

5

医療・医学における バイオインフォマティクス

東京大学大学院医学研究科
臨床バイオインフォマティクス研究ユニット

山崎 憲 山崎 力
yamazaki-tadashi@umin.ac.jp yama-ky@umin.ac.jp

バイオインフォマティクスは、当初は主としてアミノ酸配列や DNA 塩基配列情報をその研究対象としてきたが、最近では、遺伝子多型 (SNPs) 情報に基づくテーラーメイド医療 (オーダーメイド医療とも呼ばれる) や遺伝子発現、分子間相互作用等の細胞レベルへとその対象を急速に拡大しつつある。これらの研究により、個々の患者の特性に応じた、遺伝子治療を含めた疾患関連遺伝子の同定や薬剤による最小の副作用にとどめる診断・治療方法の確立に寄与することが期待されている。

バイオインフォマティクスと医学

バイオインフォマティクスと医学を語る前にまずは医学と医療の言葉の違いについて説明する。

医学とは疾患の原因解明、疾患の治療方法の解明を扱う学問で、大きく分けて基礎医学と臨床医学に分かれる。臨床医学は内科学や外科学など臨床に関する医学であり、基礎医学はそれら臨床医学の基礎となる学問、たとえば解剖学や生理学、微生物学などが含まれる。そのほか、法医学や公衆衛生学といった学問を社会医学といった別のカテゴリーに入れることもあるが、基本的にはこの2つが大きな柱となっている。もっとも病理学などのように、日本では基礎医学的な傾向が強い学問でも、欧米では臨床医学的な扱いを受けているようにその境界が曖昧な場合もある。

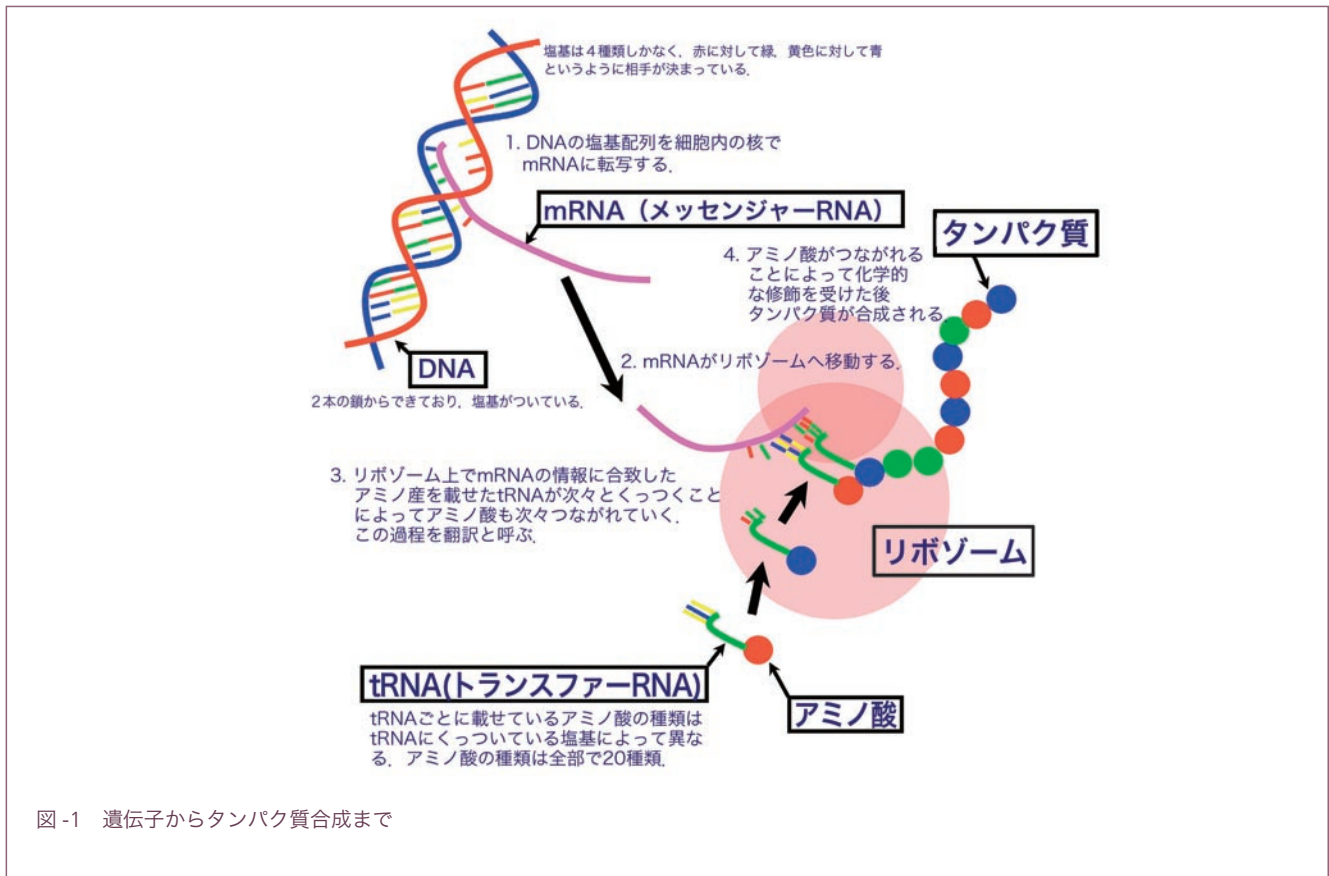
一方、医療とは疾患に対する治療行為に関連するものであり、その根元を成しているのがさまざまな分野の医学である。

1953年のJames D. WatsonとFrancis H. CrickによるDNA二重らせん構造の発見を始めとする分子生物学は医学の世界へ浸透し始め、さらにIT技術の発達はバイオインフォマティクスという新たな分野を生み出し、生物学的な分野が中心であったこれらの学問は医学分野へ進出を果たしつつある。そもそも医学の前提となるヒトの生命活動そのものがこれらの生物学と関連を持って

いる。たとえば、タンパク質の発現にしても遺伝子であるDNAからRNAへ転写された後にアミノ酸配列に翻訳され、修飾 (化学的な変化) を受けてタンパク質に合成される。そしてタンパク質は構造蛋白、酵素といった生命活動の根本を担っている (図-1)。構造蛋白は人体を構成するタンパク質、酵素はさまざまな生命活動の触媒として働いている。また、DNA塩基配列の変異によりアミノ酸が1つ置き換わるだけでタンパク質の立体構造は変化し、機能を果たさなくなることさえあり得る。したがって医学を語る上では本来、遺伝子・分子生物学は切っても切り離せない関係にあるとあってよい。

バイオインフォマティクスの歴史は1983年にDoolittleらが癌遺伝子の一種であるサル肉腫ウイルスのv-sis遺伝子のアミノ酸配列とヒトのPDGF (血小板由来増殖因子) 遺伝子のアミノ酸配列との類似性を指摘したのが最初であるといわれている¹⁾。彼らはアミノ酸配列のデータベースを用いてこれを発見した。こうしたアミノ酸配列のデータベースはDoolittleらの研究の前の1972年、Dayhoffにより作成され、やがてWeb上で公開されるようになった。DNAの塩基配列のデータベースもGoadによって1979年に作成されている。したがってこの手法はアミノ酸配列だけでなくDNA塩基配列にも応用できることはいうまでもない。

一般に遺伝子情報のことをゲノムと呼ぶが、ヒトのゲノムを解読するプロジェクトが1990年にアメリカを中心に開始され、2003年に解読が終了した。サブ解析は



まだまだ続いているが、生命活動、さらには遺伝子の違いによる発生や疾患感受性の相違などの解明の下準備が整ったといえる。これを受けて多くの医学研究施設では疾患とゲノムの関連、特に遺伝子多型に着目し、さまざまな患者の遺伝子を集め、研究が行われ始めた。ことに近年ではIT技術の発達がめざましく、以前では考えられなかったような大容量のデータベース、高速で複雑な解析ソフトウェアの開発が可能に達している。

従来の医学研究ではまず解剖がはじめにあって生理や病因などを解明しつつ治療法が開拓されたり、疫学的な調査から病因を予想していった。しかし、バイオインフォマティクスの手法を用いれば最初からデータベースを用いて治療法を予測し、生理や解剖に還元するといった従来とはまったく逆のプロセスでの治療法が開発が可能となる。またこれによって治療法開発におけるコスト低減や時間短縮が可能となることも期待される。

医療におけるIT技術の歴史

医療の世界ではITの導入は他の分野と比較して遅かった。これには複雑な医療保険システム、頻回の保険点数の改訂、忙しすぎる臨床医師のIT技術への関心の薄さ、複雑な人体のシステムがゆえの複雑な医学・医療に追いつかないハードウェア・ソフトウェアの限界などがIT

導入の遅れの原因になっているのではないかと推測される。1990年代から電子カルテやオーダーリングシステム、レセプトシステムなどでIT技術は一部普及していたが、初期の医療におけるIT技術は臨床医学研究のためでなく、あくまでも病院の医療事務的な作業を軽減するのが目的であった。ようやく近年になって厚生労働省も電子カルテをはじめとする医療へのIT導入を掲げ始め、医療現場から医学へのデータの還元のための環境づくりを始めたが、稼働しているシステムはメーカーごとに操作方法は大きく異なり、まだまだ現場医師にとって使いやすいとは言い難い。行政におけるシステム規格作りもまだ始まったばかりでゴールへの道りは遠い。

また、従来の電子カルテやオーダーリングシステム、レセプトシステム等の医療側のシステムで得られたデータを医学研究へ流用することは困難であり、これらを医学研究に流用する際は表計算ソフトウェア等といったマニュアルで取り込んでから解析を行う必要があった。データの再入力に際しては専門家であるデータマネージャを置くのが最良の方法であるが、コスト的に問題がある際には医師が業務の片手間に入力することになり、データの劣化の原因となっていた。血液データなどのオーダーリングシステムと連動した総合データベースシステムであれば、データの再入力に伴うデータの質の劣化は最小限に食い止めることが可能となる。また、検査レポートは医療法で必ず公文書として残すことが義務

- a. 患者の健康状態、症状、診断、予後および治療に関する本人を特定し得るあらゆる情報、ならびにその他すべての個人的情報の秘密は、患者の死後も守られねばならない。ただし、患者の子孫が自らの健康上の危険に関わる情報を知る権利は、例外的に認められる。
- b. 秘密情報の開示は患者本人が明確な承諾を与えるか、法律にはっきり規定されている場合のみ許される。他の医療従事者への情報開示は、患者が明確な承諾を与えていない限り、業務遂行上知る必要がある範囲内でのみ許される。
- c. 患者を特定可能なデータは保護されねばならない。データの保護はその保存形態に応じて適切になされねばならない。個人の特定が可能なデータが導き出される生体試料や標本も同様に保護されねばならない。

表-1 患者の権利宣言（リスボン宣言）第8条

1.	試料提供者の意志の尊重（インフォームド・コンセント）
2.	外部委員が過半数からなる倫理審査委員会による承認と外部調査
3.	試料提供者の人権保護
4.	個人情報管理者による厳格な個人情報の保護 個人情報保護のためのコンピューターシステムとそのセキュリティの徹底 研究遂行者に対する内規の作成
5.	遺伝カウンセリング
6.	研究の情報開示

表-2 「ヒトゲノム・遺伝子解析研究に関する論理指針」の基本的な考え方

づけられているため、検査レポートシステムを電子化して直接的なデータの取り込み・連携が可能になればデータの信頼性は上昇する。

こうして作成された臨床情報データベースを医学研究で活用するためには遺伝子データベースの構築も必要となってくるが、遺伝子データ使用の際には患者情報の秘密保持という倫理的問題が立ちはだかっている。我が国ではリスボン宣言（表-1）を批准しているため、厚生労働省における「ヒトゲノム・遺伝子解析研究に関する論理指針」（表-2）に従わなければならない。たとえば、東京大学医学部附属病院循環器内科では遺伝子採血に先立ってまず患者に説明を行い、納得していただいた後に同意を文書で得ている。その後で得られたサンプルは第三者の手によって匿名化を行った上でサンプルを業者に搬送し解析を依頼しており、サンプルの破棄に関しても厳重な規則に従って行っている。

医療用画像に関してはDICOM規格がようやく策定され、地球資源の節約などをふまえて電子化が急速に進んでいるにもかかわらず多くが画像の閲覧に使用されている程度であり、医療用画像の解析、数値化は非常に困難である。その中でも数少ない医学解析システムが心臓血管領域ですでに完成しているが、これについては後に詳しく説明することにする。

医療・医学におけるバイオインフォマティクスの利用と現状

前章で遺伝子データベースと臨床データベースが構築されれば、いよいよバイオインフォマティクスの手法を取り入れた医学研究の準備が整ったことになる。たとえば、ある疾患に注目し、遺伝子の相違による疾患感受性の違いを解析する試みが各研究施設で行われているが、我々の研究チームでは冠動脈疾患に注目し、データベースと解析ソフトウェアを構築して同様の研究を行っている。

冠動脈疾患とは動脈硬化が原因で心臓に栄養を送る血管、すなわち冠動脈が狭窄、あるいは閉塞するために起こる致命的な疾患である（図-2）。単独病名ではおそらく全世界で死因のトップであり、古くから糖尿病や高脂血症、高血圧、肥満、喫煙などが罹患のリスクを高める原因になっていることが経験的に知られていた。しかし、実際には同じ危険因子を持っているにもかかわらず患者によって罹患しやすい人としにくい人が存在することが次第に分かってきた。たとえば喫煙歴が同じであっても心筋梗塞に罹患する患者もいれば罹患しない患者もいる。近年の遺伝子工学の発達により、これらの違いは遺伝子多型から来ていると考えられ、研究が進められている。

これらの解析を行うために我々は1999年より日立製

虚血性心疾患

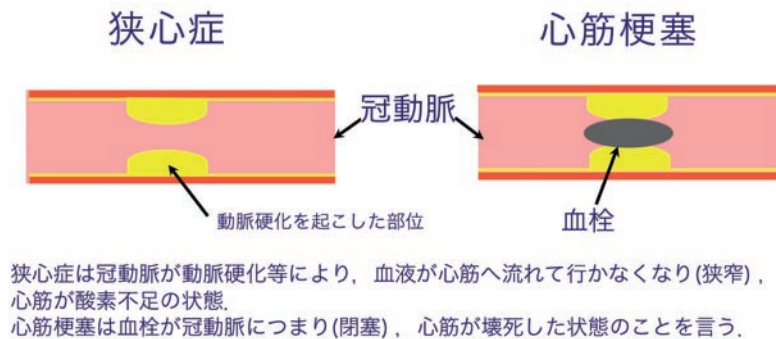


図-2 冠動脈疾患

作所と共同で東京大学医学部附属病院循環器内科にて心臓カテーテルを行った患者(2003年までに約3,000名)についてインフォームド・コンセントをとった上で遺伝子解析を行い、バイオインフォマティクスの手法を取り入れたデータベースと解析ソフトウェアを構築し、疾患と古典的危険因子(肥満、高脂血症、糖尿病、高血圧症、喫煙、男性であることなど)、遺伝子多型の関係を解析したのでその成果の一部を紹介する。

図-3(a)は各遺伝子における冠動脈疾患感受性の結果である。我々の研究ではWerner遺伝子やMMP-1などの遺伝子多型が冠動脈疾患のリスクを高めていると考えられた。しかし、図からも分かるように単一遺伝子で考えた場合いずれもリスクは2.5倍以下である。しかし、糖尿病、高血圧、肥満、高脂血症が合併するいわゆる死の四重奏と呼ばれる病態における冠動脈疾患のリスクは数倍にもなり、その数字と比較するとこれらの遺伝子による疾患の感受性は小さいと言える。したがって冠動脈疾患は単一遺伝子ではなく、複数の遺伝子多型、すなわち多因子的にリスクが増大すると考えられる。

図-3(b)は同じくABO式血液型による疾患感受性の相違である。我々の研究では日本人の一般人口比から考えて冠動脈疾患群ではAB型およびO型において冠動脈疾患の罹患リスクとの関連が示唆された。この結果からABO式血液型の遺伝子近傍に冠動脈疾患感受性を高めるような遺伝子座が存在している可能性は十分ある。しかし、サンプル数がまだ少ないためこれをもって結論づけることは困難である。今後の全国レベルでの調査に期待したいところである。

同じく図-3(c)はC型肝炎と冠動脈疾患の関連性だが、日本での一般人口におけるC型肝炎罹患率は2.4%と言われている。一方、我々の研究では冠動脈疾患群では5.8%、非冠動脈疾患群では3.8%であった。さまざま

なウイルスが宿主の遺伝子に組み込まれるという事例は数多く報告されているが、遺伝子多型による疾患感受性と併せて解析すれば感染から発病に至る遺伝子的なプロセスの解明の手助けになると思われる。ただし、非冠動脈疾患群が3.8%と一般人口からの比率よりも高い数字が出たことについて、東京大学医学部附属病院における患者層の特徴を表しているだけの可能性も否定できない。これについては他の施設とのデータを比較しながら解析を続ける必要がある。

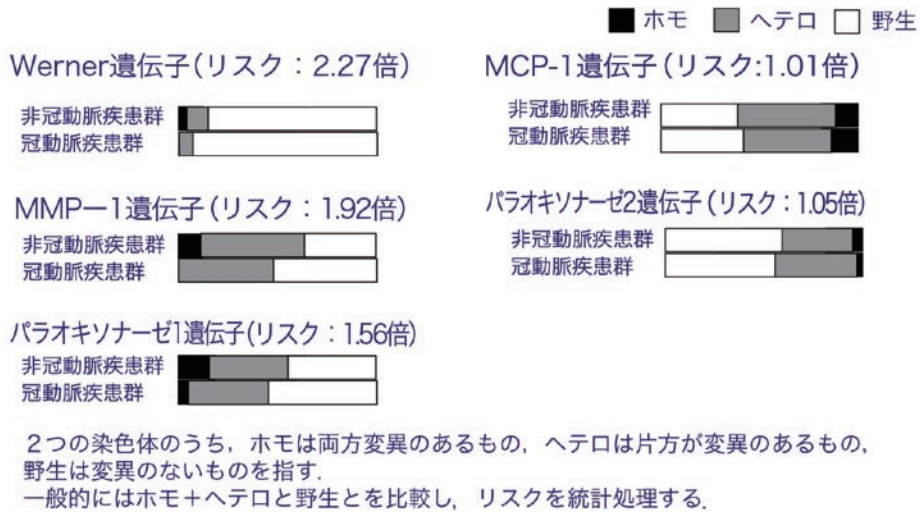
将来への展望

個人の遺伝子データベースが構築され、遺伝子多型と環境、疾患感受性の関係が明らかになれば罹患の予知もある程度可能となる。その結果、1次予防の手段を講じることが可能となるばかりか、2次予防も最小限の治療によって可能となると考えられる。また、副作用の起きやすい薬剤なども予測できるので、医療事故の発生と患者の経済負担を最小限にとどめることもできるようになる。このように患者個人に合った医療をテーラーメイド医療と呼ぶ。

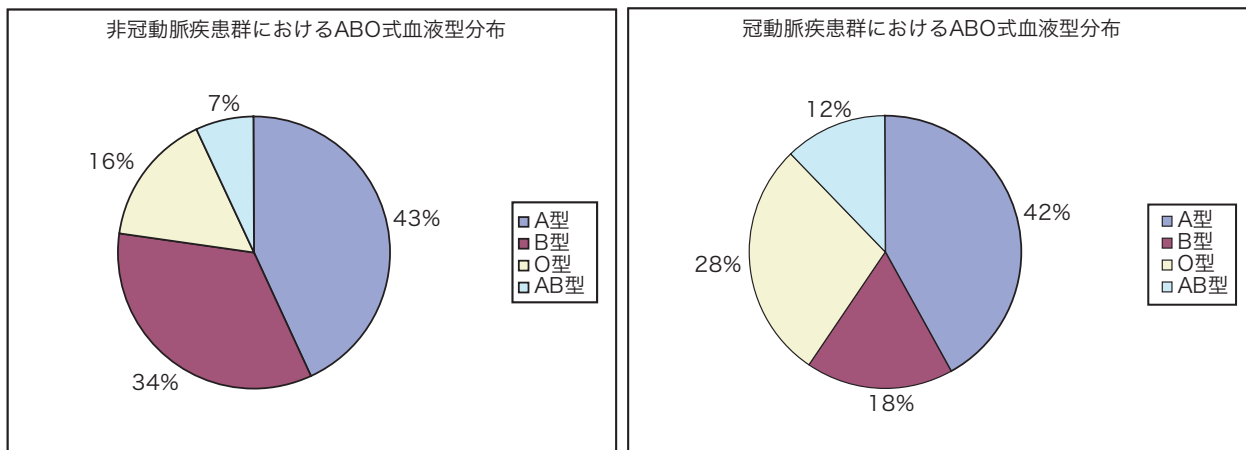
さらに薬剤開発にもバイオインフォマティクスの技術が応用される可能性もある。これまでは薬剤開発に試験管内実験(in vivo系)、生体内実験(in vitro系)を行っていたが、遺伝子データを元にしたバーチャル人間(in silico系)での実験も可能となり、安全でかつ低予算での薬剤開発が可能となるかもしれない。In silico系での遺伝子データを変更すれば遺伝子の違いによる薬剤感受性の相違も実験できるようになるであろう。

遺伝子治療への応用も期待されている。実際に日本でもADA欠損症を始め、遺伝子治療が行われ始めてお

(a) 各遺伝子における冠動脈疾患感受性



(b) 血液型別における冠動脈疾患感受性



(c) C型肝炎における冠動脈疾患感受性

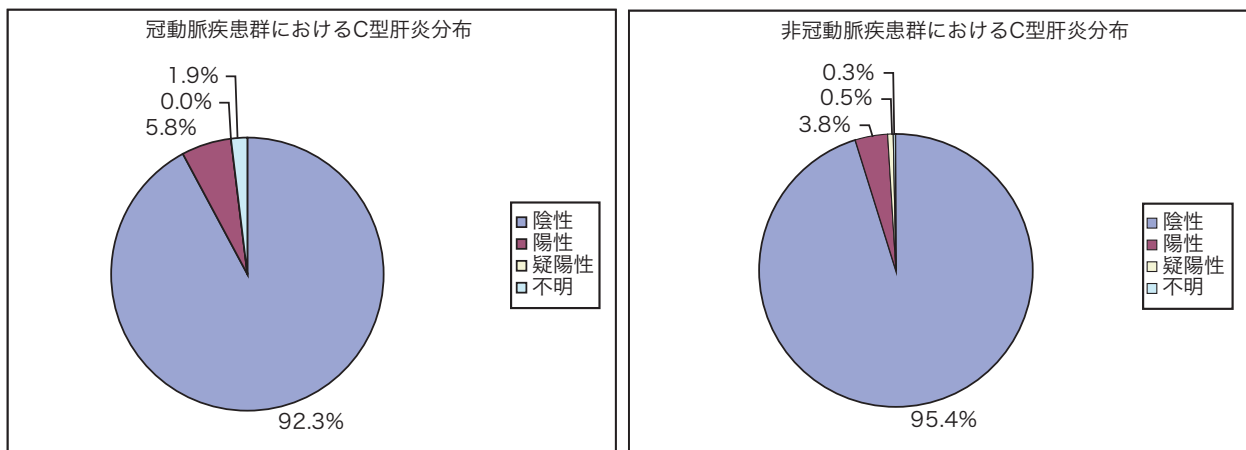


図-3 東京大学医学部附属病院循環器内科における心臓カテーテル患者におけるデータベース解析の結果

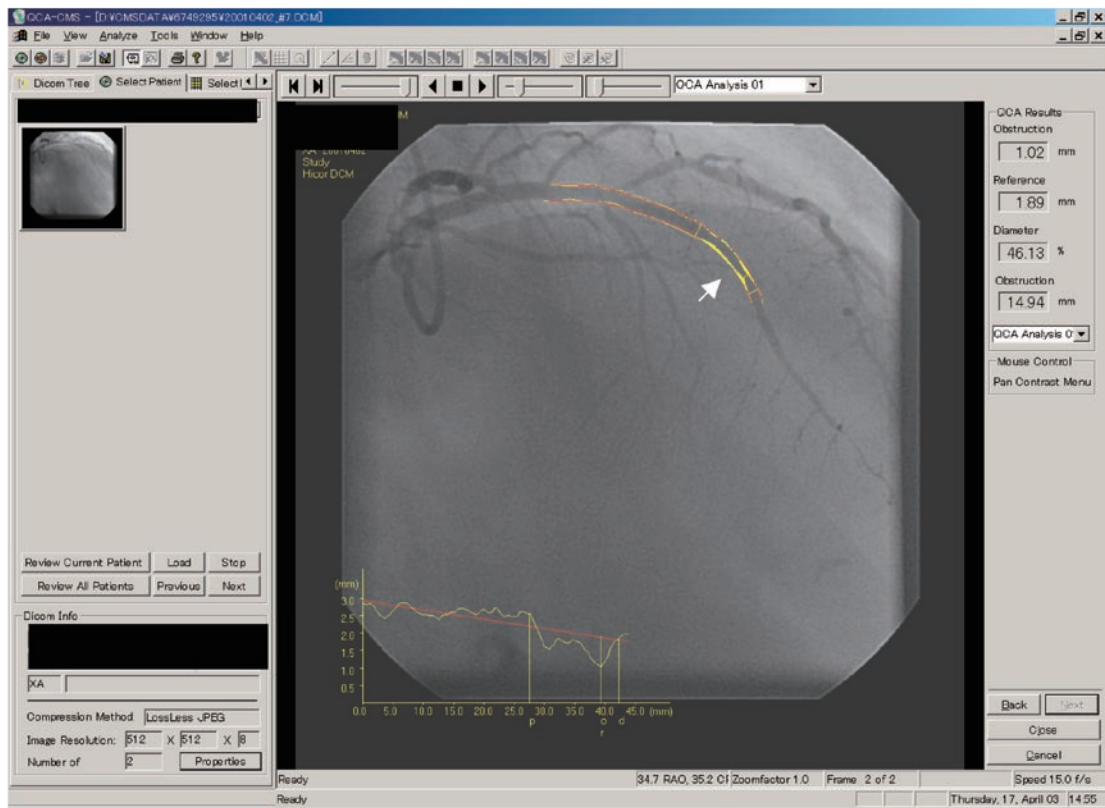


図-4 画像解析システム QCA-CMS の解析画面

冠動脈の狭窄度と虚血性心疾患の重症度は密接な関係にある。動脈硬化が進行すると狭窄も進行する。図中の矢印の部分が狭窄を起こしている部位で、その周辺の動脈硬化のない血管径と狭窄部位の血管径から狭窄度を計算し、評価する。

り一定の成果を収めている。また、糖尿病患者においては血管障害による手足の壊死が問題になることがあるが、幹細胞増殖因子の筋肉注射によって遺伝子発現を増長させることにより血管が再生することが確認されている。心筋細胞や脳細胞は一度死んでしまえば再生は不可能であるが、実際には発現していないだけで心筋細胞を増殖させる遺伝子自体は我々の細胞の中に組み込まれている。幹細胞増殖因子の遺伝子に刺激を与え、目的の部位で発現させれば心筋細胞や脳細胞の再生が可能になる。もちろん、遺伝子発現にともなうシミュレーションも可能となるため、治療の際の問題点も事前に分かるようになると思われる。実際には遺伝子を組み込むには実際の手技段階での問題が山積みしているため容易ではないが、大きく期待されている分野である。

画像の解析システムとバイオインフォマティクス研究にも切っても切り離せない関係にある。前に冠動脈画像解析ソフトウェアについて触れたが、我々の研究で冠動脈疾患の重症度判定を定量化するにあたっては実際にこ

のソフトウェアを用いて行った(図-4)。こうした画像解析システムを用いながら臨床試験を行うことが一部ではすでに始まっているが、今後も大きな伸びが期待できる。冠動脈疾患だけでなく、悪性新生物や整形外科領域の疾患の画像解析のソフトウェアが完成の域に達すれば遺伝子治療や薬剤治療などのさまざまな治療法をシミュレートすることができるであろう。

参考文献

- 1) Doolittle, R. F. et al.: Simian Sarcoma Virus Onc Gene, v-sis, is Derived from the Gene Encoding a Platelet-derived Growth Factor. *Science* 221:275-7(1983).
- 2) 山崎 憲: 電子カルテとバイオインフォマティクス. 現代医療, Vol.36, pp.1135-1142(2004).
- 3) 山崎 憲 他: バイオインフォマティクスと循環器疾患. 分子心血管病 Vol.5, p.21-(2004).

(平成17年1月25日受付)

訂 正

本誌 46 巻 3 号 (2005 年 3 月号) の特集記事「5. 医療・医学におけるバイオインフォマティクス」に誤りがありました。
お詫びして以下の通り訂正いたします。

p.262 表 -2 のキャプション

(誤) 「ヒトゲノム・遺伝子解析研究に関する倫理指針」の基本的な考え方

(正) 「ヒトゲノム・遺伝子解析研究に関する倫理指針」の基本的な考え方