

科学研究費助成事業 研究成果報告書

平成 29 年 4 月 20 日現在

機関番号：12601

研究種目：研究活動スタート支援

研究期間：2015～2016

課題番号：15H06008

研究課題名(和文) がんの薬剤耐性獲得による再発モデルの構築と薬剤応答予測の解析

研究課題名(英文) Statistical Modeling and Prediction for Therapy-induced Cancer Drug Resistance and Prediction

研究代表者

長谷川 嵩矩 (Hasagawa, Takanori)

東京大学・医科学研究所・助教

研究者番号：80753756

交付決定額(研究期間全体)：(直接経費) 2,200,000円

研究成果の概要(和文)：近年、多くのがん腫瘍において高い腫瘍内不均一性が確認されている。このようながん腫瘍に抗がん剤を投与した場合、ターゲットである優勢な感受性クローンは減少するが、感受性クローンにより抑制されていた劣勢な耐性クローンが優勢となり増殖を始め、薬剤耐性がんとして再発するメカニズムが提唱されている。本研究では、次世代シーケンス技術によって得られる腫瘍細胞のゲノム情報や血中マーカー情報と耐性がんの再発シミュレーションモデルをデータ同化により融合し統合的に解析する統計的解析手法を開発した。提案手法を用いることで、不均一性を持つがん腫瘍に対する効果的な薬剤の投与スケジュールの予測・提案が可能になる。

研究成果の概要(英文)：In recent years, high tumor heterogeneity has been confirmed in many cancer tumors. When an anticancer agent is administered to such a cancer tumor, although the target predominant sensitive clone decreases, the inferior resistant clone, which are suppressed by such predominant clone, becomes dominant and begins to proliferate as drug resistant cancer. In this study, we developed a statistical method that integrates genomic information and blood marker information of tumor cells obtained by the next generation sequencing technology and recurrence simulation model of tolerable cancer by data assimilation framework. This makes it possible to predict and suggest an effective drug administration schedule for cancer tumor with intratumoral heterogeneity.

研究分野：統計科学

キーワード：データ同化 統計科学 がん解析 シミュレーション

1. 研究開始当初の背景

日本に於けるがんの死者数は 2014 年時点で 368,103 例 (男性 218,397, 女性 149,706 例)となっており、2 人に 1 人程度が発症、3 人に 1 人程度ががんを原因にして死亡している。抗がん剤の投与はがんの治療における有力な選択肢の 1 つであるが、抗がん剤の投与によって縮小したがんが薬剤耐性を得て再発する事例は(ステージによって割合は異なるが)比較的一般的であり、がん治療における大きな障壁となっている。特に、薬剤耐性を持ったがんには以前に対処した薬剤が効かなくなっており、再発に適切に対処することがより困難になる。世界的に感染症による死者数が少なくなり、慢性疾患による死者数の割合が増えて来ていることもあり、これらの問題に対する処方法の確立は医療における一大テーマとなっている。

2. 研究の目的

近年のがん研究の進展により多くのがん腫瘍において高い腫瘍内不均一性を有することが明らかになって来た。このようながん腫瘍に抗がん剤を投与した場合、ターゲットである優勢な感受性クローンが減少するため、一時的ながん腫瘍の縮小が見られるが、感受性クローンにより抑制されていた劣勢な耐性クローンが優勢となり増殖を始め、薬剤耐性がんとして再発するメカニズムが提唱されている。本研究の目的は、次世代シーケンス技術によって得られる腫瘍細胞のゲノム情報や血中マーカー情報と耐性がんの再発シミュレーションモデルをデータ同化により融合し統合的に解析するための統計的解析手法を開発することである。これにより、がん腫瘍に対する薬剤応答を予測し、不均一性を持つがん腫瘍に対する効果的な薬剤の投与スケジュールの予測・提案することが可能となる。

3. 研究の方法

本研究計画は、(i)薬剤耐性がんの再発シミュレーションモデルの構築、(ii)シミュレーションモデルの隠れ変数とパラメータを推定する統計的解析手法の提案、(iii)提案手法のがん腫瘍への適用の三段階に分けることが出来る。初年度は主に、データの取得と文献の調査、また共同研究先である九州大学病院別府病院や東京大学医科学研究所との議論を通して再発シミュレーションモデルの構築を行い、最終年度である二年度目は、統計的解析手法の開発に着手した。申請者は既にデータ同化に関する論文を多数出版していたことから、これらの知見を利用することで速やかにシミュレーションモデルの解析手法を開発に着手することが出来た。具体的には、ABC(Approximate Bayesian Computation)並びに ABC における理論的問題の解決法として近年提案されている kernel-Approximate Bayesian Computation

(以下 kernel-ABC)をデータ同化のフレームワークに融合し汎用性を高めると共に、本研究に応用した。統計手法の開発後に、がん腫瘍データへの適用と投薬スケジュールの提案を行い、データが得られ次第解析に着手していくことになるが、がんのデータは動物実験と異なり計画通りにサンプリングすることが困難であることから、データの採取は継続して行うこととなった。

4. 研究成果

研究成果としては、既に国際学会において発表し、現在投稿中である論文に収録されている内容が成果の中心となる。実データの解析は、データが得られ次第着手していくことになる。具体的な成果を以下に述べる。

(1)本研究の候補モデルとして適切と考えられる、非線形性を持ち、同時点で複数の観測値を持つデータを扱い得る非線形の状態空間モデルを構築した。特に、観測ノイズの従う分布が一般的ではない場合に対しても適用可能な非線形状態空間モデルを考えた。ここで状態空間モデルとは以下に示される 2 つの数式(システムモデルと観測モデル)によって構築される時系列モデルの一種であり、本解析ではこのモデルを採用した。本研究では、実データを $y_{\{t\}}$ (時点 t)として与えたとき、未知パラメータ θ と未知状態変数 $x_{\{t\}}$ を推定するという問題を解くために、次に述べる統計的解析手法の開発を行った。

$$x_t = \int_{t-1}^t f(x(\tau), \theta) d\tau + x_{t-1} + v_t,$$

$$y_t = Hx_t + w_t,$$

(2)上述した非線形状態空間モデルのパラメータ推定に一般的に利用されるモンテカルロ法の一つである粒子フィルタは、観測データの時点数が増えたときに近似精度が著しく低下するパーティクル縮退の問題が知られている。また、尤度関数が解析的に計算出来なければ適用することが出来ない。本研究では、時点数が多く観測ノイズが一般的な分布に従わないようなケースを扱うこととなるため、粒子フィルタの適用は難しくなる。近年、この問題を解決するために、申請者を含むグループにおいて ABC を状態空間モデルに適用する新たな統計的解析手法が開発されているが、本研究ではこの手法を更に理論的に拡張する試みを行った。具体的には、ABC の理論的問題を解決する kernel-ABC と呼ばれる手法の適用を、状態空間モデルの潜在変数の条件付き分布を計算するフィルタリング・ステップに対して行う方法を開発した。kernel-ABC の状態空間モデルへの適用の優れた点としては、ABC においても得られた尤度が解析的に計算出来ない場合でも条件付き分布を得られるという利点に加え、以下の

2つの利点あげられる。
 (i)許容度パラメータが大きいときでもパラメータと潜在変数の理論的に正確な事後分布が計算することが可能(図1)

(ii)ABCの枠組を用いた上で、モデルの比較・評価を実施することが可能(図2)

これら既存のABCを用いたときのモデル選択における理論的問題を克服することにより、本研究課題である、より適すと考えられる再発モデルの選択とパラメータ推定が可能となる。これらの研究成果は国際会議において口頭発表を行った他、現在学術誌において投稿中である。

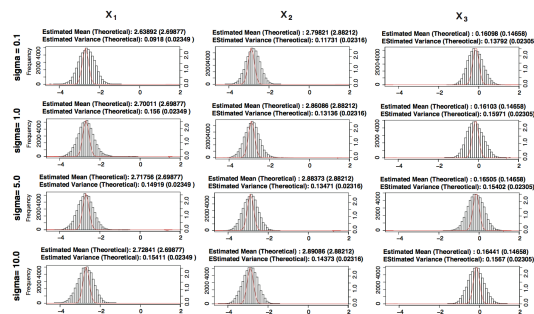


図1:許容度パラメータを0.1から10まで設定した際の、最終時点における隠れ変数 $x_{\{1,2,3\}}$ の理論分布(線分)と提案手法によって推定された確率分布(ヒストグラム)の比較図。提案手法において、理論分布を適切に表現出来ていることが分かる。

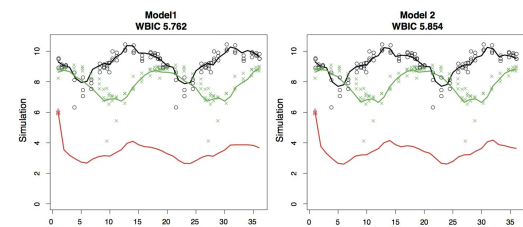


図2:提案手法を用いて、ラットのサーカディアンリズム関連遺伝子の時系列発現データを予測した結果。左図が正しいモデルで、右図が比較モデルを用いたときの予測結果である。提案手法を用いることで、正しいモデルが選択(WBICが低い値を取る)されることが分かる。

(3)シミュレーション結果の考察と各種観測データへの適用を行うため、既に研究が進んでいる前立腺がんと卵巣がんに対して提案手法を適用し、医学的解釈を行う。必要なデータは現在、共同研究先である九州大学病院別府病院において採取が計画されており、採取次第上記の提案手法を適用する。また、同手法を適用することが可能な公開データにアクセス可能になり次第、解析を行い、論文化・国際学会において発表する。特に近年では、リキッドバイオプシー(liquid

biopsy)と呼ばれる、内視鏡や針を使って腫瘍組織を採取する技術が発達しつつある。これにより、転移より十分に早い段階におけるがんを早期検出出来る可能性が高まっており、世界的に研究が活発である。今後はこのようなデータに対して適用することも検討していく。

(4)論文化とソフトウェアの公開
 本研究成果を一般に周知する上で、論文化と同時にソフトウェアでの公開を行う。ソフトウェア化は申請者がこれまでに公開しているソフトウェアと同様、Java言語での配布を考えており、所属研究機関のHP上から自由にアクセス出来るようにする。詳細に関しては、論文に記載するものとする。

5. 主な発表論文等
 (研究代表者、研究分担者及び連携研究者には下線)

〔雑誌論文〕(計 件)

〔学会発表〕(計 1 件)

(1) T.Hasegawa, K. Kojima, Y. Kawai, M. Nagasaki. Time-series filtering for replicated observations via a kernel approximate Bayesian computation. *9th International Conference of the ERCIM WG on Computational and Methodological Statistics 10th International Conference on Computational and Financial Econometrics*, 2016.

〔図書〕(計 件)

〔産業財産権〕

出願状況(計 件)

名称：
 発明者：
 権利者：
 種類：
 番号：
 出願年月日：
 国内外の別：

取得状況(計 件)

名称：
 発明者：
 権利者：
 種類：
 番号：
 取得年月日：
 国内外の別：

〔その他〕
 ホームページ等
<http://hic-imsut.jp/>

*提案手法に関するソフトウェアを順次公開
していく

6. 研究組織

(1) 研究代表者

長谷川 高矩 (HASEGAWA, Takanori)

東京大学・医科学研究所・助教

研究者番号：80753756

(2) 研究分担者

()

研究者番号：

(3) 連携研究者

()

研究者番号：

(4) 研究協力者

()