Takeaway

- AccuDetect Galileo performs significantly better than Arcadia Lab’s Galileo in terms of sensitivity and false positive rates, with over 60% reduction in false positive rate per image.
- Hospitals should pay close attention to performance details when selecting CAD systems, especially false positive rates.

Summary

Computer-Aided Detection (CAD) tools are software programs that analyze digitized or digital mammography images to find features that are associated with breast cancer. The following retrospective study compares the performance of Arcadia Lab’s Galileo and Parascript AccuDetect Galileo CAD systems.

AccuDetect Galileo is an adaptation of Parascript’s CAD software called AccuDetect for IMS Giotto Full Field Digital Mammography (FFDM) system.

What’s in this Paper?

Investigators ................................................................. 2
Background ................................................................. 2
Study Design ............................................................... 3
Results .......................................................................... 4
Conclusions ................................................................. 8
References ................................................................. 8
About University of Udine ............................................ 10
Computer-Aided Detection (CAD) tools are software programs that analyze digitized or digital mammography images to find features that are associated with breast cancer.

Background

Computer-Aided Detection (CAD) tools are software programs that analyze digitized or digital mammography images to find features that are associated with breast cancer. The purpose of CAD is to “mark” suspected findings on mammographic images, in order to help Radiologist in the detection of lesions. CAD is more suited to detect than to characterize/interpret breast lesions. In other words, CAD marks potential malignant lesions that Radiologist could have missed.

CAD is an intensive matter for research in the scenario of mammographic screening programs. Therefore, screening is a cost-effective strategy for earlier detection, and thus treatment of breast cancer. However, prevalence of disease is lower in the screening scenario (no symptomatic patients) as compared to the clinical setting (different patient categories, see below), leading to subtler and less obvious mammographic features, more difficult to identify. The rate of false-negative cases at screening (missed cancers) is up to 25%, 27-70% of lesions being visible in retrospect according to different experiences. In order to overcome the problem of false-negatives, and to improve sensitivity in the detection of breast cancer, two methods have been adopted: (a) conventional double reading strategies, involving two radiologists; (b) a single
reading by radiologist followed by the use of CAD, at lesser human cost\textsuperscript{14-5}. Despite the large use of CAD, especially in the USA screening programs, analysis of these methods is still incomplete because no randomized controlled trials have been performed to assess changes in survival\textsuperscript{6}. Since these studies are difficult to achieve, most experiments use surrogate end-points such as: (a) CAD sensitivity and the proportion of retrieved cancers (retrospective studies); (b) cancer detection rate and recall rate due to further investigations in patients with suspect findings (prospective studies)\textsuperscript{7}. Ideally, a CAD-induced increase in cancer detection should be balanced by a small, clinically acceptable increase in recall rate due to further investigations on false-positive cases\textsuperscript{8}.

Overall, available results make CAD a controversial matter. Some authors state that CAD provides no significant increase in cancer detection rate in clinical practice, at the expense of excessive recall rates\textsuperscript{7,9}. Most evidence, on the contrary, supports its use as a second reader in the screening scenario, because of an adequate intrinsic sensitivity for cancer, especially for microcalcifications, and increased readers’ sensitivity up to 23.7\%\textsuperscript{10}. In any case, the reduction of specificity invariably associated with the CAD use may result in no overall clinical benefit\textsuperscript{1}. CAD is still less sensitive to masses, especially in dense breasts. Moreover, readers who benefit from CAD are mainly inexperienced ones\textsuperscript{1}. Better CAD performances have been obtained by using full-field digital mammography, which provides digital images, compared to screen-film mammography\textsuperscript{11}.

One emerging issue is\textsuperscript{9} the occurrence of many false positive marks created by CAD systems during the screening process. The purpose of this study was to compare two different CAD systems and determine if false positive rates differ significantly from system to system.

**Study Design**

Two different CAD systems were used during the study. Both Computer Aided Detection (CAD) systems were designed for a computerized second read of mammography images in order to help radiologists confirm that a suspicious area requires further investigation. AccuDetect® Galileo is an adaptation of Parascript CAD solution called AccuDetect for IMS Giotto platform.

Galileo is manufactured by Arcadia Lab S.r.L and was specifically developed for IMS Giotto, a Bologna, Italy based manufacturer of FFDM systems.

The study was performed in collaboration with the Department of Radiology, University of Udine and focused on standalone results for malignant calcification and mass detection.

The cases for the study were collected from two IMS Giotto Full Field Digital Mammography (FFDM) units and consequently retrieved from Udine Hospital’s picture archiving system. There were 118 cancer cases (86 masses, 6 calcifications, 26 mixed) and 209 negative (normal) cases. The negative cases had a 12-month follow up radiology report to mark them as normal. The truth data was obtained by a resident, experienced
radiologist who read each case and accompanying radiology and pathology reports. The radiologist drew contours of mass lesions and outlined an area of malignant microcalcifications on the screen of a color monitor. The senior (senior) radiologist then reviewed the contours and confirmed their validity.

All retrospective normal and cancer cases were sent to Galileo and AccuDetect Galileo software and the following results were calculated:

a) Percent of mass cancer cases correctly marked by Galileo,
b) Percent of calcification cancer cases correctly marked by Galileo,
c) Percent of mass cancer cases correctly marked by AccuDetect Galileo,
d) Percent of calcification cancer cases correctly marked by AccuDetect Galileo,
e) Percent of mass cancer images correctly marked by Galileo,
f) Percent of calcification cancer images correctly marked by Galileo,
g) Percent of mass cancer images correctly marked by AccuDetect Galileo,
h) Percent of calcification cancer images correctly marked by AccuDetect Galileo.

Results

During the comparison as depicted in Table 1 (next page), AccuDetect Galileo achieved higher sensitivity per case for all malignant lesions (p<0.012).
AccuDetect Galileo achieved higher sensitivity per case for all malignant lesions (p<0.012).

Table 1: True Positive (Sensitivity) versus False Positive rate per case and per image for AccuDetect Galileo and Galileo CAD.

<table>
<thead>
<tr>
<th></th>
<th>All Lesions TP per case</th>
<th>Total FP per image</th>
<th>Masses TP per case</th>
<th>Masses FP per image</th>
<th>Calcs TP per case</th>
<th>Calcs FP per image</th>
</tr>
</thead>
<tbody>
<tr>
<td>AccuDetect Galileo</td>
<td>83.90%</td>
<td>0.62</td>
<td>83.04%</td>
<td>0.42</td>
<td>71.88%</td>
<td>0.2</td>
</tr>
<tr>
<td>Galileo</td>
<td>72.88%</td>
<td>0.85</td>
<td>69.64%</td>
<td>0.58</td>
<td>56.25%</td>
<td>0.27</td>
</tr>
<tr>
<td>AccuDetect Galileo Improvement</td>
<td>11.02%</td>
<td>-0.23</td>
<td>13.40%</td>
<td>-0.16</td>
<td>15.63%</td>
<td>-0.07</td>
</tr>
</tbody>
</table>

Figure 1: FROC curve for AccuDetect Galileo (ADG) and performance data for Galileo at 0.8 False Positive per case.
AccuDetect Galileo achieved higher sensitivity per image for all malignant lesions (p<0.0001).

<table>
<thead>
<tr>
<th></th>
<th>All Lesions TP per image</th>
<th>Total FP per image</th>
<th>Masses TP per image</th>
<th>Masses FP per image</th>
<th>Calcs TP per image</th>
<th>Calcs FP per image</th>
</tr>
</thead>
<tbody>
<tr>
<td>AccuDetect Galileo</td>
<td>68.90%</td>
<td>0.62</td>
<td>69.85%</td>
<td>0.42</td>
<td>52.94%</td>
<td>0.2</td>
</tr>
<tr>
<td>Galileo</td>
<td>56.18%</td>
<td>0.85</td>
<td>51.91%</td>
<td>0.58</td>
<td>44.71%</td>
<td>0.27</td>
</tr>
<tr>
<td>AccuDetect Galileo Improvement</td>
<td>12.72%</td>
<td>-0.23</td>
<td>17.94%</td>
<td>-0.16</td>
<td>8.23%</td>
<td>-0.07</td>
</tr>
</tbody>
</table>

Table 2: Sensitivity and False Positive rates per image.

Figure 2: AccuDetect Galileo (ADG) FROC curve for Masses when compared to Galileo.
AccuDetect Galileo achieved more than 60% lower false positive rate per image for
malignant lesion hypothesis (p<0.001) when the operating point for AccuDetect Galileo is set to match the sensitivity of Galileo.

<table>
<thead>
<tr>
<th>Total FP</th>
<th>Masses FP</th>
<th>Calcifications FP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Galileo</td>
<td>AccuDetect Galileo</td>
<td>Galileo</td>
</tr>
<tr>
<td>0.85</td>
<td>0.3</td>
<td>0.58</td>
</tr>
</tbody>
</table>

AccuDetect Galileo FP Reduction

64% | 69% | 56%

Table 3: False positive rate for all cancers (Total FP) and separately for masses and calcifications.

**AccuDetect Galileo** achieved more than 60% lower false positive rate per image for malignant lesion hypothesis.

Figure 3: AccuDetect Galileo FROC curve for calcifications when compared to Galileo.

The paired Student's T-test and Wilcoxon matched-pairs signed-rank test were used to assess the statistical significance of AccuDetect Galileo results' improvement for malignant lesions detection.
AccuDetect Galileo has significantly better overall performance than Galileo in terms of both sensitivity and false positive rates.

The paired Student’s T-test used to assess the statistical significance of false-positives reduction by 60% achieved by AccuDetect Galileo. The results of the tests are summarized in the table below:

<table>
<thead>
<tr>
<th></th>
<th>Statistical Significance of AccuDetect Galileo Sensitivity per Image Improvement</th>
<th>Statistical Significance of AccuDetect Galileo Sensitivity per Case Improvement</th>
<th>Statistical Significance of AccuDetect Galileo False-Positives per Image Reduction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student T-test</td>
<td>$P &lt;= 0.000003$</td>
<td>$P &lt;= 0.004375$</td>
<td>$P &lt;= 0.000674$</td>
</tr>
<tr>
<td>Wilcoxon test</td>
<td>$P &lt;= 0.00006$</td>
<td>$P &lt;= 0.011905$</td>
<td>N/A</td>
</tr>
</tbody>
</table>

Table 4: The paired Student T-test and Wilcoxon test results.

Conclusions

During the study AccuDetect Galileo was compared against state of the art product – Galileo CAD, which had shown better performance than widely used iCAD’s Second Look Digital according to a 2007 study. The comparison proved that:

AccuDetect Galileo has significantly better overall performance than Galileo in terms of both sensitivity and false positive rates.

AccuDetect Galileo achieved more than 60% lower false positive rate per image for malignant lesion hypothesis ($p < 0.001$) when operating point for AccuDetect Galileo is set to match the sensitivity of Galileo.

The study proved that not all CAD products are alike and in this comparison study, Parascript’s AccuDetect Galileo achieved a significant, 60% reduction in false positive rate per image and better overall performance in sensitivity. Therefore, hospitals should pay a close attention to performance details when selecting CAD systems, especially if they care about a significant false mark reduction on reviewed mammograms.

References


About University of Udine

The University of Udine was founded in 1978 as part of the reconstruction plan of Friuli after the 1976 earthquake. Its aim was to provide the Friulian community with an independent center for advanced training in cultural and scientific studies. The University currently has 10 faculties in: Agriculture, Economics, Engineering, Law, Foreign Languages, Education, Humanities, Medicine and Surgery, Veterinary Science and Mathematical, Physical and Natural Sciences. The University is actively involved in student and staff exchange projects with universities within the EU and is currently engaged in close collaboration with several universities from Eastern Europe and other non-EU countries. Moreover the University participates in many research projects at national and international level. The present number of students enrolled at the University is approx. 17000.

About Parascript

The Parascript image analysis suite extracts meaningful information from images. Employing patented digital image analysis and pattern recognition technologies, the Parascript image analysis suite improves decision quality in medical imaging, postal and payment automation, fraud detection and forms processing operations. Parascript software processes billions of digital images per year. Fortune 500 companies, postal operators, major government and financial institutions rely on Parascript products. Organizations include the U.S. Postal Service, Bell + Howell, Fiserv, Elsag, Lockheed Martin, NCR, Siemens and Burroughs. Parascript is online at http://www.parascript.com.