

令和元年5月23日現在

機関番号：32644

研究種目：基盤研究(C) (一般)

研究期間：2015～2018

課題番号：15K08473

研究課題名(和文) 薬剤耐性菌の広域拡散を可視化する方法の研究開発

研究課題名(英文) Development of Novel Algorithms for Detection and Visualization of Nation-wide ARM (antimicrobial-resistant microorganism) Dissemination

研究代表者

藤本 修平 (Fujimoto, Shuhei)

東海大学・医学部・教授

研究者番号：90241869

交付決定額(研究期間全体)：(直接経費) 3,700,000円

研究成果の概要(和文)：耐性菌の広域拡散を可視化するシステムの研究開発を行った。1)既存の2DCMの改良により、県レベルの7年分、全国レベルの3ヶ月分のVREの拡散を可視化することに成功した。2)新規の完全グラフ検索algorithmによって大型完全グラフの検索が可能になり、JANIS全国データ2017年度19,516,142件のデータから得た430,536株のMRSAおよび2008年～2017年の約1億件のデータから得た7,400株のVREの薬剤感受性によるグループ分けに成功し広域拡散を見ることができた。この方法は、遺伝子配列、質量分析を用いた解析にも利用でき、より精度の高い菌拡散の解析への応用も期待できる。

研究成果の学術的意義や社会的意義

薬剤耐性菌は全世界的問題である。動向調査は実態の把握、対策の効果判定のため特に重要である。わが国にはJANISサーベイランスが存在し、年間2千万件近いデータが集積されているが、これを用いて耐性菌の拡散状況を把握するためには耐性菌のグループ分けが必要である。今回の研究で得られた新規完全グラフ検索法は、大規模データのグループ分けを一定の曖昧さを許しながら論理的に行えるものであり画期的である。さらに、遺伝子解析データ、質量分析データにも利用可能でこれらのデータが大規模データとして集積され利用可能になれば、さらに精度の高い拡散の解析を可能にするものである。疫学的検索の強力なツールが得られたと考える。

研究成果の概要(英文)：Two related systems for visualization of nationwide AMR dissemination were developed. 1) A newly developed multi-facility 2DCM system successfully visualized nation-wide VRE dissemination for three months and seven-year long VRE dissemination at prefectural level. 2) A novel algorithm to solve large scale round robin matching ("perfect graph" in graph theory term) was developed. With the new algorithm, 430,536 MRSA strains from JANIS nationwide surveillance of 19,516,142 strains for the year of 2017 were successfully assigned groups consist of perfect graphs by antimicrobial susceptibility patterns. 7,400 VRE strains from more than 100,000,000 strains for 2008-2017 were also successfully grouped. The group assignment showed the regional accumulation of the bacteria. The algorithm can apply to analysis of DNA sequence data or MALDI-TOF mass data which would bring more precise analysis on the dissemination.

研究分野：細菌学、感染症学、感染疫学、電子疫学、感染対策、医療情報学

キーワード：薬剤耐性菌 拡散 可視化 2DCM JANIS 完全グラフ 薬剤感受性パターン

1. 研究開始当初の背景

高度先進医療の普及に伴い、多くの入院患者は生体防御能の障害された易感染状態になり、健康な人が感染しない日和見感染菌(弱毒菌)による日和見感染症を発症しやすい状態にある。日和見感染菌は身近に存在する弱毒菌である常在菌や環境菌であり、患者、医療従事者の常在菌、病院の環境菌として長時間、医療機関内に存在する。医療機関では、易感染状態にある易感染患者に対し安全に医療を実施するため、予防、治療を目的に抗菌薬が多用される。日和見感染菌は繰り返し治療濃度の抗菌薬に曝露されることになり日和見感染菌のうち抗菌薬が有効な感性菌は淘汰され、耐性菌、中でも多剤耐性菌、高度耐性菌が選択されて残ることになる。つまり、今日の医療現場で耐性菌による日和見感染症が発生することは不可避の現象であり、高度先進医療を行う者には同時に耐性菌対策を行うことが求められる。

耐性菌に対する対策として、1)抗菌薬による選択圧の抑制、2)菌の院内拡散の抑制と、現状及び、1)2)による対策の効果を評価するために、3)サーベイランス(動向調査)が重要である。1)は近年 antimicrobial stewardship として急速に推進されており、2)もいわゆる「感染対策」として20年来進められている。さらに、医療機関の役割分担による効率化に伴って患者の移動が急増したことから、感染対策の地域連携も重要になり、それを支える仕組み作りも進んでいる。サーベイランスについては JANIS(厚生労働省院内感染対策サーベイランス)による標準化とデータの蓄積が行われており、本研究者の開発した 2DCM による菌の院内拡散の可視化も JANIS のデータ還元の一翼を担う機能として JANIS に実装され、参加医療機関は自由に利用できるようになっている。

2DCM に用いられている感受性試験結果に基づく菌グループ作成 algorithm には、臨床細菌検査の誤差を考慮して、誤差範囲内であれば、同じ結果と見なす仕組みにもとづいて、2つの株の異同を判定し、3つ以上の株が同じ株かどうかは、2つの株の異同を元に、お互いに異なった株でない株同士を同じグループに割りあてて、グラフ理論の用語で完全グラフと呼ばれる総当たりで検証して同じグループに入れて良いかどうかを決める作業が存在する。この作業は、要素数が増えると検証する組み合わせが指数的に増えることが知られており、実用化を妨げていた。本研究者は、この作業を最小ステップで終了する algorithm の開発に成功し、現在利用されている 2DCM に実装し、実用性を証明した。

地域拡散を調べる重要性が増しているにもかかわらず、既存の 2DCM は、施設内での耐性菌を含む細菌の拡散を調べるには十分な 1,000 株程度までの処理こそ可能であったが、年単位の地域拡散、日本国内全体の拡散を調べることは不可能であった。そこで、JANIS の年間報告数であった 1 千万件の細菌データを解析可能な新規アルゴリズムの開発を行い、耐性菌の地域拡散、国内拡散を明らかにできるようにすることを一つの目標とした。

研究開始当初は、大型の完全グラフの検索法に全く目途がなかったため、全株処理の他に、耐性度の高い株に絞った解析、代表パターンの選出、地理的・時間的集積のある株に絞った解析などの可能性も模索することにした。

2. 研究の目的

大規模サーベイランス(JANIS)データを利用した耐性菌の地域～国内拡散の可視化を目的に

(1)大要素数の完全グラフ検索新規 algorithm の開発による地域～全国レベルでの 2DCM 解析

(2)現行の 2DCM を利用した地域～全国レベルでの耐性菌拡散の可視化

耐性度の高い菌株の絞った解析 algorithm の開発

代表株(集積株、連携する耐性を持つ株)の選定 algorithm の開発

を行った。

3. 研究の方法

(1)大要素数の完全グラフ新規検索 algorithm の開発

現行の 2DCM を元に、完全グラフの検索の際に消費されるメモリーを最少化する algorithm を検討した。

新規に開発した完全グラフ検索 algorithm を開発言語 Hot Soup Processor version 3.5(HSP 3.5)を用いて検証用のプログラム化した。(Hot Soup Processor は開発の容易な interpreter 言語であるが、実行に C++の 20～30 倍の時間が必要であることが分かっている。実行速度が遅いことは、アルゴリズムの優秀性を調べる目的には良い特性でもある。)

完全グラフの検索前の処理部分、検索後の包含される小さなグループの排除部分。菌株にグループを割りあてる部分を HSP 3.5 を用いてプログラム化した。

全体をバッチ処理するためのプログラムを作成した。

厚生労働大臣に申請を行い JANIS 検査部門の実データを入手した。

JANIS 実データと開発したプログラムを用いてアンチバイオグラムの分類と菌株のグループ化を行い、問題点を見いだした。

多くの検体が調べている抗菌薬を抽出してより実用性の高いグループ化を行うアルゴリズムを開発、プログラム化した。

JANIS データ全施設 (2008 年約 850 施設 ~ 2017 年約 2200 施設) データを用い、2017 年 1 年間 19,516,142 件から得た MRSA (メチシリン耐性黄色ブドウ球菌) 430,536 株、2007 年 ~ 2017 年約 1 億件から得た VRE (バンコマイシン耐性腸球菌のうち *Enterococcus. faecium*) 7,400 株を対象にグループ化を行い開発した algorithm の実用性を確認した。

(2) 現行の 2DCM-web 改良, PMAL (probability based microbial alert light version) と PMAL にもとづいた -alert matrix, 耐性菌の定義ファイルの利用による地域拡散の可視化

2DCM-web の内部のデータベース化を行い、これまでよりも多くのデータを安定して処理できるようにした。

耐性菌の定義ファイルを用いた処理によって CRE (カルバペネム耐性腸内細菌科菌) など任意の耐性菌を抜き出すことができるようにし、より絞り込んだ菌株での処理ができるようにした。

4. 研究成果

(1) 大要素数の完全グラフ新規検索 algorithm の開発

2DCM では、グループ化を行う全ての菌株 (n 株) について、全ての 2 株の組み合わせにそれぞれが同じ株と言えるかどうかを判定して $n \times n$ の真偽表を作成。次に 3 株以上の組み合わせについてメンバーになる 3 株以上の株が総当たりでお互いに同じ株といえるかどうかを真偽表で調べ (完全グラフの検索)、全ての組み合わせで同じ株と言える (真) であった場合 (完全グラフであった場合) その組み合わせをグループとして採用するアルゴリズムによって、もれなくグループを検索している。3 株以上の組み合わせは組み合わせが成立する限り論理的には n 株まで調べて行く。

要素数が増えると指数的にステップが増えるのはこの部分である。これについては、すでに、本研究者が、2 つの組み合わせから、組み合わせを漸増させて行く algorithm で、効率的にもれなく組み合わせを探せる方法を見だし、数学的帰納法で正当性を証明している (特許 5568816 号、特開 2013-101675; <https://www.j-platpat.inpit.go.jp/p0000> で参照可)。概要は以下の通りである。

~~~~~

同じ株と言えることになった株同士を「互いに真の株」と呼ぶことにする。2 つの株の組み合わせにおいて、ある株  $a$  に互いに真の株を全て探すことができたとしてそれが  $b_0 \sim b_k$  だったとする。  $\{a, b_0\}$  の 2 つの要素のグループに加えることができる株 (要素) をさがす。その要素は、 $a$  と互いに真でなければならぬので、 $b_0 \sim b_k$  の中にある要素以外ではない。従って、 $b_1 \sim b_k$  から  $a, b_0$  に互いに真の要素をさがせばそれが全てと言うことになる。

$j$  個の互いに真の関係の要素を持つ集合  $A$  に互いに真の関係の要素を全て探すことができそれが、 $c_0 \sim c_k$  だったとする。この場合も同様に、 $\{A, c_m\}$  にもう一つのお互いに真の要素を加えようとすると、その要素は  $c_m$  以外の  $c_0 \sim c_k$  に含まれるはずである。

2 つの要素の全ての組み合わせは  $n \times n$  の総当たりの表でできているので、2 つの株の組み合わせにおいてお互いに真の株を全て探すことはできる。従って、3 つ以上の株の組み合わせにおいて、総当たりで調べる必要はなく、一つ数の少ない組み合わせを作ったときに新たに加えた株 (要素) を参考に調べることで、もれなく、全ての組み合わせを網羅できる。

~~~~~

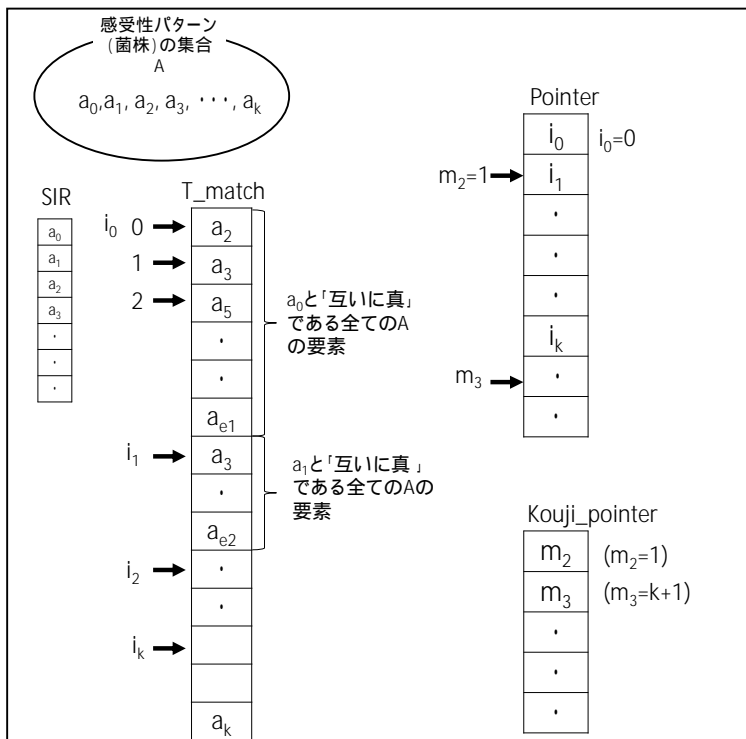
一方、このについては、株数が増加すると、 $n \times n$ の表が必要となるため、50 万株の場合、真偽に 01 の 1 ビットのみを割りあてても、30GB 以上のメモリー空間が必要になる。

新規アルゴリズム (Super SIR Grouping) では、組み合わせの作成に必要なのは、最後に加えた要素の集合であることに注目し、その情報だけをメモリー上に保持する仕組みを作った。メモリー上の配置は、3つの1次配列として、これまでの2DCMで高次配列に保持した情報を、隙間なく保持できるようにした(図1)。

乱数を用いて生成した SIR を 1:1:1 で含む dummy 感受性パターンを作成し、グループ化の動作を確認した。20 抗菌薬 100,000 件を7時間で処理できることが確認できた。

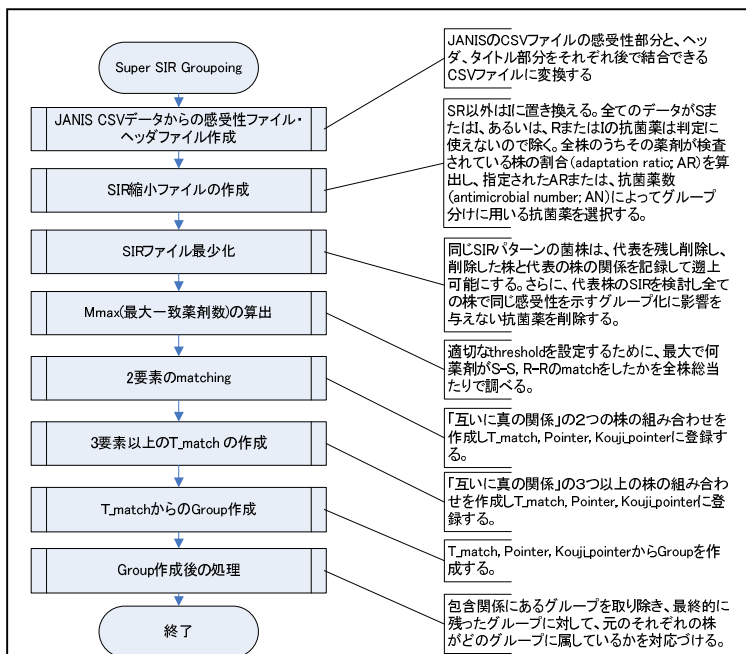
JANIS データからの菌のグループ化を可能とするためにデータの前処理部分、グループ化後の包含排除処理部分を含む全体をバッチ化した(図2)。

当初のバッチ処理には、抗菌薬数を採用率(adaptation rate; AR, 全株数に対してその薬剤の検査結果のある株数の割合)によって抗菌薬を選択する仕組みを設けなかった。JANIS データ全施設(2008年約850施設~2017年約2200施設)データを用い検証を行ったところ、3~5要素の段階でグループ数が5億を超えてしまった。原因を調べたところ、多くの未検査薬に対する結果が含まれ、これらは、調べれば、S(感性)であったかも知れず、またR(耐性)であったかも知れないために、いずれでもないI(中間)と同じ扱いで、「互いに真」と見なすSS, RRの一致数の閾値(threshold)を大きくしないと曖昧に、「偽(false)」ではない「真(true)」となってグループ数を増やす原因になっていることが推測出来た。2017年1年間19,516,142件から得たMRSA(メチシリン耐性黄色ブドウ球菌)430,536株を調べたところ194種の抗菌薬に対する感受性が調べられていた。また、2007年~2017年約1億件から得たVRE(バンコマイシン耐性腸球菌のうち *Enterococcus. faecium*) 7,400株を調べたところ216の抗菌薬に対する感受性が調べられていた。そこでARを計算し、あらかじめ指定したAR値以上の抗菌薬のみをグループ分けに使うか、あらかじめグループ分けに用いる抗菌薬数(antimicrobial number; AN)を指定してAR値の大きい順



(図1) 新規アルゴリズムでのメモリーの配置

T_matchは、グループに加えられた「互いに真」の要素のリスト。
 Pointerは、T_matchのリストが、どの集合に加えられた要素に対応するのかが示すリスト。
 Pointer(i)は、集合I(要素が一つの場合は、 a_i)に加えられた最初の追加要素のT_matchでの位置を示している。
 Kouji_pointerは、ある要素数のグループがPointerのどの位置に対応するかを示している。
 Kouji_pointer(k)は、要素数kのPointer上での最初の要素の位置(k-1個の最後の要素の次)を示している。
 SIRは感受性情報を一時的に読み込む配列で、2つの組み合わせの検索後は不要になり破棄される。



(図2) JANIS検査部門月報CSV(全検体)を利用して検証を行うためのバッチプログラム

新規アルゴリズムは、主に、「2要素のmatching」、「3要素以上のT_match作成」に活かされているが、データ構造そのものが新規アルゴリズムの本態であるため全体を新規アルゴリズムとしてSuper SIR Groupingと呼ぶことにした。

2要素のmatchingにはS-Rのmismatchによる「偽(false)」判定の他に、2DCMと同様にSS, RR matchの数による閾値(threshold)を利用できるようにした。thresholdは0に設定することも可能で、その場合は、全ての、「偽(false)」でない関係が「真(true)」となる。

薬剤感受性パターンとしてSIRを用いたが、MICによっても処理が可能で、その場合は、S-Rのmismatchの代わりに、MICで4倍以上の違いをmismatchとして扱い、2要素のmatchingに用いる。

2要素以上のmatching移行のalgorithmは、薬剤感受性パターンの分類だけでなく、2要素間で真偽が決まる全ての関係に応用可能である。

に指定した数の抗菌薬をグループ分けに採用する algorithm を導入した。

上記の MRSA の例では、AR=0.7 で 8/194 に、VRE の例では AR=0.6 で 6/216 に、抗菌薬数が減じた。今回の試験的データ解析ではこれらの値、あるいは AN である 8, 6 を指定して解析を行った。採用率の 0.7、0.6 は一定の自由度を認めた解析であり、80% 以上などもう少し厳しい条件で解析を行うのが適当であるかも知れない。また、採択率を離れて、検査無し抗菌薬を排除することを目的にするのであれば、代表パターン中の出現率で抗菌薬を選択する方法も検討する必要があると考えた。

これらの処理を行った上で、同 MRSA では、最大で 14 要素の 2,097,048 のグループが生成され、包含されるグループの排除後、3,683 グループが残った。VER では、最大で 19 要素の 2,300,801 のグループが生成され、包含排除後 212 のグループが残った。

これらのグループを最初のそれぞれ、430,536 株、7,400 株に遡って割りあて、地域集積などを見ることができた(解析結果については、今回のデータ利用申請の範囲を超えるためデータ開示せず)。

今後、多くのグループを含むデータからの特長抽出に AI を用いた解析を取り入れて行うことを検討している。また、包含されるグループの排除の処理に時間がかかっているため、この部分については、実行速度の速い C++ 言語での再コードを予定している。

(2) 現行の 2DCM-web 改良, PMAL (probability based microbial alert light version) と PMAL にもとづいた -alert matrix, 耐性菌の定義ファイルの利用による地域拡散の可視化

現行の 2DCM-web のメモリー利用の一部をデータベース化し、大型のデータを安定して処理できるように改良した。また、耐性菌の定義ファイルを利用して、任意の耐性菌定義にもとづいた 2DCM 解析を可能にした。これらの改良に伴い、耐性菌の地域拡散を 2DCM で観察できるようになったため、複数施設解析版 2DCM-web を開発して、2DCM-web 実習サイト (<http://yakutai.dept.med.gunma-u.ac.jp/project/2dcm/index.html>) で公開した。

複数施設解析版 2DCM-web によって、JANIS に 80 数施設のある県で 7 年間に検出された VRE (*E. faecium*) 896 株の 2DCM 解析が可能であった。さらに、全国データ 3 ヶ月分から抽出した 225 株についても、解析が可能であった。

PMAL、PMAL にもとづいた -alert matrix についても、2DCM-web 実習サイトで公開し、これにもとづいた 2DCM 解析を可能にした。

これら、既存アルゴリズムの改善でも、地域レベルでの耐性菌拡散の可視化が可能であることが分かったが、全国レベルになると応用は短期間に限定された。

今後、Super SIR Grouping を 2DCM-web に実装することで手軽に全国レベルで年余にわたる耐性菌の拡散を可視化できるようになるとともに、地域レベル、施設レベルでも長期間の 2DCM 解析が可能になることによって、解析期間が異なるとグループ分けが変わってしまうという既存の 2DCM の問題も解決すると考えた。

-alert matrix を地域拡散を見る仕組みとして利用することも考えられるが、そのためには、常在菌として報告しない菌の統一、あるいは、細菌検査提出ルールなど、細菌検査そのものの標準化が進むことが必要である。

今回、耐性菌の広域拡散の可視化を目的に複数の方法でアプローチを行った。最も困難と考えていた、大型の完全グラフを効率よく検索する algorithm についても開発、検証プログラムの作成を行い、当初の目標であった JANIS 検査部門全施設 1 年分以上のデータの解析に成功した。改良を行うかたちでのアプローチで得た実用プログラムに、新規の algorithm を組み込むことで、新規アルゴリズムの応用範囲が広がるのが期待でき、最終的に効率の高い研究開発になったと考える。

最後に、この研究で開発した完全グラフの検索アルゴリズムは、2 つの要素の関係を何らかの基準で真偽 2 つの値で表すことができるあらゆる集合のグループ化に用いることができるもので、遺伝子解析、質量解析などを含めて類似性のあるものをグループ化するために利用できる。比較的小さなワークステーションで効率的にグループ化ができるため、他分野での応用もひろく期待できると考えた。

5. 主な発表論文等

(雑誌論文)(計 1 件)

藤本 修平、サーベイランス、何が変わった? - 感染対策の地域連携支援システム (RICSS) は AMR 対策の情報収集還元プラットフォーム? -、Infection Control, Vol.26, 1224-1228, 2017、査読なし

〔学会発表〕(計 6 件)

藤本修平、本間操、八束眞一、大石貴幸、岩崎澄央、静野健一、荻野毅史、太田浩敏、八木哲也、村上啓雄、柴山恵吾、荒川宜親、JANIS データを活用して AMR 対策地域連携を進めよう:地域連携を支援するネットワークツール「地域連携支援ツール群」の開発と公開、第 34 回日本環境感染学会学会総会・学術集会、2019 年

藤本修平、本間操、八束眞一、大石貴幸、岩崎澄央、静野健一、荻野毅史、太田浩敏、八木哲也、村上啓雄、柴山恵吾、荒川宜親、JANIS データを利用して薬剤耐性菌の地域拡散を調べよう:複数施設解析版 2DCM-web と関連ツールの開発と公開、第 30 回日本臨床微生物学会総会・学術集会、2019 年

藤本修平、谷本弘一、富田治芳、八木哲也、柴山恵吾、荒川宜親、村上啓雄、耐性菌関連電子システム開発の現状 - 耐性菌条件/警告・案内定義メッセージの標準化・RICSS・2DCM-web の進捗 -、第 46 回薬剤耐性菌研究会、2017 年

藤本修平、村上啓雄、村木優一、八木哲也、柴山恵吾、荒川宜親、耐性菌と戦うサーベイランスシステム:日本の耐性菌対策を支える JANIS, JACS, RICSS の現状と将来、第 90 回日本細菌学会総会、2017 年

筒井敦子、松井真理、鈴木里和、柴山恵吾、藤本修平、「JANIS 検査部門のデータの活用について」、第 31 回日本環境感染学会総会、2016 年

本間操、藤本修平、薬剤感受性成績による疫学解析-2DCM-web-の活用、第 27 回臨床微生物学会総会、2016 年

〔図書〕(計 0 件)

〔産業財産権〕

出願状況(計 0 件)

取得状況(計 0 件)

〔その他〕(なし)

6. 研究組織

(1)研究分担者(なし)

(2)研究協力者(なし)

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、研究者個人に帰属されます。