

## IV. ビッグデータ研究からAIの活用へ

# 2. AI (人工知能)による 糖尿病のリスクの予測

国立国際医療研究センター臨床研究センター  
疫学・予防研究部 部長  
溝上 哲也

### [Summary]

機械学習の1つであるニューラルネットワークによって職域多施設研究(J-ECOHスタディ)の健康診断データを分析し、3年以内の糖尿病発症を予測するモデルを構築した。空腹時血糖とヘモグロビンA1cの両方を予測変数に使用したモデルのROC曲線下面積は0.91と、高い予測精度が確認できた。このモデルの社会実装として、健康診断データを入力すると糖尿病の発症リスクを表示するWebシステムを国立国際医療研究センターのホームページで公開したところ、高い関心が寄せられた。糖尿病の個別化予防を推進する動機付けツールとしての活用が期待される。

### Key Words :

糖尿病□発症□リスク予測□機械学習□  
ニューラルネットワーク

### はじめに

糖尿病の罹りやすさを規定する要因として肥満や運動不足、家族歴などが知られている。近年、こうしたリスク要因を組み合わせることで糖尿病の発症確率を推計する試みが盛んに行われている。従来の予測は統計モデルに基づいているが、国立国際医療研究センターは株式会社教育ソフトウェアと共同で機械学習による糖尿病リスク予測ツールを開発し、2018年末、同センターのホームページで公開した(<https://www.ncgm.go.jp/riskscore/>)。

このツールの対象は、糖尿病と診断されたことのない30～59歳までの方である。健康診断や人間ドックの検査データを画面に入力することで、3年以内に糖尿病を発症するリスクと、同性・同年代における相対的な位置がグラフで示される(図①)。血液データの有無によって予測モデルを選択する。

### 機械学習によるツール開発

本ツールの開発には、関東・東海地区に本社を置く10数社に勤める約10万人を対象とする職域多施設研究(J-ECOHスタディ)の健康診断データを用いた。J-ECOHスタディは糖尿病の疫学研究を行う上でいくつかの利点があ