The authors propose the DBSCAN algorithm and test its capabilities on a set of data that has already been analyzed. Another goal of this article seems to be familiarizing medical professionals and biologists with data mining techniques, terminology, and potential uses. The authors state that data mining in health care is highly important and so attempt to help the audience understand common data mining techniques and vocabulary to better understand the discussion of the DBSCAN algorithm presented. There is some speculation about the nature of confounding variables in the study of lung cancer, and the implication that perhaps data mining could decrease their effect.

The first section details uses for data mining in the medical industry. Section 2 details the literature on data mining and health care, particularly lung cancer. This leads into section 3, in which the authors explain their seven steps of knowledge discovery (Data cleaning, data integration, data selection, data transformation, data mining, pattern evaluation, and knowledge representation.) This process is differentiated from the data mining steps (problem definition, data exploration, data preparation, modelling, and evaluation and deployment) in that data mining is a part of knowledge discovery, defined as the process of turning low level data into high level data. In section 4 the lung cancer literature is specifically reviewed, focusing on the known correlations found in the data. The DBSCAN algorithm is presented and the authors present their methods in the form of the steps they described earlier. The section discussing the DBSCAN algorithm is a bit more foreign and hard to understand, however it is over quickly. The conclusions drawn are that the DBSCAN algorithm is accurate to traditional data analysis.