The University of Sydney

Generating a comprehensive disease pattern of obstructive sleep apnoea: a data mining approach

Qi Rong Huang[1], Zhenxing Qin[2], Shichao Zhang[2], Chin Moi Chow[3]

1. Discipline of Health Information Management, The University of Sydney
2. Faculty of Information Technology, the University of Technology Sydney
3. Discipline of Exercise and Sport Science, The University of Sydney

INTRODUCTION

- Obstructive sleep aponea (OSA) is a disorder characterized by repetitive, complete or partial occlusion of the posterior pharyngeal airways that results in frequent wake-up during sleep
- Sleep fragmentation often leads to lack of concentration and poor performance.
- OSA has been reported to be associated with a wide range of disorders such as obesity, diabetes, and cardiovascular diseases.
- However, a full clinical pattern of cormorbid conditions associated with OSA is not known.

AIMS

- To explore the common conditions associated with OSA
- To generate a full pattern of the comorbid conditions

METHODS

Data source

- A random dataset (1999-2004) was extracted from the Inpatients Data Collection, a clinical repository for New South Wales (NSW) Hospitals
- The data set consisted of patients' demographics and clinical diagnoses
- All clinical data have been coded with the International Classification of Diseases (ICD)-Australian Modification (AM) codes

METHODS (cont'd)

Data mining

- **A commonly used data mining tool, namely** association rule mining**, was used to identify strongly associated conditions to a given target condition.**
- **This mining method is able to find**
  - **frequent patterns (condition, sequence, etc.) of occurrence in a database and**
  - **regularities in data such as co-existing conditions**
- **The method also performs causality analysis based on the association rule:**

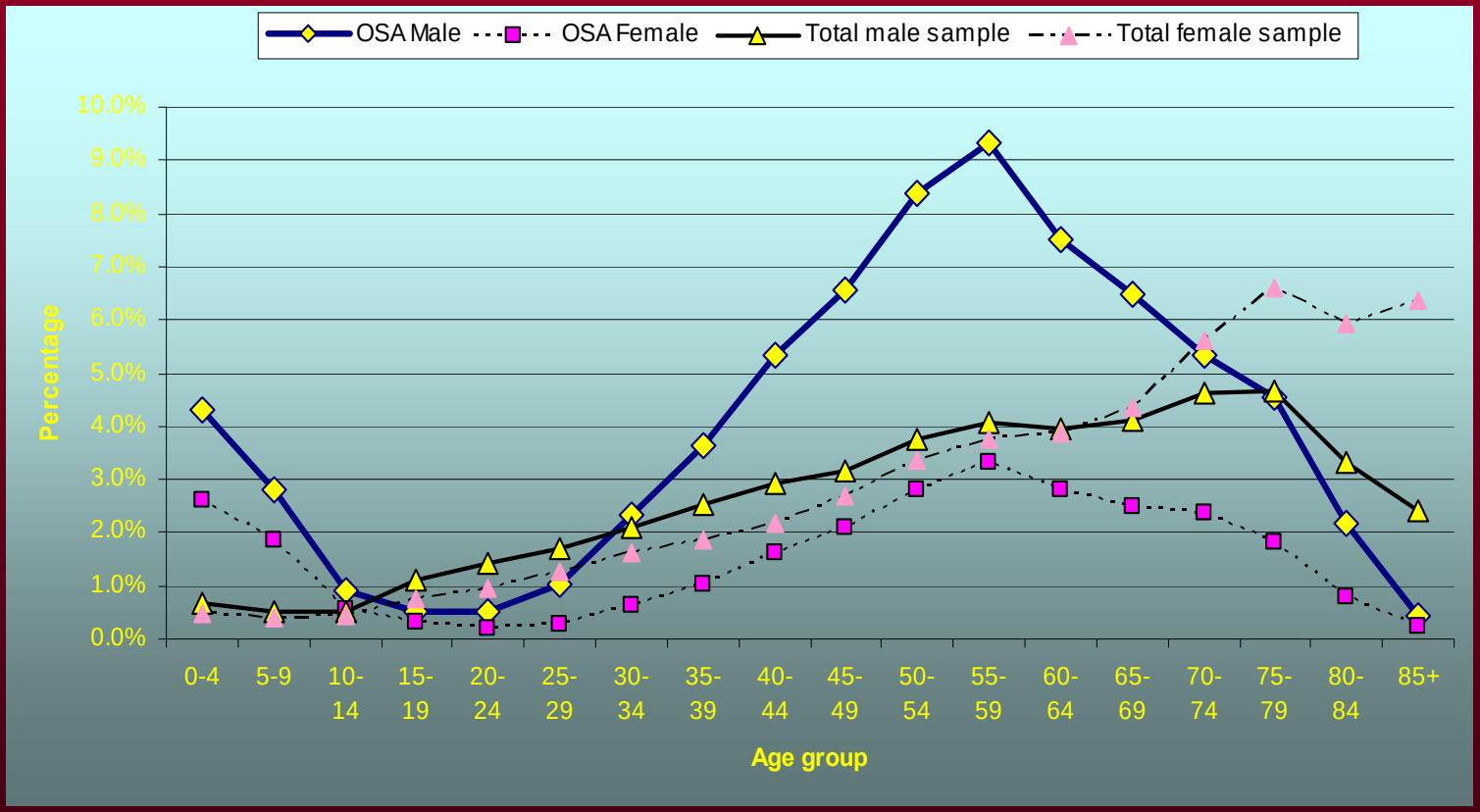
  x→y $X \subset I, Y \subset I$ $x \cap y = \phi$

  (support, confidence)
- **All association rules were searched and ranked without having to specify any minimum threshold values for support and confidence of the association rules.**

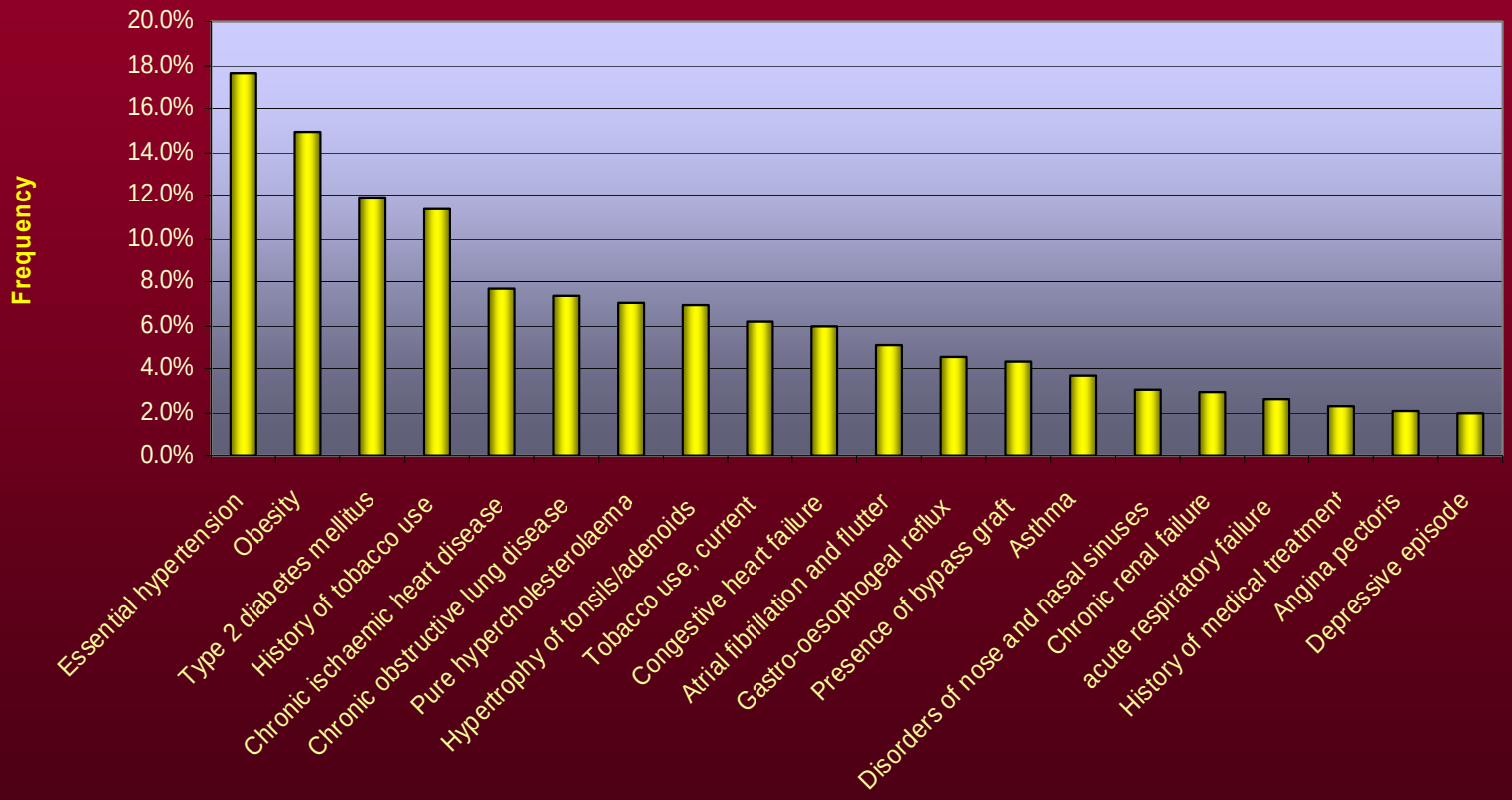
RESULTS

- There were a total of 1,511,078 patient records (47.5% males and 52.5% females) in the dataset
- A total of 60,197 (4% of the total records) had a diagnosis of OSA (72.3% males and 27.8% females)
- A complete comorbid conditions associated with OSA and their patterns were generated.

Age and gender distributions of the total sample and OSA cases

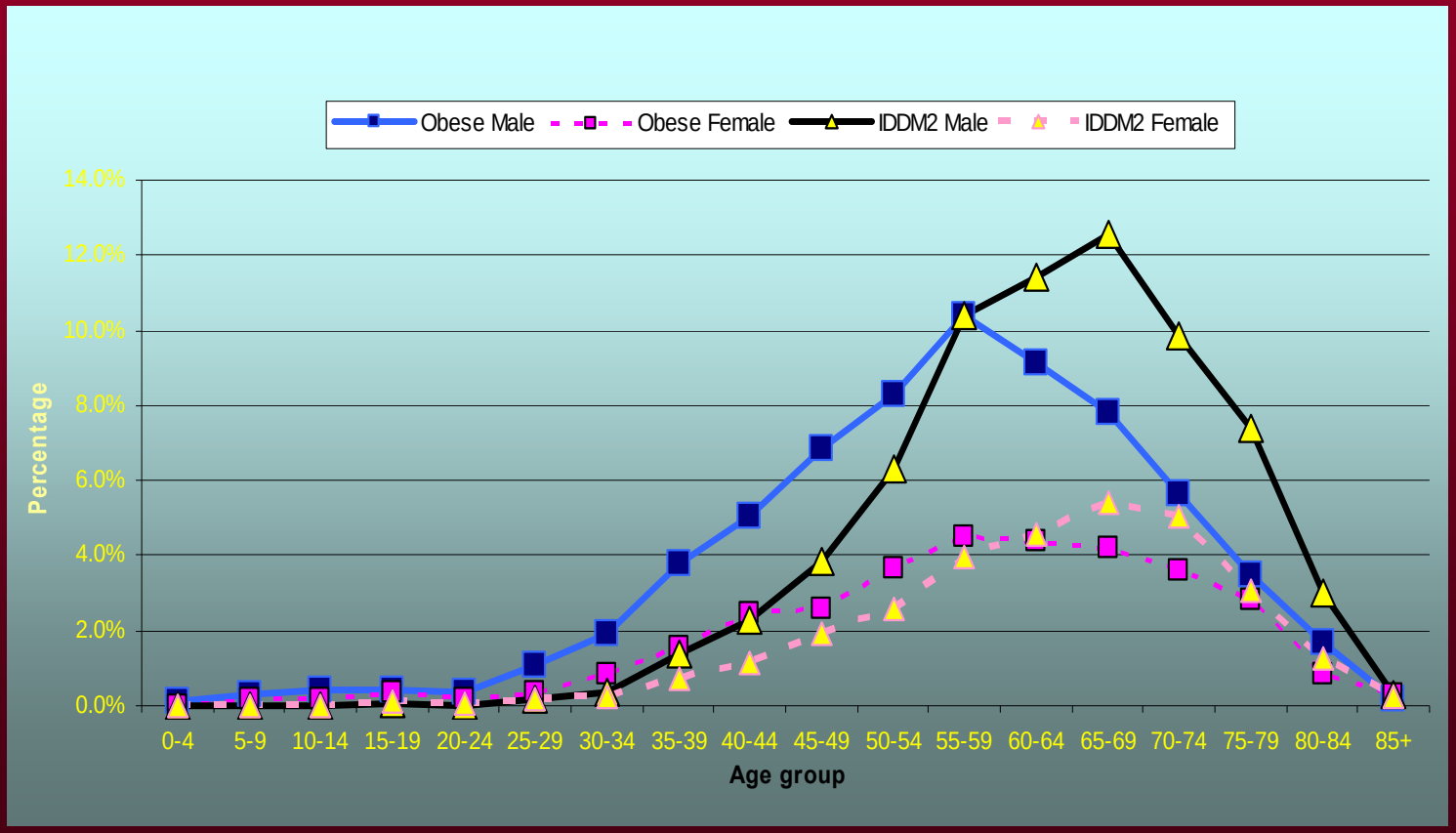


Common comorbid conditions with OSA

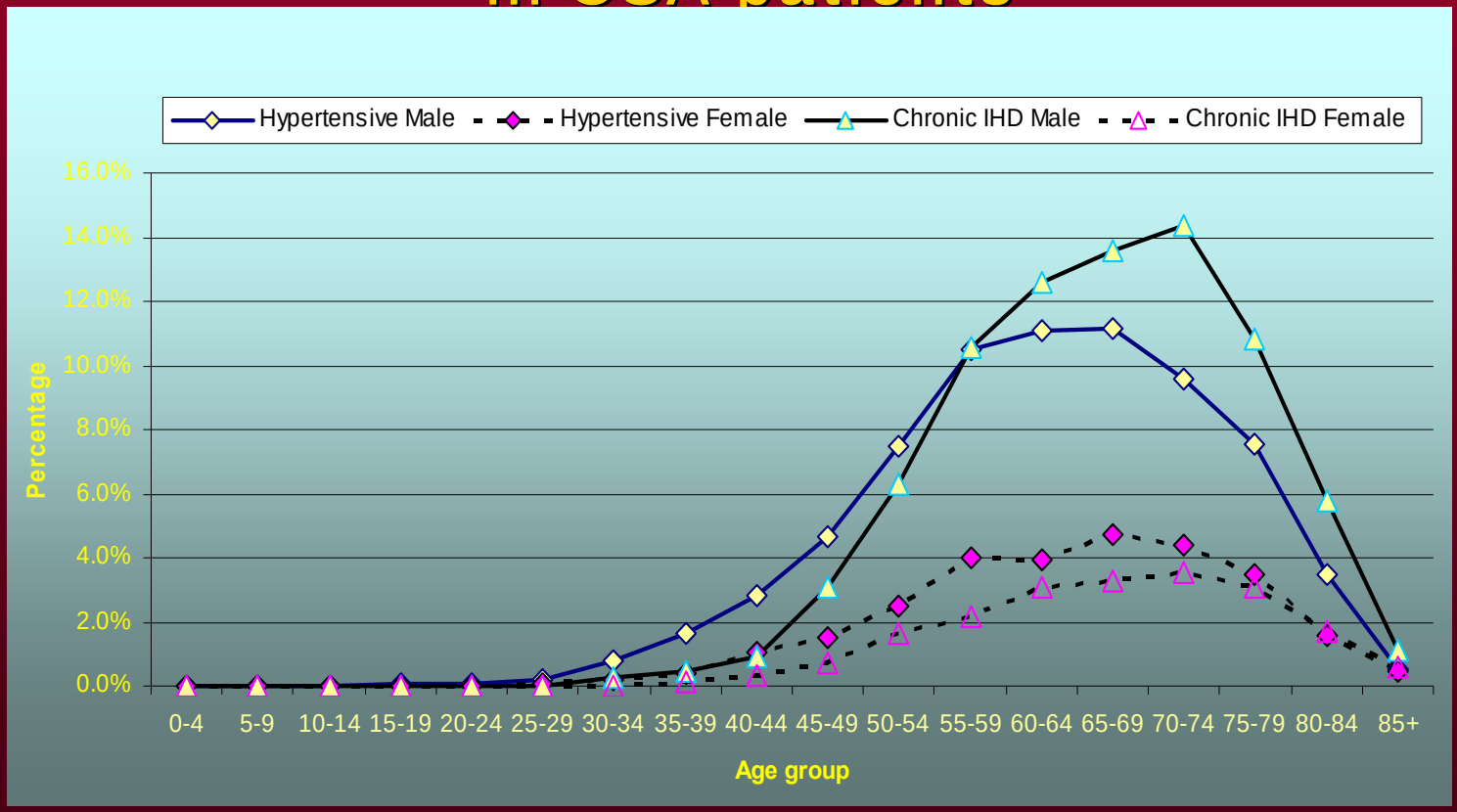


(Note: frequency is the occurrence of a condition of interest against total conditions)

Age and gender distributions of obesity and IDDM2 in OSA patients



Age and gender distributions of hypertension and ischemic heart diseases in OSA patients



DISCUSSION

- This study has for the first time generated a full clinical pattern of comorbid diseases in OSA using a data mining approach.
- It confirms early findings of a higher prevalence in males than in females, obesity as a risk factor, and associated conditions such as diabetes/ metabolic syndrome, and cardiovascular diseases (hypertension, IHD, congestive heart failure, stroke).
- It unveils other conditions that have not been previously reported including atrial fibrillation, asthma, esophageal reflux, chronic renal failure, depression, arthrosis and smoking
- It suggests that data mining, in large data repositories, is a powerful tool in uncovering clinical patterns of a disease with complex comorbidities.

- **References:**
  - Young T, Skatrud J, Peppard PE. Risk factors for obstructive sleep apnea in adults. JAMA 2004;291:2013–2016.
  - Klemettinen M, Manilla H, Ronkainen P, Toivonen H, Verkamo A. Finding interesting rules from large sets of association rules. In: Proceedings of the 3rd International Conference of Information and Knowledge Management. New York: ACM Press, 1994:401–7.
- **Acknowledgement:**
  - We would like to thank the Centre for Epidemiology and Research, NSW Department of Health, for providing the dataset.
- **Contact:**
  - Joe Qirong Huang, MD, PhD, Lecturer, Discipline of Health Information Management, Faculty of Health Sciences, The University of Sydney, email: joe.huang@usyd.edu.au