

血液検査データ解析によるデータ解析手法

The Knowledge Discovery Method by Data Analysis of Blood Test

山下 真 村岡 道明

高知大学大学院 情報科学分野 村岡研究室

1. 研究背景・研究目的

近年、医療情報の電子化に伴い、日常診療で得られた診療結果が医療情報データベースに蓄積されており、蓄積されたデータの有効活用が求められているが、有効活用があまり行われていない現状がある。本研究では蓄積されたデータの有効活用として、蓄積された血液検査データから新たな知識を発見することを目的とし、データ解析を行う。また、対象データには高知大学医学部医学情報センターのデータを用いた。

2. 提案手法

本研究では、血液検査データを解析するための手法として、血液検査正規化手法、血液検査相関モデル、血液検査比較方法を提案した。次章から各手法について述べる。

3. 血液検査正規化手法

蓄積された血液検査データには検査方法の改定により、同じ検査項目を異なる検査方法で測定した結果が蓄積されており、これらの検査値は異なる分布をもつことがわっている。データ解析を行う場合、検査方法の改定による分布の違いがノイズになる可能性がある。本研究では、検査方法の違いによる検査値の分布の違いを正規化する手法を提案した。

4. 血液検査相関モデル

血液検査はからだの状態を表す指標であるが、従来、血液検査項目間の関係性を網羅的に表現することは行われていなかった。本研究では血液検査項目の相関関係を網羅的に表現することで、新たな知識の発見を試みた。

血液検査相関モデルの構築方法を以下に示す。

1. Box-Cox 変換を用いて対象データを正規近似する
2. 検査項目間の相関係数を算出
3. 検査項目をノードとし、相関の強い検査項目間を矢印で繋いだグラフをつくる
4. 相関の強い検査項目間の回帰式を算出する

上記の方法で図 1 のようなグラフを作成する。また、本研究では相関の強い項目間の関係性を表現するために、Passing-Bablok 法を用いて回帰式を算出した。

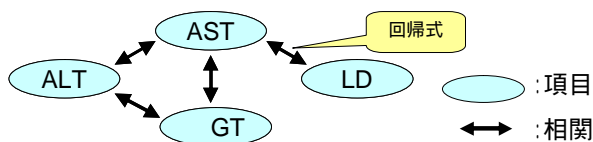


図 1 血液検査相関モデルの例

本研究では高知大学医学部附属病院総合診療部のデータを用いて血液検査相関モデルを作成した。対象とした検査項目は TP、GLU、AG などの一般的な検査項目 31 項目を対象とした。作成した相関モデルの一部を図 2 に示す。

図 2 の相関モデルは医師のもつ知識に近い相関関係を表現している。例として、従来、腎機能の指標として BUN、CRN が知られており、腎機能が悪化すると BUN、CRN の値が変化することが知られているが、図 2 の相関モデルでは BUN と CRN の間に強い相関があるため、その関係を把握できる。

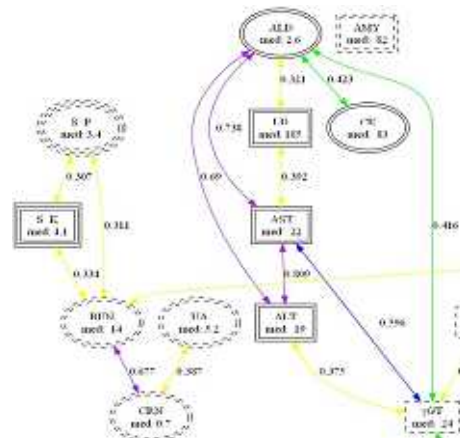


図 2 総合診療部の血液検査相関モデル

5. 血液検査比較方法

本研究では検査値の分布と血液検査相関モデルを比較することで、異なる患者群の血液検査を比較し、比較の結果から新たな知識の発見を試みた。

肝機能障害と白血病の患者の血液検査の比較では、血液の状態を表す RBC、Hb、WBC、PLT の分布と相関関係に差がみられた。また、抗うつ薬服用前後の患者の血液検査の比較では、服用前は、CK は肝機能の指標となる AST、LD と強い相関があるが、服用後は、AST、LD との相関が弱まり、血液の状態を表す RBC、Hb と強い相関をもつ変化(図 3)を発見した。

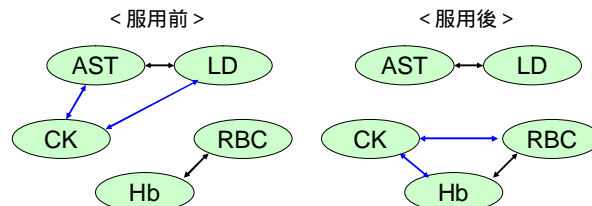


図 3 抗うつ薬服用前後の比較結果

この結果を医学的に考察すると、CK はエネルギー代謝の指標であるため、服用前は肝臓でエネルギーをつかっているのに対し、服用後は血液(脳)でエネルギーをつかっているという考察が得られた。

6. まとめ・今後の課題

本研究では蓄積された血液検査データから新たな知識を発見することを目的とし、データ解析を行う手法を提案した。提案手法を適用してデータ解析を行ったところ、新たな知識を得ることができた。今後の課題としては、様々な疾患群の血液検査相関モデルを作成し、比較を行うことで新たな知識の発見を試みることや、医学的知識からの相関モデルの考察、発見した知識を時系列解析などを用いて、さらに解析することが挙げられる。

参考文献

- [1] 片岡浩巳 “医療情報データベースからの知識発見”，高知大学大学院理学研究科情報科学専攻修士論文，2月2001年