Prediction of Chemotherapy Response in Primary Osteosarcoma by Use of the Multifractal Analysis of Magnetic Resonance Images

Goran J. Djuričić,1* Jelena S. Vasiljević,2 Dušan J. Ristić,3 Relja Z. Kovačević,4 Dalibor V. Ristić,5 Nebojša T. Milosević,6 Marko Radulović,3 and Jelena P. Sopća4

1Department of Diagnostic Imaging, University Children’s Hospital, University of Belgrade, Belgrade, Serbia
2Institute “Mihajlo Pupin”, University of Belgrade, Belgrade, Serbia
3Institute of Oncology & Radiology of Serbia, Belgrade, Serbia
4Institute of Pathology, School of Medicine, University of Belgrade, Belgrade, Serbia
5The School of Computing, University Union, Belgrade, Serbia
6Department of Biophysics, School of Medicine, University of Belgrade, Belgrade, Serbia

*Corresponding author: Goran J. Djuričić, Department of Diagnostic Imaging, University Children’s Hospital, University of Belgrade, Tirsova 10, 11000 Belgrade, Serbia. Tel: +381-641731559, E-mail: gorandjuricic@gmail.com

Abstract

Background: Due to the high level of cytogenetic heterogeneity in osteosarcoma, personalized treatment is the promising strategy for the improvement in outcomes. This is currently not possible due to the absence of targeted therapies and reliable predictors for response to induction chemotherapy.

Objectives: To investigate the predictive value of computational analysis of osteosarcoma magnetic resonance (MR) images.

Patients and Methods: Multifractal analysis was performed on MR images of primary osteosarcoma of long tubular bones prior to OsteoSa induction chemotherapy. A total of 900 images derived from 67 good and poor responder patients were classified and compared to the actual retrospective outcome.

Results: Among the six calculated multifractal features, Dqmax exerted the highest predictive value with the prediction accuracy of 74.3%, sensitivity of 72.4% and specificity of 76.2%. The obtained classification accuracy was validated by a ten V-fold split sample cross validation. The area under the curve (AUC) value for the best-performing multifractal Dqmax feature was 0.82 (95% confidence interval, 0.70 - 0.91).

Conclusion: These results suggest for the first time that measuring tumor structure by using multifractal geometry can predict an individual patient response to neoadjuvant cytotoxic therapy. Therefore, it potentially allows precise implementation of alternative treatment options. This predictive approach made use of digital data that is routinely collected but currently still underexploited.

Keywords: Magnetic Resonance Imaging, Fractals, Drug Therapy, Sarcoma, Osteosarcoma, Classification

1. Background

Primary osteosarcoma is the most common bone sarcoma in children and young adults, found predominantly in patients in their second decade of life (1).

In the prechemotherapy era, before 1970’s, osteosarcoma was a disease with very poor survival rates of less than 20% over a 5-year period. The survival rates markedly improved to a 5-year rate of over 80% in the 1970s and early 1980s upon emergence of chemotherapy regimens.

Osteosarcoma is treated with multimodal therapy comprising a combination of preoperative chemotherapy, surgical treatment and postoperative chemotherapy. The degree of tumor necrosis in response to chemotherapy is the most reliable prognostic factor of disease outcome (2). The problem with the current clinical treatment of osteosarcoma is that the plateau in patient survival attained in the 1980’s has not since changed (3). Due to the high level of the cytogenetic heterogeneity of this tumor, both between patients and within the tumors themselves, the major conceivable strategy for improvement in outcomes is based on the introduction of personalized treatment approaches in inductive chemotherapy. Although there are several therapeutic options for osteosarcoma, this is currently not possible because targeted therapies and reliable induction chemotherapy response predictors do not exist (3). This situation has been stimulating research aimed at predictive marker discovery with the main emphasis on molecular biomarkers, including proteins (4) and mRNA (5). However, these emerging prognostic tools are in the
early experimental phase and the prospect of their implementation is still uncertain.

Another evolving strategy suitable for improvement of chemotherapy response prediction is based on a structural analysis of macroscopic morphology of the tumor. It relies on computational analysis of medical images, often by use of fractal geometry which has been developed as a new geometrical concept with an intent to resolve the shortcomings of traditional geometry in structural analysis of complex and irregular natural objects (6). Methods based on multifractal spectra have been found useful for analysis of a wide range of medical images including ultrasound (7), histopathology (8-10), positron emission tomography (11) and magnetic resonance (MR) imaging (12). Tumor image analysis has been explored as a predictive tool in other cancer types (13), but never in osteosarcoma.

2. Objectives

The aim of this study was to test the hypothesis that multifractal computational analysis of routinely collected magnetic resonance images may provide clinical validity in the prediction of the response of osteosarcoma to induction chemotherapy. We employed a neoadjuvant chemotherapy model that is currently accepted as an optimal approach for assessment of in vivo chemotherapy response, since a tumor remains in situ up until the response evaluation (14). The specific objective was to identify the multifractal feature which provides the best predictive value.

3. Patients and Methods

The study was approved by the institutional review board, (Belgrade University, School of Medicine, approval #29/VI-4) and conforms with The code of ethics of the world medical association (declaration of Helsinki), printed in the British medical journal (18 July 1964) and its 7th revision in 2013. For this type of a retrospective study on archived MR images, formal consent is not required. This study conforms with the criteria specified by the Standards for the reporting of diagnostic accuracy studies (STARD) 2015 guidelines (15).

3.1. Patient Group

Cross-sectional study encompassed a group of patients suffering from primary osteosarcoma in long tubular bones, all diagnosed and treated by the national sarcoma consilium during the five-year period of 2010-2014 at the institute for oncology and radiology. Diagnostic intraexamination reliability was assessed using the single measures intraclass correlation coefficient (ICC), as previously described (16). The intra-examination agreement for region of interest (ROI) measurements of 900 images was excellent, with an ICC of 0.93 (0.89 - 0.95, P = 0.02). MRI was performed for all patients, confirming the presence and extent of the tumor formation. Patients were consequently subjected to OsteoSa MAP, a combined methotrexate (M), doxorubicin (A) and cisplatin (P) neoadjuvant therapeutic protocol (doxorubicin, cisplatin, and methotrexate). Upon completion of OsteoSa therapy, the tumor response was histopathologically evaluated. The prospective power calculation rested on a pilot experiment that included 22 patients. The parameters for the sample size calculation were target area under the curve (AUC) of 0.75, null hypothesis value of 0.50, alpha 0.05, beta 0.20, ratio of sample sizes in negative/positive groups 0.50. The required numbers were 42 patients with 28 good therapeutic outcomes and 14 poor therapeutic outcomes as calculated by the sampling for AUC function, (MedCalc software, Ostend, Belgium). The final sample size amounted to 67 patients with 36 good therapeutic outcomes. The actual proportion of patients with good response was 0.54 and AUC was 0.83.

3.2. Imaging Technique

Tumor tissue was imaged prior to the administration of chemotherapy in the transverse (TRA), coronal (COR) and sagittal (SAG) plane in T2W, using the fat suppression (FS) sequence. Nine hundred images were obtained by the diagnostic workstation (Siemens Magnetom Avanto Syngo MR B15, Siemens Healthcare, Erlangen, Germany) and images exported in size of 1920 × 1080 pixels by use of Kodak Carestream picture archiving and communication system (PACS) Client Suite version 10.2 software, Kodak, Rochester, NY. Images were cropped by Fiji/ImageJ software (17) to isolate regions of interest (ROIs) according to the borders of each individual’s tumor formation. The sizes of ROIs reflected the actual variation in tumor dimensions (18). Such a type of image processing is commonly used to obtain the optimally comparable and relevant image details for both mono- and multi-fractal analysis (19). Images were transformed from grayscale to binary values of either 0 or 255 before further processing. Scan background setting was locked to white due to the possibility of background inversion from white to black during binarization process. It is important to note that the finally processed images for multifractal analysis contained an isolated structure of the whole tumor defined by black pixels on a white pixel background.
3.3. Chemotherapy Response Evaluation

As the study was retrospective, the patient stratification into good and poor chemotherapy responder groups was done according to their actual response. Evaluation of the achieved chemotherapeutical effect was performed on the resected tumor material obtained from surgery after an induction chemotherapy along with the gold standard in the pathohistological determination of the grade, based on Huvos grading systems by the expert pathologist (J. P. Sopta) with 21 years of experience. Patients with the level of tumor necrosis exceeding 90% were considered “good responders”, while those with less than 90% of necrosis were considered “poor responders”.

3.4. Image Analysis

The images processed as described above were examined by multifractal analysis using the FracLac plugin version 2015 Marb6206 for Fiji/ImageJ software as described (20). To achieve analytical reproducibility, the range of box sizes (scale window) was set from 10 pixels to the maximal of 60% of an image, Q range from -10 to +10 while minimal density was set at 0.1 and maximal density at 0.98. Multifractal analysis undertakes measurements of the local intensity distribution that scale according to a power law. The obtained exponents thus represent different fractal structures with their own fractal dimensions. The main parameter in multifractal analysis is Holder’s exponent (21) which depends on local regularity of the observed structure (22):

$$\alpha = \frac{\log \mu (\text{box})}{\log \epsilon}$$

(1)

Where $\mu (\text{box})$ refers to the proportion of total intensity that is contained in each box as noted in detail elsewhere (22), while $\epsilon$ is the scale or box size (21). $F(\alpha)$ is defined as the multifractal spectrum. It indicates the global regularity of observed structure by providing the distribution of $\alpha$. Multifractal analysis thus describes both the local and global structure features. High values of the Holder exponent $\alpha$ thus indicate high local changes, while high values of multifractal spectrum $f(\alpha)$ specify the high frequency of specific $\alpha$ values (23). The calculated multifractal features included $\alpha f_{min}, \alpha f_{max}, f(\alpha)_{min}, f(\alpha)_{max}, D_q$ and $D_{q_{max}}$.

3.5. Image Classification

A prediction of the response to three cycles of OsteoSa MAP chemotherapy was attempted by classification of MR images obtained prior to chemotherapy through a decision tree. Images were classified from their pools in poor-and good-response folders. The single-tree model with the Gini splitting algorithm was used to calculate a stratification accuracy of each parameter (DTREG predictive modelling software version 10.3.0, Brentwood, TN (24)). The tree size control settings were set to minimum rows in a node: 1; minimum size node to split: 10; maximum tree levels: 10. The tree pruning and validation by three minimum smooth spikes was done by the ten V-fold split-sample cross validation. The accuracy of a test was the main performance measure, referring to the proportion of true results. Furthermore, sensitivity was calculated as the proportion of patients with chemosensitivity who were correctly identified by the test, while specificity is the proportion of patients who were correctly identified as chemoresistant.

3.6. Statistical Analysis

Areas under the rate-of-change curves (AUCs) were used as a quantitative measure of discrimination efficiency. AUC values were calculated and bootstrap-corrected by use of the Stata®/MP 13 software (StataCorp, TX, USA). AUC curves were produced by use of multifractal feature values obtained on individual images and also by use of average feature values for all images derived from each patient.

4. Results

The predictive potential of multifractal analysis was evaluated by using primary osteosarcoma tumors from the patient group, who were preoperatively treated with the OsteoSa MAP therapeutic protocol. The STARD flow diagram indicates the inclusions and exclusions leading to the final selection of patients and images subjected to multifractal analysis (Figure 1). Representative images for good chemotherapy responders are shown in Figures 2A and 2B and poor chemotherapy responders are demonstrated in Figures 2C and 2D. For multifractal analysis, greyscale images (Figures 2A and C) were transformed to binary images (Figures 2B and D).

The primary tumors are generally the main source of information for the prediction of their drug resistance and the risk of distant metastasis occurrence (8). The relationship between measured continuous feature values and a binary outcome is usually evaluated by the ROC analysis (8). However, in the present study, the performances of continuous multifractal predictor values were additionally assessed by the single classification tree pattern recognition model. Such model has an advantage of not making assumptions about the data structure, thus reducing the risk of model misspecification. The calculated accuracies of such stratification are presented in Table 1.
Djuričić GJ et al.

175 patients with bone tumor present on MRI
Biopsy n = 175
Enrollment
Excluded non-osteosarcoma n = 94
Histopathologically proven osteosarcoma n = 81
Excluded axial skeleton n = 14
Long bone osteosarcoma n = 67
67 tumors × 5 images × 3 planes n = 1005 images
Non-representative tumor presentation n = 105
Images of good responders n = 450
Image binarization n = 900
Images of poor responders n = 450
MR image analysis n = 900

Figure 1. Flow diagram demonstrating inclusions and exclusions of patients and tumor MR images according to the standards of reporting for diagnostic accuracy studies (STARD).

ROC analysis was performed on values of all individual images (Figure 3A) and also on the average values for images of the same tumor (Figure 3B). It is obvious from Figure 3 that a fair AUC value was already obtained on the level of individual images for a Dqmax multifractal feature (-0.65; CI95% 0.62 - 0.68). At the level of individual patients, the AUC value has improved to a range between 0.80 and 0.90, which is generally regarded as excellent (-0.82; CI95% 0.70 - 0.91).

The need for evaluation of the relative predictive power of different multifractal features derives from the fact that this type of analysis delivers numerous parameters. Based on the use of each variable as a primary and surrogate splitter, we have calculated the importance ranking score. Thereby, a variable that was selected as a primary splitter earlier in the tree was considered more important. The variable importance in classification of good and poor chemotherapy responder groups was as follows:

\[ D_{\text{qmax}} = 100.0 \]
\[ \alpha_f \max = 32.0 \]
\[ f(\alpha_f)_{\max} = 31.0 \]
\[ f(\alpha_f)_{\min} = 11.5 \]
\[ D_q = 5.1 \]
\[ \alpha_f \min = 2.9 \]

\[ D_{\text{qmax}} \] is the maximum fluctuation amplitude of \( D_q \), was identified here as the most important predictive parameter and its score adjusted to a value of 100. For better comparison, all other predictors were corrected proportionately. The indication of \( D_{\text{qmax}} \) as the best-performing predictive multifractal parameter obtained by the classification tree was confirmed by the ROC analysis of the multifractal features (not shown), while AUC curves for the best-performing \( D_{\text{qmax}} \) feature are presented in Figure 3.

**Table 1. Classification of Good and Poor Chemotherapy Responder MR Images by Multifractal Analysis**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Performance, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy</td>
<td>74.3</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>72.4</td>
</tr>
<tr>
<td>Specificity</td>
<td>76.2</td>
</tr>
<tr>
<td>Positive predictive value</td>
<td>75.3</td>
</tr>
<tr>
<td>Negative predictive value</td>
<td>73.5</td>
</tr>
<tr>
<td>F-Measure</td>
<td>0.78</td>
</tr>
</tbody>
</table>

*a Ten V-fold split-sample cross validation.

5. Discussion

Chemoresistance remains the main limit of the clinical effectiveness in osteosarcoma. This is due to the lack of tools that could reliably predict the most effective treatment consisting initially of either induction chemotherapy or amputation in larger tumors with narrow surgical margins (3). The present study thus set out to investigate the potential value of morphologic evaluation of osteosarcoma in predicting the response to chemotherapy.

Clinical validity is a criterion for how consistently and accurately the test performs. It should be noted that there are no general thresholds that define a test to be clinically valid and also no possibility of comparison with any previous similar studies on the validity of MR image analysis in osteosarcoma for prediction of the tumor response to cytotoxic therapy. A comparable study of MR images
was however performed on breast tumors with established predictive significance but did not provide any accuracy data (25). The accuracy of 74% obtained in the current study together with the good specificity and sensitivity obtained on individual images (72% sensitivity, 76% specificity; Table 1) and patients (70% sensitivity and 86% specificity for $D_{q_{\text{max}}}$; Figure 3) were an indication of good predictive value, comparable to the predictive accuracy of 82% obtained in our previous study of breast cancer patients with the similar multifractal image analysis approach (13). The predictive studies related to osteosarcoma are focused on molecular markers, with many candidate predictors including RUNX2, CDC51, MDM2, RECQL4, and CDK4. Their prognostic significance has been established, but unfortunately, quantitative measures of association such as accuracy and AUC have typically not been calculated (5), thus limiting any direct comparison with the results of the current study.

The prognosis is quite favorable for good chemotherapy responders, while it is poor for non-responders. The response of osteosarcoma patients to induction chemotherapy is therefore critical for disease outcome. Limb salvage procedures are safe and do not jeopardize the outcome of the patient, if wide surgical margins are possible. The clinical utility of the predictive test described here is most relevant for patients with borderline surgical margin width because the prediction of chemotherapy resistance in this subset could reinforce the decision to substitute the induction chemotherapy either with an experimental protocol (clinical trial) or an amputation. This could improve the disease outcome and survival by means of reducing the probability of distant metastasis occurrence during the several months’ course of an induction chemotherapy to which the tumor is resistant.

The noted predictive power of $D_{q_{\text{max}}}$ could be explained by the existence of unknown tumor histomorphological features that are typical for either responsiveness or unresponsiveness to chemotherapy. These features may arise from different growth patterns of malignant cells depending on whether they are chemoresistant or chemosensitive. The previous study considering the prediction of a chemotherapy response of breast carcinoma based on multifractal analysis of microscopic histopathology images identified $f(\alpha)_{\text{max}}$ as the best performing prediction.
parameter (13), while in prognostic studies $f(\alpha)_{\min}$ and $D_{q_{\text{max}}}$ were the best performers (8). The observed discrepancy can be explained by different tumor and imaging types and by differences in chemotherapy.

In conclusion, by using prediction of the chemotherapy response as a model, we have tested the hypothesis that computational analysis of routinely collected magnetic resonance images together with the machine-learning classification techniques, could enhance conventional strategies of individual therapy adjustments. Thereby, the multifractal analysis of primary osteosarcomas prior to chemotherapy was here shown for the first time to possess sufficient morphometric discriminating capacity to differentiate between tumor MR images based on their actual chemoresponsiveness. This is of crucial importance as osteosarcoma patients frequently develop resistance to induction chemotherapy, which is still a major obstacle to achieving more favorable outcomes. Among the employed multifractal parameters $D_{q_{\text{max}}}$ exerted the best predictive power. The generalizability of the classification model was supported by the internal ten V-fold cross validation, though additional studies in external patient groups would be needed for further characterization of the predictive clinical validity of $D_{q_{\text{max}}}$. The potential clinical benefit of improvements in chemotherapy prediction derives from the key impact of the correct choice between induction chemotherapy and amputation on the life quality and survival. Usefulness of this methodology is further potentiated by its cost-effectiveness emerging from rapid analysis of digital MR images that are routinely collected but currently still underexploited.

Footnotes

Financial Disclosure: Authors declare no financial interest related to the material in the manuscript.

Funding: This work was supported by the ministry of education, science and technological development of the Republic of Serbia, grants: #45005, TR32037, ON175068.

References


Figure 3. Receiver operating characteristic (ROC) curves for diagnostic efficacy assessment of $D_{q_{\text{max}}}$ regarding determining response to chemotherapy (per image approach and per patient approach). A, ROC diagram for $D_{q_{\text{max}}}$ obtained on the 900 individual MR images; B, ROC diagram for $D_{q_{\text{max}}}$ obtained on MR images averaged for each patient. The AUC values were transformed from below to above 0.5 for their easier comparison. A negative sign was therefore added to preserve the information of the inverse association between $D_{q_{\text{max}}}$ and chemotherapy response.


