

厚生労働科学研究費補助金（がん対策推進総合研究事業）
（分担）研究報告書

Federated Learning によるデータシェアリング

研究代表者 松田智大 国立がん研究センターがん対策研究所国際政策研究部 部長
研究分担者 杉山裕美 （財）放射線影響研究所（広島） 主任研究員
研究分担者 雑賀公美子 弘前大学大学院・客員研究員
・国立がん研究センターがん対策研究所国際政策研究部 外来研究員

研究要旨 国内、国際共同研究での個別データを外部に一切持ち出さずに、データ集約手法を導入する。既に台湾がん登録で導入しているシステム **Federated Learning** を、大阪国際がんセンター、佐久医療センター、放射線影響研究所と国立がん研究センターとオランダ **IKNL** をつないで導入し、東京、長野、広島と大阪のがん登録データに基づいて、それぞれのがん登録の年齢罹患率の算出、4 登録合計の年齢調整罹患率の算出に成功した。それをもとに、**RARECAREnet Asia** プロジェクトに基づいて、頭頸部がんを焦点を当て、日本と台湾、イタリアを結んだ三ヶ国で実際の分析を実施した。2022 年 10 月 17-20 日の **UICC World Cancer Congress**（ジュネーブ）において、本技術のセッションを設け、情報を共有する。今後は、国立がん研究センター中央病院が進める **ATLAS** 計画においても、本システムでのデータシェアについて検討する予定である。

A. 研究目的

GDPR を初めとする情報保護に関する法律が多く、多くの国で適用され、匿名であろうと、個別情報の国外提供が困難となってきた。オランダ国立総合がんセンター（**IKNL**）は、**Vantage6** というシステムを開発し、この技術は、がん登録情報の分散分析化を可能にしている（図 1）。このプロジェクトの目的は、日本のがん登録情報の、集計データの計算と共有のためのシステム利用の技術的実現性を実証することである。

B. 研究方法

本プロジェクトにおいて、業務委託先である **IKNL** から、以下のものの提供を受

け、大阪国際がんセンター、佐久医療センター、放射線影響研究所と国立がん研究センターにシステムを導入した。

- 日本のがん登録室ががんデータの分散分析を行うために、**Vantage6** インフラを利用可能にする（図 2）
- **Vantage6 Trolltunga** サーバーへのアクセス
- 日本の担当者がシステムのインストールおよび使用に関するトレーニングを実施
- 日本の **CR** から指定された分析（水平分割データの年齢調整罹患率アルゴリズムなど）の開発
- 新しいアルゴリズムや解析の開発につ

いての指導

ダミーデータにおいて、各地域からのデータを統合した形で、年齢調整罹患率を算出した後、相対生存率を算出して、統計ソフトでの答え合わせ検証を行う。

検証実施ののち、RARECAREnet Asia のプロジェクトにおいて、頭頸部がんに焦点をあて、日本、台湾、及び EU のデータで相対生存率の比較を行った。

C. 研究結果

システムの導入が各地域において実施され、東京、長野、広島と大阪のがん登録データに基づいて、それぞれのがん登録の年齢罹患率の算出、4 登録合計の年齢調整罹患率の算出に成功、国がんでサーバーを立ち上げる事にも成功した。

また、国際共同研究としての頭頸部がんの相対生存率の比較にも成功し、各国間の差異についての検討を行った。

D. 考察

Vantage6 は将来的にその解決策の一つとなりうるが、現時点では、パッケージとして完成していないため、導入に一定以上の IT の知識が必要であること、各サイトの PC 環境、ネット環境、セキュリティ環境が大き

く異なる中で、データ交換においてエラーが生じた場合の、問題解決が非常に難しい。また、基本的な解析のアルゴリズムは開発済みであるが、複雑な解析のアルゴリズム解析には非常に長い時間がかかり、検証も大きな負担がかかる。

欧州、日本だけでなく、今後アジア諸国においても、個別情報の国外提供が困難になると考えられる。しかしながら、情報を様々な方法で匿名化する、や、オンサイト解析センターを作成するなどの解決策は、利用者の視点に立てば現実的ではなく、データ利活用促進のためには積極的に推し進めるべきではない。

Federated Learning の技術が全世界的に一般化されるには、技術的なハードルを下げ、利活用のためのマニュアル等がある必要がある。

E. 結論

欧州、日本だけでなく、今後アジア諸国においても、個別情報の国外提供が困難になると考えられる。Federated Learning の技術が全世界的に一般化されるには、技術的なハードルを下げ、利活用のためのマニュアル等がある必要がある。

