

令和 6 年 6 月 2 日現在

機関番号：32620

研究種目：基盤研究(C)（一般）

研究期間：2018～2023

課題番号：18K12112

研究課題名（和文）分離分析のための成分間の重なりやノイズに強い新クラスター分析法の開発と実用化

研究課題名（英文）Development of a new clustering technique for biological component analyses

研究代表者

佐藤 正一（Sato, Shouichi）

順天堂大学・医療科学部・教授

研究者番号：90803255

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）：分離分析技術において、複数成分が重なり合うデータの分離問題を解決するために、新しいクラスター分析法を開発した。この技術は特に、白血球の分類での細胞成分間の重なりや異常細胞の出現、ノイズの混入といった問題を克服するために有効である。具体的には、画像処理を用いて細胞の種類とその分布を正確に推定し、新たに考案した自己分配方式によって、複数成分の重なりを多重正規分布理論や反復切断補正法を用いて数的に分離することによりクリアな分析結果を得ることが可能である。この技術をさらに発展させ、各種の混合分布の分離に適用可能なアルゴリズムへと改良を進め、各種の学会において報告を行った。

研究成果の学術的意義や社会的意義

本研究では、分離分析技術で問題となる成分の重なりを解決する手法として、画像処理、自己分配方式、反復切断補正法を組み合わせた新しいクラスター分析法を開発した。特に、白血球分類において、細胞成分の正確な分離とノイズ除去を行うことが可能になり、複雑な臨床データに対応する分析精度が向上することを確認した。この技術は、臨床検査の精度向上に寄与し、様々な分離分析場面に対応できるため、医療現場での診断の信頼性を高める社会的意義があると考えられる。

研究成果の概要（英文）：A new cluster analysis method has been developed to solve the problem of separating data in which multiple components overlap in separation analysis technology. This technology is particularly effective for overcoming problems such as overlapping between cell components, the appearance of abnormal cells, and the introduction of noise in the classification of white blood cells. Specifically, by using image processing to accurately estimate the types and distribution of cells, and by mathematically separating the overlapping of multiple components using multiple normal distribution theory and iterative truncation correction, it is possible to obtain clear analytical results. We have further developed this technology and improved it into an algorithm that can be applied to the separation of various mixed distributions, and have reported on it at various academic conferences.

研究分野：統計学

キーワード：クロマトグラフィー フローサイトメトリー法 クラスター解析 正規分布 シミュレーションプログラム

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属します。

1. 研究開始当初の背景

臨床検査では、高速液体クロマトグラフ (HPLC) やフローサイトメトリー (FC) を用いて分離分析が行われている。HPLC は、血液や尿中の特定の化合物を分析するために使用され、例えば HbA1c の測定では、通常のヘモグロビンと糖化ヘモグロビンを分離して定量化したり、その他ホルモンの分析などに使用されている。また、FC は細胞の種類や特性を詳細に分析するために使用される。しかし、HPLC では成分の融合部分を単純な谷値を境界値として処理しており、成分のバランスを考慮せずに、境界値のみで集団を分離する方法が取られてきたため、厳密なデータの取り扱いが不十分であった。また FC においても近接する細胞集団の適切な区分が困難な状況にある。これを解決するためには、成分のバランスや分布領域を考慮した適切な閾値を設定し、成分の重なりを計算する新たなアルゴリズムの確立が求められている。

2. 研究の目的

本研究の目的は、臨床検査における分離分析技術の精度向上を図ることである。特に、HPLC などのクロマトグラフィー (CG) や FC で発生する成分の重なりやノイズによる問題を解決するために統計手法として新たなクラスター分析法を開発することである。これまで FC の細胞成分分離の問題に対しては、(1) 実在する細胞成分数とその分布域を画像処理方式 (IP: image processing) での確に推定し、(2) 新考案の自己分配方式 (SP: self-partitioning algorithm) で、複数成分の重なりを多重正規分布理論で数理的に分離し、(3) 反復切断補正法 (ITC: iterative-truncation & correction) でノイズ処理を行う、斬新なクラスター分析法 (IP-SP-ITC: ISI 法) を開発し、白血球自動分類データに適用しその実用性を検討している。本研究ではより精度を高め、CG やその他の分離分析技術に応用し、臨床検査の信頼性を高めることを目指すものである。

3. 研究の方法

(1) 従来の統計学的分離分析手法に対する ISI 法の有効性について詳細に検討する。これまで、複数の成分の融合部分を区分する統計手法では、クラスタリングや混合モデルのパラメータ推定など、さまざまな統計モデリングに広く使用されている期待値最大化法 (Expectation-Maximization (EM) Algorithm) が有望視されてきたが、異常成分の出現、ノイズ成分の混入、成分の範囲の大小によって適切な分離が困難であることをシミュレーション技術を使い検証する。そしてより厳密な手法として、新しいクラスター分析法である ISI 法の有用性を証明する。

(2) HPLC より高精度の質量分析装置の実データを用い厳格な定量測定を実施する。具体的には、分析上近い位置に分離される成分について既知である成分量の比率を変えることに

よる EM Algorithm の不正確性と ISI 法による精確性を検証する。

(3) 多様化する分離分析技術への応用について研究を行う。最新のクラスター解析技法の多くは、画像処理技術にまで及んでいる。現在注目されている人工知能(AI)による判断技術に対する統計手法からのアプローチとして ISI 法の有効性について詳細に検討する。

4. 研究成果

(1) 2020 年に行われた第 67 回日本臨床検査医学会学術集会にて「HPLC における混合分布に対する正確な成分定量法の開発」として ISI 法を応用した HPLC の新しい定量法を紹介した。通常 HPLC の成分の分布は、正規分布に近似しているもののわずかに歪んだ分布形状を示す。このひずみに対して Box-Cox 法を利用して極力正規分布になるようにデータの補正を行い ISI 法を適応する解析方法(新 ISI 法)を実行すると、分析濃度比を 4 : 1 のアンバランス状態では、一般的に実施されている谷値法で最大 - 50.7%の誤差が生じたのに対して、統計学的定量法の ISI 法では 5.2%まで減少したことを報告した(図 1)。

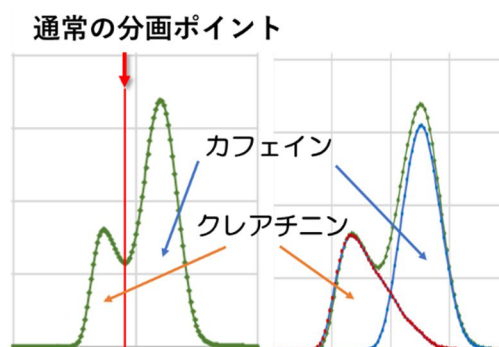


図1 クレアチニンとカフェイン(1:4 溶)のHPLC分析結果
一般的に行われる谷値法では、クレアチニンが-50.7%の過小評価となった。一方、統計学的手法を用いたISI法では誤差は5.2%まで減少した。

(2) 2020 年第 60 回日本臨床化学会年次学術集会において、機械学習による混合分布の分離方法として、「メタロバランス法を用いたクラスター分析から見るがんリスクスクリーニング検査法の有用性」を発表した。メタロバランス検査とは、血清中微量元素濃度のバランスをがんと診断されている人と、がんと診断されていない人を比較し類似度合により「がんリスク」を評価するものである。本発表においても、正規分布に変換することの重要性について提示した。

(3) 2021 年においては、HPLC の分布型検出技術を用いて、2 次元データである白血球分類で使用するフローサイトメトリー (FC) に適応した。FC の結果として表示されるスキャッタグラムから白血球中の好中球についてその分布型を検討したところ、HPLC と同様、Box-Cox 法による変数変換実施後 t-分布になることを確認した。また、白血球のスキャッタグラムは採血後の経過時間の影響で検査結果が変動することや感染時に細菌などから産生される物質によって形態変化が起こる問題があるため、人工的に形態変化を起こして分布型への影響について調査を行った。

本研究内容の一部を「次元圧縮とクラスター解析による可視化と臨床応用・臨床化学 2022 ; 51(1) : 12-20 .」、 「EBLM(evidence based laboratory medicine)の新展開クラスター分

析の臨床検査への応用 . 臨床検査 2020 ; 64(5) : 600-608 .」および「臨床検査における AI の概略-AI が臨床検査でできること . 日本臨床検査医学会誌 2021 ; 69(5) : 344-348 .」に発表した。

(4) 2023 年度の医学検査学会の生物化学分析部門シンポジウムにて、測定法の根本理論として「測定の不確かさの理論と活用」を発表した。臨床検査では、測定結果の信頼性について評価・表現する方法として「測定の不確かさ」を用いるようになっている。従来、測定値の信頼性は真度または正確さや精密さで表現してきたが、真の値は無限回の測定によってしか得られず、実際に求めることはできないという考え方から始まったものである。しかし現状では不確かさについて臨床分野での十分理解が進んでいない。今回の発表によって計測データの許容される誤差についての理解が進むものと考えられる。

(5) 2023 年度に台湾の台中山医科大学で開催された 14th Asia-Pacific Forum of Medical Laboratory Science においてシンポジストとして「臨床検査の AI の現状から実践に向けて」と題して、これまでの AI に関する様々な実験から AI の特性やクラスター解析をはじめとする分離分析技術について発表した。内容は、AI による数値的なデータ処理においては、従来の統計学的な処理に比べ僅かに有意な面があることを報告した。しかし、元データに対しては正規分布になるように新 ISI と同様な処理を実施後に分析をすることでより適切なデータ解析に繋がることを示した。また、画像処理においては現在最も注目されている分野であり、高精度に判別する能力があり、臨床分野にも応用可能であることを示した。

(6) 2023 年度に共同研究者である市原清志先生、山下哲平先生との共著として Box-Cox による正規分布変換が基準値の設定に有用であることを示した論文を発表した。あらゆるデータは、必ずしも正規分布ではなく、歪が生じている。パラメトリック（正規分布を仮定したデータ処理法）の下で歪のあるデータ処理を実行することは、統計処理上の問題を生じる可能性がある。このような問題に対して、適切なデータ変換処理法について解説と実行可能なプログラムの提示および実証結果について示した論文である。一例として中性脂肪のデータ分布について Box-Cox 変換を行って正規分布に変換された例を示す（図 2）。

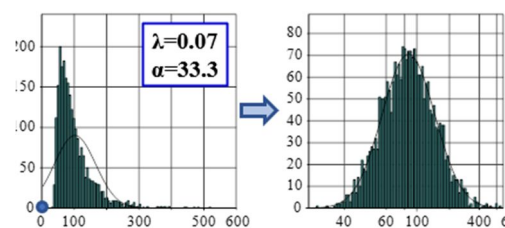


図2 健常者の中性脂肪のデータ分布を Box-Cox 変換を使って正規化した図

<引用文献>

Sato S, Ichihara K, Inaba T. A new self-partition clustering method for robust identification of subsets with heterogeneous size and density and its clinical application to leukocyte differential counting. Clin Chim Acta 445: 118-27, 2016. (査読有)

5. 主な発表論文等

〔雑誌論文〕 計12件（うち査読付論文 9件 / うち国際共著 1件 / うちオープンアクセス 5件）

1. 著者名 Ikeda Katsuhide, Sakabe Nanako, Maruyama Sayumi, Ito Chihiro, Shimoyama Yuka, Sato Shouichi, Nagata Kohzo	4. 巻 66
2. 論文標題 Relationship between Liquid-Based Cytology Preservative Solutions and Artificial Intelligence: Liquid-Based Cytology Specimen Cell Detection Using YOLOv5 Deep Convolutional Neural Network	5. 発行年 2022年
3. 雑誌名 Acta Cytologica	6. 最初と最後の頁 542 ~ 550
掲載論文のDOI (デジタルオブジェクト識別子) 10.1159/000526098	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 Maruyama Sayumi, Sakabe Nanako, Ito Chihiro, Shimoyama Yuka, Sato Shouichi, Ikeda Katsuhide	4. 巻 178
2. 論文標題 Effect of Specimen Processing Technique on Cell Detection and Classification by Artificial Intelligence	5. 発行年 2023年
3. 雑誌名 American Journal of Clinical Pathology	6. 最初と最後の頁 1 ~ 7
掲載論文のDOI (デジタルオブジェクト識別子) 10.1093/ajcp/aqac178	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 Ichikawa Akito, Neo Sakurako, Nukui Ryouhei, Yamada Yoko, Nitta Suguru, Iwaki Hidetoshi, Yanagi Yusuke, Nakayama Koichi, Sato Shoichi, Tateishi Satoko, Hisasue Masaharu	4. 巻 19
2. 論文標題 Establishment of large canine hepatocyte spheroids by mixing vascular endothelial cells and canine adipose-derived mesenchymal stem cells	5. 発行年 2022年
3. 雑誌名 Regenerative Therapy	6. 最初と最後の頁 1 ~ 8
掲載論文のDOI (デジタルオブジェクト識別子) 10.1016/j.reth.2021.11.007	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 Takeuchi Hiroaki, Yoshikane Yu, Takenaka Hirotsugu, Kimura Asako, Islam Jahirul Md., Matsuda Reimi, Okamoto Aoi, Hashimoto Yusuke, Yano Rie, Yamaguchi Koichi, Sato Shouichi, Ishizuka Satoshi	4. 巻 14
2. 論文標題 Health Effects of Drinking Water Produced from Deep Sea Water: A Randomized Double-Blind Controlled Trial	5. 発行年 2022年
3. 雑誌名 Nutrients	6. 最初と最後の頁 581 ~ 581
掲載論文のDOI (デジタルオブジェクト識別子) 10.3390/nu14030581	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 佐藤正一	4. 巻 51
2. 論文標題 次元圧縮とクラスター解析による可視化と臨床応用	5. 発行年 2022年
3. 雑誌名 臨床化学	6. 最初と最後の頁 12-20
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 -

1. 著者名 佐藤正一	4. 巻 69
2. 論文標題 臨床検査におけるAIの概略-AIが臨床検査でできること	5. 発行年 2021年
3. 雑誌名 日本臨床検査医学会誌	6. 最初と最後の頁 44-348
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 -

1. 著者名 佐藤正一	4. 巻 Vol164(5)
2. 論文標題 クラスター分析の臨床検査への応用	5. 発行年 2020年
3. 雑誌名 臨床検査	6. 最初と最後の頁 600-608
掲載論文のDOI (デジタルオブジェクト識別子) 10.11477/mf.1542202370	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 佐藤正一, 市原清志	4. 巻 Vol.57(6)
2. 論文標題 なぜ正常範囲でなく基準範囲なのか “異常” は本当に異常なのか？	5. 発行年 2020年
3. 雑誌名 Medicina	6. 最初と最後の頁 827-831
掲載論文のDOI (デジタルオブジェクト識別子) 10.11477/mf.1402226977	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 Nobuyasu Y, Yuna K, Misaki O, Wataru O, Shoichi S, Mirai Y, Takehiro N	4. 巻 7(1)
2. 論文標題 Differences in cytokine expression after a stimulation with Escherichia coli and Porphyromonas gingivalis or lipopolysaccharides from these bacteria	5. 発行年 2020年
3. 雑誌名 Journal of Life Sciences Research	6. 最初と最後の頁 13 - 20
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 -

1. 著者名 Hiroaki Takeuchi 1,2,* , Keiro Higuchi 3, Yu Yoshikane 4, Ryo Takagi 2, Shinji Tokuhito 2, Koichi Takenaka 5, Wataru Oboshi 1, Asako Kimura 1, Jahirul Md. Islam 1, Ayami Kaneko 1, Shouichi Sato 1 and Satoshi Ishizuka	4. 巻 12
2. 論文標題 Drinking Refined Deep-Sea Water Improves the Gut Ecosystem with Beneficial Effects on Intestinal Health in Humans: A Randomized Double-Blind Controlled Trial	5. 発行年 2020年
3. 雑誌名 Nutrients	6. 最初と最後の頁 26-46
掲載論文のDOI (デジタルオブジェクト識別子) 10.3390/nu12092646	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 Ichihara Kiyoshi, Yamashita Teppei, Kataoka Hiromi, Sato Shoichi	4. 巻 242
2. 論文標題 Critical appraisal of two Box-Cox formulae for their utility in determining reference intervals by realistic simulation and extensive real-world data analyses	5. 発行年 2023年
3. 雑誌名 Computer Methods and Programs in Biomedicine	6. 最初と最後の頁 107820 ~ 107820
掲載論文のDOI (デジタルオブジェクト識別子) 10.1016/j.cmpb.2023.107820	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 佐藤 正一	4. 巻 67
2. 論文標題 AI・ビッグデータ時代の臨床検査のための情報科学・5 - 検査のためのAI入門-深層学習による画像解析	5. 発行年 2023年
3. 雑誌名 臨床検査	6. 最初と最後の頁 1392 ~ 1397
掲載論文のDOI (デジタルオブジェクト識別子) 10.11477/mf.1542203482	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5件（うち招待講演 3件 / うち国際学会 1件）

1. 発表者名 佐藤正一
2. 発表標題 方法間比較のための回帰直線・偏差図の問題点と使い分け
3. 学会等名 日本臨床検査医学会学術集会（招待講演）
4. 発表年 2022年

1. 発表者名 佐藤正一
2. 発表標題 The current state of Artificial Intelligence (AI) in clinical laboratories towards its practice
3. 学会等名 The Application of Artificial Intelligence in Laboratory Medicine（招待講演）（国際学会）
4. 発表年 2023年

1. 発表者名 佐藤正一
2. 発表標題 クラスター解析と次元圧縮事例
3. 学会等名 第67回日本臨床検査医学会学術集会（招待講演）
4. 発表年 2021年

1. 発表者名 佐藤正一
2. 発表標題 HPLCにおける混合分布に対する正確な成分定量法の開発
3. 学会等名 第67回日本臨床検査医学会学術集会
4. 発表年 2020年

1. 発表者名 佐藤正一
2. 発表標題 メタロバランス法を用いたクラスター分析から見るがんリスクスクリーニング検査法の有用性
3. 学会等名 第60回日本臨床化学会年次学術集会
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	市原 清志 (Ichihara Kiyoshi) (10144495)	山口大学・大学院医学系研究科・学術研究員(寄附金) (15501)	
研究分担者	山下 哲平 (Yamashita Teppei) (50617420)	東海大学・医学部・助教 (32644)	

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