

消化器がん発生に関する遺伝子多型と生活習慣のビッグデータ解析と 革新的なリスク解析プラットフォームの開発

九州大学病院 別府病院 外科

特任准教授 杉町 圭史

(共同研究者)

九州大学病院 別府病院 外科

教授 三森 功士

九州大学大学院 システム情報科学研究院

教授 竹田 正幸

はじめに

発がん予防医学の推進は、治療費の増加に伴う医療費高騰を国家的課題とするわが国にとって最も重要な分野である。我々の研究室は、これまでに全国の多施設共同研究を主導し、大腸癌と食道癌及び対照群（健常者）併せて約7,500症例を母集団として、全ゲノム相関解析(GWAS)と生活習慣に基づく疫学因子の収集を行い、それぞれの発がん関連遺伝子多型と発がん関連疫学環境要因を明らかにしてきた^(1,2)。これらの研究より大腸癌・食道癌のがん関連遺伝子異常や、喫煙・飲酒・肉食・糖尿病などの発がんに関する基礎的な疫学因子が明らかになったが、これらのデータ量が膨大であるため発がんリスクを簡便かつ高精度に分類することが困難であった。そのため、本申請課題では医学研究者とシステム情報研究者が共同して発がん予測に有用な因子と組み合わせを統計学的に解析し、がん発症リスクを高精度かつ簡便に算出できるプラットフォームを開発し、臨床応用可能な発がんリスク評価システムを確立することを目的とした(図1)。これらの基礎データよりスクリーニング検診に応用することができれば、がんの早期発見に繋がり国民の健康増進への貢献は大きいと考えられる。

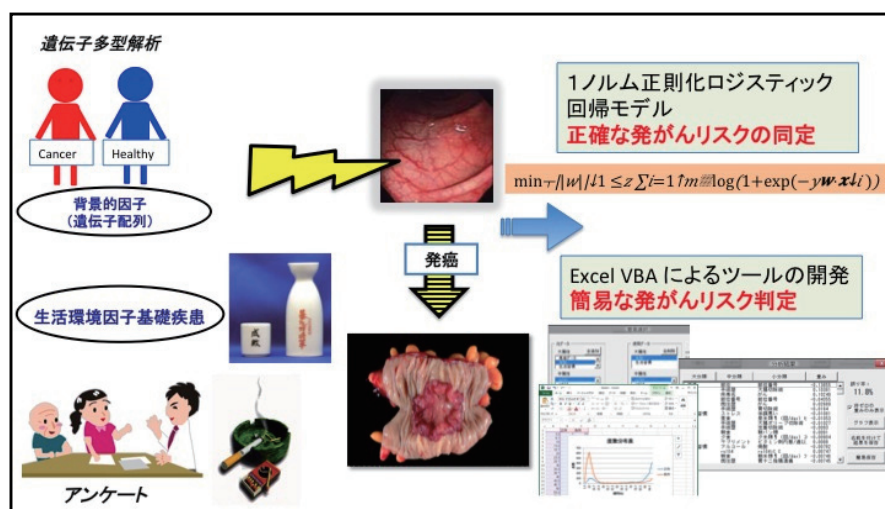


図 1. 臨床応用可能な発癌リスク評価システムの開発

結 果

①先行研究にて全4,902例（大腸がん症例群1,917例、健常人対照群2,985例）の食事・運動などの生活習慣、既往歴、家族歴などのアンケート結果を集積していた。これらのデータを今回の統計解析（1ノルム正則化ロジスティック回帰モデル）に使用するために全データの符号化を行った（図2）。

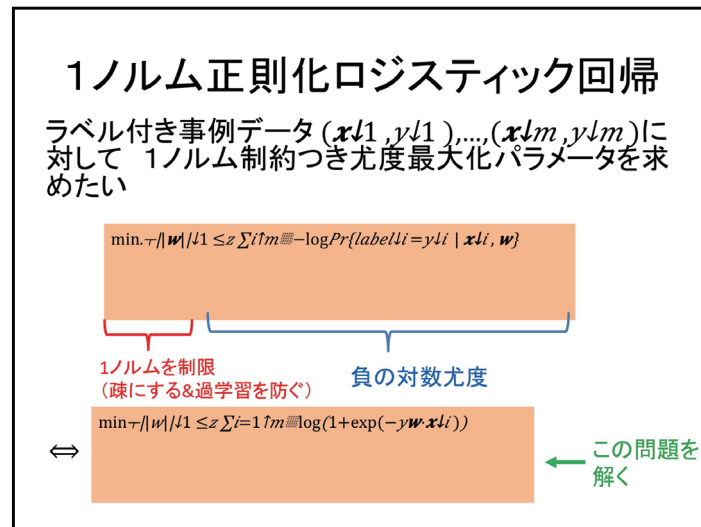


図 2. 発癌リスク評価のための統計解析

②大腸がんに関する遺伝子データベース（遺伝子多型とゲノム変異）と今回の生活習慣調査結果の統合的統計解析を行い、発がん危険因子を抽出した。手法として疎超平面が得られ、かつ確率的な解釈が可能である「1ノルム正則化ロジスティック回帰モデル」を適用した（図3）。

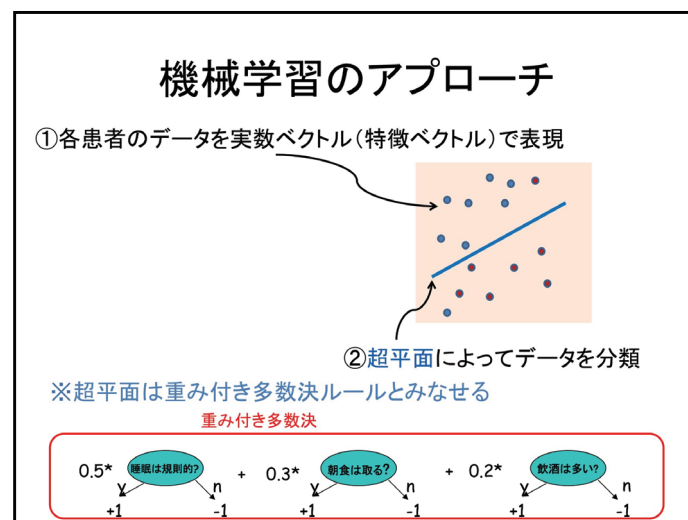


図 3. 発癌危険因子抽出のためのモデル

③これらの統計解析結果を汎用するためのExcel（Microsoft社）VBA による「がんリスク判定ツール」を製作した（図4, 5）。

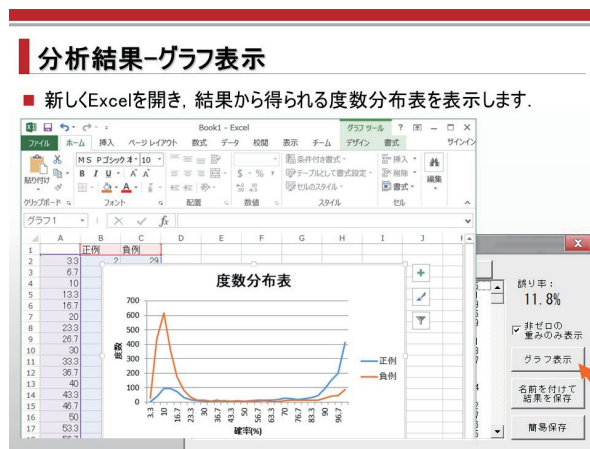


図 4. がんリスク評価ツールにおける解析結果のグラフ表示

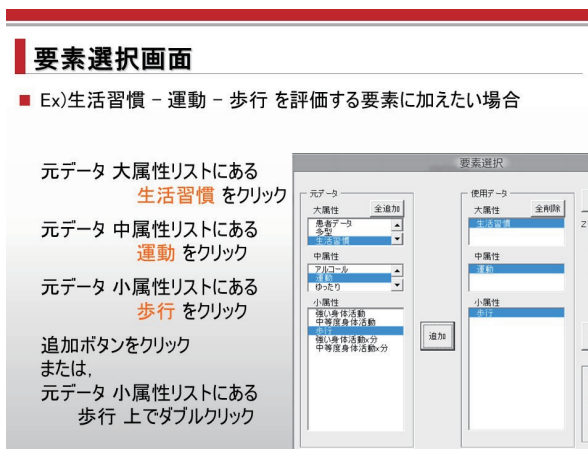


図 5. がんリスク評価ツールにおける解析要素選択画面

④今回作成して「がんリスク判定ツール」により、病歴や生活習慣などの項目を入力するだけで、大腸発がんのリスク判定を行うことができた。

重 み	項 目
0.100	大腸がん切除の既往
0.010	がんの既往
-0.018	胃切除の既往
-0.001	ストレス／体調悪い
-0.001	昼食／米類摂取

表 1. がんリスク判定ツールを利用した大腸発がん危険因子の抽出

考 察

がんの発生リスクを予測するための疫学研究においては、従来は数百例程度の観測による分析や、疾患に関わる因子の組み合わせを高々 2 項目程度にとどめての分析が主であったため、情報工学の手法を取り入れ悪性腫瘍の発症リスク診断システムを実用化した例はこれまでに無かった。今回我々が行った数千例のビッグデータを用いて、多数の因子の組み合わ

せを分析する試みは画期的であり、今回生物化学と情報工学の共同研究により画期的な発がん予測プラットフォームを開発することを目標に研究を行った。従来の診断システムでは、個性を切り捨てた形（最大公約数的）での統計分析がなされる傾向にあったが、本研究では、異種データが混在する状況で多種の分析をするデータベース工学上の問題に対する解決策が見出される可能性があると考えている。さらに本研究では、一般医療従事者や一般国民にも使いやすいように、データを入力するだけで視覚的に発癌リスクが即答されるような革新的かつ簡便なプラットフォームの構築を行うことができた。

今回作成した大腸発ガンリスク判定ツールを応用することで、大腸がん検診をうける患者のスクリーニングが可能になると考えている。発ガン高危険因子グループに対しては通常よりも検診間隔を短くしてより精度の高い検査を行うことで、癌を早期の段階で発見し早期治療を行うことにより治癒をえることができる。一方で発癌低危険因子グループに対しては、検査の密度を下げることで患者や医療者の負担を軽減し、さらに医療経済の面に対する貢献も大きいと考えられる。今回は、我々が持っている大腸がんデータベースをもとに発現リスク解析ツールを作成したが、このプラットフォームは今後爆発的に増加するであろう種々の疾患のビッグデータの解析において比較的簡便に応用転用が可能であると考えられる。

要 約

発がん予防医学の推進は治療費の増加に伴う医療費高騰を国家的課題とするわが国にとって最も重要な分野である。我々は先行研究において日本人における大腸癌および食道癌における発がん関連因子（発がん関連遺伝子多型と生活環境上の発がんリスク因子）のデータを収集した。しかし、どの発がん因子がどれだけ重要か？あるいは、どの因子の組み合わせをスコア化し調べれば、発がんを予測しうるのか？といったことが明らかではなかった。本研究では、先行研究で集積した数千例の大腸がん患者および健常人の遺伝子データ及び生活習慣に関わるデータに基づき、大腸発がん予測に有用な因子と組み合わせを統計学的に解析し、がん発症リスクを高精度かつ簡便に算出できるプラットフォームを開発し、臨床応用可能な大腸発がんリスク評価システムを作成した。Microsoft Excelを使用したシステムであり比較的簡便に大腸発癌リスクを計算することが可能になった。本研究における大腸発がんリスク評価システムを応用すれば、生活習慣のアンケートと血液検査による遺伝子多型検査のみで、大腸発がん高リスクグループを簡便かつ迅速に同定することができる。高リスクグループにがん検診等を積極的に行うことで、がんの早期発見・早期治療が可能となり、結果的にがん死亡率の低下や医療費の抑制へとつながることが期待される。また、プラットフォームを確立すれば消化器がんだけでなく他疾患のリスク評価にも広く応用可能であると考えている。

文 献

1. Tanaka F, Yamamoto K, Suzuki S, Inoue H, Tsurumaru M, Kajiyama Y, Kato H, Igaki H, Furuta K, Fujita H, Tanaka T, Tanaka Y, Kawashima Y, Natsugoe S, Setoyama T, Tokudome S, Mimori K, Haraguchi N, Ishii H, Mori M. Strong interaction between the effects of alcohol consumption and smoking on oesophageal squamous cell carcinoma among individuals with ADH1B and/or ALDH2 risk alleles. *Gut*. 59 (11) :1457-64; 2010.
2. Ishimaru S, Mimori K, Yamamoto K, Inoue H, Imoto S, Kawano S, Yamaguchi R, Sato T, Toh H, Inuma H, Maeda T, Ishii H, Suzuki S, Tokudome S, Watanabe M, Tanaka J, Kudo SE, Sugihara K, Hase K, Mochizuki H, Kusunoki M, Yamada K, Shimada Y, Moriya Y, Barnard GF, Miyano S, Mori M. Increased risk for CRC in diabetic patients with the nonrisk allele of SNPs at 8q24. *Ann Surg Oncol*. 19 (9) :2853-8; 2012.