

機関番号：62603
研究種目：基盤研究 (C)
研究期間：2008～2010
課題番号：20590599
研究課題名 (和文) 個別化医療実現のための臨床研究のデザインと統計解析に関する研究
研究課題名 (英文) Design and analysis of clinical studies for personalized medicine
研究代表者 松井 茂之
(MATSUI SHIGEYUKI)
統計数理研究所・データ科学研究系・教授
研究者番号：80305854

研究成果の概要 (和文)：

個々の疾患の遺伝的特性に基づく個別化医療の実現のためには、分子診断法の開発が不可欠である。本研究は、分子診断法の開発における臨床研究のデザインとデータの統計解析について検討した。とりわけ、膨大な数の候補分子マーカーの中から、病型、薬剤反応性、生存期間などと関連する分子マーカーを同定して、予測システムを作製し、その臨床有用性を検証するための統計的方法に関する開発研究を行った。

研究成果の概要 (英文)：

Development of molecular diagnostics is essential for establishing personalized medicine based on genetic characteristics of individual patients. In this research, we studied design and analysis of clinical studies for developing molecular diagnostics. In particular, we developed statistical methods for identifying molecular markers associated with clinical variables such as disease type, response to treatment, and survival outcomes, from a large pool of candidate molecular markers, making prediction systems using the identified molecular markers, and testing their clinical utility.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2008年度	1,200,000	360,000	1,560,000
2009年度	1,100,000	330,000	1,430,000
2010年度	1,200,000	360,000	1,560,000
年度			
年度			
総計	3,500,000	1,050,000	4,550,000

研究分野： 医歯薬学

科研費の分科・細目： 社会医学・衛生学

キーワード： 医学統計、ゲノム薬理学

1. 研究開始当初の背景

近年のヒトゲノムの全塩基配列の解読やゲノム検査技術の著しい進歩は、個別化医療の実現を大きく後押しするものである。分子診断法の開発では、医薬品の開発と同様に、

人を対象とした臨床研究を実施して、科学的根拠を示す必要がある。しかしながら、この分野での臨床研究の研究デザインとデータ解析に関する統計的方法論・方法は十分確立していない。これを解決することは分子診断

法の開発における喫緊の課題となっている。

2. 研究の目的

分子診断法の開発のための臨床研究のデザインとデータ解析の開発研究を行い、この分野での統計的方法論・方法の確立をめざす。

3. 研究の方法

臨床研究のデザインとデータ解析に関する新しい統計的方法の開発と評価を行う。とりわけ、方法が十分確立していない分子診断法の開発早期の段階での関連遺伝子のスクリーニング、予測アルゴリズムの作製と予測精度の評価の方法については重点的に検討する。開発した方法の性能評価は、実データへの適用と疑似乱数を用いた数値実験により行う。また、分子診断法の開発ストラテジーを検討するために、各開発フェーズを結びつけた統計モデルとデザインツールの構築も試みる。

4. 研究成果

(1) ゲノムワイドデータを用いた予測解析
マイクロアレーを用いたゲノムワイド研究において、臨床変数（病型、薬剤反応性、生存期間など）と関連する多くの遺伝子・マーカーのサブセットの中から特に予測に役立つサブセットを抽出する新しい予測法を開発した。この方法を多発性骨髄腫の臨床研究データに適用したところ、従来の予測法よりも優れた予測能が示された。

(2) ゲノムワイドスクリーニングでのサンプルサイズ設計

ゲノムワイド研究において、関連遺伝子のスクリーニングを行うときの偽陽性度の指標の一つに偽発見割合（false discovery proportion）がある。この指標に基づく必要症例数設計法を新たに提案した。これは、遺伝子間の相関によって生じる偽発見割合のばらつきと関連遺伝子の選択割合である感度のばらつきを同時にコントロールする方法である。数値実験により、偽発見割合の期待値である偽発見率（false discovery rate）と感度の期待値を同時にコントロールする従来法よりも、特に、遺伝子間の相関が高い場合には、より適切な方法であることが示された。

さらに、関連の強い遺伝子をスクリーニングするための必要症例数の設計法も新たに開発した。これは関連の強さに基づいて遺伝子のランキングを行うものであり、関連遺伝子を同定するだけの従来の多重検定（multiple testing）による解析とは一線を画すものである。

(3) 階層混合モデルと経験ベイズ推測

関連遺伝子の同定、関連の大きさ（効果サイズ）の推定、効果サイズに基づく遺伝子ランキング、さらに、関連遺伝子のサブセットを用いた期待予測能の推定を行うための統一した解析の枠組みをベイズ階層混合モデルのもとで構築した。関連遺伝子の効果サイズの分布にはノンパラメトリック事前分布を仮定する。このモデルによって、選択した上位遺伝子の効果サイズが過大評価されるバイアスの補正と、遺伝子間での情報共有による各種統計的推測の効率化を同時に達成することができる。前立腺がん、乳がん、肺がん、多発性骨髄腫など、多くの癌種のマイクロアレーデータ、並びに、数値実験データにおいて、多くの従来法と比較として、開発法の有効性を示すことができた。

提案した階層混合モデルと経験ベイズ推定の枠組みは、きわめて汎用性が高く、かつ、柔軟でロバストな方法であり、様々なタイプのゲノムワイドデータに適用可能である。さらに、選択バイアスの補正と効率向上の観点から、各種の統計的指標を導出するための枠組みとしても今後有用である。これより、今後のゲノムワイドスクリーニング法の開発研究において、本研究は大きなインパクトを与えるものと考えられる。

(4) がんの個体差のモデリング

がんの個体差の検出のためのモデルに基づくクラスター分析法を新たに開発した。数値実験により、一部のがんサンプルのみで発現量が変動している遺伝子のクラスターをうまく検出できることを確認した。これを白血球データに適用したところ、従来のクラスター法では検出できなかったがんの病型・予後に関わる新たながんのクラスター（分類法）を発見することができた。

(5) 臨床的有用性の評価のための臨床試験デザイン

治療法開発のための臨床試験において、分子診断法を用いて研究対象の限定や層別を伴う試験デザインについて、先行研究のレビューと実際の適用例に基づいた性能比較を行った。さらに、分子診断法の開発での統計的方法の確立に向けて、ゲノムワイドデータを用いた治療効果の予測マーカーの開発、そのバリデーション、予測マーカーを用いた治療効果の検証に至るすべてのプロセスを通して、研究デザインとデータ解析の方法について整理した。

なお、以上の研究成果の多くは、Biometrics, Biostatistics, Statistics in Medicine などの生物統計学の分野で最も権威のある国際誌に掲載されている。

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 20 件)

- ① Matsui S, Noma H. Estimation and selection in high-dimensional genomic studies for developing molecular diagnostics. *Biostatistics*, 査読有, 12 巻, 2011, 223-233.
- ② Noma H, Matsui S. Optimality of gene ranking based on univariate P-values for detecting differentially expressed genes. *Japanese Journal of Biometrics*. 査読有, 31 巻, 2010, 13-21.
- ③ Morita S, Thall PF, Müller P. Evaluating the impact of prior assumptions in Bayesian biostatistics. *Statistics in Biosciences*, 査読有, 2 巻, 2010, 1-17.
- ④ Matsui S, Oura T. Sample sizes for a robust ranking and selection of genes in microarray experiments. *Statistics in Medicine*, 査読有, 28 巻, 2009, 2801-2816.
- ⑤ Oura T, Matsui S, Kawakami K. Sample size calculations for controlling the distribution of false discovery proportion in microarray experiments. *Biostatistics*, 査読有, 10 巻, 2009, 694-705.
- ⑥ Morita S, Thall PF, Bekele BN, Mathew P. A Bayesian hierarchical mixture model for platelet derived growth factor receptor phosphorylation to improve estimation of progression-free survival in prostate cancer. *Journal of the Royal Statistical Society, Series C (Applied Statistics)*, 査読有, 59 巻, 2009, 19-34.
- ⑦ Inoue A, Kobayashi K, Usui K, Maemondo M, Okinaga S, Mikami I, Ando M, Yamazaki K, Saijo Y, Gemma A, Miyazawa H, Tanaka T, Ikebuchi K, Nukiwa T, Morita S, Hagiwara K; North East Japan Gefitinib Study Group. First-line gefitinib for patients with advanced non-small-cell lung cancer harboring epidermal growth factor receptor mutations without indication for chemotherapy. *Journal of Clinical Oncology*, 査読有, 27 巻, 2009, 1394-1400.
- ⑧ Matsui S, Zeng S, Yamanaka T, Shaughnessy J. Sample size

calculations based on ranking and selection in microarray experiments. *Biometrics*, 査読有, 64 巻, 2008, 217-226

- ⑨ Matsui S, Yamanaka T, Barlogie B, Shaughnessy JD Jr, Crowley J. Clustering of significant genes in prognostic studies with microarrays: application to a clinical study for multiple myeloma. *Statistics in Medicine*, 査読有, 27 巻, 2008, 1106-1120

[学会発表] (計 8 件)

- ① 松井茂之. がんの分子診断マーカーの臨床開発：現状と課題. 統計関連学会連合大会, 2010年9月7日, 早稲田大学 (招待講演)
- ② Terashima M, Saze Z, Hosoya A, Takahashi M, Takagane A, Hachiya O, Koeda K, Matsui S, Kusunoki Y, Oohashi W, Fukushima M, Gotoh M. Results of a phase II multicenter study of neoadjuvant S-1 and irinotecan in patients with locally advanced gastric cancer. 2010 Gastrointestinal Cancers Symposium. The American Society of Clinical Oncology, 2010年1月22日, Orlando, Florida.
- ③ 松井茂之. がんの分子診断研究：FDRは有用か？ 2008年度日本計量生物学会年会特別セッション. 2008年6月5日. 筑波大学.
- ④ Morita S. Gefitinib combined survival analysis of the mutation positives from the prospective phase II trials (I-CAMP). American Society of Clinical Oncology. 2008年5月30日, Chicago.

[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]

ホームページ等: なし

6. 研究組織

(1) 研究代表者

松井 茂之 (MATSUI SHIGEYUKI)

統計数理研究所・データ科学研究所・教授
研究者番号: 80305854

(2) 研究分担者

森田 智視 (MORITA SATOSHI)
横浜市立大学・医学研究科・教授
研究者番号：60362480

(3) 連携研究者

山中 竹春 (YAMANAKA TAKEHARU)
九州がんセンター・臨床研究部・室長
研究者番号：00325466

大門 貴志 (DAIMON TAKASHI)
兵庫医科大学・医学部・講師
研究者番号：40372156