

令和 2 年 6 月 4 日現在

機関番号：14603

研究種目：基盤研究(B)（一般）

研究期間：2017～2019

課題番号：17H01817

研究課題名（和文）呼気ガスセンシングによる病状診断と予測アルゴリズム開発

研究課題名（英文）Development of cancer diagnosis and prediction algorithm using exhaled gas

研究代表者

作村 諭一（Sakumura, Yuichi）

奈良先端科学技術大学院大学・先端科学技術研究科・准教授

研究者番号：50324968

交付決定額（研究期間全体）：（直接経費） 12,700,000 円

研究成果の概要（和文）：呼気ガスの成分を用いた機械学習により、健常者と複数の疾病の患者の診断を精度良く行うことができた。それぞれの疾病特有の呼気成分を抽出した。これらは新規バイオマーカーの候補となりうる。簡易検知器開発のため、質量分析器のデータとの診断性能比較を行ったところ、同程度の性能が出せることが分かった。これは簡易検知器の検出性能が質量分析器に劣らないことを示す。性能自体は偽陽性や偽陰性の改善が必要である。

研究成果の学術的意義や社会的意義

本研究課題の手法が確立されれば、様々な非侵襲データを統合することで、高い精度の健康診断が簡素に行うことが可能となる。その簡便性から医療現場の労力を大幅に削減できるだけでなく、被検者の観点では経済的・身体的・心的・時間的負担が大幅に軽減される。疾患・健常レベルを評価することは、疾病を速く（fast）診断するだけでなく、早い（early）段階で疾患可能性を潰す予防医学につながることもできる。予防医学が進めば我が国の医療費の削減に寄与するだけでなく、新たな産業の創出につながる可能性を秘めている。

研究成果の概要（英文）：Machine learning was applied to the components of exhaled gas to accurately diagnose healthy subjects and patients with multiple diseases. Each disease-specific exhalation component was extracted. These components may be candidates for novel biomarkers for each disease. In order to develop a simple gas component detector, we compared its diagnostic performance with that of gas chromatography, and found that it could provide the same level of performance. This indicates that the performance of the detection values by the simple detector is comparable to that of gas chromatography. The performance itself needs to be improved with high false positives and false negatives.

研究分野：計算生物学

キーワード：呼気ガス 機械学習 がん診断

様式 C - 19、F - 19 - 1、Z - 19 (共通)

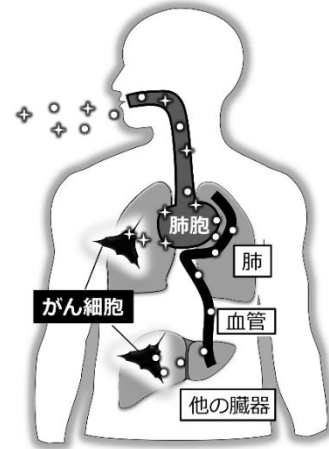
1. 研究開始当初の背景

健康維持のために現在の状態を把握することが重要である。しかし、血液検査等の侵襲的な医療行為は、特に医療診断を必要とする高齢者にとって、経済的・身体的・心的負担が大きい。非侵襲で精度の高い診断ができれば、高齢者のみならず、全ての年齢層にメリットがある。

近年、非侵襲センシングの研究が盛んである。例えば、光学技術を応用した血液検査や唾液検査、そして呼気検査) などがある。これらの中で、呼気ガスは体液の状態を大きく反映するため、病状に関する情報を多く含んでいる可能性が高い(右図)。実際、訓練されたがん探知犬ががん患者の呼気を嗅ぎ分けるという報告がある。

しかし、実用性を考えた場合、一般的な成分検出器であるガスクロマトグラフィー質量分析装置はポータブル性に欠け、診断に時と場所の制約が付く。そのため、比較的簡素な検出器の開発とその検出値で診断できるソフト開発が必要である。これまで、小型検出器と主成分分析を用いた診断法が提案されているが、精度等で問題が残る

これまで研究代表者は、細胞の定量データに対しベイズ統計による機械学習法を適用させ、データに含まれた知見抽出を行ってきた。また、愛知県・知の拠点重点研究プロジェクトの「超早期診断技術開発プロジェクト」において機械学習によるがん診断を行ってきた。一般にセンシング機器は、検出可能成分の種類と検出値の精度がトレードオフの関係にある。このプロジェクトでは、呼気成分の絞り込み、それに基づく機器の調整、その結果としての診断精度の向上を行ってきた。



2. 研究の目的

呼気ガスデータに機械学習法を適用し、健常と疾患の診断を行う。疾病対象は、肺がん・NASH(非アルコール性脂肪性肝炎)・肝がん・歯周炎である。それぞれの疾患の患者の呼気ガスを解析し、疾病特有の呼気成分を抽出し、これまでに開発しているガスセンサを調整することで、重要成分の検出精度を高める。呼気以外の身体特性値を導入して診断性能向上、被検者の現在の疾病または健常レベルを数値化し被検者の健康状態の将来予測を確認する。各疾病専用の診断器が固有の疾病のみに陽性診断がなされることを確認する。疾病特有の呼気成分からデータ間の関係性を抽出し、生理学的・医学的意味付けを行う。

3. 研究の方法

(1) データ前処理の検討

健常者と疾病患者とは生活環境に違いがあるため、呼気ガスから生活環境由来のものを排除しなくてはならない。これまで、呼気採取した部屋のバックグラウンド成分を差し引くことはもちろん、がん患者が服用する薬剤に含まれる成分(フッ素系)を診断データから排除する等の対策を行ってきた。今後も更に検討を進め、診断のアーティファクトになりうる成分を医学・工学・情報科学の各側面から綿密に検討する。

呼気成分の検出値には、ある程度の外れ値が存在する。採取された人の特殊体質や検出器のノイズに起因すると考えられる。これら少数の外れ値は、診断アルゴリズムの性能を下げる可能性があるため、適切な閾値(例えば中央値の数倍程度)により上限を設定する。

(2) 診断器の改良

これまで研究代表者は、呼気成分について全ての組み合わせ(1,048,575通り)データを非線形サポートベクターマシン(SVM)で学習させ、一個抜き交差検証(Leave-One-Out Cross-Validation; LOOCV)により真陽性率と真陰性率を評価してきた。現在のところ、肺がんの呼気ガスについては、4種の成分で86.6%の精度を得ている。その際のSVMパラメータは、過学習を防ぐため、肺がんと健常を分離する平面を規定するデータ(サポートベクター)がなるべく少なくなるように設定した。機械学習の専門家とともに、これらのパラメータの最終的な調整を行い、診断に重要な成分を疾病別に絞り込む。また、他の学習法(ランダムフォレスト、LightGBM等)の導入し、精度評価を行う。

(3) 疾患・健常者の呼気採取

愛知県がんセンターにて受診した成人肺がん患者の中から同意を得られた者を対象に、呼気を採取する。肺がんと健常者、各々約100名から行う。担当医師から本研究の説明を受け同意を得られた者とする。健常対象者はボランティアを募る。対象者は、深呼吸後約20秒間息を止めたのち採取バッグ(容量2リットル)に静かに息を吹き込んだのち、採取者がバッグの栓をする。なお、歯周病患者ならびに歯周病対照健常者については深層呼気に加え、口腔内の呼気をシリンジで別途採取する。

(4) 呼気成分に対するガスセンサの改良・開発

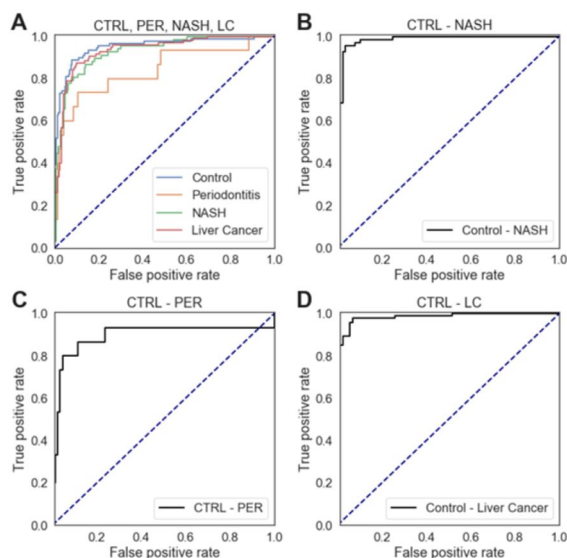
これまで呼気成分の検出機器を開発してきた。既に試作開発した室内空気をキャリアガスとする呼気の揮発性有機化合物(Volatile Organic Compounds; VOC)の検知器試作機を用いて、被験者の呼気を計測する。各々のVOCのレスポンスを実証し、その感度が疾患の判定に必要なレベ

ルであることを試験する。そのフィードバックとして必要な高感度化の開発を行う。診断解析チームが選定した肺がん、NASH、肝がん、歯周炎に由来するガス成分に合わせて、ガス分離系の改良及びガスセンサの開発を行う。ガスセンサの開発においては、ppbレベルのガスを高感度で検知しながら、複数のガスを数分間で区別できるように高速で応答可能にする。具体的には 1) センサ材料、2) センサ素子デザイン、3) センサ駆動方法を新たに設計し試作する。

4. 研究成果

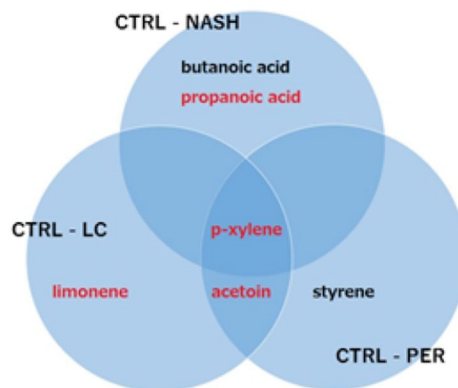
(1) 3種の疾病の診断

健康者と特定1種のがん患者との区別をする研究はいくつか存在する。1種の診断では、その特定の疾病を診断しているかどうか不明である。そのため、複数の疾病と健康者をそれぞれ区別することができる必要がある。本研究では、3種の疾病（歯周病、NASH、肝がん）と健康者の診断が可能であることを示した。右図は(A)4種の判別、(B)健康者とNASHの判別、(C)健康者と歯周病の判別、(D)健康者と肝がんの判別の性能について、ROCで評価したものである。AUC (Area Under Curve) はそれぞれ、(A)0.95、(B)0.94(C)、0.87、(D)0.94と高い値を示した。これらの結果は、呼気ガスにより特定の疾病を判別することができることを示す。



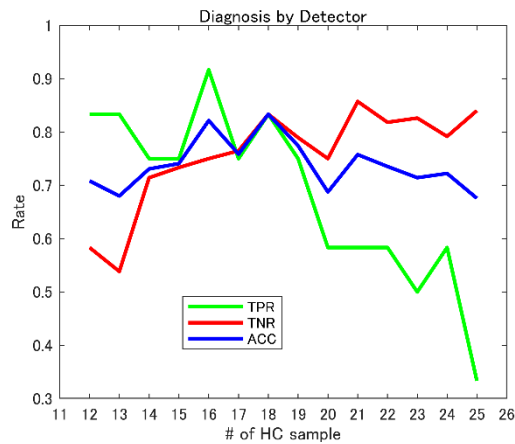
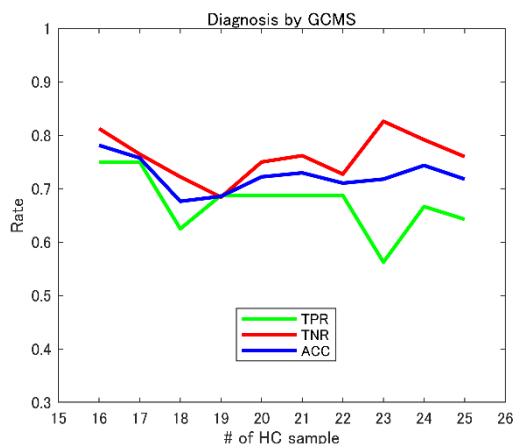
(2) 重要ガス成分の抽出

呼気ガス診断が、健康者と病気一般を判別するのではなく、特定の疾病を検出していることから、呼気ガスの中に各疾病特有の成分が含まれていることが考えられる。そこで、ゲーム理論のShapley valueを応用した重要度指標を用いて、各診断で鍵となる成分について抽出した（右図）その結果、健康者と各疾病の診断で特定の成分が、その診断で重要であることが分かった。中でも p-xylene は3つの疾病全てに関連している。歯周病、NASH、肝がんは肝臓関連の疾病と考えられており、p-xylene は肝臓一般で関係する成分である可能性がある。



(3) 簡易検知器の性能評価

質量分析器は性能が良いが大きい機器であり、将来の実用化面で問題がある。そこで、検出ガス種を絞った簡易検知器の開発とその性能評価を行った。下図は、左が質量分析器、右が4種のガスに絞った簡易検知器で計測された4種の数値に基づいてSVMにより診断性能を調べたものである。サンプルは同じ健康者と肺がん患者である。緑が真陽性率、赤が真偽性率、青が精度である。横軸は健康者のサンプル数である。検知器の方ではばらつきが大きいものの、4種での診断では、質量分析器と簡易検知器の差は見られなかった。これは簡易検知器でも質量分析器と同レベルの性能が出せる可能性を示す。



5. 主な発表論文等

〔雑誌論文〕 計1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 Sakumura Yuichi, Koyama Yutarō, Tokutake Hiroaki, Hida Toyoaki, Sato Kazuo, Itoh Toshio, Akamatsu Takafumi, Shin Woosuck	4. 巻 17
2. 論文標題 Diagnosis by Volatile Organic Compounds in Exhaled Breath from Lung Cancer Patients Using Support Vector Machine Algorithm	5. 発行年 2017年
3. 雑誌名 Sensors	6. 最初と最後の頁 287 ~ 287
掲載論文のDOI (デジタルオブジェクト識別子) 10.3390/s17020287	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 -

〔学会発表〕 計2件(うち招待講演 0件/うち国際学会 2件)

1. 発表者名 Shingo Shirone, Mutsumi Miyauchi, Kazuhisa Ouhara, Hidemi Kurihara, Takashi Takata, Hideyuki Hyogo, Toshio Ito, Woosuck Shin, Akiko Tanaka, Kazuo Sato, Kazushi Ikeda and Yuichi Sakumura
2. 発表標題 Machine learning diagnosis of multiple liver-related diseases by exhaled breath gas components
3. 学会等名 The 20th International Conference on Systems Biology (国際学会)
4. 発表年 2019年

1. 発表者名 Yutarō Koyama, Kazushi Ikeda and Yuichi Sakumura
2. 発表標題 Feature Selection using Distance from Classification Boundary and Monte Carlo Simulation
3. 学会等名 The 25th International Conference on Neural Information Processing (国際学会)
4. 発表年 2018年

〔図書〕 計1件

1. 著者名 伊藤敏雄, 作村諭一	4. 発行年 2018年
2. 出版社 電気化学会	5. 総ページ数 5
3. 書名 電気化学	

〔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	池田 和司 (Ikeda Kazushi) (10262552)	奈良先端科学技術大学院大学・先端科学技術研究科・教授 (14603)	
研究分担者	申 ウソク (Shin Woosuck) (10357246)	国立研究開発法人産業技術総合研究所・材料・化学領域・研究グループ長 (82626)	
研究分担者	宮内 睦美 (Miyachi Mutsumi) (50169265)	広島大学・医歯薬保健学研究科(歯)・准教授 (15401)	
研究分担者	樋田 豊明 (Hida Toyoaki) (80250249)	愛知県がんセンター(研究所)・分子腫瘍学分野・研究員 (83901)	
研究分担者	伊藤 敏雄 (Itoh Toshio) (90377888)	国立研究開発法人産業技術総合研究所・材料・化学領域・主任研究員 (82626)	
研究分担者	赤松 貴文 (Akamatsu Takafumi) (60635316)	国立研究開発法人産業技術総合研究所・材料・化学領域・主任研究員 (82626)	