

# Predicting the Risk of Obesity Using a Bayesian Network

<sup>(1)</sup> Christopher C. Bunn, M.A., <sup>(1)</sup> Min Du, B.S., <sup>(1)</sup> Keyi Niu, M.S.,

<sup>(1)</sup> Todd R. Johnson, Ph.D., <sup>(2)</sup> Walker S. Carlos Poston, Ph.D., and <sup>(2)</sup> John P. Foreyt, Ph.D.

<sup>(1)</sup>Department of Health Informatics, University of Texas School of Allied Health Sciences, Houston, Texas

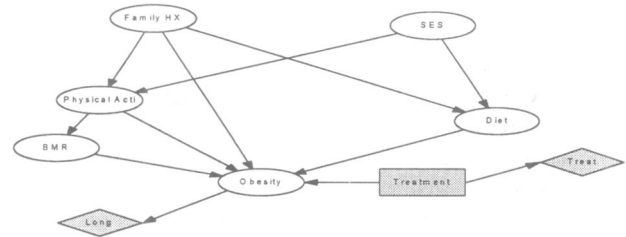
<sup>(2)</sup>Department of Medicine, Baylor College of Medicine, Houston, Texas

**Background.** Obesity has reached epidemic proportions in the United States. The current age-adjusted prevalence of obesity (i.e., BMI  $\geq 30$ ) is 19.5% for men and 25.0% for women, which represents a 35% increase over the last ten years.<sup>1</sup> Obesity is associated with increased risk for mortality and numerous other medical and health hazards in adults, including hypertension, dyslipidemia, coronary heart disease, non-insulin-dependent diabetes mellitus (NIDDM), gallbladder disease, sleep apnea, osteoarthritis, and several forms of cancer.<sup>2-3</sup>

Obesity can be defined simply as a disease in which excess body fat has accumulated to an extent that health may be adversely affected. Body mass index (BMI) provides the most useful, albeit crude, population level measure of obesity. It can be used to estimate the prevalence of obesity within a population and the risks associated with it. BMI is a simple index of weight-for-height that is commonly used to classify overweight and obesity in adults. It is calculated as the weight in kilograms divided by the square of the height in meters (kg/m<sup>2</sup>).

**System.** Bayesian theory can be used as a way of representing uncertainty in expert systems. We chose to design an expert system in predicting the risk of obesity based on current research and literature. The obesity belief network was developed using Hugin Lite 5.2. Research has shown that family history (Family Hx), socioeconomic status (SES), basal metabolic rate (BMR), diet, and physical activity (Physical Acti) are good predictors of obesity, treatment outcome, and long-term costs (Long).

The nodes in the graph correspond to relevant and uncertain variables of the problem. An arc from node Family History to Diet reflects an assertion that the probability distribution of Diet depends on Family History. Therefore, Family History conditions Diet. Thus, a node has probability



distribution for every instance of its conditioning nodes. The lack of arcs in a belief network reflects the assertion of conditional independence (SES does not condition Obesity). Given a Bayesian network, we can construct the joint probability distribution of the variables along with the chain rule of probability.<sup>4</sup>

**Evaluation.** We are in the process of evaluating the expert system using prospective cohort data.

**Conclusions.** Given the increase in mortality rates, healthcare, and socioeconomic costs, our model has potential implications for the treatment of obesity. The belief network could be used as a way to predict whether an individual with a family history of obesity is at risk for becoming obese. In addition, treatment algorithms could be developed based on an individual's obesity history and medical profile.

## References

1. National Institutes of Health (NIH), National Heart, Lung, and Blood Institute (NHLBI) (1998). *Clinical guidelines on the identification, evaluation, and treatment of overweight and obesity: The evidence report*. Washington D.C.: U.S. Government Press.
2. Solomon, C.G., Willett, W.C., & Manson, J.E. (1995). *Body weight and mortality*. In: VanItallie TB, Simopoulos AP, editors. *Obesity: New directions in assessment and management*. Philadelphia PA: The Charles Press, Publishers, 1-11.
3. Andres, R., Muller, D.C., & Sorkin, J.D. (1993) Long-term effects of change in body weight on all-cause mortality: A review. *Annals of Internal Medicine*, 119 (7 pt 2), 737-743.
4. Jensen, F.V. (1996). *An Introduction of Bayesian networks*. New York: Springer Verlag.