

複数のデータベースからの知識発見 —肝機能検査データに関する因果モデルの比較と評価—

稲田 政則[†] 寺野 隆雄[‡]

[†] 虎の門病院臨床検体検査部 〒105-8470 東京都港区虎ノ門 2-2-2
[‡] 筑波大学大学院ビジネス科学研究科 〒112-0012 東京都文京区大塚 3-29-1
E-mail: [†] m-inada@mac.email.ne.jp, [‡] terano@gssm.otsuka.tsukuba.ac.jp

あらまし 本論文では、アクティブマイニングの一例として、対話型進化計算を用いて2つの医学データベースから肝機能検査に関する共通の因果モデルを探索する。データの互換性が保証されない複数のデータセットにおいては、単なる併合やモデルの統合はできない。そこで、各データセットから因果モデルを同時に探索する問題を設定し、属性選択の手続きに対話型進化計算を取り入れたデータマイニングツールを試作した。モデルに対して人間が対話的に主観的評価を行うことで、利用者の積極的なフィードバックの下に、進化計算によって効率的な仮説空間の探索を実現した。最終的に、利用者に支持され、高い適用性の期待される因果モデルが得られた。

キーワード アクティブマイニング、対話型進化計算、構造方程式モデリング、肝機能検査データ

Active Mining from Plural Data Sources

—A Study in Quest of Common Causal Models from Two Medical Test Data Sets—

Masanori INADA[†] Takao TERANO[‡]

[†] Department of Clinical Laboratory, Toranomon Hospital 2-2-2 Toranomon, Minato-ku, Tokyo, 105-8470 Japan
[‡] Graduate School of Business Sciences, University of Tsukuba 3-29-1 Otsuka, Bunkyo-ku, Tokyo, 112-0012 Japan
E-mail: [†] m-inada@mac.email.ne.jp, [‡] terano@gssm.otsuka.tsukuba.ac.jp

Abstract This paper employs interactive evolutionary computing techniques for an active mining task. In the task, we search for common causal models in two data sets from different medical databases. Because medical test data are incompatible among hospitals, it is difficult to merge plural data sets from different sources. We therefore formalize a simultaneous searching problem to get common causal models described by structural equation modeling for each data set. To acquire common causal models, we designed a mining system, which copes with the feature selection procedure with Evolutionary Computation techniques to interactively evolve offspring. In the procedure, a user will judge individuals subjectively or interactively in a human-in-a-loop manner. Experimental results on liver function test data show the effectiveness of the proposed method.

Keyword Active Mining, Interactive Evolutionary Computing, Structural Equation Modeling, Liver Function Test Data

1. はじめに

本論文では、2つの医学データベースに蓄積された肝機能検査データから、対話型進化計算を用いて共通の因果モデルの探索を試みる。

データマイニングの医療応用においては、意外性や新規性のある知識の発見に期待が高まる一方で、EBM(Evidence-Based Medicine)[1]に貢献し得る確実な知識の獲得も重要な課題である。EBMでは、慎重に計画された臨床試験の結果を中心とした科学的根拠が求められており、医学統計学の世界では、研究目的外に取られたデータからの後ろ向き研究の危険性が論じら

れている[2]。EBMの文脈の中で、データマイニングがEvidence生成に補完的な役割を果たすには、確実な知識を提供できなければならない。

近年、医学データベースに蓄積された情報からデータマイニングによって簡単にEvidenceが得られるとのベンダーの宣伝やユーザーの報告が散見される。しかし、データの質に関する考察と得られた知識の妥当性の評価をなくしては、危険なEvidenceの氾濫を来す可能性が高い。

データの質に関する考察を行なう方法の一つとして、複数施設のデータベースの比較が挙げられる。一

般に、病院のデータベースは受診患者の偏りを含み、臨床検査データは採用する測定法によってそのデータの正確さや精密さに違いがある。患者属性や測定法が異なる施設間でのデータの互換性は保証されない場合が多い。従って、ここから得られる結論が即座に一般性を持つ Evidence になるとは考え難い。仮説生成であるとの立場に立ってみても、その仮説がどの程度有力であるかの見通しも不明な場合が多い。そこで、仮説の有力さを知る一つの方法として、複数施設のデータベースにおいて共通モデルを探索することは興味深い。

得られた知識の妥当性を評価するためには、専門家（利用者）が積極的に知識発見のプロセスに関わる必要がある。ある特定の基準に従う自動的なモデル選択では、バイアスや交絡の存在によって不適切なモデルが選択される危険がある。利用者が対話的にモデルの選択を行なえる環境が求められる。

著者らは昨年、アクティブユーザーリアクションの医療における具体的事例として、肝機能検査データの因果モデルの構築を試みた[3]。ここでは、「学習結果からの学習」という Chan らのメタ学習の定義[4]にならない、複数のデータマイニング手法を適用し、得られたマイニングに関する知見を利用して因果モデルを構築した。この事例は、一総合病院の臨床検査データベースの肝機能検査データを対象にしたものであり、得られたモデルの一般性は確認できていなかった。

今回、他施設の肝機能検査データを入手し、因果モデルを比較評価する機会を得ることができた。当初、2施設のデータセットの併合やモデルの統合を目指したものの、データの質の面から単に併合・統合することに危険があったため、共通モデルの探索問題として取り組む方針とした。

共通モデルの探索において、2つのデータセットを同時にマイニングする対話型進化計算[5]を導入した。利用者の主観的な選好を取り入れた属性選択と進化計算の組合せにより、利用者の積極的なレスポンスの下に効率のよいモデル探索を試みた。

本論文の構成は以下の通りである。2章では対象問題を考察し、その解法として対話型進化計算を考える。3章では2施設のデータセットを概観した上で問題を設定し、本問題に特化した対話型進化計算（対話型 EC）を設計する。4章ではモデル探索の結果を示すとともに、比較のために線形回帰分析による結果を示す。5章において考察を述べたあと6章にて総括する。

2. 研究の背景

2.1. 対象問題の考察

今回対象としたデータセットは、虎の門病院および千葉大学医学部附属病院の臨床検査データベースから ICG 検査の依頼のあったレコードを抽出したものであ

る。著者らの興味は、虎の門病院のデータセットから構築された肝機能検査データの因果モデルが他施設のデータでも支持されるか否か、あるいは、データの互換性の問題を越えて2つのデータセットに共通する因果モデルを見出せるか否かである。

獲得した因果モデルは予測モデルとしての利用を想定している。すなわち、肝臓の異物排泄機能試験のひとつである ICG 検査（インドシアニンググリーン試験）[6]の成績を、肝機能を反映する他の臨床検査データから予測するシステムの実用化である。ICG 検査は、生体に薬剤を静注してその排泄（停滞率）を追跡する検査であり、生体に対する負担が大きい上に、厳密な時間管理が求められる。故に過誤が比較的発生しやすい検査である。臨床検査室にとって、ICG 検査の実測値と他の検査データによる予測値との乖離から過誤を発見する業務支援システムが構築されれば有益である。

検査データの予測モデルの要件として、その予測精度ばかりではなく、モデルの内部構造に関して医学的な意義が求められる。モデルの振る舞いについて利用者が容易に理解できることが求められる。この要件に応えるためには、利用者が積極的にモデル探索に関与する必要がある。また、特定の施設だけではなく、複数の施設で利用可能な普遍的なモデルが望まれる。

以上の背景から、対象問題においては、利用者がマイニング結果に対して積極的に反応し、新たなマイニングの方向性を見出すことを支援するアクティブユーザーリアクション[7]の環境が求められる。ここでは、複数のマイニング結果の比較や複数データセットからのマイニングといったメタ学習の方法論が必要となる。

2.2. 対話型進化計算

アクティブユーザーリアクション環境を実現する技術として対話型 EC が注目される。対話型 EC は、グラフィックアートや意匠・工業デザインなど芸術の分野、音声処理をはじめとする工学分野へと幅広く普及しつつある技術[5]であり、マーケティング情報分析をはじめとしたデータマイニングにも応用されている[8][9]。対話型 EC とは遺伝的アルゴリズム (GA) [10] を利用する問題解決において、利用者が対話的にそのプロセスに介入する手法全般を指す。その特徴は、遺伝的操作における個体評価を、あらかじめ定義した評価関数の代わりに、個体の表現型の評価を人間の主観的判断で行なう点にある。

石野らが提案した SIBILE (Simulated Breeding and Inductive Learning) [8]は、帰納学習の属性選択問題向けのツールである。属性選択の問題は適切な仮説空間の選択問題であり、機械学習の文脈においては帰納バイアスの選択[11]というメタ学習の一つである。仮説空間の選択の適否はマイニングの結果に大きく影響す

ることがあり、利用者が積極的に介入すべきフェーズとも言える。

今回の対象問題の解法として、SIBILEと同様に属性選択に利用者が対話的に介入する仕組みを用いた、2つのデータセットからの同時マイニングが挙げられる。

3. 問題設定

本研究では、2つのデータセットの質的な違いから、データセットの併合やモデルの統合ではなく、共通因果モデルの探索タスクとして問題を設定する。因果モデルを構造方程式モデリング[12][13]で記述・推定し、対話型 EC の枠組みによって探索を行なう。

3.1. 2つのデータセットとデータ互換性

本研究では2つのデータセットを対象とする。各データセットは、ICG 検査データのほか、表1に示す15の生化学的検査項目[6]から構成される。

虎の門病院のデータセットは、同院臨床検査データベースから ICG 検査の検査に依頼のあった166件のレコードを抽出したものである。

1千葉大学医学部附属病院のデータセットは、同院医療情報部より提供された肝炎データから ICG 検査の依頼のあったレコードを抽出(15の生化学的検査データは ICG 検査の前後1日以内のレコードを抽出)したものである。ただし、ICG 投与量が不明なため、実際の ICG 値から標準的な投与量と思われる91レコードのみを抽出している。

両データセットにはデータの互換性に関する問題が指摘される。例えば表1から明らかのように、CHEでは単位が異なり、ALBでは単位は同じであっても平均値に有意な差が認められる。同じ検査項目であっても、測定法や検査機器、検査の手順は施設において任意に選ばれており、単位の違いばかりではなく、項目によっては正確さや精密さが異なっている可能性もあ

る。臨床検査の標準化の歴史は浅く、本データセットの中でも AST, ALT, LDH, ALP, GGT の項目が標準化されたに過ぎない[14]。しかし、これらの項目においても、今回抽出した千葉大病院のデータセットは、標準化前のデータが大部分を占めており互換性は保証されない。

そこで本研究では、単なるデータセットの併合やモデルの統合による危険を避けるために、各データセットから独立して因果モデルを探索し、共通因果モデルを発見するタスクとして問題を設定する。

3.2. 因果モデルの記述と推定

因果モデルの記述・推定には、昨年のマイニング結果[3]を踏まえて、構造方程式モデリングを用いる。構造方程式モデリングは、基本的に線形モデルであるが、潜在変数(実際には観測されていない変数)を導入できるように高いモデル記述力があり、因果関係の記述に適したツールである[15][16]。

図1には、ALB, DBIL, CHE および ICG 値による因果モデルの例を示す。このモデルは、各検査データが肝機能という一つの潜在因子と各独自因子で説明される単回帰モデルから構成される一因子モデルである。ICG 値のパス係数に1, 切片を0に指定することで、肝機能の尺度は ICG 値と同じ尺度となる。モデル内のパス係数および平均値・切片は学習用データから推定される。

推定された因果モデルを用いて、ALB, DBIL, CHE から ICG 値を推定することができる。この ICG の推定値と実測値から決定係数を計算することができ、推定に用いたデータセットを使えば学習誤差を、未知のテストデータセットを使えば予測精度を見積もることができる。今回は、データ数が少ないこと、現象記述のモデルの探索という目的から学習誤差

表1) データセットの概要

| 変数 | 検査名称 | 単位 | | 平均 | | 標準偏差 | | 最小 | | 最大 | | 欠測 | |
|------|-----------|-------|-------|-------|-------|------|------|------|------|-------|-------|-----|-----|
| | | 虎の門 | 千葉大 | 虎の門 | 千葉大 | 虎の門 | 千葉大 | 虎の門 | 千葉大 | 虎の門 | 千葉大 | 虎の門 | 千葉大 |
| ICG | ICG試験* | % | % | 1.41 | 1.41 | 0.30 | 0.30 | 0.78 | 0.60 | 1.94 | 1.93 | 0% | 0% |
| TP | 総タンパク | g/dl | g/dl | 7.35 | 7.36 | 0.66 | 0.83 | 5.04 | 5.40 | 9.83 | 9.70 | 0% | 0% |
| ALB | アルブミン | g/dl | g/dl | 3.40 | 3.97 | 0.48 | 0.44 | 1.93 | 2.80 | 4.40 | 4.80 | 0% | 0% |
| TTT | チモール混濁試験 | S-HU | U | 12.84 | 4.87 | 8.88 | 3.17 | 0.40 | 1.20 | 39.80 | 12.00 | 8% | 0% |
| ZTT | 硫酸亜鉛混濁試験 | KU | U | 11.00 | 14.21 | 4.34 | 5.93 | 1.40 | 1.20 | 24.40 | 28.50 | 8% | 0% |
| TBIL | 総ビリルビン** | mg/dl | mg/dl | 1.29 | 1.26 | 0.77 | 0.56 | 0.32 | 0.40 | 4.51 | 3.70 | 0% | 0% |
| DBIL | 直接ビリルビン | mg/dl | mg/dl | 0.71 | 0.50 | 0.42 | 0.26 | 0.16 | 0.10 | 2.71 | 1.60 | 1% | 0% |
| AST | AST | IU/l | U/l | 62 | 104 | 56 | 82 | 13 | 17 | 526 | 506 | 1% | 0% |
| ALT | ALT | IU/l | U/l | 65 | 106 | 74 | 133 | 7 | 13 | 513 | 677 | 1% | 0% |
| LDH | LD | IU/l | U/l | 171 | 193 | 39 | 69 | 92 | 101 | 326 | 459 | 0% | 0% |
| ALP | ALP | IU/l | U/l | 315 | 125 | 173 | 78 | 57 | 39 | 1136 | 382 | 0% | 0% |
| GGT | γ-GTP | IU/l | U/l | 85 | 61 | 115 | 63 | 7 | 7 | 1154 | 415 | 0% | 5% |
| LAP | LAP | U/l | U/l | 42 | 275 | 26 | 115 | 22 | 47 | 104 | 805 | 95% | 5% |
| CHE | コリンエステラーゼ | Δ p H | U/l | 0.75 | 3.48 | 0.33 | 1.19 | 0.11 | 0.96 | 1.63 | 8.02 | 0% | 5% |
| TCH | 総コレステロール | mg/dl | mg/dl | 161 | 175 | 42 | 40 | 85 | 74 | 391 | 279 | 2% | 0% |
| AG | A/G比*** | - | - | 0.88 | 1.24 | 0.20 | 0.32 | 0.44 | 0.62 | 1.40 | 1.96 | 8% | 0% |

*常用対数変換済, **DBILを含む, *** A/G=ALB/(TP-ALB)

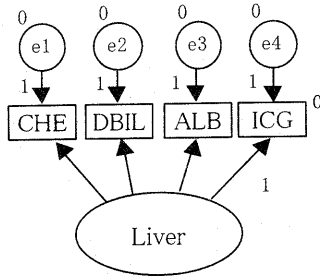


図1) 因果モデルの例

○で囲まれた変数は潜在変数、□で囲まれた変数は観測変数を表す。矢印は変数間の因果関係(パス)を表す。この図は CHE, DBIL, ALB, ICGの各変数が、共通の潜在因子 Liver と各変数独自の誤差変数から説明されることを表している。○, □, 矢印付近の数値は、平均値, 切片, パス係数が固定されていることを意味する。数値の示されていないところはデータから推定される。

をモデルの客観的な評価指標とする。

今回のモデルの探索では、図1の因果モデルにおいて、潜在変数 Liver と ICG を固定し、その他の変数のセットを変化させ、推定されたモデルの比較と ICG 値推定における学習誤差の比較を行なう。

本来、構造方程式モデリングはモデル検証のツールであるが、しばしば、モデルの構造がデータから識別できない場合や分散が負値として推定される不適解の場合があり、モデルの修正が必要となる[15][16]。モデルの修正を拡張し、広く探索に利用することは実務的に興味深い。本研究では、構造方程式モデリングを対話型 EC による属性選択の枠組みの中で探索ツールとして利用する。

3.3. 対話型進化計算による共通モデルの探索

導入する対話型 EC の概略を以下に示す。実際の対話によるモデル探索は第1著者が行なう。

個体数: 10 (初期世代はランダムに生成)

遺伝子型: 15 bit

ビット列は属性セットに対応する。各遺伝子座は各変数に対応し、その変数が選択されているときに 1, 選択されていないとき 0 を表す。

表現型: 構造方程式モデルおよび学習誤差

属性セットに対応する構造方程式モデル: パス係数推定の収束状況, 推定されたパス係数の符号と大きさ, その標準誤差, 切片, 誤差分散, 標準化パス係数, および, 学習誤差(決定係数)を評価の対象とする。

選択: エリート戦略

主観的基準と客観的基準により2つの個体を次世代に残す。主観的基準: 利用者が各個体の表現型に応じて 0~1 の選好度を与え、この選好度の最も高い個体を残す。客観的基準: 各個体毎に虎の門および千葉大データセットでの学習誤差(決定係数)に各々ラン

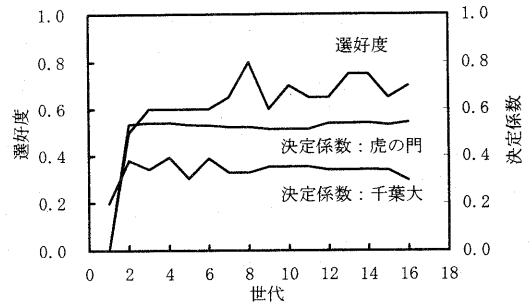


図2) 対話型進化計算の過程における選好度と決定係数のメディアン推移

クを付け、そのランクの和が最も小さい個体を残す。なお、構造方程式モデルの推定においてモデルが識別されない場合や不適解が得られた場合は、選好度 0, 決定係数 0 の評価を与える。残りの個体は主観的な選好度に応じたルーレット方式で選択する。

交差: 0.6 の確率で一点交差を起こす。

突然変異: 0.03 の確率でビットを反転させる。

世代数: 10~20

対話型 EC の終了後、利用者による主観的評価の一貫性および選好に関わる要因について、属性セット(ビット列)を説明変数、選好度を目的変数として数量化理論1類を用いて分析を行なう。

3.4. 線形回帰分析によるモデルの探索

比較のために各データセットにおいて線形回帰分析による重回帰モデルの探索を行なう。変数選択はステップワイズ法(変数投入 $p < 0.05$, 変数除去 $p > 0.10$)で行なう。なお、欠測値はペアごとに削除する。

4. 結果

4.1. 共通因果モデル

図2に対話型 EC の過程における選好度と決定係数のメディアン推移を示す。初期世代では識別不能なモデルや不適解のモデルが多勢を占めていたが、第2世代からは適切に推定できたモデルが多勢を占めている。また、最適解とは言えないものの、通常の GA に比べて少ない世代数で、利用者が支持する解を得ている点は特徴的である。

表2に利用者が最も高い選好を与えた因果モデルの概要を示す。変数として ALB, TTT, TBIL, CHE が選ばれている。各変数のパス係数を虎の門データセットと千葉大データセットで比較すると、単位の違い等も影響して両者で異なるものの、標準化パス係数ではほぼ近い値を示している。また、学習誤差を反映する決定係数では虎の門で 0.57, 千葉大で 0.28 であった。これらの値は必ずしも対話プロセス全体を通しての最大値ではなかった。

選好度に関する数量化理論1類による分析では、重

表2) 利用者が選んだ構造方程式モデルの概略

| | | 虎の門 | | 千葉大 | |
|-------------|------|--------|-------|--------|-------|
| | | 係数 | 検定統計量 | 係数 | 検定統計量 |
| パス係数 | ALB | -1.562 | -11.0 | -1.756 | -5.0 |
| | TTT | 13.625 | 4.6 | 5.283 | 2.5 |
| | TBIL | 1.433 | 5.8 | 1.310 | 3.4 |
| | CHE | -1.068 | -11.1 | -5.744 | -5.1 |
| | ICG | 1(固定) | | 1(固定) | |
| 切片 | ALB | 5.600 | 27.9 | 6.444 | 12.9 |
| | TTT | -6.491 | -1.5 | -2.579 | -0.9 |
| | TBIL | -0.726 | -2.1 | -0.592 | -1.1 |
| | CHE | 2.253 | 16.5 | 11.586 | 7.2 |
| | ICG | 0(固定) | | 0(固定) | |
| 標準化 パス係数 | ALB | -0.810 | | -0.716 | |
| | TTT | 0.340 | | 0.300 | |
| | TBIL | 0.464 | | 0.423 | |
| | CHE | -0.815 | | -0.851 | |
| | ICG | 0.835 | | 0.608 | |

相関係数が0.83であり、利用者の選好の一貫性が示唆されている。また、各変数の回帰係数を比較すると、LAPやALP、TBIL、ASTがモデルの選好に負の影響を与えている一方、CHE、ALB、TTT、GGTが正の影響を与えていることを読み取ることができる(図3)。

4.2. 重回帰モデル

表3にステップワイズ変数選択により両データセットについて得られた重回帰モデルの概要を示す。虎の門データセットでは、ALB、ZTT、DBIL、CHEが有意な説明変数として選ばれ、重相関係数が0.78であった。千葉大データセットでは、TTT、TBIL、AGが選ばれ、重相関係数は0.71であった。

5. 考察

今回用いた対話型ECは、データマイニングタスクの中で、モデルの選択(属性選択)に利用者の主観的判断が反映されるよう設計したものである。2つのデータセットに共通する因果モデルを探索するような問題では、モデルの良さを明示的に評価関数として表すことが困難であり、人間の主観的な評価が有力な手がかりになる。以下では、本事例における対話型ECの役割を論じた後、得られた共通モデルの特徴を考察する。

本事例での対話型ECの過程は16世代で完了している。明示的な評価関数を持たない対話型ECでは、評価者による任意の評価によって収束しない可能性があるが、今回の事例では、識別不能、不適解といった致命的なモデルは早期に消失し、第2世代以降では利用者の選好度は安定している(図2)。もし、ランダムに属性セットを生成する対話処理を行っていたならば、利用価値の低いモデルを繰り返し評価する必要に迫られたものと推察される。また、属性セットと選好度の関係には、数量化理論1類による分析の結果、統計的に有意な構造(重相関係数0.83)が認められ、主観的

表3) 重回帰モデルの比較

| 虎の門 | r | r ² | 調整済みr ² | | |
|---------|-------|----------------|--------------------|--------------------|------|
| | 0.780 | 0.608 | 標準誤差 | 有意確率 | |
| n = 166 | | 回帰係数 | | | |
| (定数) | | 2.029 | 0.158 | <1% | |
| ALB | | -0.212 | 0.046 | <1% | |
| ZTT | | 0.010 | 0.004 | 1.4% | |
| DBIL | | 0.209 | 0.042 | <1% | |
| CHE | | -0.209 | 0.070 | <1% | |
| 千葉大 | | r | r ² | 調整済みr ² | |
| | | 0.705 | 0.498 | 標準誤差 | 有意確率 |
| n = 91 | | | 回帰係数 | | |
| (定数) | | 1.977 | 0.147 | <1% | |
| TTT | | -0.018 | 0.009 | 3.7% | |
| TBIL | | 0.193 | 0.041 | <1% | |
| AG | | -0.582 | 0.085 | <1% | |

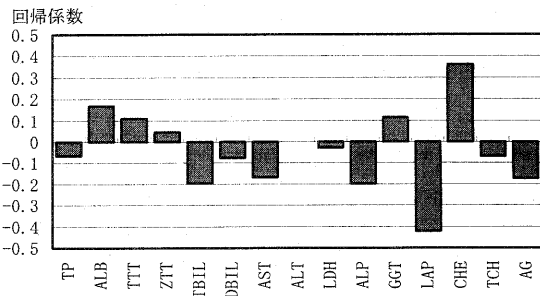


図3) 選好度に関する数量化理論1類分析の結果

評価の一貫性が示唆されている。今回の事例では、対話型ECによって効率的・効果的な探索が実現されているといえる。

利用者が最も高い選好度を与えたモデルは、ALB、TTT、TBIL、CHEを含むモデルである(表2)。ALBとCHEについては、昨年のマイニング事例[3]でも支持された変数であり、臨床的な意義としても真の意味での肝機能を反映する指標としてICG検査と同等の意義が知られている[17][18]。この2つの検査項目が利用者のモデルの選好要因として寄与していることは図3にも現われている。一方TBILは、利用者のモデル選好要因としては負の影響が示唆される(図3)ものの、対話型ECの過程の中に残されてきた変数である。臨床的意義としても、図4に示すようにALB、CHE、ICGとの関連があり注目すべき変数である。TTTは明確な反応機序が不明なまま、肝機能検査の一つとして経験的に使われてきた検査項目である。利用者もモデル選好要因として暗黙のうちに正の評価を与えている変数であり、従来から知られている臨床的意義以外の知見を示している可能性がある。

次に、比較のために行なった重回帰モデルの結果を概観する。ステップワイズ変数選択によって、虎の門ではALB、ZTT、DBIL、CHEが、千葉大ではTTT、TBIL、AGが統計的に有意な変数として採択されてい

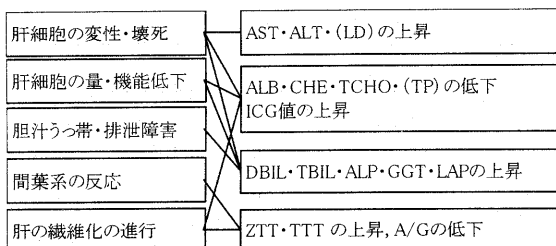


図4) 肝の病態と各検査項目との関係
(文献[17] から一部 改変して引用)

る。重回帰モデルは ICG 値が他の検査データによって説明されることを意味するモデルであり、構造方程式モデリングで記述した因果モデルとは異なる現象記述モデルである。しかしここには、興味深い結果が示唆されている。ZTT と TTT は反応原理が多少異なるものの非常に似た検査法である。DBIL は TBIL の一部の成分であり、AG は ALB の値を用いて計算する項目である(表1欄外)。すなわち、一見異なるモデルのように思われるが、比較的似たモデルとも言える。さらに、利用者が最も選好した因果モデル(表3)は、これらを融合したモデルとして解釈できる点も興味深い。

今回対象としたデータセットは、研究目的外に蓄積された過去の検査データについて、異なる2つの施設から抽出したものであり、ノイズやバイアスの存在という潜在的な問題を抱えている。しかしながら、主観的判断の下に選ばれた因果モデルは、客観的な基準で得られた重回帰モデルとも比較的似た属性から構成されるものであった。広範囲な実用化のためには検証実験を行なう必要があるものの、今回対話型 EC を経て得られたモデルには、広い適用性が期待される。

6. おわりに

肝機能検査データに関わる因果モデルの構築において、情報源の異なる2つのデータセットから共通因果モデルを探索する問題を設定し、属性選択に対話型 EC を用いてモデル探索を試みた。モデルに対する主観的評価を対話的に取り入れたことで、利用者の積極的なフィードバックの下に、進化計算によって効率的な仮説空間の探索を実現した。本事例においては、背景知識が反映された意図的な探索過程で、多くの解候補を評価した中から、利用者に支持されるモデルが得られた。得られた因果モデルは、2つの施設のデータセットにおいて支持されたものであり、単一の施設内で得られたモデルに比べ、より適用性が高いものと推察される。

謝辞

本研究を遂行するにあたり、貴重なデータとご助言をいただいた千葉大学医学部附属病院医療情報部の横井英人氏をはじめ島根医科大学医学部医療情報学講座

の津本周作氏、平野章二氏に感謝の意を表します。本研究の一部は、文部科学省科学研究費(特定領域研究(B)アクティブマイニング、ならびに、特定領域研究(C)13224007)の援助のもとに実施された。

文 献

- [1] D.L. Sackett, W.M.C. Rosenberg, J.A.M. Gray, R.B. Haynes, and W.S. Richardson, "Evidence-based Medicine: what it is and what it isn't (editorial)," *British Medical Journal*, vol. 312, pp. 71-72, 1996.
- [2] 丹後俊郎, 統計学のセンス: デザインする視点・データを見る目, 朝倉書店, 東京, 1998.
- [3] 稲田政則, 寺野隆雄, "肝機能検査データからの因果モデルの構築," *人工知能学会論文誌*, vol.17, no.6, pp.708-715, Nov.2002.
- [4] P.K. Chan, and S.J.Stolfo, "Meta-learning for multistrategy and parallel," *Proc. Second Intl. Work. Multistrategy Learning*, pp.150-165, 1993.
- [5] 高木英行, 畝見達夫, 寺野隆雄, "対話型進化計算の研究動向," *人工知能学会誌*, vol.13, no.5, pp.692-703, Sep.1998.
- [6] 黒川清, 春日雅人, 北村聖(編), 臨床検査データブック, 医学書院, 東京, 2001.
- [7] 元田浩, 沼尾正行, 山口高平, 津本周作, "アクティブマイニングの構想," *人工知能学会誌*, vol.17, no.5, pp.615-621, Sep.200.
- [8] 石野洋子, 寺野隆雄, "模擬育種法と帰納学習を適用したマーケティング情報分析," *人工知能学会誌*, vol.12, no.1, pp.121-131, Jan.1997.
- [9] K. Kira, and L.A. Rendell, "The Feature Selection Problem: Traditional Methods and a Algorithm," *Proc. AAAI'92*, pp.129-134, 1992.
- [10] D.E.Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison-Wesley, 1989.
- [11] M. Desjardins, and D. Gordon (eds.), "Special issue on Bias evaluation and selection," *Machine Learning*, vol. 20, No. 1/2, 1995.
- [12] K.G. Joreskog, "Structural analysis of covariance and correlation matrices," *Psychometrika*, vol.43, pp. 443-477, 1978.
- [13] K.A. Bollen, *Structural Equations with Latent Variables*, John Willy & Sons, 1989.
- [14] 日本臨床化学会酵素専門委員会, "血清中の酵素活性測定標準化の推進に関する指針," *臨床化学*, vol.23, pp.335-340, Sep.1994.
- [15] 豊田秀樹, SAS による共分散構造分析, 東京大学出版, 東京, 1992.
- [16] 豊田秀樹, 共分散構造分析[入門編], 朝倉書店, 東京, 東京, 1998.
- [17] 原田英治, 福井秀雄, 村井晴洋, 時田元, 上司裕史, 矢倉道泰, "慢性肝疾患の検査データの読み方," *診断と治療*, vol.89, no.10, pp.1748-1753, 2001.
- [18] シーマーケットクリニック homepage, C型肝炎の話, C型肝炎の検査, 血液検査, <http://plaza18.mbn.or.jp/~sujaku/c-ketsueki.html>, 2003.