

学 位 論 文 内 容 の 要 旨

学位論文題目	離散 Bayes 識別則と その個別化医療への応用に関する研究
--------	------------------------------------

氏 名	荻原 宏是
-----	-------

癌は生涯で日本人の2人に1人は罹患し、3人に1人はそれが原因で死亡する病気である。この癌の克服は国民的課題であり、一刻も早い治療方法の確立が望まれている。癌治療の困難さは、たとえ癌種が同じであっても患者個々によって癌が異なる多様性にある。現在、癌治療には血液検査やCTなどの多種多様な検査が行われているが、それらは癌の一側面しか診ておらず、単一で決め手となる検査はない。そこで1つの検査だけではなく複数の検査データを用いて個々の患者に最適な医療が実現されるように検討がなされている。これは「個別化医療」の考えに基づいている。

個別化医療とは、患者の遺伝情報や現在の疾患の状態を基に患者個々に最適な治療方法を選択する医療である。個別化医療では、患者が病気であると診断されると、患者個々に治療効果が見込め、かつ副作用の少ない治療を選択できる。また、治療効果が薄いと予測される薬の投薬を防ぐことで、患者の身体的、金銭的負担を抑えることにも繋がるという利点がある。癌の多様性に対処するためには、症状が同じ患者を等しく扱う従来の「平均的医療」よりも患者個々に応じた医療である「個別化医療」が適していると考えられる。

個別化医療に関する研究は、医師が構築したスコア式による予測・診断とコンピュータの機械学習による予測・診断に大別される。まずスコア式による予測・診断は、マーカーからの測定値とそのマーカーに設定したカットオフ値との大小関係でスコア値を付け、スコア値の総和により患者の状態を層別化するアプローチである。しかし、スコア式は医師の経験と直観に基づいて構築されたものであるため、用いられるマーカーの論理的根拠はなく、スコア式の最適性も不明確である。

一方、コンピュータの機械学習による予測・診断は、20世紀の分子生命科学の成果である遺伝子解析を基に、遺伝子の発現量を用いて白血病の診断が可能となった。これが引き金となり、マイクロアレイ解析による発現データを用いて癌研究が本格的に展開されるようになった。遺伝子関連のデータは数万から数十万とデータ量が膨大であることから、コンピュータを導入した機械学習による診断が注目されている。しかし、この機械学習による予測・診断にも深刻な問題点がある。それは、用いるデータが量的データ（数値データ）に限定されている点である。一般に医学データには数値で表される量的データと数値で表すことができない質的データが混在する。量的データと質的データはいずれも癌の再発予測・診断に本質的であるが、質的データは機械学習の識別器では取り扱うことはできない。機械学習と同じく、統計的パターン認識のBayes識別器もまた、質的データには適用できない。

本論文では、質的データを用いることができるように、Bayes識別則を拡張した「離散Bayes識別則」を提案する。診断アルゴリズムにBayes識別の考えを採用することで、個々のマーカーの不確実性を数量化して統計的に誤りを最小化できる。さらに、提案手法は、質的データはもちろんのこと、量的データもカットオフ値を設けて2値化することにより質的データとし、質的データと量的データの両方を取り扱うことができる。

本論文では、従来の個別化医療の問題点を解決する離散 Bayes 識別則を用いて個別化医療の実現を目的とする。

第 1 章の序論では、個別化医療の必要性と従来の個別化医療の問題点について指摘し、これらを踏まえて本論文の目的と構成を述べる。

第 2 章では、統計的パターン認識を概説し、従来の個別化医療問題を統計的パターン認識問題として定式化する。その後、離散 Bayes 識別則を提案し、その評価方法も定める。

第 3 章では、離散 Bayes 識別則を、肝癌の早期再発の予測問題、早期胃癌のリンパ節転移の予測問題、大腸癌における抗癌剤と免疫療法の併用効果の予測問題、漢方薬の処方問題に適用し、提案手法の有用性を検討する。

第 4 章では、結論として本論文の総括と今後の展望について述べる。

学位論文審査の結果及び最終試験の結果報告書

(博士後期課程博士用)

山口大学大学院医学系研究科

報告番号	医博甲 第 1498 号	氏 名	荻原 宏是
最 終 試 験 担 当 者	主 査 審 査 委 員 審 査 委 員 審 査 委 員 審 査 委 員	浜本 義彦 松藤 信哉 山口 真悟 長 篤志 藤田 悠介	
【論文題目】			
離散 Bayes 識別則とその個別化医療への応用に関する研究 (A Study on Discrete Bayes Decision Rule and its Application to Personalized Medicine)			
【論文審査の結果及び最終試験の結果】			
<p>難治性疾患の代表は癌である。癌治療の困難性は、例え癌腫が同じであっても、患者個々によって癌が異なる多様性にある。しかも癌を検出する技術は確立されておらず、どの検査も癌の一側面しか診ていない。そこで、一つの検査に頼るのではなく、複数の検査から多面的に癌を診て、個々の患者にとって最適な医療を行う個別化医療が臨床の現場から強く求められている。</p> <p>21世紀になって分子生命科学の医学への応用が本格化され、患者からのデータは膨大となり、例えば患者一人当たり遺伝子では約2万、メチル化異常に至っては数十万のオーダーとなるなど、情報爆発の状態である。このため、医師によるデータの取扱いは不可能となり、必然的に機械学習が注目されるようになった。特に近年注目されている次世代シーケンサーから取り出される遺伝子変異のデータは質的データ(記号)であり、質的データは個別化医療を実現するためには不可欠となっている。しかし、多くの機械学習は量的データ、つまり数値データしか取り扱えず、質的データは対象外となっている。</p> <p>このような背景のもと、申請者は、質的データを取り扱える機械学習の一つとして離散 Bayes 識別則を提案し、その個別化医療問題への応用を論じている。離散 Bayes 識別則は、Bayes 最適性を継承して質的データを取り扱える点にある。本論文の構成は以下となっている。</p> <p>第1章の序論では、個別化医療の必要性と従来の個別化医療の問題点を指摘し、本論文の目的と構成を述べている。</p> <p>第2章は、準備として、統計的パターン認識モデルとその個別化医療問題との対応について説明してから、離散 Bayes 識別則を提案し、識別手法の評価指標として識別率、感度、特異度、precision, F1 measure、ROC 解析まで説明している。</p> <p>第3章は、離散 Bayes 識別則の個別化医療問題への応用について述べている。最初の個別化医療問題は肝癌の早期再発の予測問題である。肝癌の怖さはその高い再発率で、術後約30%に再発がみられる。もし再発を予測できれば効果の高い先制医療を実施でき、もし再発しないとすれば抗癌剤やCT検査が不要となり、患者はこの上ない恩恵を受けるだけでなく、医療費の抑制にもなる。シミュレーション実験では独立したテストサンプルに対して再発の検出率である感度86%、特異度49%を達成した。</p> <p>次の問題は、早期胃癌のリンパ節転移の予測問題である。胃癌治療の進歩は早く、特に早期胃癌であれば内視鏡的治療によって治る癌の一つになっている。ただリンパ節転移が疑われると、外科手術が必要となる。しかし、実際にはリンパ節転移のない患者にも外科手術が行われている。もしリンパ節に転移がないことが予測できれば、不要な手術を行わなくても済む。シミュレーション実験では3マーカーを用いて独立したテストサンプルに対して感度100%、特異度86%を達成した。</p> <p>3番目の問題は、大腸癌の治療効果の予測問題である。大腸癌は、食生活の欧米化に伴い、急速に患者</p>			

数が増加している。癌治療には、外科手術、化学療法、放射線治療が代表的であるが、近年、第4の治療として免疫療法が注目されている。この免疫療法の効果を更に高めるため、抗癌剤の併用治療が行われている。もし抗癌剤の投与前に治療効果が予測できれば、効く患者に対して効果のある治療を提供できる。これによって効かない患者には別の治療を提供し、医療費の抑制にもつながる。予備実験のレベルであるが、シミュレーション実験では独立なテストサンプルに対して感度と特異度をともに100%とすることができ、本アプローチが有望であるとの見通しを得た。

4番目の問題は、漢方薬の処方問題である。現在、国内に30万人の医師がいるが、その内約9割の27万人の医師が漢方薬を処方しているほど、漢方医学は診療科を問わず、広く普及している。しかし、漢方薬の処方、医師の経験と勘に依存し、処方する医師によって個人差があるのが現状である。そこで、客観性のある処方を行うため、離散 Bayes 識別則によって患者個々に最適な漢方薬を探索する。シミュレーション実験では患者からの複数の症状に対して漢方専門医と同じ漢方薬を処方することができた。

第4章では、まとめと本研究の今後の展望について述べている。本研究の成果として、3つの個別化医療における予測問題に対しそれぞれ臨床応用可能な精度で予測を達成し、また漢方薬の処方問題では漢方専門医と同一の漢方薬が処方された。これらの成果は、個別化医療の実現に資するものと考えられる。今後の展望としては、患者層別化問題を離散 Bayes 識別則により解くことで、治験の成功確率を高める研究への応用がある。

公聴会では活発な質疑応答がなされた。主な質疑内容として

- ・最適マーカー探索について、利用できるサンプル数に応じた工夫、マーカー数の決定
- ・量的データを用いた場合との比較
- ・質的データを用いる利点
- ・提案手法とスコア式との関係

等の質問があり、いずれの質問に対しても申請者からの的確な回答がなされた。

以上より、本研究は、新規性、信頼性、有効性ともに優れており、博士（生命科学）の論文に十分に値するものと判断した。

論文内容、審査会、及び公聴会での質問に対する応答などから総合的に判断して、最終試験は合格とした。

なお、査読のある関連論文の発表状況は以下の通りである。（下記以外、国際会議会議録 2編）

(1) Yukio Tokumitsu, Takao Tamesa, Satoshi Matsukuma, Noriaki Hashimoto, Yoshinari Maeda, Yoshihiro Tokuhisa, Kazuhiko Sakamoto, Tomio Ueno, Shoichi Hazama, Hiroyuki Ogihara, Yusuke Fujita, Yoshihiko Hamamoto, Masaaki Oka, Norio Iizuka, An Accurate Prognostic Staging System for Hepatocellular Carcinoma Patients after Curative Hepatectomy, Journal of Oncology, Vol. 46, No. 3, pp. 944-952, 2015.

(2) Hiroyuki Ogihara, Norio Iizuka, Yoshihiko Hamamoto, Prediction of Early Recurrence of Liver Cancer by a Novel Discrete Bayes Decision Rule for Personalized Medicine, BioMed Research International, Vol. 2016, Article ID : 8567479, 10 pages, DOI: 10.1155/2016/8567479, 2016.