

令和 3 年 4 月 1 日現在

機関番号：11301

研究種目：若手研究

研究期間：2019～2020

課題番号：19K18343

研究課題名（和文）集中治療患者の即時転帰予測システムの開発：人工知能によるビッグデータ時系列解析

研究課題名（英文）Development of an immediate outcome prediction system for intensive care patients: Big data time series analysis with artificial intelligence

研究代表者

小林 直也 (Kobayashi, Naoya)

東北大学・大学病院・助教

研究者番号：50837036

交付決定額（研究期間全体）：（直接経費） 3,200,000 円

研究成果の概要（和文）：敗血症の転帰予測における命題として、感度、特異度の高さと予測因子数の増加に伴う多数の検査や入力による予測因子収集の煩雑化軽減の両立を指摘し、ビッグデータによる命題解決を提案した。高頻度の血圧データ、看護記録のテキストデータ、機械学習の3つの方法のうち高頻度の血圧データのみがフレームワークの構築に至り、感度、特異度の高さと、予測因子収集の煩雑さの軽減の両立を解決できる可能性が示唆された。

研究成果の学術的意義や社会的意義

本研究のような試行錯誤による経験の蓄積によって、適用可能なフレームワークの構築ができれば、近年性能が向上しているコンピュータ等の計算能力により命題解決が可能となり、さらに他の大量のデータを扱うべき命題においてもフレームワークを適用することにより命題解決が加速できる可能性があるという点で意義がある。また、ビッグデータ解析は、経済、工学、情報科学分野等との学際的研究を要するが、幅広い研究協力体制のもとフレームワークの構築に至ることができており、さらに医学以外にも、テキストデータ解析の方法論をはじめとした情報科学分野における知見も得ているという点でも当該研究には意義がある。

研究成果の概要（英文）：As a proposition for predicting the outcome of sepsis, we pointed out the need for both high sensitivity and specificity and the need to reduce the complexity of collecting predictors due to the large number of tests and inputs associated with the increase in the number of predictors, and proposed the use of big data to solve the proposition. Of the three methods, high-frequency blood pressure data, text data from nursing records, and machine learning, only high-frequency blood pressure data led to the construction of the framework, suggesting the possibility of solving both the problems of high sensitivity and specificity and reducing the complexity of collecting predictors.

研究分野：集中治療

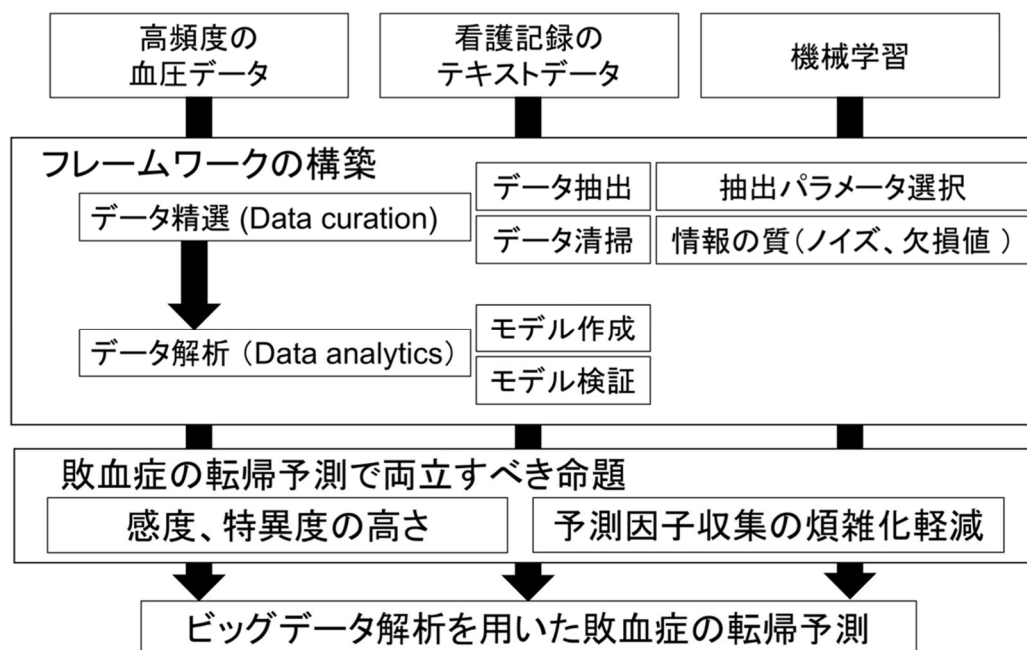
キーワード：ビッグデータ 集中治療 敗血症 転帰予測

1. 研究開始当初の背景

敗血症の治療方針決定や転帰予測のうえでは患者の重症度の把握が重要であるため、重症度スコアが用いられている。従来の重症度スコアは、予測因子数の増加に伴う多数の検査や入力により感度、特異度の一定の増加につながったが、予測因子収集の煩雑化が同時に生じている。近年、ビッグデータによる課題解決が可能となってきているが、敗血症の転帰予測に適用するためには、この命題に対してデータ抽出からデータ洗浄までの data curation（データ精選）から、モデル作成から検証までの data analytics（データ解析）までのフレームワークを構築するとともに、適切な抽出パラメータ選択や、情報の質に関わるノイズや欠損値への対処を行う必要があるが、それらに関する知見は十分ではない。

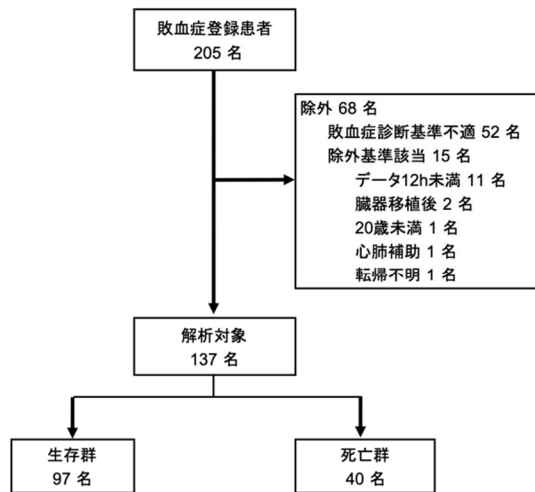
2. 研究の目的

本研究の目的は、敗血症の転帰予測における命題である、感度、特異度の高さと予測因子収集の煩雑さ軽減の両立に関して、ビッグデータ解析を適用するためのフレームワークを高頻度の血圧データ、看護記録のテキストデータ、機械学習の3つの方法で構築できるか検討し、特にデータ精選における適切な抽出パラメータの選択や情報の質に関わるノイズや欠損値への対処を明らかにすることである。

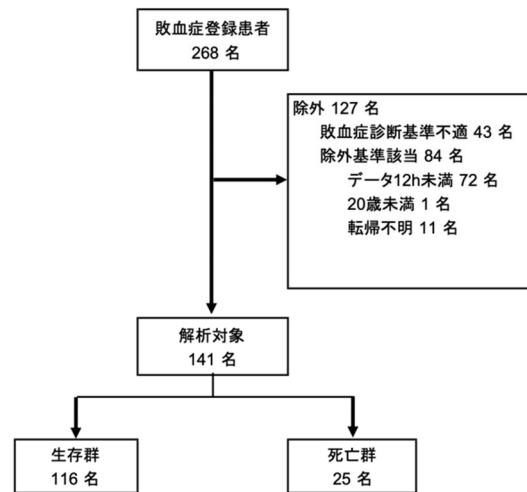


3. 研究の方法

2008年1月から2015年1月に東北大学病院の集中治療室に入室した敗血症患者137名（転帰予測のモデル作成のための導出コホート）と、同期間に仙台市立病院および大崎市民病院の集中治療室に入室した敗血症患者141名（モデル検証のための検証コホート）を対象とした。本研究では、集中治療室で得られるデータを、データ入力と構造化の観点から4つのデータ領域に便宜的に分類し、そのうち画像データに代表される自動入力の非構造化データ以外の3つのデータ領域で、データ精選からデータ解析までのフレームワークを構築し、さらにデータ精選においては適切な抽出パラメータの選択や情報の質に関わるノイズや欠損値への対処法を検討した。



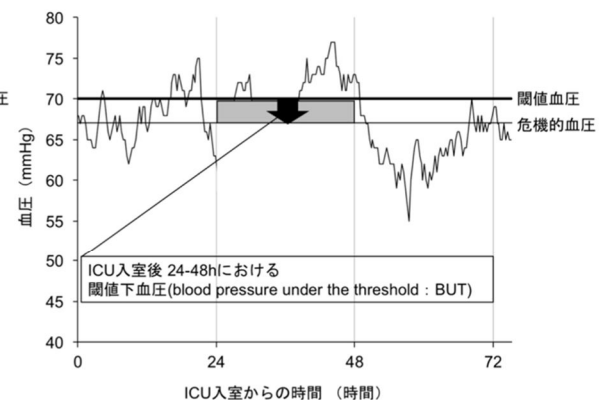
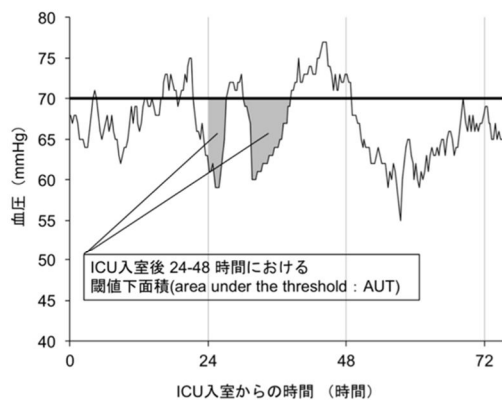
A：導出コホート（東北大学病院 137 名）



B：検証コホート（仙台市立病院、大崎市民病院 141 名）

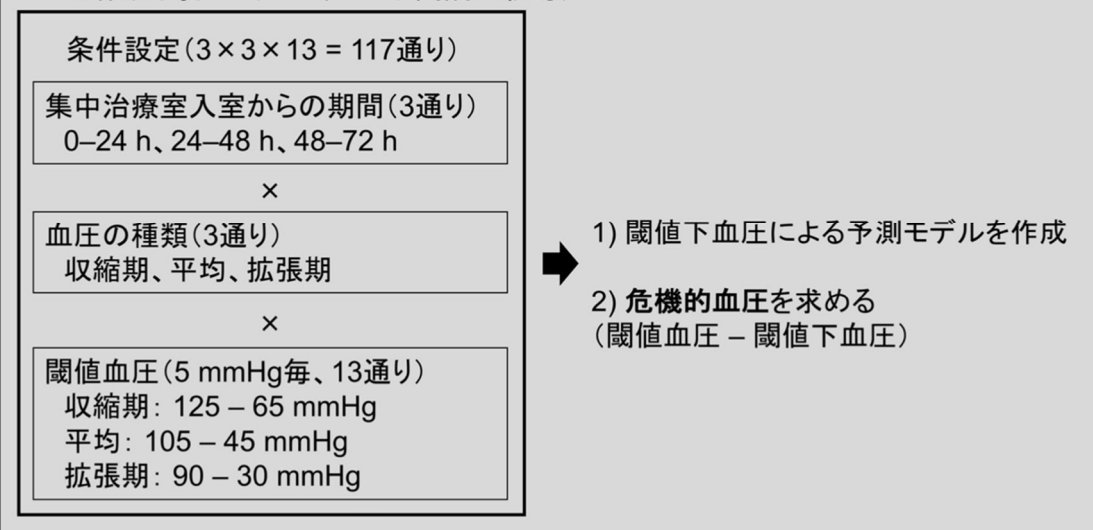
(1) 血圧の高頻度データを用いた転帰予測

データ精選：予備的検討において連続データのうち最も転帰を反映していた血圧を用いた。導出コホートにおいて、測定期間内の 3 秒毎に蓄積された高頻度の全血圧データを用い、ノイズの除去や欠損値の補完は行わなかった。データ解析：モデル作成では、集中治療室入室後 0-24、24-48、48-72 時間において、血圧低下の基準となる閾値血圧を 5 mmHg 刻みで設定した。次に、閾値血圧からの低下面積を求め、低下面積を平均した閾値下血圧による 90 日転帰予測のモデルについてロジスティック回帰分析により受信者操作特性解析の曲線下面積（Area under the curve of receiver operating characteristic：AUCROC）を求めた。モデル検証では、危機的血圧を閾値血圧 - 閾値下血圧と定義し、検証コホートで、導出コホートでの予測が最良であった場合の危機的血圧を超えたかの是非で 2 群に分けて Kaplan-Meier 曲線を作成し、ログランク検定で検証した。最後に、危機的血圧と重症度スコアの AUCROC を Nemenyi のポストホック解析を用いて検討した。

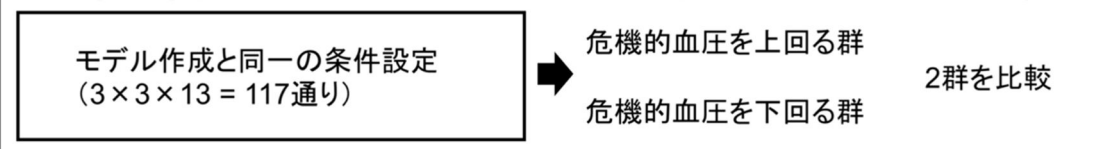


高頻度血圧データを用いた転帰予測のための次元削減

モデル作成(導出コホート、データ間隔:3秒毎)

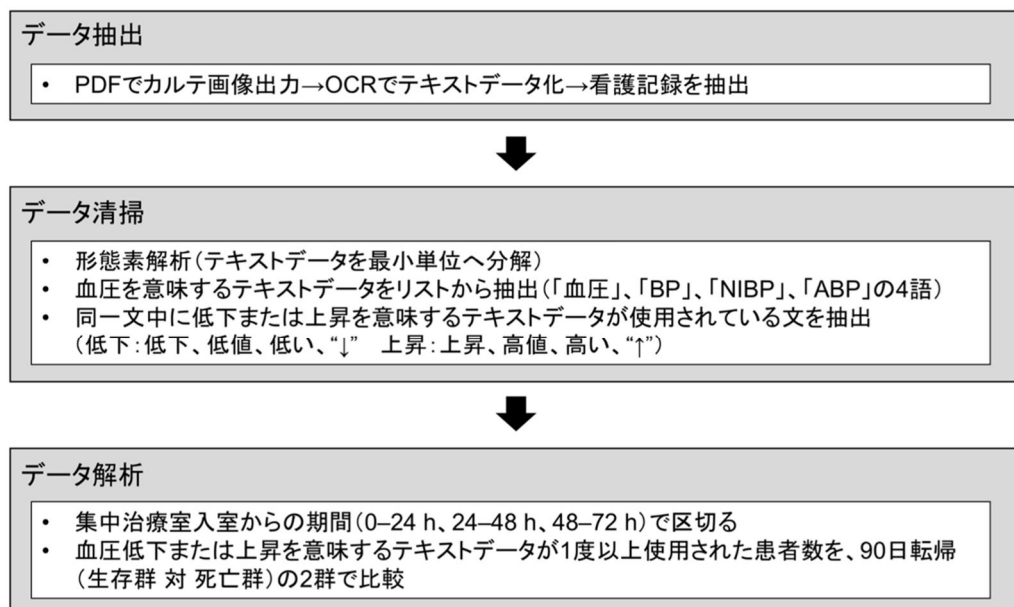


モデル検証(検証コホート、データ間隔:1分毎 ※データ保存間隔が異なるコホートで検証)



(2)看護記録のテキストデータを用いた転帰予測

データ精選: データ抽出が可能であった看護記録を用いた。ノイズは導出コホートの看護記録に対して形態素解析を用いて除去した。欠損値はテキストデータにおいては補完ができないため処理は行わなかった。データ解析: 90 日転帰で生存群と死亡群に分け、集中治療室入室後 0-24、24-48、48-72 時間において血圧低下に関するテキストデータの出現頻度によるモデル作成を検討した。統計はフィッシャーの正確確率検定を用いた。



看護記録のテキストデータの検討方法

(3)機械学習による転帰予測

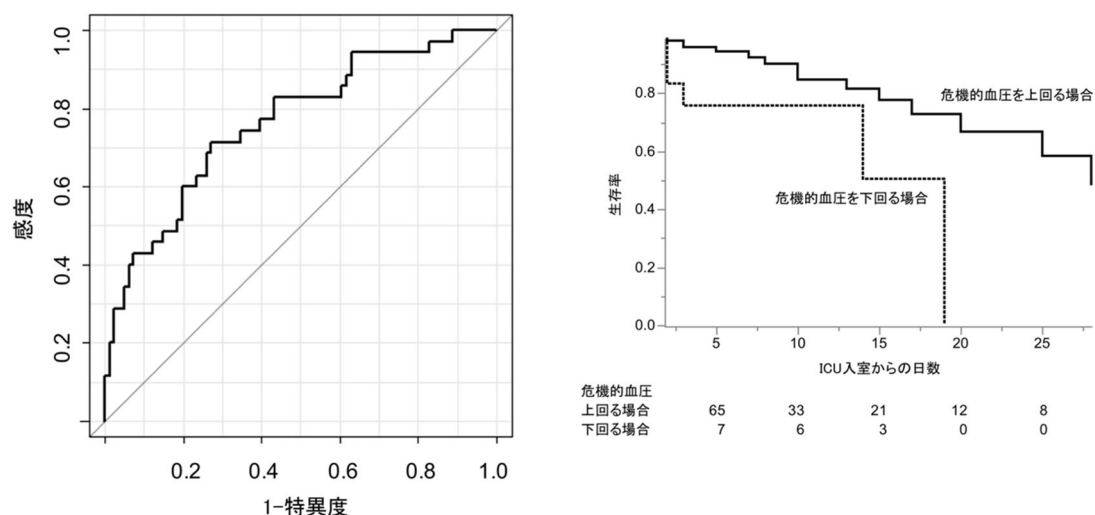
データ精選: 平均血圧、看護記録、患者背景と重症度スコアを用いた。血圧に関するノイズは数値処理による除去、欠損値は状態空間モデルによる補完を行った。看護記録に関するノイズは形態素解析で除去し、患者背景と重症度スコアのうち重症度スコアを構成する連続データは欠損

値補完を行わず、ノイズは目視により外れ値を除去した。データ解析：集中治療室入室後 0-24 時間において、患者背景および重症度スコア、平均血圧、看護記録に関するモデル作成を行い、A：患者背景および重症度スコア、B：患者背景および重症度スコアと平均血圧、C：患者背景および重症度スコアと看護記録、D：患者背景および重症度スコアと平均血圧と看護記録の組合せにおいて機械学習を用いて転帰予測の AUCROC を検討した。

4. 研究成果

(1) 血圧の高頻度データを用いた転帰予測

モデル作成では、集中治療室入室後 0-24、24-48、48-72 時間において、5 mmHg 刻みで設定した閾値血圧のうち、集中治療室入室後 24-48 時間における平均血圧 70 mmHg 未満の閾値下血圧が最も 90 日転帰と関係した（閾値下血圧：2.2 mmHg、AUCROC：0.763、 $p = 0.006$ 、危機的血圧 67.8 mmHg）。モデル検証では、検証コホートにおいて、集中治療室入室後 24-48 時間における危機的血圧である 67.8 mmHg 未満群が、67.8 mmHg 以上群に対して有意に死亡率が高かった（67.8 mmHg 未満群：4/10（40.0%）、67.8 mmHg 以上群：18/115（15.7%）、 $p = 0.018$ ）。危機的血圧を用いたモデルは、多数の予測因子を用いる 4 つの重症度スコアと同等の AUCROC であった（ $p = 0.420$ ）。



(2) 看護記録のテキストデータを用いた転帰予測

血圧低下に関連するテキストデータが使用された患者数は、集中治療室入室後 0-24 時間では生存群で有意に多く（生存群：23/32 人（71.9 %）、死亡群：9/22 人（40.9 %）、 $p = 0.028$ ）、48-72 時間後では死亡群で有意に多かった（生存群：2/32 人（6.3 %）、死亡群：10/22 人（45.5 %）、 $p = 0.002$ ）。入室からの時間により結果が異なるため、データ解析の段階においてテキストデータの出現頻度をもとにしたモデル作成が困難であった。

(3) 機械学習による転帰予測

ICU 内死亡についての AUCROC は、A：0.495、B：0.434、C：0.670、D：0.645 であり、看護記録のテキストデータを用いた組合せ（C、D）が、用いなかった組合せ（A、B）よりも AUCROC が有意に高かった（C と A： $p < 0.001$ 、C と B： $p < 0.001$ 、D と A： $p < 0.001$ 、D と B： $p < 0.001$ ）。C と D には統計学的有意差を認めなかった（ $p = 0.995$ ）。28 日死亡についての AUCROC は、A：0.480、B：0.470、C：0.607、D：0.558 であり、看護記録のテキストデータを用いた組合せ（C、D）が、用いなかった組合せ（A、B）よりも AUCROC が有意に高かった（C と A： $p < 0.001$ 、C と B： $p < 0.001$ 、D と A： $p < 0.001$ 、D と B： $p < 0.001$ ）。C と D には統計学的有意差を認めなかった（ $p = 0.995$ ）。

本研究では、高頻度の血圧データ、看護記録のテキストデータ、機械学習の 3 つの方法のうち、高頻度の血圧データのみがフレームワークの構築に至り、ノイズの除去や欠損値の補完を行わないデータ精選によって、感度、特異度の高い単一のパラメータによる転帰予測が得られたため、敗血症の転帰予測における命題である、感度、特異度の高さと、予測因子収集の煩雑さの軽減の両立を解決できる可能性が示唆された。今後の課題としては、データ抽出を迅速かつ適切に行うためのインフラ構築、特定の命題に対して適用するフレームワークを選択するための経験の蓄積、入力に際しての属人的要素の排除、および情報処理技術の向上が必要である。

5. 主な発表論文等

〔雑誌論文〕 計3件（うち査読付論文 3件／うち国際共著 0件／うちオープンアクセス 0件）

1. 著者名 Kobayashi Naoya, Shiga Takuya, Ikumi Saori, Watanabe Kazuki, Murakami Hitoshi, Yamauchi Masanori	4. 巻 11
2. 論文標題 Semi-automated tracking of pain in critical care patients using artificial intelligence: a retrospective observational study	5. 発行年 2021年
3. 雑誌名 Scientific Reports	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） 10.1038/s41598-021-84714-8	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 Kobayashi Naoya, Wagatsuma Toshihiro, Shiga Takuya, Toyama Hiroaki, Ejima Yutaka, Yamauchi Masanori	4. 巻 6
2. 論文標題 Age-related changes in factors associated with delayed extubation after general anesthesia: a retrospective study	5. 発行年 2020年
3. 雑誌名 JA Clinical Reports	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/s40981-020-00325-8	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 Kobayashi Naoya, Nakagawa Atsuhiko, Kudo Daisuke, Ishigaki Tsukasa, Ishizuka Haruya, Saito Kohji, Ejima Yutaka, Wagatsuma Toshihiro, Toyama Hiroaki, Kawaguchi Tomohiro, Niizuma Kuniyasu, Ando Kokichi, Kurotaki Kenji, Kumagai Michio, Kushimoto Shigeki, Tominaga Teiji, Yamauchi Masanori	4. 巻 24
2. 論文標題 Arterial blood pressure correlates with 90-day mortality in sepsis patients	5. 発行年 2019年
3. 雑誌名 Blood Pressure Monitoring	6. 最初と最後の頁 225～233
掲載論文のDOI（デジタルオブジェクト識別子） 10.1097/MBP.0000000000000398	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 —

〔学会発表〕 計1件（うち招待講演 0件／うち国際学会 1件）

1. 発表者名 Kobayashi Naoya	
2. 発表標題 Arterial Blood Pressure Correlates with 90-Day Mortality in Sepsis Patients: A Retrospective Multicenter Derivation and Validation Study Using High-Frequency Continuous Data	
3. 学会等名 The Kick-off Symposium of Advanced Graduate Program for Future Medicine and Health Care (国際学会)	
4. 発表年 2019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
--	---------------------------	-----------------------	----

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