

創薬基盤推進研究事業 研究開発課題 事後評価報告書

研究開発課題名	産学連携全国がんゲノムスクリーニングプログラム:SCRUM-Japan を活用したプレジジョンメディスン推進基盤構築研究
代表機関名	国立研究開発法人国立がん研究センター
職名	分野長
研究開発代表者名	土原 一哉
全研究開発期間	平成 28 年度～令和 2 年度

1. 研究開発成果

細分化した治療対象集団の被験者を迅速に選定する産学連携、臓器横断的、系統的ながんゲノムスクリーニングネットワーク「SCRUM-Japan」に即して、より効率的、臓器横断的な基盤確立のため研究事務局と参加医療機関、共同研究企業間でリアルタイムに情報共有が可能、かつ被験者保護に十分に配慮したデータベースを構築、運用した。

2016年7月に国立がん研究センター内のオンプレミスサーバを用い臨床情報と遺伝子解析情報を統合するデータベースを試験運用した。その成果に基づき、2017度からはSCRUM-Japanデータセンターと医療情報保管に関する各種ガイドライン等に準拠したクラウドサーバを virtual private network (VPN)によるデータ通信を介して接続し、参加医療機関、共同研究企業の認証されたユーザーのみがアクセス可能なデータベースの本格運用を開始した。以降、LC-SCRUM、GI-SCREENのプロトコル改訂に伴う収集データ項目の変更等に対応したデータベース改修を行い、2019年7月までに14回のメジャーリビジョンを実施した。クラウドサーバに全解析例の臨床情報(TSVファイル)、ゲノムアノテーション情報(VCFファイル)、暗号化データ・インデックス領域あわせて約50GBを格納し認証ユーザーが閲覧、ダウンロード可能としたほか、全解析例のゲノム解析 raw data (BAMファイル)約11TBもクラウドサーバに保管し、定期的なバックアップを行なった。

第一期、第二期(2015-2019年度)をあわせ、参加施設はのべ265施設、肺がん5416検体、消化器がん5740検体の計11156検体のゲノム解析結果を取得した。参加医療機関より electronic data capturing システムを介して収集した全登録症例の臨床情報について専任のデータマネージャーがデータクリーニングを行い、日米2カ所のCLIA認証検査機関で実施した遺伝子解析情報を統合して上述の通りクラウドサーバに格納した。情報は原則として毎週更新し、研究期間中に計1004回のデータ更新を行った。治療経過、転帰に関する追跡調査は年1回実施した。

参加医療機関および共同研究企業の研究者計215名にアカウントを発行した。2021年3月までに施設ユーザーから18853回、企業ユーザーから20416回のアクセスがあり、ダウンロード権限を持つユーザー(企業ユーザーおよび研究申請を行った施設ユーザー)によって計1681回のデータダウンロードが行われた。解析結果が得られた全例の解析結果計48435パリアントについて、

遺伝子変異ごとのがん種、性別、年齢層別の発生頻度を AMED が管理するデータベース MGeND に登録し 2020 年 10 月に公開された。これは同月時点で MGeND に登録されたがん関連全バリエーションの 74%を占めている。

全例の解析後の余剰 DNA、RNA 試料および未染 FFPE 薄切標本を、入室管理、各種ログ管理などが保障された検体管理システムに委託し保管した。また GI-SCREEN に登録された大腸がん症例全例について、HE 染色標本のデジタル画像を取得した。検体 2 次利用について SCRUM-Japan 事務局が定めた手順に従い、LC-SCRUM、GI-SCREEN 各研究事務局が関与して計画された研究(産業利用を目的とした委受託研究を含む)への利用に供した。保管検体を用いた診断キット開発の結果、9 種の対外診断薬・システムが製造販売承認を取得した。

研究参加施設、企業から SCRUM-Japan で解析する遺伝子異常に関連する治験等の情報を収集しウェブサイトで公開するとともに、事務局から担当医に情報提供を行った。SCRUM-Japan に関連する治験は 2020 年 11 月時点で計 57 件である。特に全国の大学、がんセンターが研究事務局となる医師主導治験は 9 拠点 20 件を数え、新薬開発のプラットフォームとして国立がん研究センターのみならず広く全国のアカデミアに活用されている。企業治験、医師主導治験のうち 16 試験で結果が報告され、うち 9 剤 11 適応が国内承認された。

2020 年度より開始した第三期のプロジェクトにおいて、肺がんグループでは 2021 年 2 月末で台湾からの登録も含め 4288 例の登録があった。従来の組織検体を用いた次世代シーケンサーを用いた遺伝子検査に加え、末梢血を用いたリキッドバイオプシー検査、検査に要する期間が短い PCR 法を応用した検査を併用している。消化器がんグループは対象がん種を肺がん以外の固形がん全般に拡大し、リキッドバイオプシー検査によるスクリーニングおよび腸内細菌叢のゲノム解析を開始し、1442 例の登録があった。第三期のプロジェクトでは各グループで収集するデータが複雑化したため、研究グループでそれぞれ独自のデータベースを構築し、一方双方のデータベースから横断的に特定の遺伝子異常や患者背景で層別化し、閲覧、ダウンロードが可能となる横断検索データベースの構築を行なった。研究期間中にアカウントを発行した 59 ユーザーから 19333 アクセスの利用があった。

研究期間を通じて研究事務局、データセンター業務に携わる研究者、技術者を雇用し on the job training を行った。SCRUM-Japan 研究事務局、データセンター、参加医療機関、参加企業の代表者により構成される SCRUM-Japan 運営委員会に、参加医療機関以外の学識経験者としてクリニカルシーケンスについて先進的な取り組みを行っている国内アカデミアの研究者を招聘し、SCRUM-Japan によって得られたノウハウを積極的に共有した。

平成 28 年 7 月に東京で開催した SCRUM-Japan 成果報告会において米国 NCI-MATCH、英国 Lung Matrix など海外で構築中の被験者選定プログラム関係者と情報交換した。同様に平成 30 年 7 月の成果報告会に米国、台湾の研究者を招聘した。令和 2 年 7 月はオンラインで成果報告会を開催した。ゲノムバイオマーカーで層別化される進行大腸癌症例に対し分子標的薬の医師主導治験等を計画している米国、欧州、韓国の研究グループとゲノム、病理情報の標準化、臨床情報の一部を共有して利用する方法について開発を進め HER2 陽性大腸がんでの国際病理診

断基準を米国・韓国の研究者グループと共同で策定した。SCRUM-Japan 等のスクリーニングプログラムで発生するバイオマーカー陽性・標的薬未治療群の臨床情報を良質な historical control として各国で集積し、それらをメタアナリシスにより共有するプロジェクトを開始した。

国際的な認知度の向上により SCRUM-Japan に参加を希望するアジア各国の研究者が増加した。平成 31 年 3 月より台湾の Chang Gung Memorial Hospital が肺がんグループに参加しデータ共有を開始した。さらにその他の Asia-Pacific 諸国も加わった国際的な臨床ゲノムデータベース構築を進めている。

今後の方針について、創薬を目指した nation-wide および international screening プラットフォームとしての SCRUM-Japan の認知度を活かし継続した活動を継続する。また、がん新薬開発に求められる生物学的データは従来のがん組織のゲノム情報に加え、トランスクリプトーム、プロテオーム、マイクロバイオーム等の多層オミックス、腫瘍浸潤リンパ球などがん微小環境の情報に拡大している。これらのニーズに応えるために、新規技術を迅速に取り入れた産学連携による新たな data collaboration を企画し実行していく。

Using the "SCRUM-Japan", a systematic, tumor-agnostic, industry-academia collaborative cancer genome screening network that immediately selects patients for a subdivided treatment target population, we constructed and operated a database that allows real-time data sharing between the researchers in the participating medical institutions and collaborating companies, and gives due consideration to the protection of the patients' privacy, resulting in more efficient infrastructure.

In the first and second phases of the project (FY2015-2019), a total of 265 institutions participated, and genome analysis was performed on a total of 11261 samples, including 6360 lung cancer samples and 4901 gastrointestinal cancer samples. A database that integrates genome analysis data and clinical information was immediately shared among collaborators (215 users in total), and a total of 39269 accesses were made. In addition, frequency information of a total of 48435 variants of detected gene mutations was registered in the AMED integrated clinical genome information database MGeND. The number accounts for 74% of all cancer-related variants registered in MGeND. As a result of the development of diagnostic kits using stored specimens after genome analysis, six in vitro diagnostics have been approved. In the meantime, 57 investigator-initiated and pharma-sponsored clinical trials related to SCRUM-Japan have been conducted. Of these, results were reported from 16 trials, of which 9 drugs and 11 indications were approved in Japan.

In the third phase of the project, which started in FY2020, the Lung Cancer Group has registered 4288 cases as of the end of February 2021, including those from Taiwan. In addition to the conventional genomic testing using next-generation sequencers on tissue specimens, liquid biopsy testing using peripheral blood and PCR-based testing, which requires less time for testing, are being used in combination. The gastrointestinal cancer group has expanded the target cancer types to include

all solid cancers other than lung cancer and has started screening using liquid biopsy tests and genomic analysis of intestinal microflora, with 1442 cases registered. We constructed a database that allows collaborators to cross-search the data accumulated by both research groups, and 59 users who were issued accounts used the database with 1933 accesses.

As a future policy, we will continue our activities by taking advantage of the recognition of SCRUM-Japan as a nation-wide and international screening platform for drug discovery. In addition to the conventional genomic information of cancer tissues, the biological data required for the development of new cancer drugs is expanding to include multilayered omics such as the transcriptome, proteome, and microbiome, and information on the cancer microenvironment such as tumor-infiltrating lymphocytes. To meet these needs, we will plan and implement new data collaborations between industry and academia that rapidly incorporate new technologies.

2. 総合評価

大変優れている

【評価コメント】

臨床ゲノム情報統合システムは新薬開発に大きな貢献をするもので、「データの収集と共有」は十分達成できている。さらに横断検索システムの充実により利用率を向上させ、プラットフォームを活用した治験と薬事承認の件数は着実に増えている。当初の目標は達成されているので、更にデータシェアリングの推進が望まれる。今後の展開が重要であり、かつ期待ができる。

以上