない。

本研究では短時間で体質を判定できる質問票を開発することを目的として、質問票既存の質問票に含まれる年齢、性別などの基本情報、病気、社会的因子、精神的因子、食習慣、睡眠状態、未病状態に関する質問272問の質問とCCMQ-J 60問に対する851名の回答を5グループに分類し、各グループに分類する際に重要だと考えられる質問のみで体質を判定できるかどうかを検証することとした。

5.2 方法

5.2.1 質問票

我々は既存の質問票として下記 9 個の質問票、合計 272 問を使用した(表 7)。心理的負荷に関する質問票として K6^{20, 21}、ミニメンタルエステート検査(MMSE)²²、厚生労働省版ストレスチェック^{23,24}、日本語版 Ten Item Personality Inventory^{25, 26}、ピッツバーグ睡眠質問票日本語版²⁷、国民生活基礎調査(平成 28~30 年度)²⁸、および国民健康・栄養調査(平成 25~29 年度)²⁹、日本語版 SNAQ³⁰、食物摂取頻度調査(Food Frequency Questionnaire Based on Food Groups; FFQg)³¹を用いた。この 9 個の質問票は、下記の 7 種類の質問群に分類できると報告されている¹⁸。

- (1) 基本情報(Basic Information; BI) : 身長、体重、年齢、性別など
- (2) 疾患(Disease; DI) : 単なるけがではなく、人間の構造または機能の障害について
- (3) 社会的因子(Social Factor; SO) : 趣味、ライフスタイル、生活環境、職場環境、社会的関係など
- (4) 精神的因子(Mental Factor; ME) : 認知、粘り強さ、感情など
- (5) 食生活(Dietary Habit; DH) : 食品、飲料の好み(料理の種類、味、サプリメント接種など)
- (6) 睡眠状態(Sleeping State ; SL) : 睡眠の質、時間、タイミングなど
- (7) 未病(Sub-health; SH) : 健康と疾患の間の状態

CCMQ-J と 7 種類の質問票を比較すると、CCMQ-J は 54 問が未病(SH)に関する質問である一方で、ヘルスケアに関する 7 種類の質問票の内訳は、BS が 4 問、DH が 111 問、SO が 46 問、ME が 52 問、SL、DH が 15 問、DI、BI がそれぞれ 4 問、SH が 29 問であった。このように CCMQ-J と他 7 種類の質問票では質問の構成が全く異なるため、CCMQ-J によって分類される 9 体質は、SH 以外の質問で予測できる可能性があると考えた。

ID	Question	Ref.
Q1	あなたのクラウドワークス ID を書いてください	
Q2	ライフイベント(配偶者・親族の死, 倒産, 冠婚葬祭, 別居, 離婚, 転職, 退職, 引っ越し, 出産, 入院など)はありましたか.	
Q3	Q2 ではいと答えた人にお聞きします. どのようなイベントがありましたか.	
Q4	Q2 ではいと答えた人にお聞きします. いつイベントがありましたか.	
Q5	あなたの国籍を教えてください.	
Q6	Q5 で日本以外と答えた方にお聞きします. どの国籍を持っていますか. (多重国籍の方は全てお答えください.)	
Q7	過去 30 日間で、どのくらいの頻度で、神経過敏だと感じましたか。	20, 21
Q8	過去 30 日間で、どのくらいの頻度で、絶望的だと感じましたか。	20, 21
Q9	過去 30 日間で、どのくらいの頻度で、そわそわ、落ち着かなく感じましたか。	20, 21
Q10	過去 30 日間で、どのくらいの頻度で、気分が沈み込んで、何が起ころうとも気が晴れないように感じましたか。	20, 21
Q11	過去 30 日間で、どのくらいの頻度で、何をするのも面倒だと感じましたか。	20, 21
Q12	過去 30 日間で、どのくらいの頻度で、自分は価値のない人間だと感じましたか。	20, 21
Q13	今年は何年ですか。(調べずにすぐ書いてください)	22
Q14	今の季節は何ですか。	22
Q15	今日は何曜日ですか。	22
Q16	今日は何月何日ですか。	22
Q17	ここは、何県ですか。	22
Q18	ここは、何市ですか。	22
Q19	今どこにいますか。(例:家、学校)	22
Q20	ここは何階ですか。	22
Q21	ここは何地方ですか。	22
Q22	あなたの仕事についてうかがいます. 非常にたくさんの仕事をしなければならない.	23, 24
Q23	あなたの仕事についてうかがいます. 時間内に仕事が処理しきれない.	23, 24

Q24	あなたの仕事についてうかがいます。一生懸命働かなければならない。	23, 24
Q25	あなたの仕事についてうかがいます。かなり注意を集中する必要がある。	23, 24
Q26	あなたの仕事についてうかがいます。高度の知識や技術が必要なむずかしい仕事だ	23, 24
Q27	あなたの仕事についてうかがいます。勤務時間中はいつも仕事のことを考えていなければならない	23, 24
Q28	あなたの仕事についてうかがいます。からだを大変よく使う仕事だ	23, 24
Q29	あなたの仕事についてうかがいます。自分のペースで仕事ができる	23, 24
Q30	あなたの仕事についてうかがいます。自分で仕事の順番・やり方を決めることができる	23, 24
Q31	あなたの仕事についてうかがいます。職場の仕事の方針に自分の意見を反映できる	23, 24
Q32	あなたの仕事についてうかがいます。自分の技能や知識を仕事で使うことが少ない	23, 24
Q33	あなたの仕事についてうかがいます。私の部署内で意見のくい違いがある	23, 24
Q34	あなたの仕事についてうかがいます。私の部署と他の部署とはうまく合わない	23, 24
Q35	あなたの仕事についてうかがいます。私の職場の雰囲気は友好的である	23, 24
Q36	あなたの仕事についてうかがいます。私の職場の作業環境(騒音, 照明, 温度, 換気など)はよくない	23, 24
Q37	あなたの仕事についてうかがいます。仕事の内容は自分に合っている	23, 24
Q38	あなたの仕事についてうかがいます。働きがいのある仕事だ	23, 24
Q39	最近1ヶ月間のあなたの状態について伺います。活気がわいてくる。	23, 24
Q40	最近1ヶ月間のあなたの状態について伺います。元気がいっぱいだ。	23, 24
Q41	最近1ヶ月間のあなたの状態について伺います。生き生きする。	23, 24
Q42	最近1ヶ月間のあなたの状態について伺います。怒りを感じる。	23, 24
Q43	最近1ヶ月間のあなたの状態について伺います。内心腹立たしい。	23, 24
Q44	最近1ヶ月間のあなたの状態について伺います。イライラしている。	23, 24
Q45	最近1ヶ月間のあなたの状態について伺います。ひどく疲れた。	23, 24
Q46	最近1ヶ月間のあなたの状態について伺います。へとへとだ。	23, 24
Q47	最近1ヶ月間のあなたの状態について伺います。だるい。	23, 24
Q48	最近1ヶ月間のあなたの状態について伺います。気がはりつめている。	23, 24
Q49	最近1ヶ月間のあなたの状態について伺います。不安だ。	23, 24
Q50	最近1ヶ月間のあなたの状態について伺います。落ち着かない。	23, 24

Q51	最近1ヶ月間のあなたの状態について伺います。憂鬱だ。	23, 24
Q52	最近1ヶ月間のあなたの状態について伺います。何をするのも面倒だ。	23, 24
Q53	最近1ヶ月間のあなたの状態について伺います。物事に集中できない。	23, 24
Q54	最近1ヶ月間のあなたの状態について伺います。気分が晴れない。	23, 24
Q55	最近1ヶ月間のあなたの状態について伺います。仕事が手につかない。	23, 24
Q56	最近1ヶ月間のあなたの状態について伺います。悲しいと感じる。	23, 24
Q57	最近1ヶ月間のあなたの状態について伺います。めまいがする。	23, 24
Q58	最近1ヶ月間のあなたの状態について伺います。体のふしぶしが痛む。	23, 24
Q59	最近1ヶ月間のあなたの状態について伺います。頭が重かったり頭痛がする。	23, 24
Q60	最近1ヶ月間のあなたの状態について伺います。首筋や肩がこる。	23, 24
Q61	最近1ヶ月間のあなたの状態について伺います。腰が痛い。	23, 24
Q62	最近1ヶ月間のあなたの状態について伺います。目が疲れる。	23, 24
Q63	最近1ヶ月間のあなたの状態について伺います。動悸や息切れがする。	23, 24
Q64	最近1ヶ月間のあなたの状態について伺います。胃腸の具合が悪い。	23, 24
Q65	最近1ヶ月間のあなたの状態について伺います。食欲がない。	23, 24
Q66	最近1ヶ月間のあなたの状態について伺います。便秘や下痢をする。	23, 24
Q67	最近1ヶ月間のあなたの状態について伺います。よく眠れない。	23, 24
Q68	あなたの周りの方々についてうかがいます。上司とどのくらい気軽に話ができますか。	23, 24
Q69	あなたの周りの方々についてうかがいます。職場の同僚とどのくらい気軽に話ができますか。	23, 24
Q70	あなたの周りの方々についてうかがいます。配偶者、家族、友人等とどのくらい気軽に話ができますか。	23, 24
Q71	あなたの周りの方々についてうかがいます。あなたが困った時、上司はどのくらい頼りになりますか。	23, 24
Q72	あなたの周りの方々についてうかがいます。あなたが困った時、職場の同僚はどのくらい頼りになりますか。	23, 24
Q73	あなたの周りの方々についてうかがいます。あなたが困った時、配偶者、家族、友人等はどのくらい頼りになりますか。	23, 24

Q74	あなたの周りの方々についてうかがいます。あなたの個人的な問題を相談したら、上司はどのくらいきいてくれますか。	23, 24
Q75	あなたの周りの方々についてうかがいます。あなたの個人的な問題を相談したら、職場の同僚はどのくらいきいてくれますか。	23, 24
Q76	あなたの周りの方々についてうかがいます。あなたの個人的な問題を相談したら、配偶者、家族、友人等はどのくらいきいてくれますか。	23, 24
Q77	あなたの満足度についてうかがいます。仕事に満足だ。	23, 24
Q78	あなたの満足度についてうかがいます。家庭生活に満足だ	23, 24
Q140	私は自分自身のことを、しっかりしていて、自分に厳しいと思う。	23, 24
Q141	私は自分自身のことを、他人に不満をもち、もめごとを起こしやすいと思う。	25, 26
Q142	私は自分自身のことを、心配性で、うろたえやすいと思う。	25, 26
Q143	私は自分自身のことを、活発で、外向的だと思う。	25, 26
Q144	私は自分自身のことを、新しいことが好きで、変わった考えをもつと思う。	25, 26
Q145	私は自分自身のことを、ひかえめで、おとなしいと思う。	25, 26
Q146	私は自分自身のことを、人に気をつかう、やさしい人間だと思う。	25, 26
Q147	私は自分自身のことを、だらしなく、うっかりしていると思う。	25, 26
Q148	私は自分自身のことを、冷静で、気分が安定していると思う。	25, 26
Q149	私は自分自身のことを、発想力に欠けた、平凡な人間だと思う。	25, 26
Q150	1週間で何人と話をしましたか。	
Q151	1週間で何時間くらい話をしましたか。(分か時間をつけて教えてください。)	
Q152	1週間で話をしたのは主に誰とですか。	
Q153	性別を選んでください。	25, 26
Q154	生年月日をご記入ください。	28
Q155	身長を半角でご記入ください(例:178cm)	28
Q156	体重をご記入ください(例:70kg)	29
Q157	現在住んでいる都道府県はどこですか。	
Q158	これまでに一番長く過ごした都道府県はどこですか。	
Q159	配偶者(夫又は妻)の有無(事実上夫婦として生活しているが、婚姻届を提出していない場合も含みます)未婚	29

Q160	普段一緒にお住まいで、生計を共にしている方は、あなたを含めて何人ですか。数字のみでお答えください。(一時的に不在の方を含めます)	28
Q161	1人(単独世帯)の方の場合は、その状況についてお答えください。	28
Q162	(単独世帯)の方は回答してください。	29
Q163	あなたのお住まいの地域についてお尋ねします。あなたのお住まいの地域の人々は、お互いに助け合っている。	29
Q164	あなたのお住まいの地域についてお尋ねします。あなたのお住まいの地域の人々は、信頼できる。	29
Q165	あなたのお住まいの地域についてお尋ねします。あなたのお住まいの地域の人々は、お互いに挨拶をしている。	29
Q166	あなたのお住まいの地域についてお尋ねします。あなたのお住まいの地域では、問題が生じた場合、人々は力を合わせて解決しようとする。	29
Q167	現在、学校に在学しているかどうかお答えください。卒業:1	29
Q168	在学中の方はその学校について、「卒業」の方は最終卒業校についてお答えください。小学・中学、高校・旧制中に○を付けた方で、特別支援学校・特別支援学級に在学中又は卒業した方はこちらにも○をつけてください。	28
Q169	この1ヶ月間の仕事の状況で収入を伴う仕事をしたことがありますか。	28
Q170	Q169で「あり」と答えた方は以下の選択肢を選択してください。	28
Q171	Q169で「なし」と答えた方は以下の選択肢を選択してください。	28
Q172	Q169で「あり」と回答した方はお答えください。先週1週間に実際に仕事をした日数を教えてください。(数字だけ記入してください)	28
Q173	Q169で「あり」と回答した方はお答えください。先週1週間に残業も含めた総労働時間を教えてください。	28
Q174	Q169で「あり」と回答した方はお答えください。現在の主な仕事の内容についてお答えください。	28
Q175	Q169で「あり」と回答した方はお答えください。主な雇用形態について、お答えください。	28
Q176	Q175で 一般常雇者(契約期間の定めのない雇用者)、一般常雇者(契約期間が1年以上の雇用者)、1ヶ月以上1年未満の契約の雇用者、日々又は1ヶ月未満の契約の雇用者と答えた方は、「勤め先での呼称」をお答えください。	28
Q177	Q177 職業希望の有無について、お答えください。	28

Q178	Q177 でしたいと思っている方に伺います。どのような形で仕事をしたいと思えますか。最もしたいと思う仕事の形の番号1つに○をつけてください。現在仕事を探していない方でも、仕事につくとしたらどのような形で仕事をしたいかをお答えください。	28
Q179	Q177 でしたいと思っている方に伺います。すぐに仕事につけますか。	28
Q180	Q179 で「つける」と回答した方に伺います。仕事を探していますか。現在採用結果を待っている方も「探している」に○をつけてください。	28
Q181	Q179 で「つけない」と回答した方はお答えください。仕事につけない理由であてはまるすべての番号に○をつけてください。	28
Q182	あなたの現在の健康状態はいかがですか。あてはまる番号1つに○をつけてください。	28
Q183	あなたは現在、健康上の問題で日常生活に何か影響がありますか。	28
Q184	「ある」と回答した方にお聞きします。それはどのようなことに影響がありますか。あてはまるすべての番号に○をつけてください。	28
Q185	あなたは現在、傷病(病気やけが)で病院や診療所(医院、歯科医院)、あんま・はり・きゅう・柔道整復師(施術所)に通っていますか。(往診、訪問診療含む)	28
Q186	「通っている」と回答した方にお聞きします。どのような傷病(病気やけが)で通っていますか。あてはまるすべての傷病名の番号に○をつけてください。	28
Q187	あなたはこれまで、どのような傷病(病気やけが)で病院や診療所(医院、歯科医院)、あんま・はり・きゅう・柔道整復師(施術所)に通いましたか。あてはまるすべての傷病名の番号に○をつけてください。	28
Q188	日常生活で体を動かす※こと(生活活動)についてお伺いします。日常生活で毎日 60 分間くらい体を動かすような生活をしていますか。あてはまる番号を1つ選んで○印をつけて下さい。※「体を動かす」とは、歩く、犬の散歩をする、そうじをする、自転車に乗る、早歩きをする、こどもと活発に遊ぶ、農作業をする、階段を速くのぼるなどと同等の動きを指します。	28
Q189	あなたは、座ったり寝転がったりして過ごす時間※が、1 日平均して平日はどのくらいありますか。(例:7時間 30 分)「座ったり寝転がったりして過ごす時間」には、机やコンピューターに向かう時間(仕事、勉強や読書などを含む)、テレビを見る時間、座って会話をする時間、車を運転する(または車に乗っている)時間、電車で座っている時間等を含みます。ただし、睡眠時間は含みません。	29

Q190	あなたは、座ったり寝転がったりして過ごす時間※が、1日平均して休日ほどのくらいありますか。(例:7時間30分)「座ったり寝転がったりして過ごす時間」には、机やコンピューターに向かう時間(仕事、勉強や読書などを含む)、テレビを見る時間、座って会話をする時間、車を運転する(または車に乗っている)時間、電車で座っている時間等を含みます。ただし、睡眠時間は含みません。	29
Q191	過去1か月間において、通常何時頃寢床につきましたか?(午前・午後を記入してください)	27
Q192	過去1か月間において、寢床についてから眠るまでにどれくらい時間を要しましたか?(単位は時間(h)になおしてあります)	27
Q193	過去1か月間において、通常何時頃起床しましたか?	27
Q194	過去1か月間において、実際の睡眠時間は何時間くらいでしたか?これは、あなたが寢床の中にいた時間とは異なる場合があるかもしれません。	27
Q195	過去1か月間において、どれくらいの頻度で、「寢床にいてから30分以内に眠ることができなかった」から睡眠が困難でしたか?最も当てはまるものを1つ選んでください。	27
Q196	過去1か月間において、どれくらいの頻度で、「夜間または早朝に目が覚めた」から睡眠が困難でしたか?最も当てはまるものを1つ選んでください。	27
Q197	過去1か月間において、どれくらいの頻度で、「トイレに起きた」から睡眠が困難でしたか?最も当てはまるものを1つ選んでください。	27
Q198	過去1か月間において、どれくらいの頻度で、「息苦しかった」から睡眠が困難でしたか?最も当てはまるものを1つ選んでください。	27
Q199	過去1か月間において、どれくらいの頻度で、「咳が出たり、大きないびきをかいた」から睡眠が困難でしたか?最も当てはまるものを1つ選んでください。	27
Q200	過去1か月間において、どれくらいの頻度で、寝る前にパソコンやスマホなどの画面を見ましたか?この質問には回答しないで[201]に進んでください。	27
Q201	過去1か月間において、どれくらいの頻度で、「ひどく寒く感じた」から睡眠が困難でしたか?最も当てはまるものを1つ選んでください。	27
Q202	過去1か月間において、どれくらいの頻度で、「ひどく暑く感じた」から睡眠が困難でしたか?最も当てはまるものを1つ選んでください。	27
Q203	過去1か月間において、どれくらいの頻度で、「悪い夢をみた」から睡眠が困難でしたか?最も当てはまるものを1つ選んでください。	27

Q204	過去1か月間において、どれくらいの頻度で、「痛みがあった」から睡眠が困難でしたか？ 最も当てはまるものを1つ選んでください。	27
Q205	上記以外の理由があれば、次の空欄に記載してください。	27
Q206	Q205の理由で過去1か月において、どれくらいの頻度で睡眠が困難でしたか？	27
Q207	過去1か月において、ご自分の睡眠の質を全体として、どのように評価しますか？	27
Q208	過去1か月において、どのくらいの頻度で、眠るために薬(医師から処方された薬あるいは 薬屋で買った薬)を服用しましたか？	27
Q209	過去1か月において、どれくらいの頻度で、車の運転中や食事中や社会活動中など眠っ てはいけない時に、おきていられなくなり困ったことがありましたか？	27
Q210	過去1か月において、物事をやり遂げるのに必要な意欲を持続するうえで、どのくらい問 題がありましたか？	27
Q211	あなたはたばこを吸いますか。○をつけて下さい。	29
Q212	吸っている方にお聞きします。たばこをやめたいと思いますか。	29
Q213	自分の歯※は何本ありますか。※自分の歯には、親知らず、入れ歯、ブリッジ、インプラント は含みません。さし歯は含みます。親知らずを抜くと全部で28本が正常ですが、28本より 多かったり少なかったりすることもあります。0本の場合は、0と書いて下さい。	29
Q214	かんで食べる時の状態について、1つ選んで○印をつけて下さい。	29
Q215	あなたの食べ方や食事の様子についておたずねします。ゆっくりよくかんで食事をしま すか。	29
Q216	あなたの食べ方や食事の様子についておたずねします。半年前に比べて固いものが食 べにくくなりましたか。	29
Q217	あなたの食べ方や食事の様子についておたずねします。お茶や汁物等でむせることが ありますか。	29
Q218	あなたの食べ方や食事の様子についておたずねします。口の渇きが気になりますか。	29
Q219	あなたの食べ方や食事の様子についておたずねします。左右両方の奥歯でしっかりか みしめられますか。	29
Q220	性別を選んでください。	
Q221	あなたは、主食(ごはん、パン、麺類などの料理)、主菜(魚介類、肉類、卵類、大豆・大豆 製品を主材料にした料理)、副菜(野菜類、海藻類、きのこ類を主材料にした料理)の3つ	29

	を組み合わせる食べることが1日に2回以上あるのは週に何日ありますか。あてはまる番号に○印をつけて下さい。	
Q222	あなたは、外食(飲食店での食事)をどのくらい利用していますか。該当する番号を1つ選んで○印をつけて下さい。	29
Q223	あなたは、持ち帰りの弁当や惣菜をどのくらい利用していますか。該当する番号を1つ選んで○印をつけて下さい。	29
Q224	食欲はありますか。	28
Q225	Q224 でほとんどない、あまりないと回答した方にお聞きします。食欲がない理由は何ですか。	
Q226	若いころと比べて食欲に変化はありますか。	
Q227	食欲がなくて困っていますか。	
Q228	今よりも痩せたいと思っていますか。	
Q229	食事を、どのくらい食べると満腹感を感じますか。	30
Q230	食事の味はいかがですか。	30
Q231	食事は、1日何回食べますか。	30
Q232	今よりも食事などを多く食べられるようになりたいですか	
Q233	食事の時間は不規則ですか？	
Q234	あなたは現在、日常生活で悩みやストレスがありますか。	28
Q235	上で「あり」とお答えした方にお聞きします。それは、どのような原因ですか。あてはまるすべての原因の番号に○をつけてください。	28
Q236	中国茶(烏龍茶、ジャスミン茶、プーアル茶など)を飲む回数	
Q237	中国茶(烏龍茶、ジャスミン茶、プーアル茶など)の1回あたりの飲む量。基準量: カップ1杯 160 mL	
Q238	水、ミネラルウォーターの飲む回数	
Q239	水、ミネラルウォーターの1回あたりの飲む量。基準量: コップ1杯 200 mL。	
Q240	あなたは現在、日常生活で悩みやストレスがありますか。	
Q241	朝、ご飯は1週間にお茶碗○杯	31
Q242	昼、ご飯は1週間にお茶碗○杯	31
Q243	夕、ご飯は1週間にお茶碗○杯	31

Q244	朝、パン・シリアルは1週間に○枚	31
Q245	昼、パン・シリアルは1週間に○枚	31
Q246	夕、パン・シリアルは1週間に○枚	31
Q247	朝、麺類は1週間に○杯	31
Q248	昼、麺類は1週間に○杯	31
Q249	夕、麺類は1週間に○杯	31
Q250	ご飯もの、1週間に○回	31
Q251	カレーやハヤシライス、1週間に○回	31
Q252	(肉・肉加工品類) 朝、一回に食べる量を選んでください	31
Q253	(肉・肉加工品類) 朝、1週間に○回	31
Q254	(肉・肉加工品類) 昼、一回に食べる量を選んでください	31
Q255	(肉・肉加工品類) 昼、1週間に○回	31
Q256	(肉・肉加工品類) 夕、一回に食べる量を選んでください	31
Q257	(肉・肉加工品類) 夕、1週間に○回	31
Q258	(魚介類) 朝、一回に食べる量を選んでください	31
Q259	(魚介類) 朝、1週間に○回	31
Q260	(魚介類) 昼、一回に食べる量を選んでください	31
Q261	(魚介類) 昼、1週間に○回	31
Q262	(魚介類) 夕、一回に食べる量を選んでください	31
Q263	(魚介類) 夕、1週間に○回	31
Q264	(卵) 1週間に○個	31
Q265	(大豆・大豆製品) 朝、一回に食べる量を選んでください	31
Q266	(大豆・大豆製品) 朝、1週間に○回	31
Q267	(大豆・大豆製品) 昼、一回に食べる量を選んでください	31
Q268	(大豆・大豆製品) 昼、1週間に○回	31
Q269	(大豆・大豆製品) 夕、一回に食べる量を選んでください	31
Q270	(大豆・大豆製品) 夕、1週間に○回	31
Q271	牛乳、1週間に○杯	31
Q272	乳製品、1週間に○回	31

Q273	(海藻) 一回に食べる量を選んでください	31
Q274	(海藻) 1週間に○回	31
Q275	(小魚) 一回に食べる量を選んでください	31
Q276	(小魚) 1週間に○回	31
Q277	(緑黄色野菜) 朝、一回に食べる量を選んでください	31
Q278	(緑黄色野菜) 朝、1週間に○回	31
Q279	(緑黄色野菜) 昼、一回に食べる量を選んでください	31
Q280	(緑黄色野菜) 昼、1週間に○回	31
Q281	(緑黄色野菜) 夕、一回に食べる量を選んでください	31
Q282	(緑黄色野菜) 夕、1週間に○回	31
Q283	(淡色野菜・きのこ類) 朝、一回に食べる量を選んでください	31
Q284	(淡色野菜・きのこ類) 朝、1週間に○回	31
Q285	(淡色野菜・きのこ類) 昼、一回に食べる量を選んでください	31
Q286	(淡色野菜・きのこ類) 昼、1週間に○回	31
Q287	(淡色野菜・きのこ類) 夕、一回に食べる量を選んでください	31
Q288	(淡色野菜・きのこ類) 夕、1週間に○回	31
Q289	(果物) 一回に食べる量を選んでください	31
Q290	(果物) 1週間に○回	31
Q291	(いも) 一回に食べる量を選んでください	31
Q292	(いも) 1週間に○回	31
Q293	ジャム・はちみつを一回に食べる量を選んでください	31
Q294	ジャム・はちみつ、1週間に○回	31
Q295	煮物料理を一回に食べる量を選んでください	31
Q296	煮物料理、1週間に○回	31
Q297	酢の物・和えを一回に食べる量を選んでください	31
Q298	酢の物・和え、1週間に○回	31
Q299	和菓子、1週間に○回	31
Q300	菓子パン・ケーキ、1週間に○回	31
Q301	スナック菓子・揚げ菓子(1/2袋 60g)、1週間に○回	31

Q302	せんべい類・クッキー(1枚 5~10g×2枚)、1週間に○回	31
Q303	アイスクリーム(1/2カップ 50g)、1週間に○回	31
Q304	キャンディ・キャラメル(1個 5g×4個)、1週間に○回	31
Q305	ゼリー・プリン(1個 100g)、1週間に○回	31
Q306	チョコレート(1/2枚 200ml)、1週間に○回	31
Q307	コーヒー・紅茶に入れる砂糖を一回に入れる量を選んでください	31
Q308	コーヒー・紅茶に入れる砂糖、1週間に○杯	31
Q309	清涼飲料・缶コーヒーを一回に飲む量を選んでください	31
Q310	清涼飲料・缶コーヒー、1週間に○本	31
Q311	アルコール飲料を一回に飲む量を選んでください	31
Q312	アルコール飲料、1週間に○回	31
Q313	栄養補助食品(1回の目安は 200kcal)、1週間に○回	31
Q314	バター・マーガリンを一回に食べる量を選んでください	31
Q315	バター・マーガリン、1週間に○回	31
Q316	天ぷらやフライの揚げ物料理、1週間に○回	31
Q317	マヨネーズやドレッシング、1週間に○回	31
Q318	炒め物、1週間に○回	31
Q319	種実類(ピーナッツやアーモンド)、一回に食べる量を選んでください	31
Q320	種実類(ピーナッツやアーモンド)、1週間に○回	31
Q321	種実類(ごま)、一回に食べる量を選んでください	31
Q322	種実類(ごま)、1週間に○回	31
Q323	塩辛いものを一回に食べる量を選んでください	31
Q324	塩辛いもの、1週間に○回	31
Q325	漬け物を一回に食べる量を選んでください	31
Q326	漬け物、1週間に○回	31
Q327	食卓しょうゆ・ソースを一回にかける量を選んでください	31
Q328	食卓しょうゆ・ソース、1週間に○回	31
Q329	味噌汁、1週間に○杯	31
Q330	すまし汁やスープ、1週間に○杯	31

Q331	麺類の汁は飲みますか	31
Q332	塩分、麺類の汁など) 1 週間に○回飲む	31
Q333	外食の味の感じ方について	31

表 7.9 つの質問票に含まれる 272 問

5.2.2 階層的クラスタリング

本研究で収集した 851 名の 332 問に対する回答を 5 グループに分類するために、4.2.2 と同様に ward 法を用いて階層的クラスタリングを実施した。

5.2.3 ランダムフォレスト

ランダムフォレストは Breiman³² によって提案された、決定木を利用したアンサンブル学習を行う手法である。決定木とは統計分析やデータマイニングに利用される予測モデルの一つであり、分類と回帰の 2 つの手法が存在するが、本研究においては体質の決定に必要な質問を見出すことを目的とするため、分類木を用いた。

決定木はデータのジニ係数やエントロピーなどの不純度を示す指標をもとに判断プロセスを構築していく。ジニ係数は誤分類する確率を平均化した指標であり、0 から 1 までの値をとる。0 に近いほど不純度は低く、1 に近いほど不純度が高い。所得の不平等さを示すために利用され、日本においても厚生労働省が実施している所得再分配調査などに所得分配の不平等さを測定するために利用されている³³。ローレンツ曲線 ($L(F)$) と均等分配線に囲まれた面積の割合の 2 倍で定義する(式 7)。

$$1 - 2 \int_0^1 L(F) \dots (7)$$

決定木では分岐前のジニ係数と分岐後のジニ係数の差を計算し、その差が最大となるような条件を探し、モデルを構築していく。また、機械学習の精度を高めるために、複数のモデルを組み合わせてより強力なモデルを構築するアンサンブル学習といった手法を用いる。アンサンブル学習では、複数の手法によって下された答えを多数決によって決めることができるため、精度が高くなると考えられる。

本研究では、複数の、少しずつ異なる決定木を作成するランダムフォレストという手法を選択し、体質判定に用いることとした。決定木のみでは学習データの偏りやノイズも学習してしまうため、過学習を起す恐れがあるが、ランダムフォレストを用いることで、複数の結果の平均値を得ることができるため、過学習の可能性を低下させることが可能となる³⁴。ランダムフォレストの利点としては、考慮するパラメーターが少ないことや、どの特徴量が重要であるかを知ることができることがあげられる。一方でノイズが多い

データやデータ自体が少ないとうまく学習できないことがデメリットとしてあげられる。ランダムフォレストはまず、データをトレーニングデータとテストデータに分け、トレーニングデータを用いて説明変数と目的変数の関係性を学習させる。次にテストデータの説明変数のみを用いて予測を行い、精度を算出する。本研究では、Python 3.8.3 を用いて、ランダムフォレストを実装した。

5.2.4 質問票のバリデーション

本研究の対象は 2017 年にクラウドソーシングにより収集した 851 名の回答者のうち、再同意を得られた 162 名に対し、CCMQ-J 60 問および、ランダムフォレストにより特徴量が高いと算出された 39 問についての回答を収集し、解析した。本研究は奈良先端科学技術大学大学院の倫理審査ならびにサントリーグローバルイノベーションセンターの期間内倫理承認を得て実施されたものである。

5.3 結果:

5.3.1 クラスタリング

332 問に対する 851 名の回答を用いて階層的クラスタリングを実施し、5 グループに分類した(図 7)。その際、各グループ数の人数はそれぞれ、155 人、257 人、156 人、151 人、132 人となった。

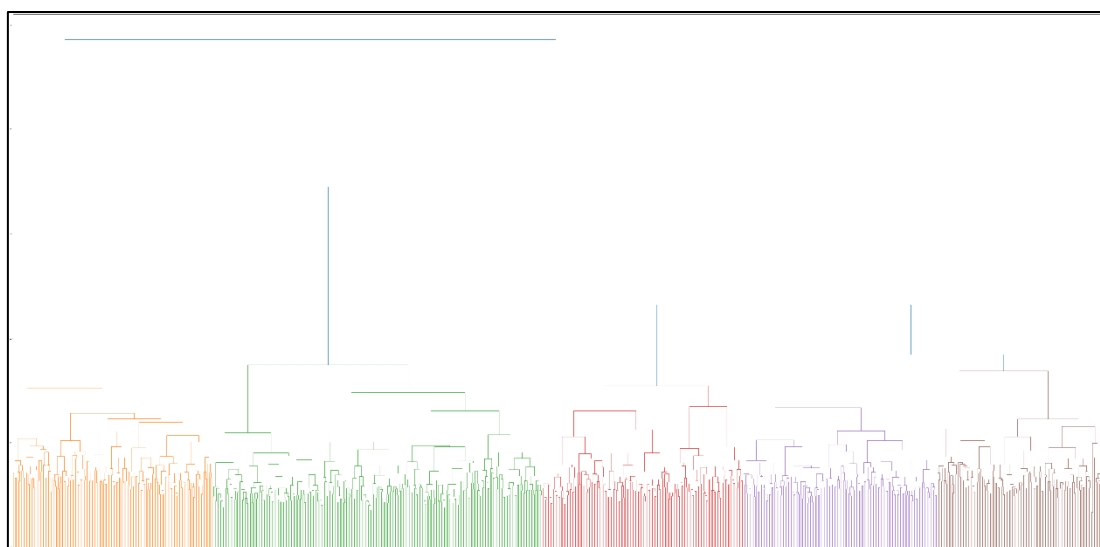


図 7. 851 名を階層的クラスタリングで 5 分類にした際のデンドログラム

5.3.2 質問数の削減

ランダムフォレストを用いて、851 名の 332 問に対する回答から、クラスタリングにより分類された体質を予測できるかどうかを検討した。この時、トレーニングデータ:テストデータ=8:2 として、モデルを作成した。また、この時、作成した決定木は 128 個、木の最大の深さは 8 とした。この時、予測精度は 59%と算出された。

次に、特徴量を算出し、特徴量が低い質問から 1 問ずつ削除していったときの精度をランダムフォレストにより求めた結果、1 問~15 問までは精度が上昇するが、それ以降は 332 問に増やした場合でも精度は 60%前後にとどまることが確認された。実際、15 問から 332 問で体質を判定したときの精度の平均は 61%であり、最小値は 52%、最大値は 66%であった。この結果より、まず特徴量の高い質問を 40 問抽出した(表 8)。40 問のうち数値で回答する質問(Q155:身長を半角でご記入ください)については削除した。

抽出された 39 問の質問について、「最近 1 か月で…」と問う質問と「過去 30 日間で…」と問う質問が混在していることや、「最近1ヶ月間のあなたの状態について伺います。気分が晴れない。」など、聞き方が不自然な質問があったため、回答者が回答しやすいように、文言の微修正を行った(表 9)。

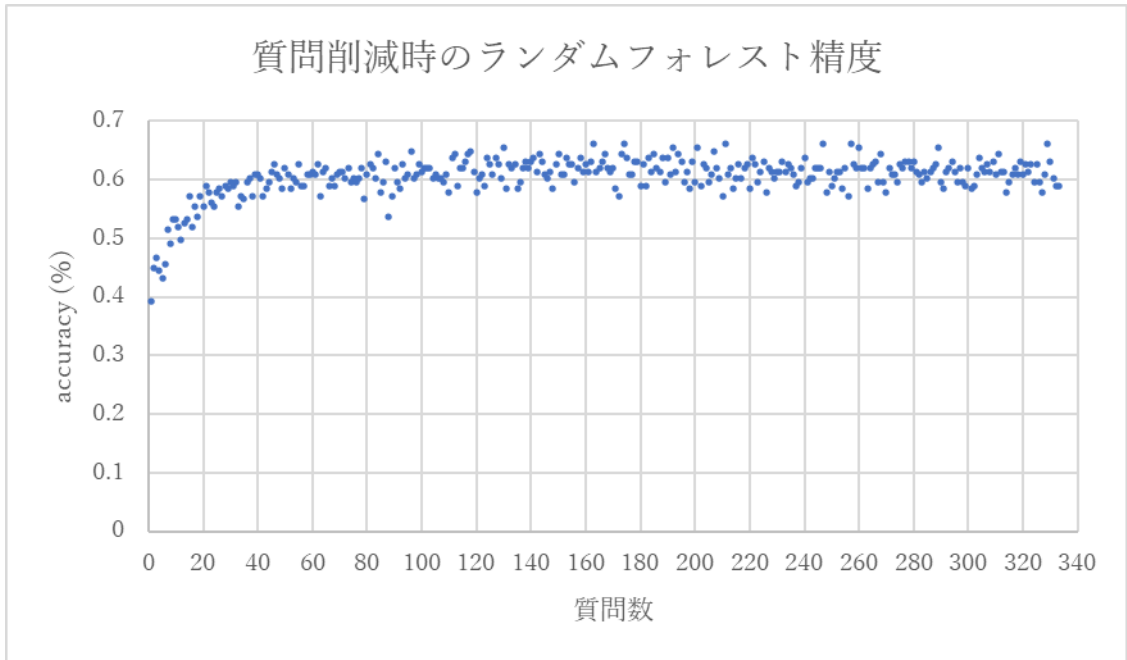


図 8. 332 問の質問から、特徴量が低い質問から 1 問ずつ削減していったランダムフォレストの精度

ID	Question
Q7	過去30日間で、どのくらいの頻度で、神経過敏だと感じましたか。
Q8	過去30日間で、どのくらいの頻度で、絶望的だと感じましたか。
Q10	過去30日間で、どのくらいの頻度で、気分が沈み込んで、何が起ころうとも気が晴れないように感じましたか。
Q11	過去30日間で、どのくらいの頻度で、何をしても面倒だと感じましたか。
Q12	過去30日間で、どのくらいの頻度で、自分は価値のない人間だと感じましたか。
Q40	最近1ヶ月間のあなたの状態について伺います。元気がいっぱいだ。
Q41	最近1ヶ月間のあなたの状態について伺います。生き生きする。
Q45	最近1ヶ月間のあなたの状態について伺います。ひどく疲れた。
Q46	最近1ヶ月間のあなたの状態について伺います。へとへとだ。
Q47	最近1ヶ月間のあなたの状態について伺います。だるい。
Q49	最近1ヶ月間のあなたの状態について伺います。不安だ。
Q50	最近1ヶ月間のあなたの状態について伺います。落ち着かない。
Q51	最近1ヶ月間のあなたの状態について伺います。憂鬱だ。
Q52	最近1ヶ月間のあなたの状態について伺います。何をしても面倒だ。
Q53	最近1ヶ月間のあなたの状態について伺います。物事に集中できない。
Q54	最近1ヶ月間のあなたの状態について伺います。気分が晴れない。
Q56	最近1ヶ月間のあなたの状態について伺います。悲しいと感じる。
Q69	あなたの周りの方々についてうかがいます。職場の同僚とどのくらい気軽に話ができますか。
Q74	あなたの周りの方々についてうかがいます。あなたの個人的な問題を相談したら、上司はどのくらい聞いてくれますか。
Q79	元気がありますか。
Q80	疲れやすいですか。
Q82	動悸（脈が速い）がしますか。
Q87	憂鬱になったり、気が沈む感じがしますか。
Q88	精神が緊張したり、イライラしますか。
Q90	怖がりやすく、不安感がありますか。
Q93	無意識によくため息をつきますか。
Q94	体や手足が重だるく感じますか。
Q104	季節の変わり目、温度の変化、嫌な匂いなどによって息苦しくなることがありますか。
Q108	過敏症ですか。（薬、食べ物、花粉、季節、気候など）
Q119	顔色が暗く、シミがでやすいですか。
Q125	口や喉が乾きますか。
Q133	不眠がちですか。
Q136	便が硬く、便秘しやすいですか。
Q143	私は自分自身のことを、活発で、外向的だと思う。
Q153	性別を選んでください。
Q155	身長を半角でご記入ください（例：178cm）
Q159	配偶者（夫又は妻）の有無(事実上夫婦として生活しているが、婚姻届を提出していない場合も含みます)未婚
Q182	あなたの現在の健康状態はいかがですか。あてはまる番号1つに○をつけてください。
Q210	過去1か月において、物事をやり遂げるのに必要な意欲を持続するうえで、どのくらい問題がありましたか？
Q285	(淡色野菜・きのこ類) 昼、一回に食べる量を選んでください

表 8.5 体質の分類に寄与が大きい上位 40 問

ID	question
Q7	最近1ヶ月間で神経過敏だと感じましたか。
Q8	最近1ヶ月間で絶望的だと感じたことがありましたか
Q10	最近1ヶ月間で気分が沈み込んで、何が起ころしても気が晴れないと感じたことがありましたか。
Q11	最近1ヶ月においてどのくらいの頻度で、何をするのも面倒だと感じましたか。
Q12	最近1ヶ月においてどのくらいの頻度で、自分は価値のない人間だと感じましたか。
Q40	最近1ヶ月間においてどのくらいの頻度で元気がいっばいと感じましたか。
Q41	最近1ヶ月間においてどのくらいの頻度で生き生きすると感じましたか。
Q45	最近1ヶ月間においてどのくらいの頻度でひどく疲れたと感じたことがありましたか。
Q46	最近1ヶ月間でへとへとだと感じたことがありましたか。
Q47	最近1ヶ月間でだるいと感じたことがありましたか。
Q49	最近1ヶ月間で不安と感じた事がありましたか
Q50	最近1ヶ月間においてどのくらいの頻度で落ち着かないことがありましたか。
Q51	最近1ヶ月間で憂鬱と感じたことがありましたか。
Q52	最近1ヶ月間においてどのくらいの頻度で何をするのも面倒だと感じましたか。
Q53	最近1ヶ月間で物事に集中できないと感じたことがありましたか。
Q54	最近1ヶ月間で気分が晴れないことがありましたか。
Q56	最近1ヶ月間においてどのくらいの頻度で悲しいと感じるたことがありましたか。
Q69	あなたは周りの方々とどのくらい気軽に話ができますか。
Q74	あなたの周りの方々はあなたの個人的な問題を相談したら、どのくらいきいてくれますか。
Q79	元気がありますか。
Q80	疲れやすいと感じることがありますかですか。
Q82	動悸がしますか。
Q87	憂鬱と感じたり、気が沈むと感じたりすることがありますか。
Q88	精神が緊張したり、イライラしたりしますか。
Q90	怖がりやすく、不安感がありますか。
Q93	無意識によくため息をつきますか。
Q94	体や手足が重だるく感じますか。
Q104	季節の変わり目、温度の変化、嫌な匂いなどによって息苦しくなることがありますか。
Q108	過敏症ですか。(薬、食物、花粉、季節、光など)
Q119	[119] 顔色が暗く、シミができやすいですか。
Q125	口や喉が乾きますか。
Q133	不眠がちですか。
Q136	便が硬く、便秘しやすいですか。
Q143	私は自分自身のことを、活発で、外向的だと思う。
Q153	性別を選んでください。
Q159	結婚経験がありますか
Q182	現在の健康状態はいかがですか。
Q210	最近1ヶ月において物事をやり遂げるのに必要な意欲の持続が困難だったことがありましたか
Q285	淡色野菜（ダイコン・タマネギ、レンコン・ゴボウ）・きのこ類を昼食一回に食べる量を選んでください。

表 9. 質問票の文言変更

5.3.3 バリデーション

抽出した 39 問について、文言の修正を行ったため、再度クラウドソーシングを用いて回答を収集した。本検討では、851 名の回答者のうち、再同意が得られた方を対象とし、162 名から回答を得ることができた。162 名の 39 問に対する回答を置き換え、再度グループ数 5 となるよう、階層クラスタリングを行った結果、各グループを新 A~E とおくと、人数はそれぞれ新 A:9 人、新 B:38 人、新 C:18 人、新 D:72 人、新 E:25 人となった。

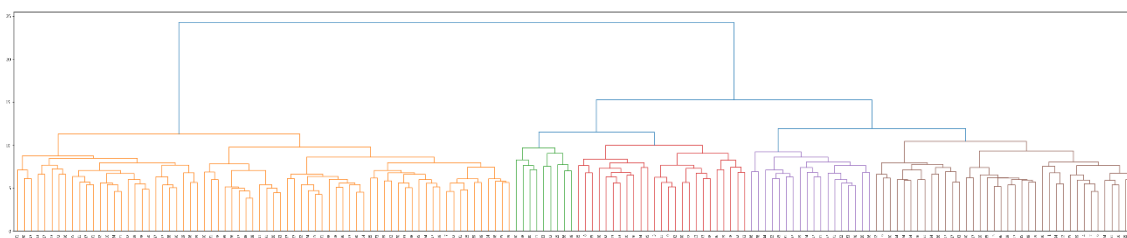


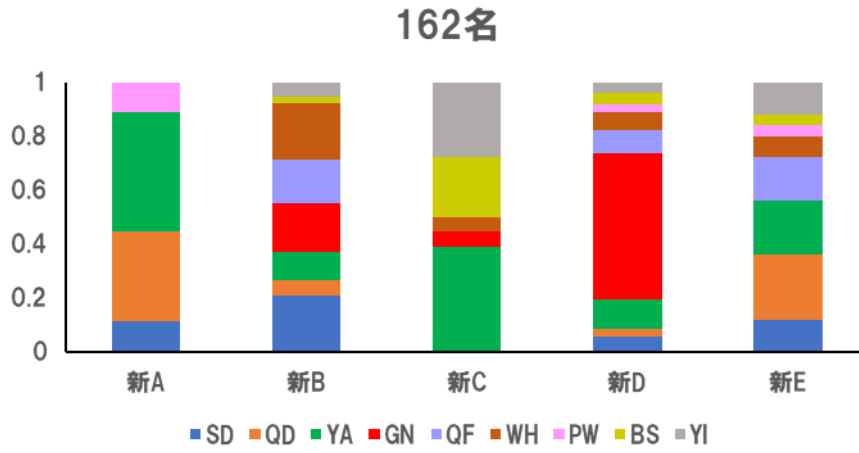
図 9. 162 名を階層的クラスタリングで 5 分類にした際のデンドログラム

5.3.4 中医体質の比較

第 4 章の 4.3.1 において、851 名の 332 問に対する回答を k 平均法により求め、最適体質数が 5 である可能性を見出した。またその際、5 グループと中医体質との関連から、日本人が陽虚質と気鬱質によって特徴づけられることが確認された。本研究において得られた 162 名の中医体質の分布と、5 グループの関連解析を行った。

4.3.1 で得られた 5 グループ KA~KE と、バリデーショndata で得られた新 A~新 E における中医体質の割合およびピアソンの順位相関係数を算出した(図 11、表 10)。それぞれの相関係数はそれぞれ、0.928、0.360、0.679、0.990、0.594 であり、平均は 0.710 となり、高い相関が確認された。以上より、5 分類の再現性が確認されたといえる。

(A)



(B)

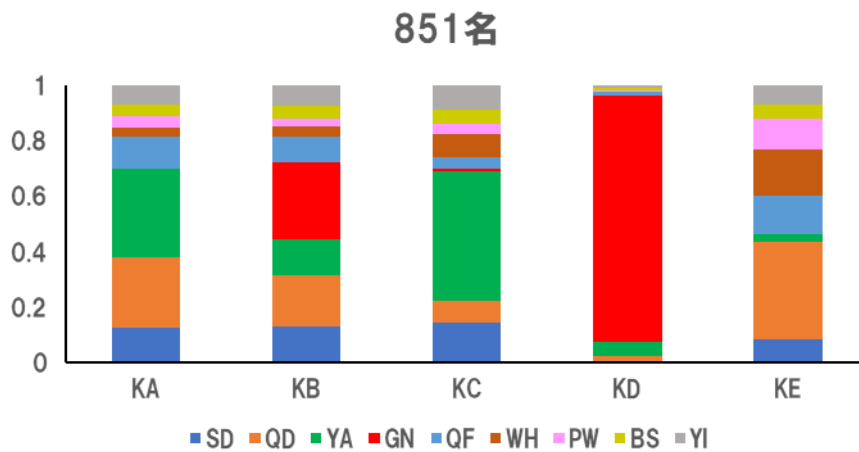


図 10. 5 グループにおける中医体質の相対比

(A)バリデーションで得た 162 名の回答データを階層的クラスタリングにより 5 グループに分類した際の中医体質。

(B)4.3.1 で実施した k 平均法による 851 名の CCMQ-J 60 問への回答データを k=5 となるようにクラスタリングした際の中医体質。

9 体質の略号をカッコ内に示す. 平和質[GN], 気虚質[QF], 陽虚質[YA], 陰虚質[YI], 痰湿質[PW], 湿熱質[WH], 血瘀質[BS], 気鬱質[QD], 特稟質[SD]。

新 A~E はバリデーションで得た 162 名の 18 問に対する回答データを階層的クラスタリングにより分類した 5 グループであり、KA~KE は 4.3.1 で実施した k 平均法による 851 名の CCMQ-J 60 問への回答データを k=5 となるようにクラスタリングした際のグループを示す。

	KA	KB	KC	KD	KE
新 A	0.928	0.249	0.776	-0.209	0.291
新 B	-0.110	0.360	0.042	0.328	-0.146
新 C	0.327	-0.132	0.679	-0.109	-0.492
新 D	-0.310	0.774	-0.160	0.990	-0.453
新 E	0.898	0.093	0.528	-0.499	0.594

表 10. 851 名の CCMQ-J への回答と 162 名の 18 問に対する回答から得られる中医体質の相関係数.

新 A~E はバリデーションで得た 162 名の 18 問に対する回答データを階層的クラスタリングにより分類した 5 グループであり、KA~KE は 4.3.1 で実施した k 平均法による 851 名の CCMQ-J 60 問への回答データを k=5 となるようにクラスタリングした際のグループを示す。

5.3.5 質問票の作成

ランダムフォレストにより抽出した 39 問のうち上位 18 問に対する 162 名の回答から、クラスタリングにより分類された体質を分類する決定木を作成した(図 10)。18 問とした理由は、本研究で開発した質問票を質問紙として観察研究で利用する際に、A4 1 枚で実施できるようにするためである。作成された決定木は図 10(B)のように「Q8: 最近1ヶ月間においてどのくらいの頻度で、絶望的だと感じましたか」に対して、「全くない」、「少しだけ」と回答した人は左に、「ときどき」「たいてい」「いつも」と回答した人は次の質問に進む。この決定木の作成により、一人当たり 4, 5 問の回答で体質判定が可能になる質問票が作成できた。

5.4 考察

5.4.1 ランダムフォレストを用いた質問数の削減

本研究では、第4章により、日本人の体質が5つである可能性を見出した。また、先行研究において、中医体質は、ほかのヘルスケアに関する質問から予測可能であることが確認された。以上より、本章では、クラウドソーシングを用いて取得した851名の332問に対する回答からランダムフォレストを用いて、体質を5分類にする際に重要な上位39問を抽出した。また、特徴量が低い順に質問を減らして言った際に、分類の精度この際、精度は約60%と高くなかったが、質問数を増やしても特徴量が大きく増加しないことから、60%前後がこのデータから作成されるモデルの精度の限界であると判断した。

5.4.2 39問によるバリデーション

これまで解析してきた質問は、既存の質問票(ストレスチェックやピッツバーグ睡眠質問票など)であり、「過去1か月のあなたについて伺います」など、時間が指定されている質問とそうでない質問が混合している状態であった。実際の観察研究やサービスとして質問票を作成していく際に、位相のそろっていない、回答者が答えにくい質問は適切でないため、質問の意図を変えないよう、質問の文言を修正した。そこで、修正した質問でも同様の結果を得られるかを確認するため、再度バリデーション試験を実施した。バリデーション試験では851名の回答者のうち、再同意が得られた162名の方を対象とした。バリデーション試験の結果、851名を5グループに分けた際のCCMQ-Jによる中医体質の分布と、バリデーションで得られた162名のデータを5グループに分けた際の中医体質の分布の相関係数の平均が0.7を超えた。このことから、文言を変更の影響はなかったと考えられ、また39問においても再現良く5グループへ分類がされることが示唆された。さらに、バリデーションデータによる5グループのうち1つは平和質の分布が高い、健康体質であることが確認され、それ以外の体質には陽虚質を含むそのほかの未病体質の割合が多いことが確認された。

第4章では、851名のCCMQ-Jに対する回答を解析し、陽虚質と気鬱質によって日本人の体質は特徴づけられると報告した。バリデーションデータにおいても新A、新C、新Eは陽虚質、気虚質により特徴づけられることが明らかになった。一方で、162

名の解析の場合、新 C は KC と比較し、陽虚質のほかに血瘀質や陰虚質の割合が高いことが確認された。851 名の解析と分布が異なった理由として、人数が少なくなったことが考えられる。今回は同じ人に 2 度回答を得ることで再現性を確認したが、別の集団においても 5 グループに分類された際の中医体質の分布を確認し、同様に分類されるかは検証すべきであると考えられる。

5.4.3 決定木による簡便な質問票の作成

抽出された質問を今後観察研究で利用していく際に、回答者が苦痛を感じない質問数であることが望ましい。そのため、質問紙として利用することを考慮し、39 問のうち上位 18 問を用いることとした。また、18 問を決定木形式とすることで、1 人当たり 4、5 問で体質判定が可能な質問票を作成することが可能であるため、上位 18 問を用いて決定木を作成した。決定木の上位にある質問は「Q8:最近1ヶ月間においてどのくらいの頻度で、絶望的だと感じましたか」、「Q7:最近1ヶ月間においてどのくらいの頻度で、神経過敏だと感じましたか」、「Q87:憂鬱と感じたり、気が沈むと感じたりすることがありますか」、であり、精神の健康を問う質問であった。第 4 章にて、日本人はこころの病気にかかる人数が増えており、実際に気鬱質によって特徴づけられる体質もあったことから、精神的因子を問う要素により体質が分類されることはこれまでの結果と相違ないといえる。実際に、最初の質問に「Q8:最近1ヶ月間においてどのくらいの頻度で、絶望的だと感じましたか」があるため、Q8 に「全くない」、「少しだけ」と回答した人は新 D:健康体質(平和質)、「ときどき」「たいてい」「いつも」と回答した人は他の未病体質グループに分類される可能性が高いことが確認できた。

本研究では、332 問の質問を 18 問に圧縮し、簡便な質問票を作成することができた。今後健康な人を対象とした観察研究などで利用し、日本人の未病体質の把握、ひいては体質と食品摂取の関連性などを検証していくことで、日本人の健康維持増進に貢献できることが期待される。

5.5 課題

本研究によって、日本人の体質を 5 グループに分類するために CCMQ-J を含む 332 問の質問をランダムフォレストにより 18 問に圧縮することができた。また、質問を

回答しやすい文言に変更し、再同意を得た方に 18 問の質問を再度回答してもらった結果、再現性を確認することができた。しかし、別の母集団での検証や、今後も CCMQ-J と関連性は確認していく必要がある。

第6章 総括

6.1 第4章の日本人の体質分類数に関するまとめ

本研究の被験者は 851 名であるものの、クラウドソーシングという調査手法を活用することで、日本人を対象とした体質グループの統計的推定を実施できる可能性を示すことができた。また、2 種類の手法により設定された 5 グループは「境界型(グループ A)」、「境界型(グループ B)」、「陽虚質型(グループ C)」、「平和質型(グループ D)」、「気鬱質型(グループ E)」と定義することができた。また、朱燕波らの著書によると陽虚質は冷え性タイプ、気鬱質は情緒不安定な鬱タイプであり、現代の日本人にうつ病患者が増加傾向にあることと一致している。現代社会に直面する心身、精神にかかわる未病が特徴づけられた 5 グループは超高齢化社会となった日本国民の健康維持増進を目指すうえで、考慮すべき体質であると考えられる。

6.2 日本人の体質を簡便に判定する体質質問表開発に関するまとめ

第4章により、日本人の体質を 5 分類する簡便な質問票を作成するために、CCMQ-J を含む 332 問についてランダムフォレストを用いて、332 問の質問から特徴量の高い 39 問を選定することができた。また、39 問について答えやすい文章に変更し、再度 162 名の方に回答いただきバリデーションを行った結果、再現性を確認することができた。さらに上位 18 問を用いて決定木を作成し、1 人当たり 4-5 問で体質を判定できる質問票を作成することができた。

6.3 今後の展望

第4章において、日本人の体質が 5 分類であることを見出し、第5章では日本人の体質を 5 つに分類することができる簡易的な質問票を作成した。本研究では、今まで中医体質は 9 分類であったのに対し、日本人の体質が 5 体質であることが示唆された。日本人の体質が 5 分類であることのメリットとしては、少ない体質となることで、その人の体質に合った、より効果的な健康食品の提供が可能になることや、判定される側へのわかりやすさが挙げられる。一方で、5 体質に分類されることのデメリットとしては、

今後の社会情勢や生活習慣の変化によって 5 体質に当てはまらない人がでてきた場合に、その人に対して効果的な解決策を提供できないことが挙げられる。特に近年はコロナウイルスの影響により、日本においても在宅勤務の増加や、外食の機会減少など生活習慣は大きく変化した。特に在宅勤務の増加は体の不調に大きく影響している³⁵。しかし、コロナウイルスの収束後も企業の約 6 割が在宅勤務を続けると回答しており³⁶、今後も在宅勤務による不調はますます顕在化していくと考えられる。そのため、コロナ収束後においても本研究によって見出された 5 体質が適当であるかは確認していくべきである。

本研究においては 300 問以上の質問の回答収集にクラウドソーシングを用いた。来所を伴う観察研究の場合、会場準備や日程調整、運営スタッフの手配や質問票の回答収集に時間がかかる。また、来所に伴う交通費を含む数千円～数万円の謝礼が必要になる。しかしクラウドソーシングの場合、来所に伴う準備は不要であり、また、一人当たり数十円～数百円の謝礼で回答収集が可能である。また、来所を伴う観察研究の場合は、会場に来所可能な方のみしか参加できないが、クラウドソーシングの場合、日本全国の方が参加可能であるため、地域によるばらつきを抑えることが可能となる。本研究において、クラウドソーシングは多くの質問に対する大人数の回答収集に有用であり、今後の e-Health の基盤となると考えられる。特に最近ではコロナウイルスの影響により、1 か所に同時に集まることが推奨されないため、積極的に使用していくべきだと考える。

謝辞

本研究を進めるに当たり、主指導教員である奈良先端科学技術大学院大学 先端科学技術研究科 計算システムズ生物学研究室 教授 金谷重彦先生に終始懇切丁寧な御指導および御鞭撻を賜りました。厚く御礼を申し上げます。また、研究発表及び博士論文作成に当たり、貴重な御助言を賜りました副指導教員である奈良先端科学技術大学院大学教授 松本健一先生、准教授 小野直亮先生、准教授 MD.ALTAf-UL-AMIN 先生、助教 黄銘先生に厚く御礼を申し上げます。最後に、本研究にご協力して下さった対象者の皆様、データ集計や研究のご指導をして下さった奈良先端科学技術大学院大学 教授 中村哲先生、助教 田中宏季先生、中医体質に関して様々な御助言を戴きました金沢大学大学院 教授 鈴木信孝先生、助教 許鳳浩先生、共同研究者であるサントリーグローバルイノベーションセンター株式会社 村山宣人氏、渡辺齊志氏に心より御礼を申し上げます。

研究業績

査読付き論文

1. 佐藤真梨子、川崎聡大, 黄銘, 許鳳浩, 小野直亮, 江口遼平, Md. Altaf-Ul-Amin, 徳田(角谷)佐紀, 渡辺斉志, 村山宣人, 中村哲, 山口栞, 田中 宏季, 金谷重彦, 朱燕波, 戴 昭宇, 王琦, 鈴木信孝,. クラウドソーシングを活用した日本全体の CCMQ-J 問診票調査における最適体質数とその体質特性, 日本補完代替医療学会誌, 16, 2, 105-112, 2019, 第4章

査読付き国際学会発表

1. Mariko Sato, Determination of Optimum Number of Groups on Japanese participants recruited from crowdsourcing Interpreted by Physical Constitution Defined by CCMQ-J, 24th World Congress on Nutrition and Food Sciences, 2021, 第4章

参考文献

1. 統計トピクス No.126 統計からみた我が国の高齢者―「敬老の日」にちなんで―
総務省 報道資料 (<https://www.stat.go.jp/data/topics/pdf/topics126.pdf>)
2. 令和元年の簡易生命表の概要 厚生労働省
(<https://www.mhlw.go.jp/toukei/saikin/hw/life/life19/index.html>)
3. 健康寿命の延伸・健康格差の縮小 第11回健康日本21(第2次)推進専門委員会
(https://www.mhlw.go.jp/file/05-Shingikai-10601000-Daijinkanboukouseikagakuka-Kouseikagakuka/0000166297_5.pdf)
4. Yanbo Z, Origasa H, Uebaba K, et al. Development and Validation of the Japanese version of the constitution in Chinese Medicine Questionnaire (CCMQ). *Kampo Med.* 2008; 59: 783-792.
5. 朱 燕波, 折笠 秀樹, 上馬場 和夫ら. 体質に関する CCMQ 質問票日本語版の臨床開発とその妥当性の検証. 2008 年 59 巻 6 号 p. 783-792
6. 朱燕波, 許鳳浩, 宋函ら. 健康体質づくり-スマートライフの実現に向けて. 金沢. 未病体質研究会. 2014.
7. 許鳳浩, 上馬場和夫, 朱燕波ら. 中医体質理論を用いた黒部地域住民の健康関連情報のデータ解析. *日本補完代替医療学会誌.* 2020; 17(2):145-153.
8. 濱舘直史, 許鳳浩, 松本祥幸ら. 黒酢含有食品による血圧, QOL, 体質への効果. *日本補完代替医療学会誌.* 2014; 11: 95-102.
9. 許鳳浩, 上馬場和夫, 阿部哲朗ら. 乳酸菌飲料 ANP71 の排便状況に及ぼす影響―オープン臨床試験. *日本補完代替医療学会誌.* 2017; 14: 39-44.
10. 許鳳浩, 大野木宏, 橋本慎太郎ら. 中医体質九分類による明日葉青汁の評価. 第21回日本補完代替医療学会 学術集会 一般口演.
11. 濱舘直史, 許鳳浩, 松本祥幸ら. 黒酢含有食品による血圧, QOL, 体質への効果. *日本補完代替医療学会誌.* 2014; 11: 95-102

12. 山森明弘, 許鳳浩, 渡辺知幸ら. 多変量解析法にもとづいた日本語版中医体質調査票(CCMQ-J)における各質問間相関の評価ならびに BMI, 年齢の推定モデルの構築. 日本補完代替医療学会誌.2016; 13; 43-56
13. Peter Naur. Studentlitteratur, Lund, Sweden, 1974, Concise Survey of Computer Methods, 397 p.
14. Tibeshirani R, Walther G, Hastie T. Estimating the number of clusters in a data set via the gap statistic. J.R.Statist.Soc.B.2001;63, 411-423.
15. Rees L. The physical constitution and mental illness. Eugean Rev. 1947;39:50-55.
16. 平成 30 年版厚生労働白書－障害や病気などとの向き合い、全ての人が活躍できる社会に (<https://www.mhlw.go.jp/stf/wp/hakusyo/kousei/18/backdata/01-01-02-09.html>)
17. 江崎グリコ株式会社 プレスリリース.
(https://www.glico.com/assets/files/%E3%80%90%E3%82%B0%E3%83%AA%E3%82%B3%E3%80%91%E3%80%8C%E5%86%B7%E3%81%88%E6%80%A7%E5%AF%BE%E7%AD%96%E3%80%8D%E8%AA%BF%E6%9F%BBPR_1.pdf)
18. 時光, 許鳳浩, 川崎聡大, 金谷重彦ら. ヘルスケアにかかわる 7 種の質問票による CCMQ-J の体質の推定と解釈法, 日本補完代替医療学会誌, 16, 2, 79-93, 2019
19. 回答所要時間/質問数が回答完遂率に与える影響 (<https://quest-research.co.jp/column/2020-1201>)
20. Furukawa TA, Kawakami N, Saitoh M, et al. The performance of the Japanese version of the K6 and K10 in the World Mental Health Survey Japan. Int J Methods Psychiatr Res. 2008;17(3):152-8.
21. Dadfar M, Vahid MKA, Lester D et al. Kessler Psychological Distress Scale (K6): Psychometric testing of the Farsi form in psychiatric outpatients. Advances in Bioresearch. 2016; 7(2):105-108.
22. Sugishita M, Hemmi I, Takeuchi T. Reexamination of the Validity and Reliability of the Japanese Version of the Mini-Mental State Examination (MMSE-J). 2016; 18; 168-183.

23. Ota A. Scientific base for the Japanese Stress Check Program. J Occup Health. 2018 Jan 25;60(1):1-2.
24. Inoue A, Kawakami N, Shimomitsu T, et al. Development of a short version of the new brief job stress questionnaire. Ind Health. 2014;52(6):535-40.
25. Iwasa H, Yoshida Y. Psychometric evaluation of the Japanese version of Ten-Item Personality Inventory (TIPI-J) among middle-aged, and elderly adults: concurrent validity, internal consistency and test-retest reliability. Cogent Psychology. 2018; 5: 1426254.1-10
26. 小塩 真司, 阿部 晋吾, Pino Cutrone. 日本語版 Ten Item Personality Inventory (TIPI-J) 作成の試み. パーソナリティ研究 2012; 21(1), 40-52.
27. Doi Y, Minowa M, Uchiyama M, et al. Psychometric assessment of subjective sleep quality using the Japanese version of the Pittsburgh Sleep Quality Index (PSQI-J) in psychiatric disordered and control subjects. Psychiatry Res. 2000; 97(2-3):165-72.
28. 国民生活基礎調査(平成 28-30 年)
<https://www.mhlw.go.jp/toukei/list/20-21.html>
29. 国民健康・栄養調査(平成 25-29 年)
https://www.mhlw.go.jp/bunya/kenkou/kenkou_eiyou_chousa.html
30. Nakatsu N, Sawa R, Misu S, et al. Reliability and validity of the Japanese version of the simplified nutritional appetite questionnaire in community-dwelling older adults. Geriatr Gerontol Int. 2015;15(12):1264-9.
31. Yoshimura Y, Takahashi K. エクセル栄養君食物摂取頻度調査 FFQg (建帛社).
32. Breiman L. Radom Forests. Machine Learning volume 45, pages5–32 (2001)
33. [平成 29 年所得再分配調査報告書. 厚生労働省. 29hou.pdf \(mhlw.go.jp\)](#)
34. Liaw A, et al. Classification and Regression by RandomForest.
35. ニチバン株式会社 プレスリリース. (https://www.nichiban.co.jp/news/2021/07_01/)
36. 日本経済新聞.2021/3/23.
(<https://www.nikkei.com/article/DGXZQODZ224U40S1A320C2000000/>)