Optimisation of the quantitative analysis of inflammatory cell infiltrates in breast cancer

Optimizarea analizei cantitative a infiltratului celular inflamator în cancerul mamar

Anca-Raluca Vrânceanu1*, Cristina Claudia Tănăcieriu2, Daniela Jitaru3, Cristina Terinte4, Florin Zugun-Eloae3, Eugen Carasevici3

1. Department of Oncology, Regional Institute of Oncology, Iași, România; 2. Department of Anatomy Sciences - “Grigore T. Popa” University of Medicine and Pharmacy, Iași, România; 3. Laboratory of Immunology and Genetics, Regional Institute of Oncology Iași, România; 4. Department of Pathology, Regional Institute of Oncology, Iași, România.

Abstract

In this study we aimed to determine the optimal cut-off point for the quantitative analysis of inflammatory infiltrates in breast cancer, using the HistoQuest system. We used samples of tumour breast tissue which were IHC stained with CD68 and CD8 and subsequently tested with automated systems on three regions: intratumoral, invasive front and peritumoral, using the HistoQuest system. In order to delimit between positive and negative cells on histograms and scattergrams, we need to set a cut-off value. We compared 5 cut-off types for optimisation of the quantitative analysis. The results obtained statistically for the CD8 marker for all 5 types of cut-offs applied on IT, PT and IF regions did not show statistically significant differences (p > 0.05). As for the CD68 marker, we found statistically significant differences (p < 0.05) between manual cut-offs (C2 – manual and C3 – manual, arithmetic mean) and automated cut-offs placed by the software (C1 – automated, C4 – negative region, and C5 – automated, arithmetic mean), which suggests that the use of an automated cut-off should be preferred in order to remove the subjective factor. The automated cut-off setting generates objective and reproducible data and can be used in subsequent quantitative analyses.

Keyword: the quantitative analysis, cut-off, inflammatory infiltrates, breast cancer

Rezumat

În acest studiu am analizat modalitatea optimă de poziționare a cut off-ului în cadrul analizei cantitative a infiltratului inflamator în cancerul mamar cu ajutorul sistemului HistoQuest. Am utilizat specimene de ţesut mamar tumoral, care au fost marcate imunohistochimic cu markerii CD68 (pentru macrofaje) și CD8 (pentru limfocitul T citotoxic) și ulterior analizate automat pe regiunile IT (intratumoral), PT (peritumoral) și FI (front de invazie) cu ajutorul sistemului soft HistoQuest. Pentru evidențierea populației de celule pozitive pe histograme și

*Corresponding author: Anca Raluca Vrânceanu, Regional Institute of Oncology Iași, România, e-mail: vranceanuancaraluca@yahoo.com
scatergrame, este necesară poziționarea cut off-ului (discriminator). Pentru optimizarea tehnicii, am comparat 5 modalități de poziționare a cut off-ului. Rezultatele obținute statistic pentru markerul CD8 pentru toate cele 5 tipuri de cutoff aplicate pe regiunile IT, PT și FI nu au înregistrat diferențe semnificative statistic (p < 0,05). Pentru markerul CD68 s-au înregistrat diferențe semnificative statistic (p<0,05) între cut off-urile alese pe criterii manuale (cut off-ul C2-manual și cut off –ul C3-medie manual) versus cut off-urile stabilite pe criterii ce ţin de soft (cut off-ul C1-automat, cut off-ul C4-regiune negativă și cut off-ul C5-medie automat), utilizarea cutoff-ului automat fiind preferat pentru a îndepărta factorul subiectiv. Utilizarea modalității de poziționare automată a cut off-ului generează date obiective și reproductibile și poate fi utilizat în analize cantitative ulterioare.

Cuvinte cheie: analiză cantitativă, cut off, infiltrat inflamator, cancer mamar

Received: 9th April 2014; Accepted: 1st September 2014; Published: 8th September 2014.

Introduction

Breast cancer is a public health problem and the second leading cause of death in women [1]. Cancer results from a series of genetic changes, but interactions between tumour and stromal cells are also involved [2]. Currently, the only markers used for identification of the subset of patients who can benefit from specific treatments are estrogen receptors (ER), progesterone receptors (PR) and human epidermal growth factor receptor 2 (HER2); however, the predictive and prognostic value of these markers when combined needs to be defined better [3].

Many studies have proved the existence of a relation between immune cell infiltrates, prognosis, and treatment response for a range of types of human carcinomas [4]. The analysis of immune cell infiltrates in breast cancer has suggested a correlation between immune cell infiltrates and prognosis, but with conflicting findings which can be partially explained by the use of different methodologies or reactants [5]. The study of immune cell infiltrates is complex and involves quantification of many cell types which may have reached different levels of maturation or activation (the CD4+ T-lymphocyte can be anergic, activated or regulatory). Another impediment in the study of cell infiltrates is that different cell types can share the same specific marker. Technology advances like PCR, flow cytometry and microarray cannot deliver a detailed analysis of the tumour immune infiltrate in relation to its location from the tumour site [6]. Recent investigations have shown that many tumours are strongly infiltrated by myeloid and lymphoid cells which focus in the tumour bed and at the invasive margin, while lymphoid cell islands are found close to the tumour site [7]. There is evidence that the type, number and location of cells have prognostic value, therefore the quantitative analysis of infiltrates can be considered as a potential prognostic factor [8].

Computer-aided quantitative analysis of histopathologic specimens is under continuous development and improvement FARSIGHT, analySIS [9,10]. The advantages of automated systems are the operation speed and the minimal interference of the human factor, thus reducing the risk of inter-observer and intra-observer variability. New software products even more easy to use are continuously being developed and released every year [11-15]. One of the latest systems designed for the quantitative analysis of IHC staining is HistoFAXS, which we used for our current research. The first version of the software was released in 2004. The quantification of myeloid and lymphoid infiltrates needs to be optimised in order to obtain objective and reproducible results. The aim of current research is to optimise the workflow for the digital quantitative analysis of IHC stainings. To that effect, immunohistochemical evaluation of myeloid and lymphoid infiltrates was performed on a set of cases with breast neoplasm. Optimisation and standardisa-
tion of the quantitative analysis method are particularly focused on the determination of cut-off values in order to avoid errors which can lead to altered results.

**Material and method**

The quantitative analysis of inflammatory infiltrates on the breast neoplasm set was performed in the Molecular Biology Laboratory from the Regional Institute of Oncology, Iaşi. The quantitative analysis comprised of the following steps: immunohistochemical staining, image acquisition and digital quantitative analysis using a cell analysis system – HistoFAXS – with its two components: TissueFAXS and HistoQuest. A statistical analysis was performed on the findings.

**Tissue Staining**

The samples from 10 cases of breast neoplasm were IHC stained with CD68 macrophage and CD8 cytotoxic T-lymphocyte specific markers. Paraffin-embedded tumour blocks were cut in sections of 4 µm and mounted on adhesive-coated slides. They were treated with Dako retrieval solution, pH=9 (code S2367) for 30 minutes before staining. Antibodies used for immunohistochemistry: Monoclonal Mouse Anti-Human CD8, Clone C8/144B, Dako, (code No. M 7103) with 1:100 dilution and 1h incubation at room temperature and FLEX Monoclonal Mouse Anti-Human CD68, Clone KP1, Ready-to-Use, Dako, Code IS609, with overnight incubation at 4ºC. For visualisation we used DAKO LSAB™+/HRP kit, code No. K 0679.

**Tissue Imaging**

The stained tissue sections were analysed with HistoFAXS, which includes an image-acquiring system (TissueQuest) and uses a microscope Zeiss Observer Z1 area scan, colour camera Pixelink PL-623 CF and a software for cytometric analysis (HistoQuest). Images were obtained using a 20x magnification objective on the regions of interest, after acquisition of the whole section with a 2.5x objective. The scope was to observe the tissue section overall, after obtaining the image with a 2.5x objective. The overall analysis intends to identify and define the regions of interest. Our analysis focused on the definition of regions, grouped as follows: intratumoral, peritumoral, and invasive front. The regions were identified based on the defined criteria, marked and annotated. The definition and analysis of these regions will help us to interpret the intensity and structure of myeloid and lymphoid infiltrates in the respective regions in the hope that we can partially decipher the tumour process dynamics, which will lead to an understanding of the tumour growth and metastasis mechanism. The inflammatory infiltrates were analysed in 9 regions of interest, 3 for each section: 3 intratumoral, 3 peritumoral, and 3 from the invasive front.

**Intratumoral regions** (IT) are chosen from the tumour bed, where tumour cells are well represented and exhibit the specific tumour changes: small-sized nucleus with visible nucleoli, ill-defined cytoplasmic borders (Figure 1B).

**Peritumoral regions** (PT) are located at the tumour periphery. The selected section comprises tumour tissue and peritumoral tissue in a 50:50 ratio. The peritumoral tissue should be free of tumour invasion, so as to compare it against the invasive front (Figure 1C).

**The invasive front** (IF) is also peripheral to the tumour. It comprises tumour tissue and peritumoral tissue in a 50:50 ratio, however it shows evidence of tumour invasion (as opposed to the peritumoral area).

When defining the tumour invasion, we considered data from specialized literature [7], which describe it as the presence of a group of 3-4 cells with tumour characteristics in the peritumoral (fat, conjunctive, glandular) tissue (Figure 1A).
In order to obtain objective data, we chose a standard-sized region from all analysed regions of interest. All the regions of interest (3 intratumoral, 3 peritumoral, 3 from the invasive front) were round and had a fixed size of 0.3760 mm$^2$ (Figure 1A, 1B, 1D). The obtained images were stored and subsequently imported into the HistoQuest analysis software.

The software performs a cell delimitation which may be optimised by modelling certain parameters. These parameters define certain cell characteristics specific to nuclei (Nuclei size, Remove small-sized objects, Remove weakly stained objects, Automatic Background Threshold, Virtual channel, Post Processing Order, Remove Labels, Use Merging Rules) and to the markers used (Use Ring Mask, Use Identified Cell Mask, Automatic Background Threshold). This method is under continuous developing, and new variants are already available. For the current quantitative analysis we used HistoQuest version 3.5.5.0171. Optimisation of these parameters is necessary to obtain a proper delimitation of cell elements and to minimize potential errors due to IHC staining, which are inherent to this technique (e.g. the background noise). After this step, the software automatically generates statistical data. Histograms and scattergrams of the regions of interest comprise all the cells from the respective regions (Figure 1C). In order to delimit between positive and negative cells, we need to place a cut-off value. This cut-off can be generated automatically or manually. Automated cut-off has variable values from one region to another, while manual values can be unique for all regions or can be variable based on initially defined criteria. In order to identify the cut-off type which generates objective data with no systematic errors, we suggested to compare 5 cut-off types. The cut-off value will be subsequently used for the analysis of myeloid and lymphoid infiltrates on an extended set of patients, with a range of markers which define its composition and function in further detail.

**Cut-off 1** – automated cut-off, with variable value for each region of interest.

**Cut-off 2** – manual cut-off, initially placed at random and subsequently modified by creating a narrow gate around the random cut-off and moving it such as 50% positive cells and 50% negative cells are obtained – variable value from one region of interest to another.

**Cut-off 3** – the cut-off value is the arithmetic mean of all manual cut-off values (cut-off 2), and applies to all regions – a unique value for all regions.

**Cut-off 4** – is obtained by using a negative region (all cells are negative for the used marker), and positioned such as max. 1% of the cells are interpreted as positive – a unique value for all regions.

**Cut-off 5** – is the arithmetic mean of all automated cut-off values (cut-off 1), re-applied to all histograms of the regions of interest – a unique value for all regions.

Automated values obtained after successive placing of all 5 cut-off types on each region of interest were inserted in Excel tables, showing the number of CD68-positive and CD8-positive cells per mm$^2$.

The obtained results were statistically analysed based on the two used markers: CD68 and CD8. Correlation between cut-offs based on the two IHC markers was made using a statistical analysis by Wilcoxon test.

**Results**

The above methods are intended to identify the differences in the three types of regions (intratumoral, peritumoral and invasive front) based on each cut-off type. The final scope of the analysis is to see if possible differences between regions or even the homogeneous results obtained for a cut-off value are valid for all the
Figure 1 A) CD68-stained invasive front region from breast tumour tissue, obtained with a 20x objective. The 0.3760 mm$^2$ peritumoral region of interest is adjacent to the tumour: it comprises of tumour tissue and peritumoral tissue in a 50:50 ratio, and it shows evidence of tumour invasion. Nuclei-specific variables are shown on the left side. Graphic representations of the region of interest are included on the right side: histograms and scattergrams. B) CD68-stained breast tumour tissue, obtained with a 20x objective, shows a 0.3760 mm$^2$ round tumour region of fixed size taken from the tumour bed. C) CD68-stained peritumoral region from the breast tumour tissue, obtained with a 20x objective. The 0.3760 mm$^2$ peritumoral region of interest comprises of tumour tissue and peritumoral tissue in a 50:50 ratio. The clear delimitation between the two types of tissue suggests that there is no invasion at this level.
other cut-off values. If the cut-offs show no statistically significant differences, we suggest the use of the most accessible variant – the automated cut-off generated by the software, which varies with the analysed region. If statistically significant differences are found, these should be analysed and explained in order to support the use of a certain type of cut-off.

The statistical results obtained for the CD8 marker for all 5 types of cut-offs applied on IT, PT and IF regions showed no statistically significant differences (p>0.05). The arithmetic means of CD8-positive cells from the three regions of interest are not significantly different for each considered cut-off (C1, C2, C3, C4, C5) as statistical results by nonparametric Wilcoxon test and the obtained results are graphically presented (Figure 2).

Statistically significant differences were found for the CD68 marker (p<0.05) using nonparametric Wilcoxon test. The differences were found between cut-offs C1 (automated cut-off) and C3 (manual cut-off, arithmetic mean) for the invasive front (p=0.047) and the intratumoral regions (p=0.022), respectively. Statistically significant differences between C1 and C2 cut-offs were obvious for the intra-tumoral regions (p=0.017), as also C2 and C3 cut-offs were for the same region (p=0.017). The results were significantly different between C2 and C5 again in the intratumoral regions (p=0.007), as well as C3 and C5 cut-offs (p=0.017). All these data are presented in Table I and are graphically shown in Figure 3. Statistical analysis show a significant power effect on cut-offs C1_Fi – C3_F (r = 0.628), C1_PT – C3_PT (r = 0.725), C1_IT – C2_IT (r = 0.757), C2_IT – C3_IT (r = 0.757), C2_IT – C5_IT (r = 0.854), C3_IT – C5_IT (r = 0.757) (Table I).

![Figure 2](image_url)

**Figure 2** The graphic representation of the values obtained from quantitative analysis of sections of CD8-stained breast tumour tissue for the three types of regions of interest (IF, PT, IT) with the 5 suggested cut-offs.
Table I. Statistical data for CD68 show that there are statistically significant differences (p<0.05) between the compared pairs of values, i.e. the arithmetic means of CD68-positive cells from the three regions of interest (IF, PT, IT) are significantly different between C1 and C3 for the invasive front and peritumoral region and between C1 and C2 for the intratumoral region, also between C2-C3, C2-C5 and C3-C5.

<table>
<thead>
<tr>
<th>Test Statistics(^c)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C2_Fi - C1_Fi</td>
</tr>
<tr>
<td>(Z)</td>
</tr>
<tr>
<td>Asymp. Sig. (2-tailed)</td>
</tr>
</tbody>
</table>

\(a.\) Based on negative ranks.  
\(b.\) Based on positive ranks.  
\(c.\) Wilcoxon Signed Ranks Test

<table>
<thead>
<tr>
<th>Test Statistics(^c)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C2_PT - C1_PT</td>
</tr>
<tr>
<td>(Z)</td>
</tr>
<tr>
<td>Asymp. Sig. (2-tailed)</td>
</tr>
</tbody>
</table>

\(a.\) Based on negative ranks.  
\(b.\) Based on positive ranks.  
\(c.\) Wilcoxon Signed Ranks Test

<table>
<thead>
<tr>
<th>Test Statistics(^c)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C2_IT - C1_IT</td>
</tr>
<tr>
<td>(Z)</td>
</tr>
<tr>
<td>Asymp. Sig. (2-tailed)</td>
</tr>
</tbody>
</table>

\(a.\) Based on negative ranks.  
\(b.\) Based on positive ranks.  
\(c.\) Wilcoxon Signed Ranks Test
Figure 3 The graphic representation of the values obtained from quantitative analysis of sections of CD68-stained breast tumour tissue for the three types of regions of interest (IF, PT, IT) with the 5 suggested cut-offs.

Figure 4 Comparative graphic representation of the mean quantitative values obtained by quantitative analysis of the three types of regions, using two cut-offs (automated C1 and manual C2).
Considering the statistically significant differences found between the two cut-off types (C1 and C2), we present them graphically for comparison. (Figure 4)

Discussions

Recent investigations of the tumour micro-environment have shown that many tumours are richly infiltrated by a complex range of myeloid and lymphoid cells. During our current research we have found that immune cells appear as dense infiltrates in the tumour bed, and at the invasive front, and as lymphoid islands adjacent to the tumour (peritumoral), a fact which is also mentioned by other recent researches [7]. There is more and more evidence that the number, type and location of the myeloid-lymphoid infiltrates in the primary tumour have prognostic value, which leads to a new possible approach of “Immunoscore”. The “Immunoscores” can be used as prognostic factors [16]. The methods used for quantitative analysis in current research varied from the use of manual methods by different observers or the re-reading at different times by the same observer [17, 18], to the use of digital methods which are continuously improved [10, 19]. This proves the current interest for the quantification of the myeloid-lymphoid infiltrate using a quantitative analysis method. In our opinion, one should prefer the digital quantitative analysis performed by software which analyses the images and generate automated statistical results, with a minimal interference of the subjective human factor (unavoidable when manual quantification methods are used). Even if the digital quantitative analysis system can generate automated statistical results, the human factor interferes when placing the cut-off value in order to delimit positive and negative cell populations. The current study used the digital quantitative analysis system – HistoQuest version 3.5.3.0171 – which places the cut-off based on the analysis of a negative region. The negative region is defined as a tissue fragment where the IHC staining reveals no positive cells, and the cut-off is placed such as at most 1% of the evaluated cells are interpreted as falsely positive. This cut-off is propagated with the same value for all regions of interest from the same section. The quantitative analysis system performs an automated cut-off, but the cut-off values vary between regions from the same section. On many occasions, the obtained sections did not allow for identification of a negative region as control, which led to the suggestion of other cut-off placing methods such as to allow for an objective differentiation of the two cell populations (positive vs. negative cells).

We started from this idea when we proposed to compare the 5 types of cut-off described above. The data obtained by statistical analysis did not show statistically significant differences in the 5 cut-off types, based on CD8-staining. The statistical analysis studied the automated cut-off (C1) to which the other 4 cut-off types were related, based on the three regions of interest (IF, PT, IT): the statistical p-value varies between 0.189 – C1_Fi-C4_Fi and 0.937 – C1_IT-C5_IT, which are statistically insignificant values.

Statistically significant differences were found in the 5 cut-off types, based on the CD68-staining. The differences were found between C1 (automated cut-off) and C3 (manual cut-off, arithmetic mean) for the invasive front (p=0.034, statistically significant) and the peritumoral region (p=0.014, statistically significant), respectively. Additionally, differences were found between C1 (automated) and C2 (manual) for the intratumoral region (p=0.006, statistically significant). Comparing these data, we find that there are statistically significant differences between manual cut-offs C2 and C3 and automated cut-offs placed by the software (automated C1, C4 – negative region, and C5 – automated, arithmetic mean).
Considering that the statistical analysis for CD8-staining did not show any statistically significant differences in the cut-off types, while the CD68-staining shows differences between manual cut-offs and software-generated automated cut-offs, we propose the use of automated cut-offs in order to minimize the subjective factor. Considering the three types of automated cut-offs which generated statistically significant differences, i.e. C1 (automated, with a variable value for each region of interest), C4 (negative region, unique value) and C5 (arithmetic mean of all automated cut-offs, unique value subsequently applied to all graphic representations), we propose the use of C1 (automated cut-off generated by the software), as being the easiest to use. The quantitative analysis studies from specialized literature bring relatively few data on the optimisation of the quantitative analysis. A similar optimisation study showed that the automated data generated by the software are not different from those quantified manually by anatomical pathologists [20]. Studies which strictly refer to cut-off placing [21] are insufficient, which makes our study important for the optimisation of the method by generating statistically analysable objective data. It is scientifically demonstrated that the cut-off supported by the current study, i.e. the automated cut-off, generates objective data and minimizes human interference, which makes it adequate for use in subsequent research.

Conclusions

The present study demonstrates that the use of automated quantitative analysis of images is preferable because it minimizes human interference, hence the objectivity of data. An important step for the quantitative analysis is the adequate placing of the cut-off. The automated cut-off, although variable from one region to another, does not interfere with final data and is adequate for use in future research.

References

11. Chang H, Yang Q, Parvin B. Segmentation of heterogeneous blob objects through voting and level set for-


