FRactal Parameters of Tumour Microscopic Images as Prognostic Indicators of Clinical Outcome in Early Breast Cancer

Jelena Pribic¹, Jelena Vasiljevic², Ksenija Kanjer¹, Zora Neskovic Konstantinovic¹, Nebojsa T. Milosevic³, Dragica Nikolic Vukosavljevic¹ and Marko Radulovic¹ and Natasa Zivic⁴

¹Department of Experimental Oncology, Institute of Oncology and Radiology of Serbia, Pasterova 14, University of Belgrade, Belgrade, Serbia
²Institute “Mihajlo Pupin”, Volgina 15, University of Belgrade, Belgrade, Serbia
³Department of Biophysics, School of Medicine, University of Belgrade Visegradska 26/2, Belgrade, Serbia
⁴University of Siegen, Germany

ABSTRACT

Research in the field of breast cancer outcome prognosis has been focused on molecular biomarkers, while neglecting the discovery of novel tumour histology structural clues. We thus aimed to improve breast cancer prognosis by fractal analysis of tumour histomorphology. This study included 92 breast cancer patients without systemic treatment. Fractal parameters-fractal dimension and lacunarity of the breast tumour microscopic histology possess prognostic value comparable to the major clinicopathological prognostic parameters. Fractal analysis was performed for the first time on routinely produced archived pan-tissue stained primary breast tumour sections, indicating its potential for clinical use as a simple and cost-effective prognostic indicator of distant metastasis risk to complement the molecular approaches for cancer risk prognosis.

KEYWORDS

fractal dimension, image analysis, lacunarity, breast cancer, prognosis, chemotherapy

1. INTRODUCTION

The importance of breast cancer metastasis risk prognostication derives from the fact that metastasis occurrence is exceptionally variable and that death by breast cancer is mainly caused by metastatic relapse at distant sites. Such importance of metastasis for the disease outcome means that the prognostication of metastasis risk is the central factor for decisions on the...
individual therapeutic regime. For this reason, besides local surgery and radiotherapy, patients at elevated risk are increasingly treated with postoperative systemic therapy with an aim to eradicate distant micrometastatic deposits.

Breast cancer metastasis risk biomarkers have been a field of intensive research in the past decade with a focus on molecular prognosticators including transcriptional profiling [1], microRNA analysis [2], detection of circulating tumour cells in blood [3], proliferation [4], and stem cell markers [5], while research for novel histomorphological prognostic clues as a source of prognostic information has been largely neglected. The need for new prognostic approaches derives from the fact that molecular risk biomarkers often outperform the established clinicopathological prognosticators, but regrettably still exhibit insufficient prognostic accuracy, with the remaining unreliable therapeutic guidance. A therapeutic bias towards over-treatment arises as a consequence of such insufficient risk prognosis accuracy based on standard clinicopathological features such as tumour size, nodal status, metastasis (TNM stage), histological grade, steroid receptor status, age, menopausal status and the molecular prognosticators including gene expression arrays [6, 7]. The downside of over-treatment strategy is that patients which do not develop distant metastases needlessly suffer from systemic toxic side effects of chemotherapy [8]. Taken together, there is an urgent need for distant metastasis prognostic tools with improved accuracy, in order to achieve a more effective therapeutic management.

The accurate prognosis of breast cancer outcome is exceptionally challenging based on the fact that malignant transformation is a random process rendering every tumour unique. The primary tumour of patients with early breast cancer is used as the main source of information for assessment of distant metastasis risk. Apart from the above mentioned molecular traces of a tumour, its histology reveals valued prognostic clues due to the fact that the tissue structure reflects the individual patterns of malignant cell growth. A demand for improvement of breast cancer risk prognosis has triggered the idea to implement alternatives to molecular approaches, with a resulting expansion of the digital pathology as a structure examination tool for medical images [4].

2. MATERIALS AND METHODS

Samples were obtained from surgically removed invasive primary breast tumours. Tissue was formalin-fixed, paraffin-embedded and cut to produce 4 µm whole sections which were mounted on slides, deparaffinised and rehydrated. Storage time of archival embedded samples was up to 1 month.

Tissue sections were subsequently stained with haematoxylin and eosin (H&E), the most common type of staining in routine pathological examination of the tumour tissue. Haematoxylin stains the cell cytoplasm and collagen pink, while eosin stains cell nuclei blue.

Digital photographs were captured at x400 magnification under the Olympus BX-51 light microscope with a mounted Olympus digital camera. At least five representative non-overlapping colour images were captured from each section in 3638 × 2736 pixel resolution and saved as TIFF. Pixel size on image was 145 nm. Photographs were taken to comprise tumour areas with large numbers of malignant cells showing typical patterns of malignant growth for each tumour.
These images were also screened by a pathologist and those including mostly non-malignant cells, untypical growth or artefacts, were deleted.

Fractal analysis of histology images was based on the regular non-overlapping box counting method by use of the freely available ImageJ software (http://imagej.nih.gov/ij/) and its FracLac plugin [11]. The reproducibility of the analysis was based on the settings for the range of box sizes (scale window) defined as 2, 4, 8, 16, 32, 64, 128, 256, 512, 1024 pixels. With the pixel size of 145 nm the grid lengths ranged between 0.29 m and 148 m. Furthermore, scan background setting was locked to white in order to avoid the inversion of the binary image from white to black background during analysis.

Box counting delivers two main parameters: fractal dimension (FD) and lacunarity as respective measures of complexity and heterogeneity. Three control tumour histology images with known fractal dimensions and lacunarities were analysed in all runs in order to control for reproducibility. The box-counting method involves covering of the digital image with a grid of boxes with size (scale) expressed as the box size relative to image size. The count (N) refers to the space filling properties of the image measured as the minimum number of grid boxes needed to cover all parts of the image containing foreground pixels [32]. Equation (1) for FD represents the negative limit of the ratio of the logs of the number and size of boxes [33]:

\[ FD = \lim_{\varepsilon \to 0} \frac{\log N(\varepsilon)}{\log(\varepsilon)} \]  

(1)

However, such limit cannot be calculated for digital images due to the lack of infinite scaling as the resolution restricts the range of available scales to one pixel as the smallest scale. Furthermore, the largest available scale is also constrained by image width and height. For these reasons, the fractal dimension of natural objects is estimated as a negative slope of the straight part of the regression line (fractal window) at the above (1) double logarithmic plot of box-size and box-count [12].

Lacunarity was also obtained by use of the regular box counting method as a measure of the heterogeneity within an object. Patterns having larger or more numerous gaps generally have higher lacunarity, while low values of lacunarity imply homogeneity based on similarly sized gaps and little rotational variance. Lacunarity is estimated on the basis of the pixel mass distribution probability determined from the number of pixels per box as a function of box size (\(\varepsilon\)). Lacunarity at a particular \(\varepsilon\) (\(\lambda_\varepsilon\)) is calculated as:

\[ \lambda_\varepsilon = \left( \frac{\sigma}{\mu} \right)^2 \]  

(2)

where \(\sigma\) is the standard deviation and \(\mu\) is the mean of the foreground pixels per box at \(\varepsilon\). To obtain a single number, the variance calculations are expressed as mean lacunarity (\(\Lambda_{av}\)) for the total number of box sizes used:
\[ \Lambda_{av} = \sum \frac{\lambda_x}{E} \]  

Measured continuous values are converted to categorical values in order to place patients into risk strata. For the present analysis, such categorization was performed by use of "optimal" cutpoints which define the two groups with minimal P-value, calculated by the X-tile 3.6.1 software (Yale University, New Haven, CT). The prognostic accuracy for each variable was calculated based on categorisation performed by the internal split-sample cross-validation procedure. The original sample was randomly split into a training set of 44 patients and a validation set of 48 patients by use of =RAND() function in Excel, for three validation cycles [14]. The discriminating cutpoint with minimal P-value for categorisation was selected in the training partition and applied to the test partition. The training and test partitions were subsequently swapped and the process repeated.

Fractal and lacunarity features were adjusted for standard clinicopathological variables by multivariate Cox analysis in order to identify independent prognostic factors. Variables were added to a model using a training partition of 44 patients with a forward selection entry criterion of P<0.05 in univariate analysis. Features were removed by use of backward elimination according to a selection stay criterion of P<0.05. The procedure was based on internal split-sample cross-validation as described above in detail for feature categorization.

Spearman’s rank correlation test was employed to calculate the correlation between variables. Analyses were completed by SPSS version 20 (IBM, Chicago, IL) and Stata/MP 13 software. Adjuvant! Online score for Breast Cancer (Version 8.0) was calculated at https://www.adjuvantonline.com/breastnew.jsp as a 10 year risk of relapse with no additional therapy based on age, tumour grade, ER status and tumour size.

3. RESULTS

Fractal analysis presents a morphometric technique useful for measurement of irregular and complex shapes [12]. Fractal dimension and lacunarity were calculated for individual histology sections as parameters which may quantify the risk diversity of primary breast tumours (FIGURE 1). Estimation of fractal dimension is a multiscale measure of space filling, based on the fact that complex objects fill more space than simpler ones. A 2-dimensional image can theoretically have a maximal FD of 2 which means that the obtained value range for FD (1.66 -1.86) indicates a geometrically rather complex tissue structure that is highly space-filling. The value ranges for lacunarity were wider at 0.21 - 0.68. Whereas mathematical fractals are invariant in shape over an unlimited range of scales, biological fractal patterns are self-similar only statistically in terms of complexity within a fractal window, with upper and lower scaling limits spreading at least two orders of magnitude, a condition that must be experimentally established for a particular type of a sample. Therefore, we examined the statistical self-similarity based on the linearity of the double logarithmic box-number/box-size curve for the range of analysed scales (2 - 1024 pixels), as previously explained in detail [15]. The linearity was thereby established for the whole range of scales of nearly three orders of magnitude, meaning that scaling and fractal windows overlapped from 0.29 \( \mu \)m to 148 \( \mu \)m, based on the pixel size of 145 nm.
Several significant correlations were evident between fractal and clinicopathological variables, yet with low Spearman coefficients. FD correlated significantly only with the pathological tumour size (Spearman coefficient of 0.27), while lacunarity correlated with Adjuvant! (-0.22) and with the pathological tumour size (-0.25). Such low correlations suggested a possible prognostic independence of FD and lacunarities, to the advantage of their prognostic values. In agreement with this, fractal dimension was indicated as independent prognostic factor by the multivariate Cox regression.

4. DISCUSSION

We approached the task of breast cancer risk prognosis improvement by exploiting the prognostic value of tumour microscopic histology at the time of surgery. Based on the established ability of fractal geometry to quantify irregular structures of tumour microscopic histology [37, 38], we hypothesized that this approach may efficiently acquire prognostically relevant structural information and thus provide a valuable addition to existing clinicopathological and molecular prognosticators in breast cancer.

This study was performed on patients with a natural course of disease due to interference of systemic therapy with metastasis occurrence. Such patients are increasingly difficult to find due to a trend of widening systemic therapy use [16]. The patient group was assembled from a period of over 20 years ago when low metastasis risk patients have not been prescribed systemic therapy at our institution. Lower risk classification was based on a tumour grade < 3, pT < 3 and absence of lymph node metastases.

Morphological complexity is the property revealed by the retention of shapes of high structural complexity even after magnification (scaling, zooming). FD quantifies such complexity by measuring the rate of addition of structural detail with increasing magnification within the linear part of the plot designated as fractal window [17]. Our data indicate that high FD associates with high distant metastasis risk. With FD regarded as a formal means of quantifying shape complexity [11, 16], it can be considered that high structural complexity of H&E stained tumour tissue associates with high risk. This result is consistent with the previously reported prognostic value of FD obtained on pan-cytokeratin stained breast tumour tissue sections [14]. By marking epithelial cells, this kind of staining directs the fractal analysis primarily to the outline of malignant tissue growth patterns, while our analysis was performed on unspecifically stained H&E tissue sections. Furthermore, high tumour grade as an indicator of increased risk also reflected high histomorphological disorder and complexity [18].

On the other hand, lacunarity associated negatively with metastasis risk. Mandelbrot originally proposed lacunarity as an adjunct to fractal analysis because it proved useful in discerning amongst images that have similar fractal dimensions [19]. Lacunarity assesses the texture qualities of the structure based on distribution and size of the empty domains, while fractal dimension specifies how completely a fractal-like structure fills the space for decreasing scales [20]. Beyond being a measure of gappiness, lacunarity quantifies additional features such as "rotational invariance" and more generally, heterogeneity. It follows that high risk is prognosticated by low heterogeneity and high complexity of microscopic histology images, as respectively indicated by lacunarity and FD.
Results obtained in this study are the first to indicate the association between FD and lacunarity of the whole H&E tumour tissue microscopic histology images and disease outcome. Cost efficiency is an advantage of H&E tissue section analysis, as such histology samples are routinely produced following tumour extraction surgery for use in standard pathological diagnostic and prognostic assessments, without any need for immunohistochemical staining and associated expenditure for chemicals.

We investigated the prognostic potential of fractal properties in a hypothesis-free fashion as previous studies mainly analysed tumour tissue specifically stained with known prognostic markers [42]. By targeting specific features, primarily nuclei [21], even the studies using non-specifically H&E-stained tissue sections [22-26] were in fact not analytically hypothesis-free. Although pathologists commonly use the pan-tissue H&E staining for pathologic assessment of histology specimens, it has been considered as suboptimal for fractal analysis27. This opinion was based on the theoretical consideration that only specific highlighting of malignant epithelial cells may provide structures of interest for optimal image analysis, as non-malignant components stained by H&E such as stromal cells, extracellular collagen and elastic fibres could diminish the extraction of useful prognostic clues. However, our current results indicate that image analysis aimed at cancer risk prognosis can in fact be simplified by omitting the specific staining step, without sacrifice of prognostic power. On the basis of this result it is concluded that an overall tissue structure, without any specific markings of malignant cells, provides optimal prognostic clues which can be extracted by fractal image analysis. This is a tremendous cost-advantage relative to the established immunohistochemically-stained prognostic markers and especially the expensive gene expression profiling prognostic tools such as Oncotype DX and MammaPrint [28]. Remarkably, even these expensive tests actually present cost-effectiveness through improved low-risk stratification with the consequent reduction in chemotherapy utilization [1].

Prognostic value of fractal analysis is acquired in a different way than a visual grading of histological clues by a pathologist. The underlying tumor biology that fractal analysis may reflect include yet unidentified microscopic structural patterns of a tumour, such as tissue growth, cellularity apoptosis, hypoxia, angiogenesis and distribution of mitotic cells [29-31]. Due to a stochastic nature of carcinogenesis, each tumour is a different entity with unique growth patterns that may be based in large part on fractal geometry [32], thus rendering fractal analysis a particularly relevant readout of tumour heterogeneity.

The important limitation of the regular box counting used in this study is that its methodological validity is restricted only to statistically self-similar signals. Furthermore, the requirement for signal binarization leads to a loss of information [33] adaptation of the method called differential box-counting was developed to overcome this limitation by allowing analysis of grey-scale images. A limitation in scaling is another general issue for image analysis, with a range between one pixel as the smallest scale and the size of the whole image as the largest scale. In addition to limitations in image resolution and size, natural objects are intrinsically limited in scaling due to the finite size of their structural units [12]. A limit to observable increase in detail for the biological samples becomes evident at the critical magnification of ×130,000. Furthermore, the regular box counting method is limited by its calculation of generalized parameters for the whole image without providing any insight into the regional variations. Multifractal analysis as an extended version of monofractal analysis provides regional measurements of several multifractal parameters across an image [13]. Fractal sub-scans and local connected fractal dimension (LCFD) also provide regional information by calculating and visualising the distribution of local fractal dimensions within the image. All of these methods are a variation of a common calculation of the
pixel numbers for various lengths (scales) of the image. In case where fractal properties are not constant in all parts of the image, such regional analyses may perform better than fractal parameter values averaged across the whole image. The lack of full objectivity is a general important limitation of this type of image analysis based on the fact that although fractal approach enables to objectively describe irregular morphologic components and ultrastructural features, the selection of the representative tissue segments for analysis is still done subjectively to the best knowledge of a pathologist.

Based on discrimination efficiencies indicated by Kaplan Meier plots and effect sizes by hazard ratios, the prognostic performance of FD and lacunarity was comparable to established clinicopathological prognostic parameters such as pathological tumour size, estrogen receptor (ER) status and clinicopathological Adjuvant! composite score. AUCs as discrimination measures were underperformers for fractal parameters, while validated accuracies obtained in this study (64-74%) surpassed the specific prognostic factors previously investigated by us such as Cathepsin D expression [28], trefoil factor 1 and comparable to the established and costly FDA-cleared 70-genes expression Mammaprint molecular prognostic tool with the reported accuracy of 65%

Internal validation was successfully performed by bootstrap and split-sample cross-validation, suggesting that the prognostic parameters tested are generalizable. Whereas further validation is needed on external and expanded group of patients, this preliminary study indicates for the first time the potential use of fractal analysis of primary breast tumour histology as the cost-effective prognostic indicator of distant metastasis risk.

5. CONCLUSIONS

By use of fractal analysis we measured the structural characteristics of 92 breast tumour histological specimens in patients who did not receive systemic therapy. On the basis of obtained results it is concluded that histology fractal dimension and lacunarity correlate with biological properties of a tumour and present a promising strategy to assess the risk of distant metastasis. The original contribution of this study to the field of prognosis of breast cancer outcome is based on providing the first evidence that fractal characteristics of native tumour histology highlighted by the H&E pan-tissue stain, without any focus on specific structures or molecular biomarkers, can prognosticate an individual's risk of metastasis within a long median follow-up period of 150 months. The potential clinical use of FD and lacunarity is based on their prognostic performance which was comparable to standard clinicopathological prognosticicators. Furthermore, the fractal prognostic biomarkers are simple and inexpensive to measure from routine archived H&E-stained histology, thus permitting both prospective and retrospective validation studies as well as possible clinical application without additional costs.

ACKNOWLEDGEMENTS

This work was supported by the Ministry of Education and Science, Republic of Serbia, Science and Technological Development grants ON175068, III 45005 and TR32037.
Papers of special note have been highlighted as:
* of interest
** of special interest

REFERENCES


[31] Ieee Computer Society., Institute of Electrical and Electronics Engineers.: IEEE transactions on pattern analysis and machine intelligence. v.


AUTHORS

Jelena Vasiljevic, PhD of Electrical Engineering, University of Belgrade, Institut Mihajlo Pupin, Docent for Intelligent Systems on Faculty for Computer Science. Jelena Vasiljevic works as Science Assistant in Institute Mihajlo Pupin in Belgrade, Serbia and also as Docent for Intelligent Systems teaching on Faculty for Computer Science. She is teaching Intelligent Systems (Artificial Intelligence), Genetic Algorithms, Neural networks, Fuzzy Logic, Bioinformatics, Knowledge based Systems on PhD, Master and Primary studies. She is a graduate of Faculty for Electrical Engineering on University in Belgrade, Serbia. PhD research was application of multifractal analysis of digital medical images in cancer diagnostic and for it she has been several times awarded: The First Award on International Medical Fair 2014, The Special Award on International Technical Fair 2014 and many others (also many interviews for TV and newspaper, in Oxford Encyclopedia). She is also an author and co-author of many journal and conference peer reviewed.