97th Scientific Assembly and Annual Meeting
November 27 – December 2 | McCormick Place, Chicago

Scientific Informal (Poster) Presentations

CODE: LL-INS-TU4B
SESSION: Informatics

Automatic Extraction of Breast Density Information from Mammography Reports

DATE: Tuesday, November 29 2011
TIME: 12:45 PM - 01:15 PM
LOCATION: Lakeside Learning Center

PARTICIPANTS
Bethany Percha
Houssam Nassif
Jafi Lipson MD
Elizabeth Burnside MD, MPH
Daniel Rubin MD
 Presenter

 DISCLOSURES
B.P. - Nothing to disclose.
H.N.- Nothing to disclose.
J.L. - Nothing to disclose.
E.B. - Nothing to disclose.
D.R.- Grant, General Electric Company

SUBSPECIALTY CONTENT
 ▷ Informatics

CITE THIS ABSTRACT

PURPOSE

Breast tissue composition is an important risk factor for susceptibility to breast cancer, and information on breast density could play an important role in the development of classification systems for the early and accurate diagnosis of malignancy. Our goal was to develop an automated method to detect and extract the breast density assessments from free-text mammography reports.

METHOD AND MATERIALS

We studied many different mammography reports to understand the variety of ways breast density is reported. We developed an algorithm, based on pattern matching and regular expressions in free text, to automatically detect and extract BI-RADS breast density classes. Using radiologists' unstructured descriptions of mammograms as its input, the algorithm classifies each report into one of four density classes: fatty (1), fibroglandular (2), heterogeneously dense (3), or extremely dense (4). We evaluated the algorithm's performance on a total of 600 reports from two different institutions. Two different radiologists reviewed the reports to establish the gold standard.

RESULTS

The algorithm achieved 100% and 99% classification accuracy on the reports from the two institutions, respectively.

CONCLUSION

Our automated method to extract breast density composition from unstructured mammography reports has high accuracy in our testing, and appears promising. This method could enable epidemiological research by facilitating data-mining of large-scale data sets to correlate breast density with other covariates.

CLINICAL RELEVANCE/APPLICATION

Our method for extracting breast density from reports could enable large-scale coding of mammography reports to enable epidemiological research and decision support based on breast density.

QUESTIONS ABOUT THIS EVENT EMAIL:
 dlrubin@stanford.edu