

02-025

自己組織化マップを用いた健康診断データの解析

○藤末 浩司¹⁾、阿部 尚美²⁾、本田 純久³⁾、DOOSUB JAHNG¹⁾

九州工業大学大学院 生命体工学研究科 脳情報専攻¹⁾、財団法人 日本予防医学協会 産業保健研究開発部²⁾、長崎大学 熱帯医学研究所 国際保健学³⁾

【はじめに】自己組織化マップ (Self-Organizing Map : SOM) は、Kohonenにより提唱されたニューラルネットワークの一つであり、近年データの可視化やデータマイニングに応用されている。特に高次元のデータを低次元のデータに写像し、可視化する特徴を持つ。平成 20 年 4 月から施行された高齢者の医療の確保に関する法律により、特定健康診査と特定保健指導の実施が保険者に義務付けられることになった。これらの健康診断で集められるデータは、検査値、生活習慣、不定愁訴などで構成される高次元のデータである。そこで、自己組織化マップを用いてデータ解析を行い可視化することで、個人の健康状態を分類することを試みた。【対象】某健診機関で 2006 年 1 月から 12 月に健康診断を受診した人のうち、収縮期血圧、拡張期血圧、尿糖、尿蛋白、AST (GOT)、ALT (GPT)、 γ -GTP、総コレステロール、中性脂肪、HDL コレステロール、血糖、赤血球数、白血球数、血色素量の検査をすべて受診し、自覚症状に関する質問 24 項目、生活習慣に関する質問 26 項目にすべて回答した 110,044 人 (男 82,567 人、女 27,477 人) を解析の対象とした。【方法】項目ごとの測定値の単位の違いから、重みに偏りが生じることを防ぐために、データの平均値と標準偏差を用いて規準化をおこなった。マップサイズは、解析に要する時間や視覚的問題を考慮し、8×8 (ユニット数 64) とした。自己組織化マップによる解析には、C 言語で書いたオリジナルプログラムを用いた。【結果】自己組織化マップによる解析で 1 つのユニットに対し、およそ 600~5,300 のデータが配置された。各ユニットに配置されたデータ群について検査値との相関を分析したところ、BMI、収縮期血圧、拡張期血圧、中性脂肪、HDL コレステロール、血糖との間に特徴的な相関がみられた。同様に運動習慣、食習慣、喫煙、ストレスなどの生活習慣との間にも特徴的な相関がみられた。【まとめ】健康診断で得られた情報に対して、自己組織化マップを用いてデータ解析を行ったところ、個人の健康状態を分類することができた。また、同じような特徴を持つデータ群は互いに近接したユニットに配置され、検査値や生活習慣と相関がみられた。