

### 6-3. DPC データマイニングに適した手法の検討

データマイニングには、上述したように Decision Tree などのクラス判別、予測、クラスター解析、相関関係分析、時系列順パターン分析、類似時系列分析などがあるが、本研究では、DPC 分類ツリー図の解析に Decision Tree 解析が、DPC 分類基準等を含む定義表の解析に予測解析およびクラスター分割解析が適していると考えられた。

特に予測解析は、「教師つき」クラスター解析とも捉えられ、ある特定の指数、例えば、在院日数や1日あたりの医療費、1入院あたりの医療費等を指標にグループ分けを試みる解析に適していると考えられた。この方法により、年齢を含めた患者の諸因子が診断群分類にどのような影響を与えているかを検討することが可能となると考えられた。

年齢に関しては、年齢階級を解析に含める方法と、年齢階級毎にデータを分割してそれぞれ独立した解析を行う2方法を併用し、さらに、患者年齢構成と医療機関間差異、地域格差等の影響を検討することとした。

以上の要点をまとめると下記のようなになる。

1. 診断群分類の精緻化のための分析においては、診療関連情報、診断情報等の様々な要因の中から、ケースミックスを決定づける情報を見いだすためには、グループの分岐条件を探求するという観点から、クラス判別分析 (Predicting Classification) の適用が有用であると考えられる。
2. 一方、在院日数、診療報酬点数等を決定づける因子から診断群分類の精緻化を図ろうとするアプローチにおいては、診療関連情報、診断情報等の様々な要因のうち何が在院日数あるいは診療報酬点数を決定づけているかを探るという観点から、予測 (Predicting Values) の適用が有用であると考えられる。
3. さらに必要に応じて、非常に多数の因子を集約して、分析の効率を高めるためなどの目的によって、相関関係分析 (Discovering Association)、クラスター分析 (Clustering) 等を併用して利用していく必要があると考えられる。

## 7. 考察

DPC データ解析データマイニングシステムの設計と構築について、データマイニング用データセットの構築、収集されたデータのマイニングデータへの変換、データマイニングに不適當なデータのクレンジング作業は順調に完了され、データマイニングシステムの稼働環境は十分に整えられた。

マイニングロジックについては基礎的な検討は完了したので、次年度以降、実効性の高いマイニング結果を求めていく方向で検討を進める必要があると考えられた。