

# Using an artificial neural network to identify the population at risk of obesity in the United States

Y Fei, Q Feng, LYF Wong, BMY Cheung

Department of Medicine, The University of Hong Kong, Hong Kong



HKU  
Med

LKS Faculty of Medicine  
Department of Medicine  
香港大學內科學系

## Introduction

- Obesity is an escalating problem in the United States.
- Artificial neural network (ANN) has been increasingly used in disease classification and prediction.
- We therefore aim to build an obesity classifier for identifying patients with high obesity risk and novel predictors of obesity risk using this method.

## Methods

- Participants of the US National Health and Nutrition Examination Survey 2013-2016 aged  $\geq 20$  years with records of body mass index were used to develop a multilayer perceptron ANN model in R.
- Pregnant participants and those with data missing on key variables were excluded.
- Demographic, clinical, and questionnaire data were used to build an ANN for identifying patients with high obesity risk.
- Obesity was defined as BMI  $\geq 30$  kg/m<sup>2</sup>.

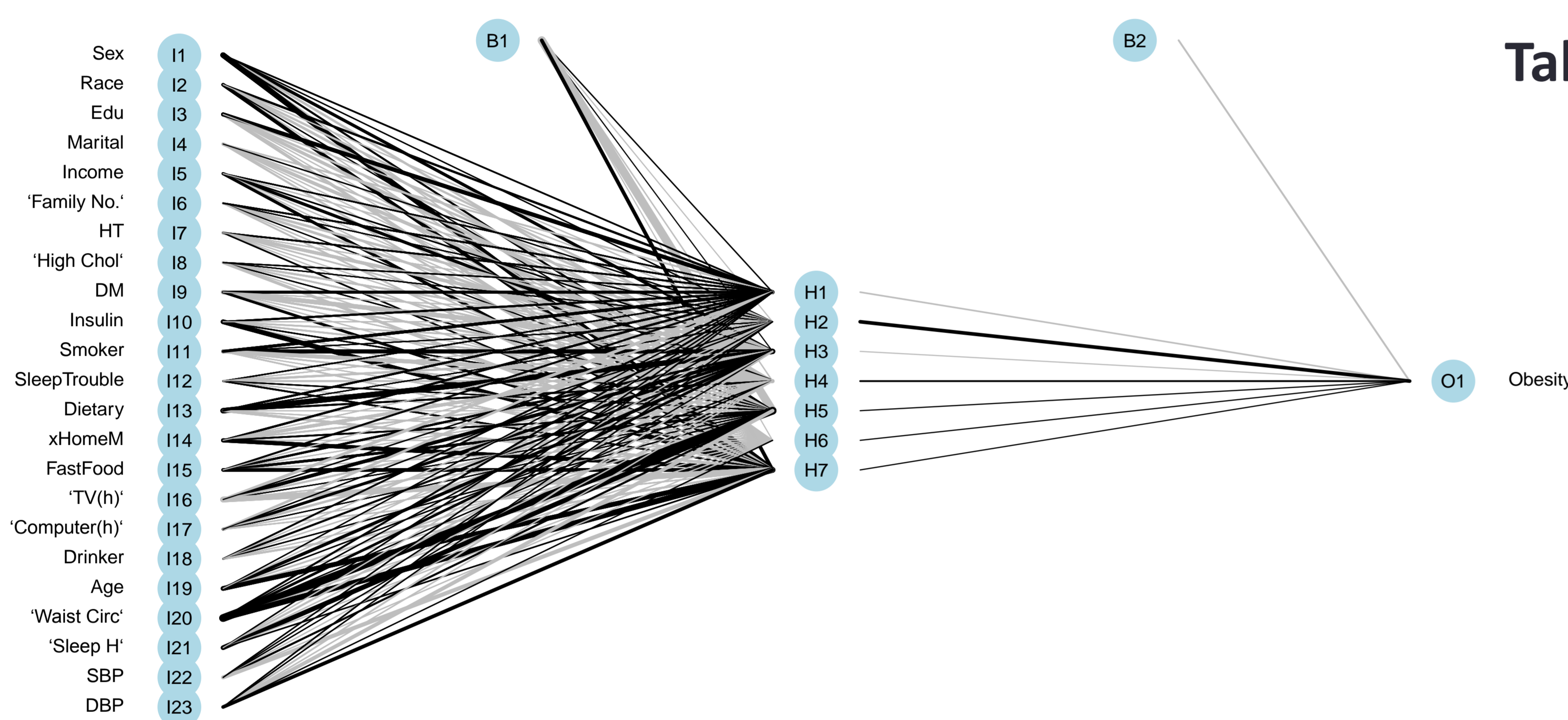


Table 1. Performance of the ANN model in obesity classification

Parameters	Model performance
Specificity	0.894
Positive predictive value	0.840
Sensitivity	0.876
Negative predictive value	0.919
Accuracy (95% CI)	0.887 (0.873-0.899)
Kappa	0.764
F1	0.860
Area under curve (AUC)	0.880

Figure 1. Architecture of the three-layer perceptron ANN

## Results

- There were 3836 (40.2%) obese and 5701 (59.8%) nonobese people eligible to be used to build the ANN.
- A three-layer multilayer perceptron ANN model composed of an input layer of twenty-three variables, a hidden layer with seven neurons, and an output layer representing the probability of obesity, was finally established (Figure 1, Table 1).
- In a decreasing order of importance, waist circumference, smoking, time spent on watching TV, education level, dietary quality, sex, sleep hours, diabetes mellitus, having meals not home-prepared, and time spent using a computer were identified as the 10 leading predictors of obesity risk (Figure 2).

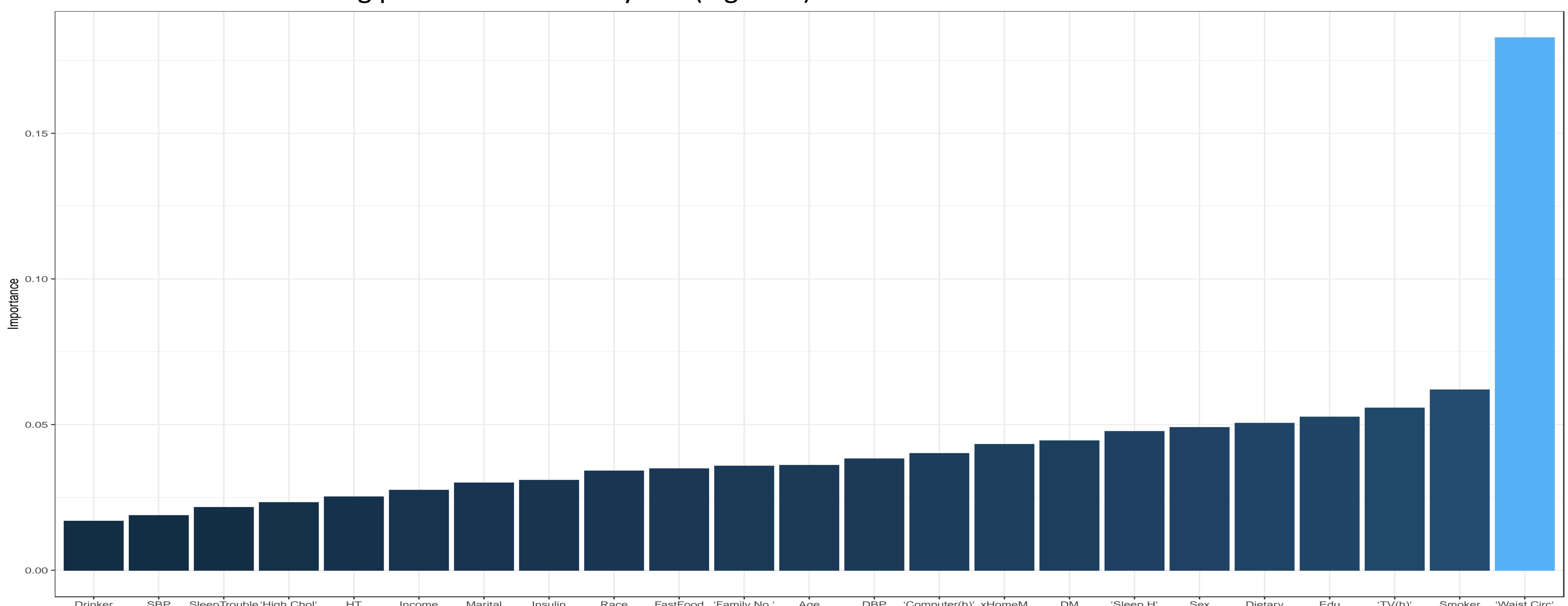


Figure 2. Variable importance on obesity risk from ANN using Garson's algorithm

## Conclusion

- ANN is useful in risk stratification of obesity.
- Our model identified several lifestyle risk factors associated with obesity that can be used to improve intervention programs. Modifying these lifestyle factors would be effective, safe and inexpensive, which can be implemented at the population level to reduce the burden of morbidity and mortality associated with obesity.
- Our findings can readily be implemented, both in population screening and in clinical management of obesity.

## References

1. Samaranayake NR, et al. *Ann Epidemiol.* 2012;22(5):349-53.
2. Heydari ST, et al. *Med Syst.* 2012;36(4):2449-54.
3. DeGregory KW, et al. *Obes Rev.* 2018;19:668-685.
4. Ergün U. *J Med Syst.* 2009;33(1):67-72.