

多型情報、病名、有害事象である。

C. 研究結果

本年度においては、がん予防薬として 140 種がリストにまとめられた。データソースの表記の例として、薬物名のクエン酸タモキシフェン 10mg 錠をあげると、HOT7 コードとして <HOT7>1092059</HOT7>となる。代謝酵素名としてはチトクローム P-450、遺伝子多型情報としては、代謝亢進として<代謝>CYP3A4 </代謝>、病名の項目として<効能・効果> 乳がん <MedDRA>10006192</MedDRA> <EBM>1a</EBM> </効能・効果>というように、乳がんの MedDRA コードが 10006192 とわかるように記し、またこの効能効果のエビデンスレベルとしては 1a であるとなっている。さらに有害事象について、相互作用で代謝促進がリファンピシンとの間であることから、<相互作用>代謝促進 <MedDRA>10013678</MedDRA> <薬物名>リファンピシン <HOT7>1112771 </HOT7> </薬物名>というような記述になり、このエビデンスレベルは <EBM>4</EBM>とのように記述された。

D. 考察

データベース化するにあたりコードなどの標準化が重要であることから、MedDRA や HOT7 によりコード化を行ったが、かなり時間と労力を要することとなった。もともと

添付文書などで表記される病名等は標準化されていれば本作業は容易となる。また相互作用などで、一方の添付文書では記載されているが対応するもう一方の添付文書では記載されていないなかったり、薬剤分類で表記されているなど、これも標準化されていない状況である。これを補う意味で本研究は役立つものと考えられた。今後はさらにサプリメント情報の拡充などを行うこととする。これまで様々ながん関連情報が散在しており、特に添付文書においても効能効果や副作用などの病名の表記が標準化されておらず、検索などの利用が困難となっていた。しかし本研究により、病名などを国際標準である MedDRA を用いて標準化することで、今後のデータベースなどでの検索が効率的かつ質が高くなることが期待される。また相互作用などで、薬物相互の記述も対応取れるようにしたため、知識ベースとしての相互の抜け落ちもなくシステムとしての臨床支援も組みやすいと考えられる。

E. 結論

がん予防薬の薬理遺伝学情報データベースの基礎となる予防薬を検索してリスト化した。リストには薬物名、代謝酵素名、遺伝子多型情報、有害事象、病名等を含め、その中で使用される特に病名(効能・効果や有害事象名など)について国際標準のMedDRAに基づ

いてコード化し、薬物名についてはまだ国際標準がないことから国内標準のHOT7を用いた。

がん予防情報の解析についてはEBMの手法に基づいた質的評価も加えた。

F. 研究発表

1. 論文

- 1) 小出大介, 山崎力: エビデンスに基づく

臨床ガイドライン. 臨床と薬物治療
23(1):68-71.2004.

2. 学会発表 準備中

G. 知的財産権の出願・登録状況 (予定を含む)

1. 特許取得 なし
2. 実用新案登録 なし
3. その他 なし

厚生労働科学研究費補助金(第3次がん総合戦略研究事業(分野3))

(分担)研究報告書

分担研究課題:「がん予防対象患者の病院情報システムからの抽出法に関する調査研究」

分担研究者:大江 和彦

所属: 東京大学大学院医学系研究科医療情報経済学分野

研究要旨:悪性腫瘍患者、がんの初発、治療の実施、治療による軽快または寛解、の複合条件検索によって一般の病院情報システムから患者抽出する方法を検討した。病名は標準病名マスターが普及しつつあり、これを調査したところ、悪性腫瘍患者であることは、理論上はICD10コードにより抽出可能である。しかし悪性腫瘍であるにもかかわらず悪性腫瘍のICD10コードを振られないケースが少なからず存在することも明らかとなった。これらは悪性・良性の区別を明示しないICD10コードが割り当てられるケースであり、具体的には前立腺腫瘍、卵巣腫瘍、皮膚腫瘍、骨腫瘍のように臨床診断名登録時には悪性・良性の区別がつかず、その後の病理検査もしくは術中・術後病理診断によって悪性が判明する場合であって、その判明結果が最終的に病院情報システムに入力されないままとなるケースがあることがわかった。

A. 研究目的

一般的な病院情報システム（オーダーリングシステムまたは電子カルテ）における初発がん治療患者抽出の可能性と解決方法の提案を目的とした。

B. 研究方法

悪性腫瘍患者、がんの初発、治療の実施、治療による軽快または寛解、の複合条件検索によって一般の病院情報システムから患者抽出する方法を検討した。

C. 研究結果

病名は標準病名マスターが普及しつつあり、これを調査したところ、悪性腫瘍患者であるこ

とは、理論上はICD10コードにより抽出可能である。しかし悪性腫瘍であるにもかかわらず悪性腫瘍のICD10コードを振られないケースが少なからず存在することも明らかとなった。これらは悪性・良性の区別を明示しないICD10コードが割り当てられるケースであり、具体的には前立腺腫瘍、卵巣腫瘍、皮膚腫瘍、骨腫瘍のように臨床診断名登録時には悪性・良性の区別がつかず、その後の病理検査もしくは術中・術後病理診断によって悪性が判明する場合であって、その判明結果が最終的に病院情報システムに入力されないままとなるケースがあることがわかった。

D. 考察

これを解消するには、1) 臨床診断名登録時に性状不明の腫瘍病名が登録されているケースにおいて、退院時または病理検査結果報告時に再度確定診断名を登録するように自動的にシステムで促す仕組みを導入する、2) 病理検査結果から自動的に悪性か良性かを抽出する仕組みを導入する、などが必要である。次に初発かどうかについては、既存のシステム内情報だけでは困難であり、明示的に悪性腫瘍病名登録時に入力させる必要がある。治療の実施とそれによる軽快については、腫瘍のICD10分類ごとに行われる治療手技をマスターに設定しておき、これをもとにその治療が実施されたかどうかを抽出すること、および治療結果を評価可能な特性値（たとえば腫瘍マーカー）を事前登録しておきその推移により判定することが考えられた。しかし特定の評価特性値が存在しないような腫瘍の場合には、退院時または治療終了時に治療効果を一定の区分で入力できるようなシステムとすることが必要であった。

病院情報システムのデータのうち病名登録機構の整備は必須である。すでに標準病名マスターが整備され普及しつつあるが、これを病院情報システムに導入する際に必須の機能として、悪性・良性不詳の病名を登録した場合に最終診断を再度登録させるようなシステム機能の必要性が明らかになったことは、今後

の標準病名マスターの普及方策に大きな示唆を与える。また別に進められている標準的電子カルテ機能のモデル化にも影響を与えるものと期待される。また、病理診断報告書からの良性、悪性の診断結果をICD10もしくはSNOMEDコードにより自動コーディングする機構の開発が必須であったことから、これを提案することによって今後の病理診断データベースの設計に影響を与えるものと考えられる。

E. 結論

一般的病院情報システムからのがん予防対象患者抽出上の課題として悪性腫瘍(例:前立腺腫瘍、卵巣腫瘍など)の病名登録機能の課題を明らかにした。

F. 研究発表

1. 論文

- 1) Onogi Y, Ohe K, Tanaka M, Nozoe A, Sasaki T, Sato M, Kikuchi Y, Shinohara T, Suzuki H, Kaihara S, Seyama Y: Mapping Japanese Medical Terms to UMLS Metathesaurus. in Proceedings of the 11th World Congress on Medical Informatics. Sep.7-11,2004, San Francisco, California, USA. (EDS) M. Fieschi et al. p406-410, 2004.
- 2) 大江和彦: 標準的電子カルテの考え方と機能モデル. 新医療 総特集 社会資本としての電子カルテ 標準化の方法論, Vol.31(7), 73-76, 2004.

- 3) 澄田有紀、高田真美、梁瀬鐵太郎、神代知範、大江和彦: 電子カルテシステムの機能モデル表現の検討. 第24回医療情報学連合大会論文集, 532-533, 2004.
- 4) 高田真美、澄田有紀、梁瀬鐵太郎、神代知範、大江和彦: 電子カルテ稼働病院の調査に基づくシステム機能の分析. 第24回医療情報学連合大会論文集, 1140-

1141, 2004.

2. 学会発表 発表予定

G. 知的財産権の出願・登録状況 (予定を含む)

1. 特許取得 なし
2. 実用新案登録 なし
3. その他 なし

厚生労働科学研究費補助金(第3次がん総合戦略研究事業(分野3))

(分担)研究報告書

研究課題名:「がん検診画像データからの予防情報抽出に関する研究」

分担研究者 氏名: 若尾 文彦

所属:国立がんセンター中央病院

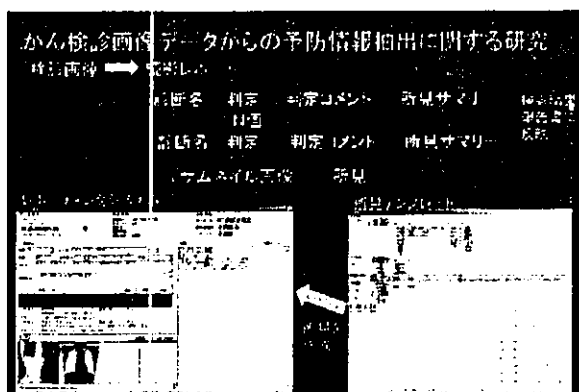
研究要旨: 検診の画像診断報告書の効率的作成をを支援する所見テンプレートを考案し、所見テンプレートを実装した検診レポートシステムを構築し、実際に、検診の診断レポート作成に用いて、その利用状況について、解析を行った。診断-所見サマリー-所見テンプレートによる所見入力-サムネイル画像を用いることで、検診の際に発生する大量の画像情報を効率的にデータベース化する事が可能となり、画像情報の抽出に有用であると考えられた。

A. 研究目的

検診の画像診断報告書の効率的作成をを支援する所見テンプレートを考案し、所見テンプレートを実装した検診レポートシステムを構築し、実際に、検診の診断レポート作成に用いて、その有用性について評価することを目的とした。

B. 研究方法

がん検診で発生する大量の画像情報から



簡便な操作で、診断報告書を作成する「検診レポートシステム」および、診断報告書から検診結果報告書を作成する「判定登録システム」を構築し、検診業務の中で利用し、がん検診画像情報の登録状況を解析した。

検診レポートシステムでは、入力の手間を軽減するために、臓器-診断-判定-判定コメント-所見サマリーで構成されるセットを作成し、診断-判定のセットを選択することにより、他の項目が選択され、簡単な操作で入力できる設計した。さらに、詳細な所見を入力するために、所見テンプレートとして、病変の詳細情報を簡単に記載できるシステムを構築した。

C. 研究結果

本検診レポートニングシステムを検診業務の中で利用し、その利用状況について解析を行った。対象は、2004年2月～2005年2月に実施された肺CT検診3,911件で、登録された総診断数は4,222件で、診断数1個のものが3,703件、2個のものが150件、3個のものが37件、4個のものが12件、5個以上のものが9件で、最大で9個の診断がつけられていた。また、3,911中、3,634件93%に1個以上のサムネイル画像が添付され、画像1個が713件、画像2個が904件、画像3個が750件、4個が512件、5個以上は755件で最大10個で、総計11,505個の画像が貼り付けられていた。また、所見が記載されたものは、3,444件(88.1%)であった。つまり、検査1件あたり平均1.08個の診断、及び、平均2.94個のサムネイル画像がつけられ、88.1%に所見が記入されていたことになり、複雑な所見について、所見テンプレートを利用することで、多くの所見入力がされていることが確認された。ただし、今回のシステムでは、所見フィールドが1検査について、1つであり、各診断と所見との紐付けが、困難で今後の改修が必要であると考えられた。診断別では、肺内の結節影(5mm以上)が、1,345件、肺内の結節影(5mm未満)が、1,080件、胸膜肥厚像433件と多かった。

所見テンプレートを有する検診レポート

ニングシステムを用いて、肺CT検診時に発生する大量の画像情報を診断+所見サマリー+所見+サムネイル画像としてデータベース化することが可能となった。

D. 考察

今後、肺結節の経過観察の結果を結節データベースとして登録し、解析することは、CT検診における肺野小結節のフォローアップ指針作成に有用であると考えられた。

E. 結論

診断-所見サマリー-所見テンプレートによる所見入力-サムネイル画像を用いることで、検診の際に発生する大量の画像情報を効率的にデータベース化する事が可能となり、画像情報の抽出に有用であると考えられた。

F. 研究発表

1. 論文

- 1) Yamashiro K, Kawamura N, Matsubayashi S, Dota K, Suzuki H, Mizushima H, Wakao F, Azumi N.: Telecytology in Hokkaido Island, Japan: result of primary telecytodiagnosis of routine cases. *Cytopathology* 15:221-227,2004.
- 2) 若尾文彦: 画像検査部門システムの概要と導入の考え方. *IT vision* 6:12-16,2004.
- 3) 飯沼 元、内山菜智子、村松幸男、森山紀之、立石宇貴秀、宮川国久、若尾文彦、佐竹光夫、赤須孝之: CT, MRI検査における大腸癌の術後再発診断と新たな診

- 断技術。早期大腸癌 8:139-144,2004.
- 4) 飯沼 元、内山菜智子、宮川国久、若尾文彦、立石宇貴秀、佐竹光夫、村松幸男、森山紀之：大腸癌術前における multi-detector row CT colonography の有用性。臨床放射線 49:409-418, 2004.
- 5) 飯沼 元、富松英夫、斎藤 博、村松幸男、森山紀之、宮川国久、若尾文彦：消化管造影検査におけるFPD-DR。カレントセラピー 23:17-21, 2005.
2. 学会発表 発表予定
- G. 知的財産権の出願・登録状況（予定を含む）**
1. 特許取得 なし
 2. 実用新案登録 なし
 3. その他 なし

厚生労働科学研究費補助金(第3次がん総合戦略研究事業(分野3))

(分担)研究報告書

研究課題名:「がん予防薬の臨床試験支援用情報システム構築に必要な機能仕様に関する調査研究」

分担研究者氏名: 山本 精一郎

所属:国立がんセンター予防・検診研究センター情報研究部

研究要旨: 米国国立がん研究所の行っている標記システムについて検討した。米国では、実験室から臨床試験まで、即ち、薬剤発見から、臨床応用までを一貫して国がサポートしている。システムとして、前臨床から臨床試験に移行するために必要なデータ収集、手続き支援、早期臨床試験のマネジメント、後期臨床試験の組織化をサポートしている。また、恒常的に研究計画をレビューし、結果を評価する国の機関が存在し、そこでは、国内外の薬剤開発状況を把握し、情報を研究者と共有し、研究の重複を防ぎ、研究が予防・治療が必要な疾患をあまねくカバーするよう調整していたことが判明した。

A. 研究目的

米国では、実験室から臨床試験まで、即ち、薬剤発見から、臨床応用までを一貫して国がサポートしている。システムとして、前臨床から臨床試験に移行するために必要なデータ収集、手続き支援、早期臨床試験のマネジメント、後期臨床試験の組織化をサポートしている。また、恒常的に研究計画をレビューし、結果を評価する国の機関が存在し、そこでは、国内外の薬剤開発状況を把握し、情報を研究者と共有し、研究の重複を防ぎ、研究が予防・治療が必要な疾患をあまねくカバーするよう調整している。

本研究の目的を、米国国立がん研究所の

行っている標記システムについて調査検討し、がん予防に関する国としてのシステム化されている内容とその機能について明らかにすることとした。

B. 研究方法

本年度は2003年7月～2004年12月まで米国国立がん研究所に滞在し、extramural division のひとつである Division of Cancer Prevention による国家的ながん化学予防薬開発のサポートについて調べた。当該 division の担当者数人をインタビューするとともに、solicitation や protocol review の会議に出席することによって、化学予防薬開発の戦略、方法、

実施について調べた。

C. 研究結果

NCIには大きく分けて intramural division と extramural division がある。ともに臨床、基礎、疫学部門があるが、わが国に存在しないのが extramural division である。NCI 全体の研究費のうち、9 割が NCI 以外で行われるがん研究に投下され、それを仕切っているのが extramural division である。研究費は基本的に peer review で採択が決められるので、extramural division は研究費の採択を決定する部門ではなく、各研究機関による研究の進捗管理と進捗の手助け、そしてがん研究の方向付けをする部門である。本報告では extramural division の一つの division of Cancer Prevention (以下 DCP)による化学予防薬開発を紹介する。

まず背景として、日本と大きく異なることは予防薬が健康保険で認められることである。これによって製薬会社に新規予防薬を開発する incentive が生じる(しかし現実には治療薬に比べて製薬会社の incentive はかなり低い)。これを反映して、化学予防薬の開発も DCTD (Division of Cancer Treatment and Diagnosis)で行われている治療薬の開発と同じ仕組みが用いられている。DCP のうち、Chemopreventive Agent Development Research Group が preclinical の開発を担当し、

臨床早期開発を Organ Systems Group、後期開発を Community Clinical Oncology Program (CCOP)が担当し、各大学や研究機関、大小の製薬会社、あるいは NCI 自体で考案された予防薬の候補薬剤のその後の開発を援助する。まず、これら候補薬剤に対してまず Preclinical に毒性や in vivo や vitro での効果を調べることになるが、自分ですべてを調べる能力がないところに対しては NCI が代わりに調べてくれる。次にこの中から人間でも効果がありそうなものについて臨床開発を行うわけであるが、これら候補薬剤に対して DCP が研究課題の公募(solicitation)が行う。このシステムは最近始まったばかりで、これまでの DCP 側から決め打ちで計画を立て募集する方式からよりより緩い形(研究者に計画自体を求める形)で募集を行うようになった。ただし本当の意味での公募ではなく、予防研究に興味と実績のある米国の 6 施設の Consortia に対して、候補薬剤と大まかな臓器を決めて募集を行う方式をとっている。各薬剤に対して数個程度のアイデア(LOI, letter of Intent)の応募があり、これらに対して、「科学的な情報と研究デザイン」「組織、施設、これまでの経験」「登録能力」「予算はあるか」などの面から評価を行う。よい評価を得た LOI にはプロトコールを書く権利が与えられるが、プロトコール作成の段階で reject されることもある。プロトコールが

認められれば実際に研究を行うことができる。ここまでの臨床早期開発である。

NCI におけるすべての臨床開発は IND (Investigational New Drug, FDA に対する新薬利用申請)のもとで行われる。IND により、FDA に対して様々な手続きと報告義務が生じ、それを満たすことによって研究の質が保たれるということになる。IND に必要な面倒な手続きは全て NCI がやってくれる。ただし、予防薬の臨床開発は、必ずしも抗がん剤の開発と同じようには進まない。抗がん剤の場合は、かなりの毒性が予測され、腫瘍縮小効果というある程度信頼できる効果の surrogate endpoint (代替指標)があるのに対し、予防薬の場合には、毒性はかなり低い場合が多いものの、罹患の surrogate endpoint となるよい指標や biomarker がないからである。したがって、予防薬に対しては phase I で毒性、phase II で腫瘍縮小効果、phase III で生存を endpoint として RCT を行う、という抗がん剤の典型的なデザインを取ることができない場合が多い。その代わりに、早期開発では毒性に加えて feasibility を調べたり、biomarker を探索したり、妥当性を確認したりすることが必要となる。後期開発でも、罹患を endpoint とした大規模 RCT を組む前に、プラセボや concurrent control を対象とし、biomarker を endpoint とした小規模 RCT を組む場合もある。臨床開発

の補助として、LOI の solicitation や protocol の審査を行うだけでなく、薬剤の作成(一部)、貯蔵、shipping、RCT では blinding などを行う (IND なので GMP 対応)。しかし、治療薬に比べより大規模でお金もかかり、対象者をリクルートするのも難しいため、実際に罹患やその surrogate を endpoint とした RCT が行われることはそれほど多くない。これは、年間数百の IND 薬があり、数十の phase III が行われる治療の世界とは大きな違いである。臨床後期開発は数が少ないことと一つ一つが大規模であることから治療分野ほどシステムティックになっていないようで議会の要請とともに計画が進められることもあるようである。

DCP では、早期開発の結果を受けて効果がありそうだったものを臨床後期開発につなげるが、これは Community Clinical Oncology Program (CCOP) が担当する。CCOP は、CCOP と呼ばれる地域の医師や病院が NCI sponsored の臨床試験や予防試験へ対象者をエントリーすること、Research base と呼ばれる多施設共同研究グループや各地のがんセンターが予防研究を計画・実施することをサポートするための grant であり、DCP 内にそれを管理する部門がある。

現在、地域の病院か病院組織で構成された 50 の CCOP と 11 の minority-based CCOP があり、NCI の臨床試験や予防研究に

対象者を登録することで grant をもらっている。CCOP 施設は臨床試験、予防試験にそれぞれ年間 50 例以上登録しなければならないが、それを満たすと約\$1900/例の grant がもらえるため、それを用いて臨床試験や予防研究専属の Nurse、administrative worker の給料などを賄う。CCOP は、多施設共同試験グループや NCI が audit や教育を受け、疾患別グループの会議に参加することにより、治療や data management の質を保つ。CCOP はこれまでに NCI 治療試験に 10 万人、予防試験に 9 万人を登録しており、これらに大きな貢献を行っている。さらには、臨床試験や予防試験に地域の病院組織が参加することによって標準治療や予防の最新情報を学ぶという教育的な側面もあり、これがかなり重要な意味を持つ。

D. 考察

NCI の予防薬開発は、個々の研究者が自分たちですべてを担わなければならない日本とは大きな差がある。予算の額や国の事情も異なるため、同じことをする必要はないが、新規候補薬剤から臨床開発まで組織的にサポートを行うという考え方は参考になる。わが国のがん治療開発システムは米国に 20-30 年遅れているといわれるが、がん治療を世に広め、予防研究を実施するための組織である CCOP が米国にできたのが 1986 年であることを考えると、組織的な予防薬開発プロセスと CCOP に対応する Network を作る具体的

なアクションを早急に取り始めないと予防開発システムの遅れは治療開発システムの差以上のものになってしまうだろう。

E. 結論

治療開発、予防法開発の先進国である米国の方法を学ぶことによって、わが国において必要なシステムの案を考察できる。本邦は予防薬剤は保険適応にならないが、米国では保険適応となる可能性がある、それに関連して、規制当局の認可の方法が異なる、など違いがある。今後は両国の違いを考慮に入れ、本邦にあてはまるようないくつかのモデルを提案する予定である。

F. 研究発表

1. 論文

- 1) Horstmann E, Yamamoto S, et al. Risks and Benefits of Phase 1 Oncology Trials: 1991-2002 New Engl J Med (in press).
- 2) Abe M, Yamamoto S, et al. CpG island methylator phenotype is a strong determinant of poor prognosis with neuroblastomas. Cancer Research (in press).
- 3) Hanaoka T, Yamamoto S, et al. Active and passive smoking and breast cancer risk: observational cohort study. Int J Cancer. Nov 2004 (advanced online publication).

- 4) Ishida K, Yamamoto S, et al. Phase II Study of Cisplatin and 5-Fluorouracil with Concurrent Radiotherapy in Advanced Squamous Cell Carcinoma of the Esophagus: a Japan Esophageal Oncology Group (JEOG) /Japan Clinical Oncology Group Trial (JCOG9516). *Jpn J Clin Oncol* 2004; 34 (10) 615-619.
- 5) Sano T, Yamamoto S, et al. Gastric cancer surgery: results of morbidity and mortality of a prospective randomized controlled trial (JCOG 9501) comparing D2 and extended para-aortic lymphadenectomy. *J Clin Oncol* 2004; 22, 2767-2773
- 6) Sasazuki S, Yamamoto S, et al. Green tea consumption and subsequent risk of gastric cancer by subsite: the JPHC Study. *Cancer Causes Control*. 2004 Jun; 15 (5): 483-91.
- 7) Yamao T, Yamamoto S, et al. . Phase II Study of Sequential Methotrexate and 5-Fluorouracil Chemotherapy against Peritoneal Disseminated Gastric Cancer with Malignant Ascites. A Report from the Gastrointestinal Oncology Study Group of the Japan Clinical Oncology Group, JCOG9603 Trial. *Jpn J Clin Oncol* 2004; 34 (6) 316-322.
- 8) Yamamoto S, et al. RESPONSE: Re: Soy, Isoflavones, and Breast Cancer Risk in Japan. *J Natl Cancer Inst* 2003; 95: 1881-1882.
- 9) Yamamoto S, et al. Soy, isoflavones, and Breast cancer in Japan. *J Natl Cancer Inst* 95: 906-913, 2003.
- 10) Iwasaki M, Yamamoto S, et al. JPHC Study Group . Background characteristics of basic health examination participants: the JPHC Study Baseline Survey. *J Epidemiol*. 2003 Jul; 13 (4): 216-25.
- 11) Hironaka S, Yamamoto S, et al. Nonrandomized comparison between definitive chemoradiotherapy and radical surgery in patients with T(2-3)N(any) M(0) squamous cell carcinoma of the esophagus. *Int J Radiat Oncol Biol Phys*. 2003 Oct 1; 57 (2): 425-33.
- 12) Otani T, Yamamoto S, et al. Alcohol Consumption, Smoking, and Subsequent Risk of Colorectal Cancer in Middle-aged and Elderly Japanese Men and Women: JPHC Study . *Cancer, Epidemiology, Biomarkers, and*

- Prevention. 2003; 12, 1492-1600.
- 13) Watanabe H, Yamamoto S, et al. Tumor Response to Chemotherapy: the Validity and Reproducibility of RECIST Guidelines in NSCLC Patients. Cancer Science 94; 1015-20; 2003.
 - 14) Ishihara J, Yamamoto S, et al. Demographics, lifestyles, health characteristics, and dietary intake among dietary supplement users in Japan. Int J Epidemiol. 2003 Aug; 32 (4): 546-53.
 - 15) Nashimoto A, Yamamoto S, et al. A Randomized Trial of Adjuvant Chemotherapy with Mitomycin C, 5-Fluorouracil and Cytosine arabinoside followed by Oral Fluorouracil in Serosa Negative Gastric Cancer (JCOG9206-1) J Clin Oncol 21; 2282-2287, 2003.
 - 16) Ohtsu A, Yamamoto S, et al. Randomized Phase III Trial of 5-fluorouracil Alone versus 5-fluorouracil plus Cisplatin versus Uracil and Tegafur plus Mitomycin C in Patients with Unresectable Advanced Gastric Cancer: the Japan Clinical Oncology Group Study (JCOG9205). J Clin Oncol 2003, 21: 54-59.
 - 17) Ohtsu A, Yamamoto S, et al. A phase II study of irinotecan in combination with 120-hour infusion of 5-fluorouracil in patients with metastatic colorectal carcinoma: Japan Clinical Oncology Group Study (JCOG9703) Jpn J Clin Oncol 33: 28-32, 2003
 - 18) 山本精一郎, 臨床試験のデザイン, 臨床腫瘍学, 癌と化学療法社, 2003.
 - 19) Ishihara J, Sobue T, Yamamoto S, Yoshimi I, Sasaki S, Kobayashi M, Takahashi T, Itoi Y, Akabane M, Tsugane S. Validity and Reproducibility of a Self-administered Food Frequency Questionnaire used in the 5-year Follow-up Survey of the JPHC Study Cohort II: Study Design, Participant Profile and Results in Comparison with Cohort I. J Epidemiol 2003; 13 (suppl): S134-S147.
 - 20) 石塚直樹, 山本精一郎, 大規模臨床試験成績連載 第 17 回「大規模臨床試験の解析の仕方」血液・免疫・腫瘍 Vol. 8 No. 3 2003-7, p84-88.
 - 21) 河本 博, 佐藤暁洋, 山本精一郎, 他 : Cons. メタアナリシスは「レベル 1」のエビデンスか? . 臨床研究・生物統計研

- 誌. 23(1):97-105, 2003.
- 22) 山本精一郎, 臨床試験のエンドポイント, 肺癌:診断・治療の最前線, 癌の臨床, 第 49 巻第 9 号, 篠原出版新社, 2003.
- 23) Kaba H, Fukuda H, Yamamoto S, Ohashi Y. Reliability at the National Cancer Institute-Common Toxicity

Criteria version 2.0 Gan To Kagaku Ryoho. 2004 (8): 1187-92.

2. 学会発表
発表予定

G. 知的財産権の出願・登録状況 (予定を含む)

1. 特許取得 なし
2. 実用新案登録 なし
3. その他 なし

厚生労働科学研究費補助金(第3次がん総合戦略研究事業(分野3))

(分担)研究報告書

研究課題名:「がん予防薬の薬物動態関連酵素とSNPとの関連データベースの開発」

分担研究者氏名: 日紫喜 光良

所属:産業技術総合研究所 生物情報解析研究センター

研究要旨: (1) がん予防薬と遺伝子・タンパク質の関係のテキストからの収集: a) 既知のがん予防薬を含む薬剤のリスト、及び、遺伝子・タンパク質のリストを作成し、b) 両リストの要素を共に含む文を文献データベース(PubMed)から抽出。c) 抽出した文を調査して、薬剤-遺伝子関係の典型的な表現パターンを収集し、d) それを用いた薬剤-遺伝子関係の網羅的検索を実施した。(2) がん予防薬と関連する遺伝子のSNPの調査: a) 文献データベースから収集したがん予防薬関連遺伝子・タンパク質上に存在するSNP情報を、データベース統合の技法を用いて収集してデータベースを構築、b) 蓄積した情報を相互に参照可能なWWWベースのユーザインターフェース“GenoCache”のプロトタイプを開発した。

A. 研究目的

既知のがん予防薬を含む薬剤のリストを、標準的な英語医学生物学用語語彙(Metathesaurus)等から作成し、また、パブリックドメインのヒト遺伝子データベースを複数用いて遺伝子・タンパク質名のリストを作成した。薬剤と遺伝子との関係を網羅的に収集するために、これらをPubMedデータベースのタイトルならびにアブストラクト全文にマッチさせることによって、両者の共起(同じ文に存在すること)を網羅的に収集した。その中から、がん予防薬とその代謝酵素の関係をあらわす可能性のある表現を調査した。さらに、上記の方法で収集した薬剤

関連遺伝子のSNP情報を、データベース統合の技法を用いて収集を目的記とした。

B. 研究方法

遺伝子辞書の作成:パブリックドメインのヒト遺伝子データベース(HUGO、LocusLinkならびにUniProt)を用いて、遺伝子のIDとその代表的な名称、ならびに代替の名称を収集した。それらのうち、NCBIのdbSNPに対応のあるものを用い、およそ23,000の遺伝子座に対する遺伝子名を収集した。

物質辞書の作成:UMLS Metathesaurusには、用語の意味カテゴリー(Semantic Type)が付与されているが、がん予防薬をあらわすもの

はない。もっとも近いものは、たとえば次の Semantic Type をもつものと考えられる:T121 (Pharmacologic Substance)ならびに T200 (Clinical Drugs)。そこでまず、がん予防物質についての総説を収集し、その表から物質名を収集した。次に、事項に示す、がん予防薬と代謝酵素の関係をあらわす文のパターンを用いて、代謝の様式がわかっているようながん予防薬をテキストから収集することを試みた。次に、Metathesaurus 中の上記 Semantic Type をもつ用語で、がん予防についての文献に出現するものから候補を選択した。

がん予防薬の代謝を示唆する文脈の収集:乳がんの再発予防に用いられている Tamoxifen を例にとつて、遺伝子との共起文を収集し、そのうちから代謝を示唆するものを選択した。それらの文を参考にして想定される、遺伝子名の周辺の文のパターンを作成し、PubMed 全体に適用した。ノイズが大きく、がん予防物質を効率よく収集するには至らなかった。

がん予防物質と遺伝子との共起文の収集:上記の遺伝子名辞書とがん予防薬(物質)辞書に含まれる用語を PubMed 全件(およそ 1300 万レコード)にマッチさせ、共起を収集した。共起に対して上記の代謝を示唆するパターンを探索した。

がん予防薬(物質)とゲノム情報との関連についての Web サイトプロトタイプを作成

:薬剤等によるがん予防(ケモプリベンション)についての総説を参考にして、がん予防薬(物質)とゲノム情報との関連についてまとめた Web サイトを試作した。

C. 研究結果

1) がん予防薬に関連する遺伝子・タンパク質のテキストからの収集

既知のがん予防薬を含む薬剤のリストを、標準的な英語医学生物学用語語彙(Metathesaurus)等から作成し、また、パブリックドメインのヒト遺伝子データベースを複数用いて遺伝子・タンパク質名のリストを作成した。薬剤と遺伝子との関係を網羅的に収集するために、これらを PubMed データベースのタイトルならびにアブストラクト全件にマッチさせることによって、両者の共起(同じ文に存在すること)を網羅的に収集した。その中から、がん予防薬とその代謝酵素の関係をあらわす可能性のある表現を調査した。

2) がん予防薬と関連する遺伝子の SNP の調査

上記の方法で収集した薬剤関連遺伝子の SNP 情報を、データベース統合の技法を用いて収集した。

遺伝子辞書の作成:パブリックドメインのヒト遺伝子データベース(HUGO、LocusLink ならびに

UniProt)を用いて、遺伝子の ID とその代表

的な名称、ならびに代替の名称を収集した。それら
のうち、NCBI の dbSNP に対応のあるものを用い、およそ 23,000 の遺伝子座に対する遺伝子名を収集した。

物質辞書の作成:UMLS Metathesaurus には、用語の意味カテゴリー(Semantic Type)が付与されているが、がん予防薬をあらわすものはない。もっとも近いものは、次の Semantic Type をもつものと考えられる:T123 ならびに T231。そこでまず、がん予防物質についての総説を収集し、その表から物質名を収集した。次に、事項に示す、がん予防薬と代謝酵素の関係をあらわす文のパターンを用いて、代謝の様式がわかっているようながん予防薬をテキストから収集することを試みた。次に、Metathesaurus 中の上記の Semantic Type をもつ用語で、がん予防についての文献に出現するものから候補を選択した。

のテキストからの収集・データベース構築フロー。

2) がん予防薬の代謝を示唆する文脈の収集: 乳がんの再発予防に用いられている Tamoxifen を例にとり、遺伝子との共起文を収集し、そのうちから代謝を示唆するものを選択した。それらの文を参考にして想定される、遺伝子名の周辺の文のパターンを作成し、PubMed 全体に適用した。ノイズが大きく、がん予防物質を効率よく収集するには至らなかった。

がん予防物質と遺伝子との共起文の収集: 上記の遺伝子名辞書とがん予防薬(物質)辞書に含まれる用語を PubMed 全件(およそ 1300 万レコード)にマッチさせ、共起を収集した。共起に対して上記の代謝を示唆するパターンを探索した。

がん予防薬(物質)とゲノム情報との関連についての Web サイトプロトタイプ作成: 薬剤等によるがん予防(ケモプリベンション)についての総説を参考にして、がん予防薬(物質)とゲノム情報との関連についてまとめた Web サイトを試作した(図2)。

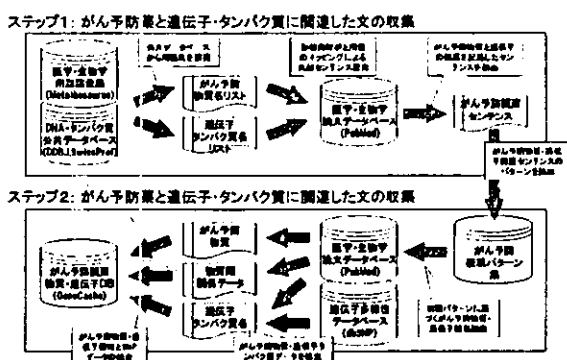
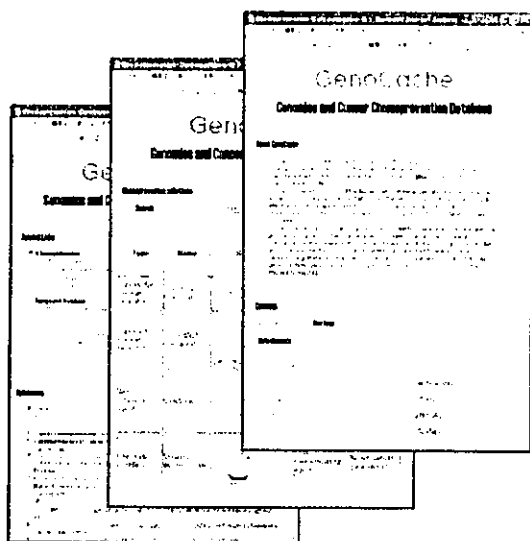


図1. がん予防薬と遺伝子・タンパク質の関係



● 図2. がん予防薬の薬物動態関連酵素とSNPの関連データベース・プロトタイプの開発。

GenoCache は、がん化にかかわる物質と遺伝子との関係を格納したデータベース&利用インターフェースシステムである。医学生物学論文(テキスト)より抽出した がん予防薬と遺伝子・タンパク質情報との関連、該当遺伝子上にコードされている遺伝子多型(SNP)の情報、また、テキストや公共データベースに存在する生体分子間相互作用情報(マップ情報)を関連付けた統合データベースシステムである。

D. 考察

薬剤等によるがんの予防(がんのケモプリベンション)と遺伝子との関連に関する Web サイト

(GenoCache Web サイト)をつくり、関連する

文献から抜き出した情報をわかりやすく提供することを考えている。その情報を増強するために、研究の成果を利用することを想定している。この Web サイトは、がん予防にかかわる物質の遺伝子あるいは遺伝子産物との相互作用(物理的な結合だけでなく、他のメカニズムを介して結果的に遺伝子の機能を調節するものを含む)ならびに、それらとがんの発生や増殖との関係を収集したデータベースである。

E. 結論

- (1) がん予防薬と遺伝子・タンパク質の関係のテキストからの収集: a) 既知のがん予防薬を含む薬剤のリスト、及び、遺伝子・タンパク質のリストを作成し、b) 両リストの要素を共に含む文を文献データベース(PubMed)から抽出。c) 抽出した文を調査して、薬剤-遺伝子関係の典型的な表現パターンを収集し、d) それを用いた薬剤-遺伝子関係の網羅的検索を実施した。
- (2) がん予防薬と関連する遺伝子のSNPの調査: a) 文献データベースから収集したがん予防薬関連遺伝子・タンパク質上に存在する SNP 情報を、データベース統合の技法を用いて収集してデータベースを構築、b) 蓄積した情報を相互に参照可能なWWW ベースのユーザインターフェース “GenoCache” のプロトタイプを開発した。

F. 研究発表

1. 論文

1) Hishiki T, Tamura T, and Oyama H.:
GenoCache: a genomics and cancer
chemoprevention portal. AMIA annual
symposium, Oct 22-25 (2005),
Washington DC (submitted).

2) 大久保公策, 日紫喜光良: ポストゲノム時代に高まるバイオ自然言語処理への期待: バイオ自然言語処理最新事情 5. ゲノムデータの機械解釈. 情報処理 (2005) 46(2).

2. 学会発表

1) 日紫喜光良: 配列データの関連情報: 文献からの抽出と相互運用性. 国立遺伝学研究所共同研究会「生物情報資源の相互運用性」, 2005年3月3日, 三島.

2) 日紫喜光良, 長田直樹 他: H-Invitational Disease Edition における疾患-遺伝子関連データベースの構築. 第27回日本分子生物学会年会, 2004年12月9日, 神戸.

3) Nagata N, Hishiki T, et al: Cataloguing of Disease-Associated Genetic and Environmental Factors by a Text-mining Approach. ISMB/ECCB 2004, Jul. 31-Aug. 4 (2004), Glasgow, UK.

G. 知的財産権の出願・登録状況 (予定を含む)

1. 特許取得 なし
2. 実用新案登録 なし
3. その他 なし