

京都大学	博士（ 医 学 ）	氏 名	高 田 正 泰
論文題目	Prediction of axillary lymph node metastasis and the pathological response to neoadjuvant chemotherapy in patients with primary breast cancer using a decision tree-based model (データマイニング手法を用いた原発性乳癌の腋窩リンパ節転移の予測ならびに術前化学療法の病理学的治療効果の予測に関する数理モデルの開発)		
(論文内容の要旨)			
乳癌の診療は、癌の不均質性や治療の多様性から複雑化しており、意思決定を要する場面が多い。病態の把握・治療効果予測は全身薬物療法の適応判断などに必要であるため、ノモグラムをはじめとする多変量ロジスティック回帰を用いた数理モデルが開発されてきた。しかし本法は、１）多くの予測因子を同時に扱えない、２）欠損値があるデータでの精度低下が著しい、３）罹患率の違うコホートでの予測精度が大きく劣化する、などの問題点がある。 これらの問題に対処するため、決定木アルゴリズムの一種である Alternating decision tree(ADTree)を用いてモデル化を試みた。研究対象として、腋窩リンパ節転移の有無の予測および術前化学療法の病理学的治療効果の予測を選択した。 １）腋窩リンパ節転移の有無を予測するモデルの開発と評価 国内施設からの 148 例について、93 項目の臨床病理学的データを訓練用データとしてレトロスペクティブに収集し、変数選択および ADTree のパラメータ設定を行った。モデルの訓練用データへの過剰適合を抑制し新たなデータに対する予測能(汎化能力)を維持するためにリサンプリング法を組み合わせた。精度向上のため複数モデルを統合して予測するアンサンブル学習を取り入れた。次に別の国内施設からの 143 症例のテスト用データを用いて、最も予測精度の高いモデルを選択した。選択されたモデルの汎化能力を独立データで評価するため、韓国ソウル大学病院で収集した 174 例のデータにあてはめた。 ADTree により 15 の臨床病理学的因子がモデルに選択された。予測精度は、テスト用データで受動者動作特性(ROC)曲線下面積(AUC)=0.770、独立評価用データで AUC=0.772 と同等の予測精度を示した。モデルは欠損値がある場合でも予測値を算出可能であり予測精度劣化は軽度であった。同一データで比較した場合、ADTree モデルが既存モデルより高い予測精度を示した。 ２）術前化学療法による病理学的治療効果を予測するモデルの開発と評価 乳癌の術前化学療法は、手術の縮小化、治療効果による予後の層別化、新たな治療法・バイオマーカーの探索などを目的に行われる。したがって、術前化学療法による病理学的完全奏効(pCR)を治療開始前に予測することは、臨床的に意義が高い。 訓練用データとして国内 3 施設でアンスラサイクリン系およびタキサン系抗癌剤の逐次併用レジメンを術前に施行された 150 例のデータを収集し、腋			

窩リンパ節転移予測モデルと同じ手法を用いて、pCR を予測アウトカムとしたモデルを作成した。モデルの汎化能力の評価には、訓練用データと同様のレジメンを用いた前向きランダム化比較試験から得られた 173 例のデータを用いた。			
ADTree により 15 の臨床病理学的因子がモデルに選択された。モデルは、訓練用データで AUC=0.766、評価用データで 0.787 の予測精度を示した。欠損値をランダムに変化させた場合のモデルの AUC の 95%信頼区間幅は 0.002 であり、欠損値への耐性が示唆された。乳癌のサブタイプ別の解析では、luminal type (エストロゲン受容体陽性) 乳癌で AUC=0.779 の精度を示した。同一データの比較では、ADTree モデルは従来法によるモデルよりも高い予測精度を示した。 いずれの予測モデルも既存モデルより多くの変数を統合しながら高い汎化能力を示していた。ADTree を中心とした手法により高精度の予測モデルを構築することが可能であり、新規予測因子の導入によって予測精度の向上を図ることができる。			
(論文審査の結果の要旨)			
癌診療の複雑な意思決定過程の定式化を目的に、乳癌の病態や治療効果を予測する数理モデルを開発した。モデル化には予測精度が高く可解釈性を保つモデルを構築可能な新しい決定木アルゴリズムである Alternating decision tree(ADTree)を用いた。 腋窩リンパ節転移の有無の予測によりモデル構築手法を確立した。ADTree モデルは 15 の臨床病理学的因子を含み、ROC 曲線下面積(AUC)はテストデータ 0.770、独立評価データ 0.772 であった。同一データの比較では、ADTree モデルは既存モデルより高精度であった。術前化学療法の病理学的完全奏効の予測で本手法の有用性を確認した。15 因子を含むモデルの AUC 値は、訓練データ 0.766、独立評価データ 0.787 であり、臨床的抗腫瘍効果を統合すると AUC=0.855 と精度が向上した。サブタイプ別では luminal type にて AUC=0.779 の精度を示した。 ADTree により構築されたモデルは、既存モデルより多くの因子を統合しながら高い汎化能力を示した。本手法により高精度の予測モデルを構築することが可能であり、新規予測因子の導入により精度の向上を図ることができる。			
以上の研究は癌診療の意思決定過程を定式化する最適な手法の解明に貢献し、個別化医療のための新たな診療アルゴリズム形成に寄与するところが多い。 したがって、本論文は博士（ 医学 ）の学位論文として価値あるものと認める。 なお、本学位授与申請者は、平成 24 年 8 月 15 日実施の論文内容とそれに関連した試問を受け、合格と認められたものである。			
要旨公開可能日： 年 月 日以降			