Can semi-quantitative evaluation of uncertain (type II) time-intensity curves improve diagnosis in breast DCE-MRI?

Roberta Fusco¹, Salvatore Filice¹, Vincenza Granata¹, Ylenia Mandato², Annamaria Porto², Massimiliano D’Aiuto³, Massimo Rinaldo³, Maurizio Di Bonito³, Mario Sansone⁵, Carlo Sansone⁵, Antonio Rotondo³, Antonella Petrillo¹

¹Department of Diagnostic Imaging, National Cancer Institute of Naples “Pascale Foundation”, Naples, Italy
²Department of Diagnostic Imaging, Second University of Naples (SUN), Naples, Italy
³Department of Breast Surgery and Oncology, National Cancer Institute of Naples “Pascale Foundation”, Naples, Italy
⁴Department of Diagnostic and Laboratory Pathology, National Cancer Institute of Naples “Pascale Foundation”, Naples, Italy
⁵Department of Electrical Engineering and Information Technology, University “Federico II” of Naples, Naples, Italy

Email: msansone@unina.it

Received 17 January 2013; revised 20 February 2013; accepted 4 March 2013

ABSTRACT

Objective/Background: Qualitative assessment of uncertain (type II) time-intensity curves (TICs) in breast DCE-MRI is problematic and operator dependent. The aim of this work is to evaluate if a semi-quantitative assessment of uncertain TICs could improve overall diagnostic performance. Methods: In this study 49 lesions from 44 patients were retrospectively analysed. Per each lesion one region-of-interest (ROI)-averaged TIC was qualitatively evaluated by two radiologists in consensus: all the ROIs resulted in type II (uncertain) TIC. The same TICs were semi-quantitatively re-classified on the basis of the difference between the signal intensities of the last-time-point and of the peak: this difference was classified according to two different cut-off ranges (±5% and ±3%). All patients were cytological or histological biopsy proven. Fisher test and McNemar test were performed to evaluate if results were statistically significant (p < 0.05). Results: Using ±5% cut-off 16 TICs were reclassified as type III and 12 as type I while 21 were reclassified again as type II. Using ±3% 22 TICs were reclassified as type III and 16 as type I while 11 were reclassified again as type II. The semi-quantitative classification was compared to the histological-cytological results: the sensitivity, specificity, positive and negative predictive values obtained with ±3% were 77%, 91%, 91% and 78% respectively while using ±5% were 58%, 96%, 94% and 68% respectively. Using the ±5% cut-off 26/28 (93%) TICs were correctly reclassified while using the ±3% cut-off 34/38 (90%) TICs were correctly reclassified (p < 0.05). Conclusions: Semi-quantitative methods in kinetic curve assessment on DCE-MRI could improve classification of qualitatively uncertain TICs, leading to a more accurate classification of suspicious breast lesions.

Keywords: Breast Cancer; Dynamic Contrast Enhanced-MRI; Time Intensity Curve; Tracer Kinetics; Semi-Quantitative Analysis

1. INTRODUCTION

Despite the low specificity, dynamic contrast-enhanced MR imaging (DCE-MRI) has a high sensitivity in detecting and characterizing breast disease and it has evolved into an important adjunctive tool in Breast Imaging [1-6].

The vascularisation of the malignant lesions resulting from neo-vessels, that are the basis of exponential tumour growth, is one of the major reasons that dynamic features of MRI have played a crucial role in breast cancer for a decade [7-19].

DCE-MRI has been widely used to improve the sensitivity of MRI [4], adding information derived from kinetic-curve type to architectural and morphologic features. In fact lesions detection and characterization depend upon a combination of morphologic and kinetic observations [7,15].

Breast lesion enhancement can be characterized by assessing the enhancement kinetic curves obtained by plotting the signal intensity values over time after contrast material injection with time-intensity curves (TICs) [4,19]. The classification of the TICs is operator dependent and some overlaps are evident between benign and malignant lesions [7].
Kinetic curve analysis can be performed qualitatively (visual inspection of the curve shape), semi-quantitatively (by means of empirical parameters of signal intensity changes as gradient of the upslope of enhancement curves, maximum signal intensity and wash-out gradient) or quantitatively through pharmacokinetic modelling techniques [5-24].

In qualitative analysis is visually classified the initial and last enhancement of TIC. The initial enhancement can be described as fast, medium and slow, but this distinction is somewhat arbitrary and there are no set definitions [4-6].

The late patterns are defined as persistent, plateau and wash-out [4-6] and have been described by Kuhl [4]: type I (Ia-Ib) curve is a slow steady enhancement curve and is a strong indicator of benignancy (sensitivity and specificity of 52% and 71%); type III curve is associated with wash-out of signal intensity and is a strong predictor of malignancy (sensitivity and specificity of 20.5% and 90.4%); type II demonstrates plateau signal intensity and represents an intermediate probability of malignancy (sensitivity and specificity of 42.6% and 75%) [4-5,19].

Nevertheless it is suggested both type II and III curves should be considered suggestive of malignancy [4] intermediate are the most unspecific.

Since DCE-MRI was developed, the qualitative assessment has been considered one of the most important approaches for diagnosis. However, due to large operator dependency, many experimental studies have just proved this analysis assessment cannot be enough alone in distinguishing benign from malignant lesions [5].

Quantitative DCE-MRI can be achieved applying an adequate pharmacokinetic model to the TIC. This approach can yield parameters having a direct physiological interpretation. However, quantitative DCE-MRI involves many critical issues: accurate measurement of the arterial input function, accurate quantification gadolinium, choice of an adequate model, and accurate estimation of tracer kinetics parameters [6].

The simplest way of assigning pathological significance to TICs is to provide a description of the initial enhancement (by 1 - 2 min) followed by evaluation of the late enhancement pattern (semi-quantitative approach). It should be analyzed the first phase of signal intensity, how and how much it increases and the slope of late phase to obtain a careful evaluation.

The aim of this work is to evaluate if a semi-quantitative assessment of quantitatively uncertain (type II) TICs could improve overall diagnostic performance.

2. MATERIALS AND METHODS

2.1. Patients and Protocol

44 women (from 24- to 65-year-old, median age 46 years) underwent breast DCE-MRI examination at our institution. They were affected by: Invasive Lobular Carcinoma (ILC) (6 subjects), Lobular Carcinoma in Situ (LCIS) (2), Invasive Ductal Carcinoma (IDC) (12), Ductal Carcinoma in Situ (DCIS) (6), phylloides giant tumor (1), fibroadenomas (13), focalatypical forms of hyperplasias (10). They were all cytological or histological proven.

MRI was performed with a 1.5 T dedicated breast scanner (Aurora, USA), with an integrated coil designed specifically for 3-D bilateral breast imaging. The Spiral-Rodeo sequences fat-sat, in axial planes, were used (TE: 4.8 ms; TR: 29 ms; Matrix: 512 × 512; Slice Thickness: 1.12 mm; Gap: 0; Flip Angle: 45°).

Dynamic study involved intravenous paramagnetic contrast media injection (Gd-BOPTA, Bracco Milan Italy, 0.1 mmol/kg; flow rate 2 ml/s; 20 ml of saline solution) and consisted of five measurements with an interval of 90 sec. The first frame was acquired before contrast injection immediately followed by the four other measurements. Maximum intensity projection (MIP) reconstructions and TICs were realized in the post-processing.

2.2. ROI Placement

Two expert radiologists in consensus manually delineated region of interests (ROI) along tumor contours: care was taken in covering the whole lesion excluding artefacts and blood vessels. All ROIs were drawn on the basis of the first or second subtraction image. One single ROI per lesion was placed.

2.3. Qualitative Analysis

ROI-averaged TICs were examined using either a qualitative or a semi quantitative analysis. A total of 49 TICs were analysed.

By qualitative analysis all TICs showed the same qualitative kinetic pattern: initial increasing enhancement, followed by plateau, thus they were all classified as type II according to classification reported in [4].

2.4. Semi-Quantitative Analysis

The semi-quantitative assessment was inspired by the work of El Khouli et al. [19]. Each TIC was classified on the basis of the difference between the percentage-enhancement at the last time point and the peak percentage-enhancement: if the difference was in the range −5% to 5% the TIC was classified again as type II; for a difference greater than +5% the TIC was classified as type I; for a difference less than −5% the TIC was classified as type III.

2.5. Statistical Analysis

To evaluate the statistical performance of this method we
analysed also other cut-off points in order to construct a ROC curve: maximization of the Youden Index [25] resulted in an optimal cut-off value.

The cut-off obtained by means the ROC analysis was used in order to reevaluate the sensibility, specificity, positive and negative predictive values (PPV and NPV) of semi-quantitative analysis in comparisons of the cut-off range proposed by El Khouli et al. [26].

The number of TICs having a different classification after semi-quantitative analysis was calculated as follows: the percentage of lesions that changed from type II curve in type III; the percentage of lesions that changed from type II curve in type I; the percentage of correctly and incorrectly classified was also calculated and the chi-square test was used in order to evaluate if those percentage differences were statistically significant.

In order to evaluate sensibility (SEN), specificity (SPE), positive predictive value (PPV), negative predictive value (NPV) and accuracy (ACC), two different criteria for TIC malignancy were used: first, a TIC was considered malignant if it was type III (Sq1); second, a TIC was considered malignant if it was type III or type II (Sq2). Fisher’s exact test was used to investigate the statistical significance of the Decision Matrix [27]. For comparing the Sq1 and Sq2 criteria the Mc-Nemar test was used [28]. A P value less than 0.05 was considered significant. The whole analysis was performed using the Statistic Toolbox of Matlab R2009b.

3. RESULTS

All the 49 lesions were classified as type II using qualitative analysis.

Figure 1 shows the ROC curve for various cut-off ranges. The optimum cut-off found maximizing the Youden Index was ±3%.

Table 1 reports the number of TICs per each category after re-classification according to the ±5% cut-off (proposed by [19]) and the optimal value of ±3% found in our study.

Table 2 reports SEN, SPE, PPV, NPV and ACC for both Sq1 and Sq2 criteria, using either ±5% or ±3% as cut-off ranges, with the corresponding P value.

In Table 3 the percentages of correctly and incorrectly re-classified TICs after the semi-quantitative analysis were reported with the corresponding P value.

In Figures 2-4 we report a few exemplificative cases of FP and FN. In particular, Figure 2 shows a FN case of IDC that was classified as type I either using ±3% or ±5% as cut-off; Figure 3 shows a FP case of fibroadenoma that was classified as type III either using ±3% or ±5%; Figure 4 shows a case of fibroadenoma that was correctly classified as type I.

The differences in SEN and SPE between Sq1 and Sq2 criteria (using either ±5% or ±3%) were statistically significant (McNemar Test, P < 0.05). Considering ±5% as cut-off range the Sq1 methods had higher SPE and Sq2 had higher SEN. Considering ±3% as cut-off range the Sq1 methods had higher SPE and SEN compared to Sq2. Using ±3% as cut-off range higher SEN was achieved (McNemar Test, P < 0.05).

### Table 1. Number of TICs per each type after re-classification.

<table>
<thead>
<tr>
<th>TIC type</th>
<th>cut-off ±5%</th>
<th>cut-off ±3%</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>12</td>
<td>16</td>
</tr>
<tr>
<td>II</td>
<td>21</td>
<td>11</td>
</tr>
<tr>
<td>III</td>
<td>16</td>
<td>22</td>
</tr>
</tbody>
</table>

### Table 2. Accuracy of the two criteria for malignancy (Sq1 and Sq2) with either cut-off (±5% and ±3%).

<table>
<thead>
<tr>
<th>Method</th>
<th>SEN (%)</th>
<th>SPE (%)</th>
<th>PPV (%)</th>
<th>NPV (%)</th>
<th>ACC (%)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sq1 (±5%)</td>
<td>58</td>
<td>96</td>
<td>94</td>
<td>68</td>
<td>95</td>
<td>0.0001</td>
</tr>
<tr>
<td>Sq2 (±5%)</td>
<td>89</td>
<td>39</td>
<td>62</td>
<td>75</td>
<td>53</td>
<td>0.0441</td>
</tr>
<tr>
<td>Sq1 (±3%)</td>
<td>77</td>
<td>91</td>
<td>91</td>
<td>78</td>
<td>91</td>
<td>0.0001</td>
</tr>
<tr>
<td>Sq2 (±3%)</td>
<td>89</td>
<td>57</td>
<td>70</td>
<td>81</td>
<td>64</td>
<td>0.0018</td>
</tr>
</tbody>
</table>

Copyright © 2013 SciRes.

OPEN ACCESS
Table 3. Number of correctly and incorrectly diagnosed cases.

<table>
<thead>
<tr>
<th>Method ±</th>
<th>Correct (%)</th>
<th>Incorrect (%)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>5%</td>
<td>26/28 (93)</td>
<td>2/28 (7)</td>
<td>&lt;0.05</td>
</tr>
<tr>
<td>±3%</td>
<td>34/38 (90)</td>
<td>4/38 (10)</td>
<td>&lt;0.05</td>
</tr>
</tbody>
</table>

Figure 2. A false negative case of IDC (a) with a curve type I (b) using 3% or 5% as cut-off.

4. DISCUSSION

Qualitative assessment of uncertain (type II) time-intensity curves (TICs) in breast DCE-MRI is problematic and operator dependent. The aim of this work is to evaluate if a semi-quantitative assessment of uncertain TICs could improve overall diagnostic performance.

To this aim we retrospectively evaluated 49 lesions that were histologically or cytologically proven but were qualitatively classified as uncertain (type II).

The method we used for semi-quantitative re-classification was inspired by the work of El Khouli et al. [19].

As reported in Tables 2 and 3, after re-classification a number of TICs was correctly classified.

It should be noticed that our work has been inspired by [19]. However, our work is different from [19] and is an extension of it because we considered several cut-off ranges in order to estimate the optimal cut-off; moreover, we concentrated only on uncertain cases, while [19] analysed a mixed group of patients. The study by El Khouli et al. [19] showed SEN, SPE, NPV and PPV of 92.3%, 64%, 70% and 90% respectively evaluating each type of curve (I-II-III), while plateau curves showed a PPV of 67% [19]. It should be emphasised however, that they analysed a large group of subjects including type I type II and type III TICs.

Of course, large difficulties in the differentiation of benign or malignant lesions are found when a rapid contrast increase is followed by a plateau phase. This is why we systematically analysed type II TICs to understand if the qualitative method could be improved by a semi-quantitative approach.

Being greatly operator dependent, qualitative assess-
Figure 4. A case of fibroadenoma (a), (c) with correct classification semi-quantitative analysis: curve type I (b).

The major problem consists, in fact, in missing breast cancer detection and multiple reports have documented false-negative cases, not only of non-invasive cancer, but also of invasive ductal and lobular cancer with rates from 4% to 12% [41].

To overcome this limit in heterogeneous lesions there were placed different ROI in maximum contrast enhancement different points for each patient.

Other false negative cases were caused by atypical enhancement pattern of some lesions. ILC with a diffuse growth pattern appearing as a non mass-like enhancement may exhibit low-magnitude and persistent-enhancement kinetics possibly associated with weak angiogenic activity. Similarly, the patterns of enhancement kinetics are unreliable for diagnosis of DCIS; only about 70% of DCIS exhibit fast, initial enhancement, with variable delayed-phase enhancement patterns [13,14].

Difficulties arise with the diagnosis of “borderline” lesions (lesions of uncertain malignancy) according to the United Kingdom National Health Service Breast Screening program or those that are “probably benign” according to the Breast Imaging Reporting and Data System lexicon (MRI-BIRADS) [31]. Typical borderline lesions are atypical ductal hyperplasia (ADH), atypical lobular hyperplasia (ALH), lobular carcinoma in situ, papillary lesions, radial sclerosing lesions, fibroepithelial lesions, mucocele like lesions and columnar cell lesions [10].

These lesions appear classically with an intermediate kinetic curve type and thus can be classified as MRI
BI-RADS category III. The problem consists in doubt persistence and uncertain diagnosis. In our analysis these lesions presented a type II curve at qualitative analysis, but they became type I or type III at semi-quantitative analysis obtaining a more selective classification of these lesions.

Therefore the semi-quantitative approach could achieve more specific attribution of curve type confirmed by the hystopathologic diagnosis.

These results demonstrate it was possible contribute to a more precise assignation in differential diagnose of borderline lesions also and in reducing the number of excision biopsies of these lesions.

Limits of our study are the small number of cases and those intrinsic of a retrospective study and a semi-quantitative assessment of data. Moreover it was adopted only a semi-quantitative parameter (the absolute wash-out percentage-enhancement difference) compared with other authors that used the wash-out slope also [17]. Instead in the same way, we did not include morphologic features of lesions, to isolate the effect of the semi-quantitative method.

A future application should be a new classification of kinetic curves through pharmacokinetic models that might be more accurate in evaluating response chemotherapy and monitoring follow-up of breast cancer.

5. CONCLUSIONS

Our study shows that semi-quantitative assessment represents a method in DCE-MRI kinetic curves evaluation and could improve diagnostic performance of type II TICs.

Exclusion of cancer on qualitative or quantitative kinetic curve enhancement assessment may lead to high false negative rates, therefore the most accurate diagnostic approach on breast DCE-MRI is the combined analysis of morphologic and dynamic pattern, since neither alone is decisive to make differential diagnosis in breast cancer. However also this matched evaluation could be not conclusive and could become a real challenge in those lesions with intermediate characteristics of malignancy.

REFERENCES

DCEMRI of breast lesions: Is kinetic analysis equally effective for both mass and nonmass-like enhancement? Medical Physics, 35, 3102-3109. doi:10.1118/1.2936220


