Claudins as prognostic factors of breast carcinoma

Synopsis of Ph.D. Thesis

Dr. A. Marcell Szász

Semmelweis University,
School of Doctoral Studies

Supervisor:
Janina Kulka, M.D., Ph.D.

Critical Examiners:
Katalin Boér, M.D., Ph.D.
Béla Molnár, M.D., Ph.D. D.Sc.

President of the University Examination Committee:
György Bodoky, M.D., Ph.D.
Members of the University Examination Committee:
Péter Nagy, M.D., Ph.D.
András Lászik, M.D., Ph.D.

Budapest
2011
1. **INTRODUCTION**

Breast cancer is one of the most frequently diagnosed malignant tumors in women in Hungary and worldwide. Almost every third tumor is breast cancer. Every year approximately 1 million women are affected and newly diagnosed with breast cancer. The incidence was growing until the ‘90s, and a slow decrease is seen in the mortality since then, because of early diagnosis and improving clinical protocols.

The tumors recognised by organized screening or palpable masses may be diagnosed by biopsies according to recent guidelines. Depending of the latter results multidisciplinary team decides about the best therapeutical protocol lead by an oncologist. This work incorporates more specialists into the „oncoteam”, who need to take the patients conditions, pathological and molecular diagnostic results into consideration.

Prediction of prognosis has always been a major challenge for physicians treating tumorous patients. Majority of breast carcinomas originate from epithelial tissue. The time of tumor formation is of prognostic importance, younger patients have poorer outcomes. The prognosis of malignant tumors can be predicted by tumor size, existence or non-existence of lymph node metastases. Histological appearance of tumors provides information for identification of histological grade, vascular invasion, axillary lymph node status, Nottingham Prognostic Index which are all used in the everyday routine work with their limitations. The clinical and pathological stage at time of diagnosis is an accepted basis for administration of therapy. Further refinements can be supported by mitotic index, S-phase fraction, and Ki-67 expression.

Expression of hormone receptors has gained predictive importance in estrogen receptor positive (ER+) and negative (ER-) breast tumors. The biological behavior of these subgroups can be quite different, supported by gene expression analyses in the past decade. Perou et al identified breast cancer subtypes based on cDNA microarray studies, which is the cornerstone of state-of-the-art treatment in breast cancer.

Mutations in several genes (p53, BRCA1/2, CHEK2, KIP) have gained causative importance in carcinogenesis and prediction of prognosis, but with the progression of personalised medicine and development of genomical methods further tests may support clinical decision making: Agendia Mammaprint, Genomic Health Oncotype DX, Ipsogen Mapquant DX. The main drawback in the application of these complicated assays is double: high prices and requirement of frozen tissue are compromising their wide use (Table 1).
Table 1. Various methods for prediction of prognosis and therapy response in breast cancer.

<table>
<thead>
<tr>
<th>Technique</th>
<th>Method</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sequencing (DNA)</strong></td>
<td>BRCA1/2</td>
</tr>
<tr>
<td><strong>FISH (DNA)</strong></td>
<td>HER2</td>
</tr>
<tr>
<td></td>
<td>3-gene(CYP24, PDCD6IP, and BIRC5)</td>
</tr>
<tr>
<td><strong>Expresszio Array (RNA)</strong></td>
<td>70-gene Amsterdam</td>
</tr>
<tr>
<td></td>
<td>76-gene Rotterdam</td>
</tr>
<tr>
<td></td>
<td>97-gene Genomic Grade Index</td>
</tr>
<tr>
<td></td>
<td>163-gene Stroma-derived Prognostic Predictor</td>
</tr>
<tr>
<td></td>
<td>459-gene Wound-response signature</td>
</tr>
<tr>
<td><strong>qPCR (RNA)</strong></td>
<td>21-gene (Recurrence Score)</td>
</tr>
<tr>
<td></td>
<td>4-gene (MYBL2, KPNA2, CDC2, CDC20)</td>
</tr>
<tr>
<td></td>
<td>5-gene (HOXB13, IL17BR, CHDH, MIB1, MKI67)</td>
</tr>
<tr>
<td></td>
<td>8-gene (TOPFOX)</td>
</tr>
<tr>
<td></td>
<td>14-gene (Metastasis Test)</td>
</tr>
<tr>
<td><strong>Nanostring (RNA)</strong></td>
<td>50-gene (PAM50)</td>
</tr>
<tr>
<td><strong>IHC (Protein)</strong></td>
<td>HER2</td>
</tr>
<tr>
<td></td>
<td>Ki67</td>
</tr>
<tr>
<td></td>
<td>IHC4 (ER, PgR, HER2, Ki67)</td>
</tr>
<tr>
<td></td>
<td>5-marker (p53, NDRG1, CEACAM5, SLC7A5, HTF9C)</td>
</tr>
</tbody>
</table>

Importance of cell-cell, cell-matrix junctions is underlined by the fact that 10% of all the genetical substance codes junctional molecules, which are of fundamental importance in the development and maintenance of the organization. In epithelial cells the tight junctions are situated at the apical/basolateral regions, playing primary role in sealing of intercellular gaps, and controlling paracellular selective diffusion ("gate function"). These structures also maintain the different protein and lipid composition between the apical and basolateral regions ("fence function"). Furthermore, they have been found to control cell growth and differentiation.

There is a lack of in depth investigations of junctional proteins and their relationship to adhesional molecules in the literature. Furthermore, the expression of junctional proteins was rarely investigated in regional lymph node and distant metastases of solid tumors.

Claudins were described in 1998, these molecules possess four trans membrane domain. Their function has not been fully elucidated, mostly claudins 1, 2, 3, 4, 5 and 7 have been investigated.

Their description in breast tissues has subsequently started after their exploration. Kramer et al evaluated 96 breast samples in 2000, and found claudin-1 protein expression to be decreased in carcinomas. Kominsky investigated the expression of claudin-3 and -4 in 2004 in 21 breast tumors, and described the elevated expression of both in carcinomas and
also emphasized the possible role of Clostridium perfringens enterotoxin in the therapy of breast cancer, as claudin-4 is the receptor of this toxin. Tőkés et al has investigated 56 breast samples both RNA and protein expression level for claudin-1, -3 and -4, and found that their expression is correlated with the differential state of the tumor. Hewitt et al evaluated 47 breast tumors in 2006, and provided the description of claudin expression pattern confirming the elevated expression of claudin-3 and -4, and a slight increase in the expression of claudin-7.

While the purely descriptive investigations have taken place, there were several workgroups after the causative role of this family of proteins. Elevated expression of claudin-2 was found to promote development of liver metastases. Herschkowitz was the first to identify the emerging “claudin-low” subtype among triple negative breast tumors in 2007, which subclass proved to be resistant to hormonal and chemotherapy to date. The investigation of claudins from the prognostic aspect has not been performed throughoutly, and there are only preliminary results with a lower number of cases published.

Cadherins are members contributing to the structure of the adherens junctions, all of them function in a Ca-dependent manner. They can be classified according to the organic location of their expression. In breast tissue, E-cadherin is expressed and its mutation or loss has been described in carcinogenesis and tumor progression. In the routine pathological setting it is used to distinguish ductal vs. lobular breast carcinomas. Its role has been identified in the process of epithelial-to-mesenchymal transition, where the decreased expression or loss of E-cadherin contributes to the mobility and metastatic spread of cancer cells.
2. AIMS AND OBJECTIVES

1. Does the expression of claudins and E-cadherin differ in different histological types of breast carcinomas at protein level?
2. Do the expression of claudins differ at levels of dedifferentiation (grade) in breast tumors at protein level?
3. Does the expression of claudins and E-cadherin differ at protein level in the various breast cancer subtypes as reflected by the routine immunohistochemical profile?
4. Does the expression of claudins differ in primary breast tumors and their consecutive lymph node metastases at protein level?
5. Does the expression of claudins and E-cadherin either alone or together as signatures provide prognostic information at mRNA level when publicly available datasets are investigated?
6. Does the expression of claudins and E-cadherin provide prognostic information at protein level in primary breast tumors?
7. The expression of claudins in lymph node metastases does provide any prognostic information regarding outcome of breast cancer?
8. Can the index of resulting prognostic claudin profiles from the investigation of genomical and training sets be validated in an independent cohort of breast carcinomas?
9. Does the expression index of claudin-4 and E-cadherin provide independent prognostic information when compared to routinely used clinicopathological factors?
3. METHODS

*Patients and datasets.* The study was approved by the Institutional Review Board of the Semmelweis University (#85/2007, #185/2007, #7/2008). A training cohort of 249 selected breast carcinomas diagnosed between 1982 and 2007 with known relapse and cases with no recurrence of disease on long-term follow-up was set up. Selected cases of formalin-fixed and paraffin-embedded tissues underwent histological re-examination, and regions of interests were marked for core punching. Expression of claudin-1, -2, -3, -4, -5, -7 and E-cadherin proteins were evaluated by immunohistochemistry. The validation cohort consisted of 387 unselected breast cancer samples diagnosed between 1999 and 2002. The samples have been analyzed for claudin-4 and E-cadherin protein expression in the same fashion as the training set.

An *in silico* analysis of publicly available microarray data from 13 datasets was performed. The refined and established database was based on the gene expression data (Affymetrix HGU133A and HGU133+2 microarrays only) and survival information of 1809 patients downloaded from Gene Expression Omnibus database. Expression data of CDH1, CTNNB1, CLDN1, CLDN3, CLDN4, CLDN5, CLDN6, CLDN7, CLDN8, CLDN9, CLDN10, CLDN11, CLDN14, CLDN15, CLDN16, CLDN17, CLDN18 was analyzed. All available probes were used for the analysis, but those were excluded which were not reliable across datasets.

*Tissue microarray (TMA) construction.* TMAs of all FFPE tumors were composed with the Tissue Micro-Array Builder instrument (Histopathology Ltd., Pécs, Hungary). The cores were 2.0 mm in diameter and all the samples were investigated in duplicates.

*Immunohistochemistry.* The IHC reactions were performed on 4 µm thick sections obtained from the TMA blocks. After deparaffination, the slides were heated in a microwave oven in retrieval solution. The automated Ventana ES immunostainer system was used according to the protocol provided by the manufacturer. The E-cadherin and claudin antibodies from Zymed, Inc. and DAKO were used. Positive controls and negative control tissues were applied in all IHC runs.

*IHC evaluation.* The immunohistochemical reactions were considered positive when the claudins (except claudin-2, which was granular cytoplasmic, predominantly submembranous), and E-cadherin were expressed at the cell membranes. Claudin-5 was not
only expressed in the tumor cells but in the endothelial cells as well. The stained slides were
digitalized with a slide scanner (Mirax Scan), and intensity of the reaction (0 - negative
reaction, +1 - weak positivity, +2 - moderate positivity, +3 - strong reaction) and frequency of
positive cells (1: 0-5%, 2: >5-10%, 3: >10-20%, 4: >20-33%, 5: >33-50%, 6: >50-66%, 7:
>66-80%, 8: >80-100%) were integrated in a scoring system using the TMA module software
(v.11.49) resulting in a numerical variable (ranging 0-11) consisting of the sum of the
intensity and frequency scores above (Figure 1). The indices were created from the arithmetic
means of the respective claudins and E-cadherin.

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Intensity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unsuccessful(-1)</td>
<td>+1 (1)</td>
</tr>
<tr>
<td>Negative (0)</td>
<td>+2 (2)</td>
</tr>
<tr>
<td>0-1% (1)</td>
<td>+3 (3)</td>
</tr>
<tr>
<td>1-5% (2)</td>
<td></td>
</tr>
<tr>
<td>5-10% (3)</td>
<td></td>
</tr>
<tr>
<td>10-33% (4)</td>
<td></td>
</tr>
<tr>
<td>33-50% (5)</td>
<td></td>
</tr>
<tr>
<td>50-66% (6)</td>
<td></td>
</tr>
<tr>
<td>66-80% (7)</td>
<td></td>
</tr>
<tr>
<td>80-100% (8)</td>
<td></td>
</tr>
</tbody>
</table>

*Figure 1. The expression of investigated proteins was analyzed according to their frequency and intensity in a two dimensional scale. Unsuccessful cases and negative cases were also annotated.*

*Fluorescent in situ hybridization (FISH).* FISH was performed to stratify breast
carcinomas into subclasses of HER2 amplified cases vs. others. The HER2 and CE17 probes
were utilized from the Poseidon kit (Kreatech).

*RNA purification and cDNA transcription.* RNAs were isolated with the Qiagen
RNaseasy FFPE kit according to the protocol provided by the manufacturer. Nanodrop ND-
1000 system was used to quantify amounts of RNAs. For the transcription 1 µg RNA was
used which was performed using the High Capacity RNA-to-cDNA Transcription kit.

*Quantitative real-time polymerase chain reaction (PCR).* The PCR reactions were run
in 25 µl volume using ABI Power SYBR® Green PCR Master Mix in ABI 7000 PCR system.
The expression of CLDN4 was measured at mRNA level normalized to the GAPDH97
housekeeping gene.
Statistics. The analyses for the datasets were performed in the R statistical environment (v.2.10.1) using the Prediction Analysis of Microarrays (PAM, v.2.19). PAM uses soft thresholding to produce a shrunken centroid, which allows the selection of genes with high predictive potential. This made it possible to derive the simplified junctional signature from the complex juntional profile including all investigated mRNA transcripts. The data analysis and statistical evaluation of tissue samples was performed with SPSS 15.0 Family Pack. In the training cohort, all proteins were included in the first analysis testing the complex junctional profile. The latter score was reached when considering all scores of all proteins and taking their arithmetic average as a numerical variable. All individual values of the distinct scores of IHC were included in the statistical evaluation. Mann-Whitney U test was performed for pair-wise statistical analysis. Hierarchical clustering was used to rank the protein expression data according to their contribution to the identification of prognostic groups. Similarly to the gene expression signature, this helped to define those proteins which were able to define the prognostic subgroups by themselves. Combining the overlapping members of the gene and protein expression methods resulted in the Claudin-4 and E-cadherin score (CC index) which was tested to predict survival in the validation cohort. The contribution of expression data to predict relapse-free survival, distant metastasis-free survival and overall survival was displayed according to Kaplan-Meier method, statistically supported with log-rank test. For multivariate analysis Cox regression method was applied. ‘p’ values of less than 0.05 were considered to be statistically significant.
4. Results

We were able to describe the following new results based on the investigations:

1. Expression of claudins differ in the lobular (ILC) and ductal type (IDC) of carcinomas. Claudin-1 expression was lower in IDCs as compared to the ILCs. Analysis yielded significance between the ductal and the lobular carcinoma group considering claudin-1 and -2 expression: both were lower in the ductal carcinoma group. As expected, E-cadherin showed expression mainly in ductal type of carcinomas and less expression was seen in the majority of lobular carcinomas.

2. The expression of claudin-1 showed decrease with increasing grade, as opposed to claudin-4 and -7 which was expressed more as grade increased in tumors.

Figure 2. Claudins and E-cadherin expressed at different levels in the phenotypical subgroups reflected by IHC. Arrows show significantly differing protein expression levels.
3. Expression of claudin-1 was lowest in the Luminal B subtype, and claudin-4 was less expressed in Luminal A cancers. E-cadherin showed highest expression among HER2+ tumors (Figure 2). Regarding TNBCs, there was no single marker distinguishing them from other subgroups.

4. The claudin expression of lymph node metastases differs significantly of those expressed by the primary breast tumors at protein level. Claudin-5 is likely to be elevated and claudin-3,-4 and -7 decreased in lymph node metastatic cancer cells.

5. Based on the investigation of public datasets, claudins and E-cadherin in themselves provide prognostic information at mRNA level, which can be further strengthened by an identified robust signature of claudin-3, -4, -7 and E-cadherin).

6. Based on the investigation of the training cohort, claudins and E-cadherin in themselves provide prognostic information at protein level, which can be further strengthened by an identified robust expression pattern of claudin-2, -4 and E-cadherin.

7. Expression of claudin in lymph node metastases provides prognostic information on outcome in breast carcinomas: lower expression of claudin-1, and higher expression of claudin-4 is reflecting poor prognosis (Figure 3).

8. The CC index which was resulting from the prognostic analysis of claudin and E-cadherin expression at mRNA level in public datasets and at protein level in the training set: both univariate and multivariate tests were passed successfully (Figure 4, Table 2).

9. The expression of claudin-1 and -4 and E-cadherin thus provides further prognostic information besides the routinely utilized clinicopathological parameters.
Figure 4. Kaplan-Meier curve showing the distinguished groups according to expression of the claudin-4/E-cadherin score (CC index) in the tissue samples of validation cohort.

Table 2. Results of Cox multivariate analysis with hazard ratio and 95% confidence interval.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>p</th>
<th>HR</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>lower</td>
<td>upper</td>
</tr>
<tr>
<td><strong>CC index</strong></td>
<td>0.040</td>
<td>1.503</td>
<td>1.02</td>
</tr>
<tr>
<td><strong>Immunophenotype</strong></td>
<td>0.001</td>
<td>1.064</td>
<td>0.59</td>
</tr>
<tr>
<td><strong>vascular invasion</strong></td>
<td>0.011</td>
<td>0.704</td>
<td>0.31</td>
</tr>
<tr>
<td><strong>Lymph node metastases</strong></td>
<td>0.091</td>
<td>0.791</td>
<td>0.6</td>
</tr>
<tr>
<td><strong>Grade</strong></td>
<td>0.767</td>
<td>1.037</td>
<td>0.82</td>
</tr>
<tr>
<td><strong>Necrosis</strong></td>
<td>0.232</td>
<td>2.233</td>
<td>0.72</td>
</tr>
</tbody>
</table>
5. Conclusions

Expression of claudins and E-cadherin changes in numerous pathological states especially malignancies. Their role has been implicated in tumors of the breast, gastrointestinal, gynecological and urological organs.

In this series of investigations we were able to demonstrate that the expression of claudins provides prognostic information. Based on bioinformatic analysis the expression of juntional molecules may provide the basis for further prognostic classification especially using them in a combinatorial manner. Unique expression and the signature of claudin-3, -4, -7 and E-cadherin was identified to show the most powerful predictive potential compared to other variables. This profile may be validated at protein level utilizing immunohistochemistry, although, the markers only show a partial overlap. In formalin-fixed, paraffin-embeded tissues, protein expression of claudin-2, -4 and E-cadherin has proven to be of utmost prognostic importance. Thus, the overlapping members of the bioinformatic and the tissue-based evaluation (CC index corresponding to the profile integrating the expression of claudin-4 and E-cadherin) were analyzed for their prognostic potential in an independent validation set. The CC index has proven to be a powerful prognostic predictor in univariate and multivariate analysis besides the immunophenotypical characterization and presence of vascular invasion. Furthermore, the CC index was able to refine prognosis in the molecular subtypes of breast cancer reflected by immunohistochemical characterization. The claudin-cadherin index can be used in a routine laboratory setting to further classify breast tumors according to clinical outcome.

The identification of „claudin-low” breast tumors warrants investigations e.g. considering stem cell markers in a larger set of primary breast carcinomas.

The expression of claudins in lymph node metastases of breast tumors also provides additional prognostic information, lower expression of claudin-1 and higher expression of claudin-4 is associated with poor prognosis in breast tumors. Investigations in paired primary tumors and their consecutive distant metastases is ongoing in numerous workgroups, which may provide further data for understanding breast cancer and may help refine prognosis in metastatic disease.
PUBLICATIONS RELATED TO THE DISSERTATION


QUOTABLE ABSTRACTS RELATED TO THE DISSERTATION


PUBLICATIONS NOT RELATED TO THE DISSERTATION


18) Baranyai Zs, Josa V, Szasz AM: A műtői hatékonyság javítása (Improving operating room efficiency). IME 2009;9:15-21


