

厚生労働科学研究研究費補助金

医療技術評価総合研究事業

一般内科外来における結果予測の
効果的還元法の開発と調査
(H14-医療-023)

平成14年度 総括研究報告書 / 分担研究報告書

主任研究者

京都大学医学部附属病院 総合診療科 岡田好一

分担研究者

京都大学医学部附属病院 総合診療科 小山弘
京都大学医学部附属病院 総合診療科 福井次矢

平成15(2003)年3月

平成14年度厚生労働科学研究、医療技術評価総合研究事業

総括研究報告書 / 分担研究報告書

一般内科外来における結果予測の効果的還元法の開発と調査
(H14-医療-023)

主任研究者 京都大学医学部附属病院 総合診療科 岡田好一

研究要旨

本研究の目的は、一般内科の日常の診療において、科学的根拠の効果的提示法を開発し、患者・医師・医療関係者の満足度や行動に、どのような影響をもたらすかを明らかにすることである。その際、特定の疾患の一般的な診療ガイドラインをそのまま適用するのではなく、患者の個々の特性に合わせた表示法の開発を試みる。

本年度は、医療効果の予測の提示法につき、広く普及可能なシステムを開発し、次年度以降の实地調査に向けた実験環境の調査を行った。

治療効果を何らかの方法で数値化することは古くから行われている。さらに理解度向上のための判断ツールは、パンフレット、ビデオ、コンピュータプログラム等が種々提案されている。

本研究で用いた表示方法は、全体としては全く新しいものであるが、個々の要素については過去の研究によって検討されているものである。

本年度は患者教育が日常的に行われている高脂血症を取り上げ、米国のガイドラインを検討した。患者に応じたシミュレーションのための生存モデルとして、信頼性工学で十分に検討されているワイブル分布を採用した。ワイブル分布のパラメータはガイドラインと日本人の生命表から抽出した。

システム開発に関しては、インターネット技術を応用し、院外や院内のネットワークで容易に利用できるものとした。内部表現にも国際標準を採用し、将来、電子カルテ等と連動させることも可能となるよう検討した。

具体的には、自動セッション管理が可能なWebアプリケーション開発ソフトウェアの選択、グラフやイラスト表示に次世代標準のSVG (Scalable Vector Graphics)の採用、インテルアーキテクチャによる経済的で安定したサーバコンピュータの選択である。将来の電子カルテとの連携のために、情報交換形式が標

準化されているデータベース開発システムの一つを採用した。

表示内容について、相対危険度減少(RRR: Relative Risk Reduction)や治療必要人数(NNT: Number Needed to Treat)などの数値を採用し、理解度を増す目的で生存曲線に基づくグラフと人形による表現を用いた、治療効果の判断ツールを開発した。

本年度の評価の指標は理解度とした。理解度はコンプライアンスや治療選択行動と関連していると考えられ、しかも必要であれば普遍的な評価が可能だからである。

医師による事前評価と改良の後、模擬患者による評価を行った。

その結果は以下の通りである。

(1) インターネット系のサーバを用意し、動作確認した。また、入院・外来端末等に接続されている院内LAN上にサーバを置き、同じシステムが動作することを確認した。院内端末の整備を開始した。

(2) 判断ツールの開発を行った。新しい技術を採用したために、開発期間は多少要したものの、模擬患者による予備実験が可能となる水準に達した。表示に関する技術開発は一段落したと考えられた。

(3) 判断ツールを用いた模擬患者による理解度の試験を実施した。主観的理解度と正答率の関係、医師と模擬患者の理解の違い、表現法の優劣に関する結果が得られ、考察を加えた。

次年度以降は、さらに臨床現場に密着したシステム開発を行い、実地調査を検討する。また、評価項目として患者や医療関係者の満足度等を検討する。

分担研究者

京都大学医学部附属病院 総合診療科 小山弘
京都大学医学部附属病院 総合診療科 福井次矢

A. 研究目的

A.1 概要

本研究の目的は、一般内科の日常の診療において、科学的根拠に基づく医療の効果的提示法を選択し、実験的に効果を調査することにより、患者・医師・医療関係者の満足度や行動への影響を明らかにすることにある。その際、特定の疾患の一般的な診療ガイドラインをそのまま適応するのではなく、患者の個々の特性に合わせた表示法の検討を行う。

A.2 歴史的背景

A.2.1 治療効果の解釈

インフォームドコンセントの考え方にに基づき、医師は患者に治療効果の予測をできる限り正確に理解してもらう必要がある。そうすることにより、患者の治療に対する遵守 (compliance / adhesion) が向上する。EBM(根拠に基づく医療)の観点からは、患者と医師は過去の研究の成果(根拠)を共有し、両者の価値観を踏まえて、その時点で最良の治療を選択する必要がある。

日常診療において、医師は診断や治療に関する決断の理由を患者に説明しなくてはならないが、通常、その説明は口頭で行われる。しかしながら、「ときに」や「まれに」などの量的な表現の不確実さは良く知られている。そのような言語表現の食い違いは言語間にも存在するから、治療等の判断の根拠を大きく海外に依存しているわが国の現状からは、二重の不確実性が発生する恐れがある。

また、明確な研究成果が得られていたとしても、治療効果の程度の解釈は医療従事者にとっても必ずしも容易ではない。そこで、比較のための指標がしばしば用いられる。近年作成されている EBM の手順に基づいた診療ガイドラインでは相対危険度(RR: Relative Risk)、治療必要人数(NNT: Number Needed to Treat)などの理解されやすくする工夫が行われている。

ただし、一般に臨床研究の多くは、理想的な母集団に対してランダムサンプリングされず、任意の母集団内でランダム割付されているのみなので、EBM で行うような批判的吟味(critical appraisal)が必要である。

さらに、患者の特性や置かれた状況はさまざまであるから、研究成果(エビデンス)の個々の患者への適用は、医師が責任を持って慎重に行う必要がある。特定の医療施設が適用過程に対して少しでも支援するためには、その医療施設に訪れる患者集団の特性を把握し、

診療ガイドライン等の適用時の注意を周知することが有用であろう。

A.2.2 判断ツールの現状

医師または患者の理解度の向上のための判断ツールは、パンフレットやビデオ教材やコンピュータプログラムに至るまで、多数提案されている。その内容と評価の質は年々向上している。

しかしながら、我々の検索範囲では、医師等が直接患者に説明するツールとして、電子カルテ等と連動できる判断ツールは見当たらなかった。電子カルテの調査でも、解剖図などの一般的な説明ツールは備えられているが、診断支援ツールの搭載例は無い。

A.3 研究の手順

本研究では、開発と評価の2段階を相互にフィードバックしながら並行して進める。

治療効果の表現方法には、(i) 言葉による表現、(ii) RRR (Relative Risk Reduction)、ARR (Absolute Risk Reduction)、NNT などの数値、(iii) 単純な棒グラフ等、(iv) 生存曲線とNNT曲線、などが考えられる。本研究では、代表的な表現方法を選び、判断ツールを開発する。この段階での評価項目は理解度とする。理解度とコンプライアンスと治療選択行動の変化は互いに関連していると考えられるので、最も分かりやすく、かつ客観的な設計もできる理解度を評価尺度として選択した。評価は、模擬患者を対象として行う。

次年度以降は、従来から広く使われている説明のためのツールである紙のパンフレットとの比較を計画している。評価項目は、理解度に加えて、満足度、可能であればコンプライアンスと医師・患者の治療行動の変化とする。評価は、当面模擬患者を対象とするが、実用に耐えると判断した時点から、倫理委員会の承諾を得て、患者を対象とする予定である。

B. 研究方法

B.1 対象とされた疾患 —高脂血症—

慢性疾患で、かつ、患者教育が日常的に行われている、高血圧と高脂血症を初期の目標とした。開発される判断ツールに一般性を持たせるため、他に糖尿病、骨粗鬆症、ホルモン補充療法、さらに、急性疾患の代表として上・下気道炎（かぜ）、の各ガイドラインを参照した。

今回、主に参照したのは、米国の NIH (National Institutes of Health)による成人の高コレステロール血症のガイドライン、通称 ATP-III と呼ばれるものである。ATP-III には、高脂血症患者の動脈硬化性疾患のリスクの評価と治療効果が具体的な数値として記述されている。他にもガイドラインは存在するが、今回は実験に必要な数値を抽出するために ATP-III を採用した。

B.2 生存曲線の理論分布 —ワイブル分布—

開発されるべき判断ツールでは患者の状況に応じたシミュレーションを行うのであるから、表示される生存曲線のグラフは実際の患者集団のデータではなく、適切な理論分布から導出される曲線を採用することとした。

一般に、製品の寿命や故障率(ハザード)を扱う分野は信頼性工学と呼ばれ、米国では 1950 年代から信頼性理論が発達したと言われている。

工業製品でも、人口でも、ハザードの時間分布は、いわゆるバスタブ曲線を描くとされている。実際、厚生労働省がまとめた日本人の生命表の数値を解析すると、20 才までの幼児期・青年期はハザードが減少する初期故障期、20 才から 40 才まではハザードが一定(実際にはハザードは若干加速する)の偶発故障期、50 才以上はハザードが加速する摩耗故障期のパターンをとる。

このような場合、信頼性工学では、まずはワイブル分布を当てはめ、いわゆるパラメトリック解析をするのが通常の手順である。

ワイブル分布の式は、以下の通りである。ただし、 $\exp()$ は指数関数を、 $^{\wedge}$ は累乗を、 $*$ は掛け算を強調する場合に示す。

$$\begin{array}{ll} \text{生存関数} & \exp(-(t/\eta)^m) \\ \text{ハザード} & (m/\eta)*(t/\eta)^{(m-1)} \end{array}$$

ワイブル分布には 2 つのパラメータがある。1 つは形状パラメータ(m)と呼ばれ、ハザードが減少するか、一定か、増加するかを決める。形状パラメータが 1 より小さい場合はハザードは時間とともに減少し、1 の場合はハザード一定で、1 よりも大きい場合はハザード増加であり、それぞれ初期、偶発、摩耗の故障期に該当する。

もう一つのパラメータは尺度パラメータ(η)と呼ばれ、累積故障率が $1-e$ (e は自然対数の底)つまり約 63.2%になる期間で、いわゆる時定数に相当する。ワイブル分布で形状パラメータが一定の場合、尺度パラメータによって横軸の尺度が規定される。つまり尺度パラメータは時間経過の加速率を表すので、ワイブル分布は加速モデルの 1 つと位置づけされている。

ワイブル分布の生存関数のグラフの一例を図 1 に示す。

本研究で開発したグラフの生存曲線には、ワイブル分布の生存関数を用いられている。形状パラメータは日本人の生命表からワイブル確率紙を用いて抽出した。

なお、疫学と信頼性工学では統計用語に若干の相違がある。

生存関数は、累積生存率、信頼度関数とも呼ばれる。信頼度関数を 1 から引いたものは故障分布関数と呼ばれる。

生存関数の確率密度関数は、死亡密度と呼ばれる。本実験ではイベントは死亡ではなく脳卒中や冠動脈疾患の罹患であるので、「イベント率」の用語も使用している。従って、故障分布関数に相当するものを「累積イベント率」と呼んでいる。

疫学で言うハザードは、死力や故障率とも呼ばれる。

扱う統計手法は同じであるので、言葉の違いだけであり、数学的定義に違いはない。

B.3 各資料からのパラメータの抽出

生存曲線を描くためには、基礎データとして、日本の一般人口の動向が必要となる。これには厚生労働省がまとめた第 19 回生命表を参照し、ワイブル確率紙からワイブル分布の形状パラメータを抽出した。男性では 20~40 才が約 1.6、50~80 才が約 5.5、女性では 20~40 才が約 1.3、50 才から 80 才が約 5.2 である。

これは疾患を問わない日本の全人口からの数値であるので、臨床試験の論文等に発表されているグラフとは見た目の印象が異なることがある。特定の患者集団の生命表が手に入るのならば、そちらを優先して使用すべきであろう。

尺度パラメータは、ATP-III ガイドラインのリスク評価用紙から計算で求めた。つまり、ATP-III ガイドラインで採用されたフラミンガム点数により 10 年後のリスクが分かるから、形状パラメータを固定すれば、尺度パラメータが生存関数から逆計算できる。

フラミンガム点数は 5 つの因子、すなわち、年齢、血漿総コレステロール値、喫煙の有無、血漿 HDL コレステロール値、収縮期血圧と高血圧治療の有無、による表の数値を足して求められる。フラミンガム点数に対応して 10 年後のリスクが決まる。なお、ATP-III ガイドラインのリスク評価用紙は男女別になっている。

B.4 インターネット技術

B.4.1 インターネット技術の採用

本研究の目的の 1 つは広く応用可能な判断ツールの開発である。現在においては、時間と距離を克服する情報技術として、インターネット技術が大きくクローズアップされ、その応用が世界中で検討されている。

応用例の一つとして、イントラネットがある。イントラネットはインターネット技術を企業・団体内の情報処理システムに応用するものである。つまり、インターネットで稼働するシステムは、企業・団体内でも使用可能となる。

本研究でも、医師や患者に対して判断ツールが身近に利用できるよう、特定の OS や言語を選ばないインターネット技術の活用を選択した。

本研究でのネットワーク構成を図 2 に示す。

インターネット技術の採用には、開発側の都合だけでなく、利用者側に受け入れられやすい、という利点もある。研究者らの過去の経験から、BASIC 等で独自の画面を作成するよりも、Web ページで入力画面を作成すると、利用者への教育が行いやすい利点があると考えられた。

B.4.2 Web 応用アプリケーション開発システムの普及

初期において、イントラネットを情報処理システムに採用する際の障害の1つが、セッションの扱いの問題であった。

インターネットで発達した World Wide Web (WWW)技術は単純な情報提供が基本であるために、Web ページの表示順はさほど問題ではなく、1 ページの表示ごとにサーバとの接続がとぎれていた。

しかし、オンラインシステムなどの通常の情報処理システムでは、一連の作業の流れが重要である。たとえば、診療予約を行う際に空きを検索して登録するまでに他の端末からの登録を許すと、利用者に対して操作の困難感をもたらすだろう。

そこで、一連の端末上の作業、つまりセッションを陽に管理する必要があるが、自明な方法はなく、従来はシステム作成者が独自の方法で管理の仕組みを作成するしかなかった。

現在では Web 技術の発展により、セッションを自動的に管理する仕組みが利用できる。Web の自動セッション管理はごく最近に普及したもので、数年前には本研究で行ったようなシステム構築方法は、はなはだ困難を伴うものであった。

自動的にセッション管理が行われる開発システムはオンラインシステム単独で利用可能なものもある。しかしながら、本研究では将来の広範な利用を考え、データベース管理システムと連携の取れるものを選択した。データベースシステムを利用することにより、将来、さまざまな疾患に判断ツールを応用する際に、データの扱いで柔軟な対応が可能となる。また、電子カルテ等との連携も容易になると考えられた。なぜなら、データベースシステムのインターフェースには SQL 等の標準があり、情報交換が標準化されているからである。

B.4.3 SVG の採用

本研究で開発した判断ツールにはグラフ表示が含まれている。本研究の申請時には、グラフを直接表示する適切な方法はなかった。そこで、技術者のアドバイスを元に、グラフの画像をサーバ内で作成し、インターネットで標準的な画像圧縮方法を用いて端末側に表示させる計画であった。

現在、Web の標準化は World Wide Web Consortium (W3C)という団体が行っている。その規格の中に、SVG (Scalable Vector Graphics)がある。これは直線や円などを xml (eXtended Markup Language)に合致するテキスト形式で表現する情報交換形式である。グラフや線画のイラストなどは高圧縮で情報が送れるばかりでなく、利用者側で拡大や縮小も可能であり、拡大しても画像品質の劣化がない。ただし、写真等の一般画像の情報交換には全く向いていない。

最近のパソコン用の Web ブラウザには、標準で SVG が表示できるものが普及しつつある。また、従来のパソコンの OS を使用していても、SVG の表示のための組込用ソフトはインターネットから無料で手にはいる。

ただし、残念なことに、SVG を Web ドキュメントに直接埋め込むことが出来る XHTML (eXtended HTML)対応の Web ブラウザは普及していない。そのため、本研究では SVG ファイルを別にサーバー内に用意し、Web ページの HTML (HyperText Markup Language) 文書にはリンクで埋め込みを指示することにした。将来、XHTML が普及すれば、SVG が HTML 文書と一体となった、いわゆる複合ドキュメントとしての扱いが期待できる。

本研究では、普及の現状と将来性から、図の表示には SVG を採用した。SVG の開発システムは今回は採用せず、データベースから直接 SVG テキストを作成することにした。これは、本研究で表示するグラフや図が比較的単純であるから採用できた方法であり、将来、複雑な図が必要な場合に開発用ソフトがなければ困難が予想される。広く普及しているわけではないが、現状でも複数の有力な SVG 開発ソフトが利用でき、また、SVG 自身の将来の普及も見込まれるので、研究者らは楽観的に発展を予想している。

B.5 Cache Server Pages (CSP)の採用

Web アプリケーションの開発者側の利点の 1 つに、画面設計とデータベース設計が分離できる点がある。Web 画面の作成ツールとして良質なソフトウェアが多数普及している。研究者らの経験でも、画面の作成はツールを用いてデザイナーが行い、データベースに必要なプログラムコードは後から HTML に埋め込むことが可能であった。

今回の開発には、米国 Intersystems 社の Cache と呼ばれるデータベース開発システムを採用した。採用の理由は、将来予想される負荷に十分に耐え、なおかつ研究者らが扱い慣れているデータベース言語の MUMPS 言語を含んでいるからである。

Cache には 2 つの Web アプリケーション開発系がある。

1 つは WebLink Developer と呼ばれ、従来はこれを使うしかなかった。Web ページの設計は通常のホームページ作成ソフトで行い、出力の HTML ドキュメントに、必要なデータベース操作コードを MUMPS 言語を用いて ASP (Active Server Pages) と呼ばれる形式で追加する。このファイルの拡張子が asp のファイルは WebLink Developer の「コンパイラ」によって MUMPS 言語のルーチンに機械翻訳される。Web ページが参照されると、CGI (Common Gate Interface) として MUMPS のルーチンが実行される。

もう 1 つが Cache Server Pages (CSP) と呼ばれる、HTML ベースの言語である。他の Web アプリケーション開発系と同様に、繰り返しや条件分岐などの制御構造を HTML 内にタグで指定する。データベース操作もタグで指定することができる。ただし、データベースの細かな操作は埋め込みの MUMPS 言語または BASIC 言語を使う必要がある。Intersystems 社は、将来的には CSP を推進するという。

CSP の利点はコンパイルが自動化されていて、セッション管理もより高度になっている点である。不利な点としては、ライセンス料が高くなることである。本研究は三年計画であるため、CSP を用いることにした。

CSP と SVG を用いたシステムの応答を図 3 に示す。

今回作成された CSP ファイルの一例をリスト 1 に示す。一画面について、同様のリストが一つ対応する。CSP ファイルが更新されると、自動的に MUMPS のルーチン等が生成される(リスト 2、システム応答の(3))。

生成された MUMPS ルーチンの中から、グラフを描くための MUMPS ルーチンが呼び出される。グラフに対応する MUMPS ルーチンの一例をリスト 3 に示す(システム応答の(4))。グラフの種類ごとに同様のルーチンが用意された。これらのルーチンが自動作成した SVG ファイルの一例をリスト 4 に示す。

B.6 実行環境の選択

以上の選択により、今回の研究のためのシステムの仕様が確定した。

ハードウェアにはサーバ専用機を用意した。Cache は通常のパソコンでも動作する。実

際、従来は通常のパソコンを WebLink Developer の動作環境として使用していた。しかしながら、別の Web アプリケーションの応用ソフト利用では、数十の利用者のアクセスで OS が停止してしまう事態が頻発した。そのため、三年計画で、将来の利用者が数百名以上と予想される本研究には、CSP が十分に動作する機械環境を用意する必要が生じた。

ソフトメーカーとの打ち合わせにより、Intel アーキテクチャの CPU を用い、Microsoft 社の Windows 2000 Server が動作する小型のサーバを採用した。Cache は UNIX 系の OS でも動作するのだが、研究者が扱い慣れている Windows をとりあえず利用することとした。

Web ページの提供に必要な HTTP デーモン(httpd)には Microsoft 社の IIS (Internet Information Server)を採用した。研究者らは別の普及 httpd である Apache にも慣れているのだが、OS である Windows との相性を重視した。Cache は IIS にも Apache にも対応している。

B.7 画面設計

研究者間の意志疎通を図るため、上述の本調査にも使える開発システムを用いていくつかのグラフ表示プログラムを作成した。

ワイブル分布を用いた生存曲線は、医師側の理解が容易であることが予想されたため、最初に用意した。さらに、過去の研究では人形の絵を用いた説明図が評価されていたため、コホートを想起させる人形の図と、NNT を表す人形の図を用意した。

B.8 理解度について

本年度の模擬患者を用いた試験問題実施の目的は、システムの動作確認と、将来の本調査に向けた質問項目の吟味である。当面の測定項目として、理解度を選択した。

理解度とコンプライアンスには関連があることが示唆されている。また、満足度や患者・医師の行動と理解度の関係は本研究でいずれ解明したい話題である。したがって、試験結果から直接得られ、かつ、代替の結果(surrogate outcome)として適当と思われた理解度を最初の測定項目とした。

今回の調査では、「理解度」を 2 つの手段で測定する。一つは効用(utility)と同様に、100 点満点の数値で試験の回答者に答えさせるものである。これによって、各図に対して「自分はどれほど理解しやすいと考えたか」の数値が得られる。本研究では、こちらの理解度を強調する場合は「主観的理解度」と記述することにする。

もう一つは、試験の成績を直接測定するものである。試験問題は、各図から必要な数値を読み取って答えさせるものである。回答を「ほぼ正解」、「正解とは言いがたいが、評価の方向は間違っていない」、「効果の評価が逆向きの答え」に分類し、項目反応理論(IRT: Item Response Theory)が扱える形式とする。今回は予備調査であるので、被検者数が少なく、しかも各質問項目の項目特性を得るのは目的ではない。しかし仮に、将来大規模な調査、たとえば、質問項目数を 20 程度以上に増やし、被検者数を 200 程度以上に増やすなどの工夫、が行われれば項目特性が得られるはずである。得られた質問項目特性から逆向きに、図の理解度が得られる。

B.9 試験問題の作成

理解度測定のための仮の試験を画面の詳細設計の前に用意した。この当初の試験では、生存率、罹患率、NNT のそれぞれにおいて、数値と曲線グラフと人形の表示に対する理解

度を回答の正確さから評価することになっていた。

仮の試験の項目の中から、表示に特徴のある5つの図形を選択した。(1) 生存曲線。治療を行った場合と行わなかった場合を色分けし、2つの生存曲線を1つのグラフに描いたもの。(2) 累積イベント曲線。治療を行った場合と行わなかった場合を色分けし、累積イベント数の割合の2曲線を1つのグラフに描いたもの。(3) NNT 曲線。NNTの経時的変化をグラフに表したもの。NNTは一般には単独の数値で示され、本研究で用いた経時的なグラフ表示は珍しい。(4) コホート人形。設計当初は、1,000体の人形を表示し、治療の有無に関わらず死亡する割合を灰色の人形で、治療すればイベントが起きない割合を青または赤の人形で、治療の有無に関わらずイベントが起きない割合を白の人形で表した。(5) NNT 人形。人形の一体を青または赤で示し、治療の益を表した。治療必要人数を表す追加の人形は灰色表示とした。つまり、たとえばNNTが7.1の場合、灰色の6.1体の人形を追加表示した。

患者のシナリオは、図から読み取るべき生存率が適当にばらつく条件で、一例を用意した。つまり、高脂血症のガイドラインから、10年後の自然生存率(脳卒中や心血管イベントが起こらない率)が約70%、スタチンで10年間治療すれば生存率が約84%になる条件を導き出した。ワイブル分布の形状パラメータは日本の男性成人に対応する1.6を採用した。ワイブル分布の尺度パラメータは、以上の条件から算出できる。

試験では、10年後以外に、3年後の罹患あるいは持続的に健康な状態の割合を問うている。上述の条件では、3年後の自然生存率は約95%、3年間の治療後の生存率は約97.5%と計算できる。幸いなことに、これらの確率のオッズの対数を取ると、適当にばらつくことが確認できた。つまり、概算でも正答かどうかを判別できる。

B.10 医師による図と試験項目のチェック

研究者以外の医師によるチェックの際、(4)のコホート人形が分かりにくいとの指摘が出た。つまり、試験問題では百分率で問うているが、人形が1,000体あるので、単位換算が必要なのと、グラフでは2曲線になる情報を1つの図に盛っているのが、情報が詰め込まれすぎて理解しにくい、ということであった(図12)。そこで、模擬患者を用いた試験には、100体の人形で、治療群と未治療群を2つの図に分けて表現することとした。

さらに、(3)のNNT曲線は見慣れないため、(5)のNNT人形は試験の問いが割合で答えるので計算が必要なため、理解困難との指摘が出たが、本年度の実験期間中には有力な代替案が出なかったため、そのままの表示を使用することにした。

当初の案では、生存曲線等からイベント率等を数値で答えさせるものであったが、この計算は医師にとっても困難を伴うものであった。そこで、グラフには注目すべき場所を色丸で示し、その点のy軸の値をテキストでも表示することにした。この親切な機能は将来本開発されるであろう判断ツールでも採用可能である。

さらに、具体的な解答の数値を8つ挙げ、最も近い数値の記号を答えさせることにした。これによって、回答時間の節約と、無回答の割合が減ると考えられた。

B.11 模擬患者に対する試験の実施

非医師14名(学生、一般事務職)を被検者として、試験を実施した。回答時間は無制限であったが、全員が20分以内に回答可能であった。

被検者はディスプレイ上の図を見ながら、試験問題に答えた(図5~11)。模擬患者の設定や質問のテキストは、ディスプレイでも見られるが、別に用意した問題用紙にも同じ文章が書かれている。解答用紙を別に用意し、無記名で提出させた(資料1)。また、図やテキ

ストに対する感想を自由記述で回答させた。なお、前のページに戻るリンクを設けたので、回答者はいつでも任意の図を見ることができ、いったん回答用紙に記入した後も回答を変更することができる。問題用紙と解答用紙に図は描かれていない。

主観的理解度については SPSS 10.0J で統計処理を行った。探索的統計処理の後、対応のある t 検定で互いの平均の比較を行い、主観的理解度間の相関係数(スピアマン)を計算した。

B.12 倫理面への配慮

本年度はシステム開発と模擬患者を用いた予備調査までを行った。

次年度以降は実際の患者の参加を計画している。これは疫学的実験研究であるので、本学の倫理委員会の承認が必要であろう。インフォームドコンセントを指定の書式で得た上、患者のプライバシーの保護に関して厳重に管理する予定である。

C. 研究結果

C.1 システムの整備

本研究の遂行に必要なサーバコンピュータを用意し、インターネットに直接つながる大学ネットワークに接続し、動作を確認した。また、院内 LAN には医療情報部が用意したサーバパソコンが接続され、外来・病棟端末からのイントラネット利用が可能であることを確認した。

開発したシステムは、インターネット標準を採用している。しかしながら、端末側は古い OS を継続して使用しているものもあり、SVG のグラフが表示されないとの連絡をたびたび受けることとなった。インターネット系に関しては、SVG 表示機能は急速に普及しつつあるので、今後の改善が見込まれる。

院内 LAN 上の端末は、導入後 OS の更新がされていないので、SVG はそのままでは利用できない。OS または Web ブラウザの更新を行えば表示可能なはずで、来年度に向けて院内 LAN の管理者と改善を交渉中である。具体的には、院内 LAN では業務のための多数のソフトウェアが稼動しているので、SVG が表示できる環境への移行に伴う予期しない影響があるかどうかを検討している。

C.2 プログラム開発

最初の表示に 2 ヶ月間の開発期間を要した。これは、SVG も CSP も研究者らには初めての経験であるためと考えられる。また、その後も研究者間や情報提供者との話し合いで表示内容を少しずつ改良したために、試験問題の画面の完成までにさらに約 2 ヶ月を要した。しかし、これらの経験によって年度末にはシステム作成の技能と知識が獲得できたので、今後はより短期間での開発が可能と考えられた。

今回開発された、図を含む試験問題の Web ページは問題なく動作している。将来開発予定の判断ツールにも同じ技術が用いられるので、研究続行に必要な基本技術の開拓は一段落したと考えられる。

C.3 判断ツールの評価

学生、一般事務職の 14 人が被検者となり、ホームページに表示された生存曲線等の治療効果の説明図に対する試験問題に答えた。

C.3.1 正答率

正解を 1、不正解だが方向が同じ解答を 0.5、不正解または無回答を 0 の重み付けをすると、5 つの図の正答率は、

生存曲線	94.6%
累積イベント曲線	98.2%
NNT 曲線	62.5%
コホート人形	99.1%
NNT 人形	75.0%

正答率で並べると、

コホート人形 > 累積イベント曲線 > 生存曲線 > NNT 人形 > NNT 曲線

となった。

無回答はNNT曲線で3名、NNT人形で1名見られた。無回答の被検者は、他の3つの図に関しては正解を回答している。逆に、1名はNNTの2図に関しては正答を答えたにもかかわらず、生存率曲線で正解を逸している。

C.3.2 主観的理解度

主観的理解度(100点満点)の箱ひげ図を図4に示す。以下の略記を用いている。

生存曲線	seizon
累積イベント曲線	ruiev
NNT曲線	nttcur
コホート人形	chofig
NNT人形	nttfig

生存曲線と累積イベント曲線の主観的理解度の平均に差は認められなかった(対応のあるt検定)。スピアマンの相関係数では、生存曲線と累積イベント曲線間、NNT曲線とNNT人形間で有意な相関が見られた(表1)。なお、正規性の検定では、NNT曲線の主観的理解度だけが正規分布と明らかに異なる、との結果であった。

主観的理解度の算術平均と標本標準偏差は以下の通りであった。

	平均	標準偏差
生存曲線	73.6	20.8
累積イベント曲線	74.6	18.6
NNT曲線	20.4	21.6
コホート人形	85.4	11.8
NNT人形	29.6	24.8

各図の分かりやすさを算術平均で並べると、

コホート人形 > 累積イベント曲線 > 生存曲線 > NNT人形 > NNT曲線

となり、主観的理解度は正答率と一致した。

事前に研究者を含む3人の医師に主観的理解度を尋ねたところ、順序は一致していて、

生存曲線 > 累積イベント曲線 > NNT曲線 > コホート人形 > NNT人形

であった。ただし、コホート人形は旧版の表示である。

興味深いことに、この順で答えた被検者は皆無であった。

C.3.3 試験問題の自由記載欄に書かれた意見

(以下、順不同。元の文章とは、てにをはなどを変更しているものがある)

「病気にかかる見込み」と「心臓発作の見込み」はどう違うのだろうと思いました。NNT曲線は分かりそうで、なんだか分かりませんでした。

設問の意味を理解するのが大変でした。普段使わないような言い回しが多く難しかったです。

NNT 曲線と NNT 人形について、示していることは分かるが、設問の言い回しがわかりづらかった。

NNT 曲線と NNT 人形共にグラフや図は理解できるが、質問の意味がさっぱり分からなかった。

NNT 曲線と NNT 人形に関しては意味が分かりません。利益とは何のことですか。問題の言葉がよくわからないので人形のほうが分かりよいイメージを持ちますが、やはり文章の説明が分かりませんでしたので安易に予想して回答を書くことをやめました。

NNT 曲線と NNT 人形は難解で理解不能

NNT の定義が良く分からなかった。イメージもつかめなかった。自分が治療を受けなかったら NNT の数も変わるのか? 「治療を受けても受けなくても」で、治療を受けなかった場合の利益の確率は、この NNT 曲線から分かるのか? など疑問に思った。

NNT 曲線と NNT 人形は最初の説明文と設問がどうつながっているのか全く分からなかった。NNT 人形は人形で表していたので、ますます%とつながらず、わけがわからなくなった。あとはどれも大差なく分かりやすかったと思う。

NNT 曲線と NNT 人形は問題の意味がよく分かりませんでした

NNT 曲線は表の説明文を見てすぐに何を表しているのか分からなかった。さらにそれを分析するのに割合を計算するのが面倒に感じた。しかし、この設問の後にコホート人形と NNT 人形の設問があったので、これらは理解しやすかった。自分の病気のことだし家族のことも心配なので分かりにくくても色々な角度から分析してある資料はありがたいと思う(客観的に判断するという意味で)

グラフより人形のほうが分かりやすい。NNT 曲線の設問は理解するのにかなりの時間を要した。NNT 曲線で考えるのに結局頭では NNT 人形と同じことを考えていた。生存曲線は発作が発生していない人数だったので各設問とグラフの見方が混乱した。

曲線で考える場合も、頭の中では人形のような形におこかえて考えました。ですから、曲線のみで判断するのは難しい。しかし、人形のみだと割合の変化が見えてこないのが、最初から両方をつかって考えるほうが良いのではないのでしょうか。また質問のされ方がストレートでないもの(NNT 曲線の 2 番など)は患者には分かりにくいと思います。

D. 考察

D.1 目標となる疾患の選択

ガイドラインが整備されており、患者教育が必要であり、一般の関心も高い疾患として、本研究では最初に高脂血症を取り上げることとした。高脂血症の治療の目的は、動脈硬化性疾患、つまり脳卒中や心血管イベントの防止にある。そうしたイベントの発生率は、血漿コレステロール値や高血圧の程度などの危険因子からある程度予想できるほど、高脂血症は研究されている。

今回の研究の端緒としての高脂血症の選択に対する反論は、研究者の周りからは聞かれなかった。また、模擬患者にも自然な状況設定として受け入れられたようである。

ただし、材料と方法の項で述べたように、糖尿病などの、他の疾患のガイドラインも検討し、本研究が他疾患へ応用可能であるよう、極力注意してシステムを開発した。

D.2 ワイブル分布

本研究で開発した判断ツールの中に出てくる生存曲線は、信頼性工学でよく検討されているワイブル分布に基づく理論分布である。実験で用いられた滑らかな生存曲線は、医師にも模擬患者にも、ほとんど違和感なく受け入れられた。

一般に、医学分野の生存解析では、カプラン・マイヤー法などのノンパラメトリックな解析が用いられ、パラメトリック解析であるワイブル分布への当てはめは、ほとんど行われない。そのため、医学分野でのワイブル分布の認知度は低い。

しかしながら、よりよい理解のために、さまざまなシナリオを検討するシミュレーションには、ワイブル分布などの、何らかの生存に関するモデルが必要である。今回の経験から、医学分野においてもワイブル分布の採用がそれほど突飛でないことが明らかになった。

D.3 インターネット技術

本研究では、できるだけ特定の技術に頼らないで済むように、インターネット標準を可能な限り採用した。材料と方法の項でも述べたように、インターネット技術には親しみを覚える人が多く、システムに関する教育が容易であるという副次効果もある。また、対応機器やソフトウェアも豊富で、経済的で信頼のおけるものが容易に手に入る。たとえば、本研究では、Web ページのセッション管理が必要となると予想されたが、セッション管理を自動で行う基本ソフトは容易に手に入った。一例ではあるが、インターネット技術の普及がもたらした恩恵と考えられた。

しかしながら、インターネット技術には、現在も発展しつつある、つまり必ずしも安定していない技術、という側面もある。

たとえば、本研究で採用した SVG がある。SVG 自身は厳格な規格であり、逸脱したシステムはない。しかし、SVG の記述はユニコードを前提としており、UTF-8 と呼ばれる情報交換形式が要求されている。ユニコードの情報交換形式は UTF-8 をはじめとする数種類のバリエーションがあり、そのためにユニコードの編集ソフト(エディタ)はかなり複雑なことになっている。また、SVG と HTML との統合の道筋も、必ずしもはっきりしない。本研究では極力主流となりそうな規格を将来も取り込むことで対応することとした。

もう一つ、CSP のような Web アプリケーションの開発系の問題がある。Web アプリの開発系は、経済的で信頼できるものが多数あるが、どれもが独自規格であり、開発したアプリケーションを他に流用することができない。プログラム言語の FORTRAN や COBOL の

ような規格統一の話は、調査範囲では見当たらなかった。したがって、現状ではいずれかの開発系を選択することになる。本研究では、研究者が慣れているという理由で MUMPS 言語が利用できる CSP を採択するに至った。

D.4 図の表示方法の検討

SVG の採用を決断したのは、初年度も後半になってからである。それ以前は、MUMPS からパラメータ付きで描画ソフトを起動させ、図をビットマップ形式でファイルにいったん格納させてから、別のソフトを起動させ、インターネットで使われる画像圧縮の一つである PNG (Portable Network Graphics)形式に変換させ、Web ページの HTML にはリンクで絵を埋め込む計画であった。PNG は JPEG や GIF に次いで普及している可逆画像圧縮形式である。描画ソフトは簡単なものを C 言語で作成する予定であった。

SVG は現在でも対応していない Web ブラウザが多く残っている。しかしながら、円や直線を xml 準拠のテキスト形式で表すため、グラフや線画のイラストなどは非常に高圧縮になる。拡大しても画像の劣化が無く、SVG を知っている利用者なら色や線の太さを変えるなどの加工も容易である。また、規格上はマウスのクリックへの反応や、アニメーションの機能もある。xml 形式で表されるので将来は HTML に完全に埋め込むことも可能である。

PNG による図の埋め込みも魅力的ではあったが、将来性と拡張性を考え、研究の初期段階から SVG を採用することとした。

	SVG を用いる図の表示	PNG を用いる図の表示
圧縮原理	基本図形の記号化	同じパターンの冗長性除去
圧縮率	基本図形は高圧縮	イラスト・線画で高圧縮
Web ブラウザの対応	普及進行中	ほとんどのブラウザが対応済
拡大時の画像の劣化	なし	設計時の解像度のまま
線の太さ等の加工	容易	困難
HTML との統合	将来の可能性あり	不可能

D.5 サーバ専用機の導入

理屈としては、通常のパソコンでもデータベースの運用は可能である。しかし、通常のパソコンでは、周辺機器などの機器構成が予想も付かず、結果として基本ソフトウェアの動作が不安定になりがちである。実際、能力的には今回導入したサーバと比べても遜色の無いパソコンを Web アプリケーションに使用していたが、アクセス数が数十に増えると処理が追いつかなくなると全く動作しなくなることが 1 週間に 1 回程度の頻度で発生した。

本研究では将来において安定動作が求められたので、サーバとして設計された計算機で、必要十分な機能を有するものを採用することとした。サーバと言えども、基本ソフトの構成はパソコンと類似しているため、操作に迷うことは無かった。

D.6 表示されるグラフの選択

文献的には癌の再発リスクに関して、数値による表現、棒グラフによる表現、十体の人形による表現、千体の人形による表現などが検討されている。

医師は医学研究論文等で生存曲線を見慣れており、患者に説明する際に医師側の違和感が少ないと考えられる。今回実施した模擬患者による理解度調査でも、生存曲線の評価はまずまずであり、患者にとっても苦痛を感じるほどではないと感じられた。模擬患者の意見でも、経時変化が分かりやすい、といった感想も聞かれた。

千体の人形による表現は、事前の意見調査で、百体にしたほうが百分率に換算しやすい

との意見があったため、模擬患者での実験には百体の人形表現を用いた。

D.7 NNT について

NNT は一つの数値で治療効果が表されるので分かりやすい数字とされている。しかしながら、NNT は経時的に変化するし、相対危険度の情報が失われているので、必ずしも簡便さに賛同しない医師もいる。今回の実験でも、試験問題の日本語表現に苦慮するありさまで、被検者には当方の苦悩が伝わってしまったようである。

NNT も人形表現すれば理解の助けになるかと考えたが、試験の正答率は比較的高いものの、主観的理解度は全体的に点数が低く、被検者間のばらつきも大きかった。ただし、一部の被検者の意見からは、人形表現は NNT の理解に多少役立ったようである。

D.8 理解度について

本年度の評価実験では理解度を主に測定した。理解度はコンプライアンスや満足度と関連しており、かつ測定が容易と考えられたからである。

材料と方法の項で述べたように、2種の「理解度」を用意した。

一つは、効用と同様の方法で被検者に記載させる、主観的理解度である。もう一つは試験の点数である。

今回の実験結果から、各図で得られた主観的理解度と試験の点数の順は一致していることが示された。

効用はスタンダードギャンブル等の導出法から主観確率と解釈できるので、本研究で用いた主観的理解度も主観確率の一種と解釈可能である。

一方、試験の点数は普遍的な客観確率ではない。項目反応理論の帰結として、テストの平均点数は問題の出しようによって、いくらでも調整可能だからである。しかし、他施設での同様のテストと比較可能という観点からは、普遍性のある数値である。たとえば今回の実験では、生存曲線とコホート人形の質問は全く同一なので、生存曲線とコホート人形の理解しやすさは被検者の理解度とは独立に、直接比較できる、ということである。

本研究では今後も両者の特性を考慮して、併用することになるだろう。

D.9 実験結果の解釈

開発された判断ツールの可能性評価のため、14人の大学生及び一般事務職に模擬患者になっていただき、試験に答えていただいた。おそらく、この被検者集団では、試験問題の図から数値を読み取ったり割合を計算することは、さほど困難ではないと推察される。

この実験に先立って、臨床疫学に詳しい3名の医師にも問題のチェックのために同様の試験問題を解いていただいた。ただし、コホート人形の図は医師に対しては千体の図で、しかも治療と未治療を一つの図で表したものであった。被検者には百体の人形の図を治療と未治療の2枚並べて表示している。

主観的理解度では、生存曲線と累積イベント曲線間に平均の差は認められなかったし、有意の相関係数も得られた。両曲線は鏡像になっているので、この結果は当然と言える。医師には生存曲線の方が見慣れているので、次年度以降の実験では生存曲線を使用することになるだろう。

NNT 曲線と NNT 人形の相関係数も有意であった。これは試験の時間中に NNT が理解できたか、できなかったかの違いが反映されていると考えられた。傍証として、被検者間で主観的理解度の点数が大きくばらついている。NNT 人形の方が点数が高いので、NNT

が理解できる人には人形表示が役立ったと考えられる。

NNT 曲線と NNT 人形表示は模擬患者にはまったく人気がなかった。ところが、医師からの回答では、NNT 曲線はコホート人形よりも理解しやすいとなっている。これは、臨床疫学では NNT がしばしば話題になるために、内容の予備知識があったからだと考えられる。NNT は平文で意味を記述しようとする、内容がかなり難しく、被検者に戸惑いを与えたようである。一方、医師の NNT に対する不満は、相対危険率減少(RRR)の情報が無くなることにある。

コホート人形の人気が高いのは、示唆に富む。最も安易な解釈は、試験問題の問いが百分率で尋ねているので、直接の数値が問題内に書いてあったため回答しやすかった、であろう。もう一つの解釈は、実際の患者集団を想起しやすかったから、であり、これは図の作成者の意図に合致している。将来の実験系では、この違いを明らかにする設計が必要と考えられた。一方、医師の人気がやや低いのは、経時的な変化が読み取りにくいからと考えられ、被検者の一人もそのように感想で述べている。

図の作成者からすると、NNT 人形の人気の無さは残念である。解説の日本語表現に問題があったかもしれないので、NNT は NNT 独自の調査を要すると考えられた。

コホート人形の印象の強さと、生存曲線の情報量の多さが両立するような表現法があればそれに勝るものは無い。しかし当面はこの両者を並行して発展させることになる。

臨床疫学に関心のある医師にとっては、生存曲線の意味は非常に分かりやすい。したがって、患者への説明も円滑に進むと考えられる。しかし、患者からすると、時間変化があるという印象しか残らない可能性がある。その点では、コホート人形の方が真実味を与えると考えられる。

D.10 今後の展開

判断ツールとして役立つためには、パンフレットがいつでも好きな時に見られるように、インターネットで参照できる、パソコンでの表示ソフトを組み込んだフロッピーなどで渡すなどの、患者の身近な存在になる必要があると考えられる。

医師からの要請は、外来端末などで診療中に解説できる環境整備であろう。できれば電子カルテに組み込まれた説明ツールとしての位置づけが望ましいと考えられる。次年度の研究では、電子カルテの部品としての展開も調査範囲に含めたい。

これら、患者と医師からの要請は、必ずしも矛盾するわけではないので、両立するシステムの開発を目指したい。

E. 結論

医療効果の予測の提示法を調査・選択し、システムを開発した。

システムに必要なパラメータは、厚生労働省の公開資料やガイドラインから取得可能であった。

システムはインターネットでも病院内 LAN でも動作確認できた。つまり、広く普及可能な方式を選ぶことができた。さらに、電子カルテ等との統合も視野に入れることができた。

システムには専門家から批判的吟味が加えられ、模擬患者を用いた実験を行うことができた。本年度の評価項目は理解度とした。

模擬患者での実験では、主観的理解度と正答率が一致した。また、各表示法について考察を行うことができた。これらにより、次年度からの本調査に必要な情報が得られた。