Tissue microarray in prognostic studies on vulva cancer
Fons, G.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Chapter 8

General discussion and conclusions
The studies described in this thesis aim at improving the prognostic accuracy in patients with vulva cancer, by the assessment of a possible additional predictive value of protein markers and by the re-examination of well established clinical pathologic parameters. Improving the prognostic accuracy is of considerable importance as it will benefit the selection of the most appropriate treatment modality and will provide patients with more reliable information on the course of their disease and their future.

In our study on protein markers we were not able to demonstrate an additional value of 16 protein markers for the assessment of the prognosis of vulva cancer patients. We tested initially a panel of 16 protein markers. The prognostic significance of 10 markers had been described in the literature as mentioned in the introduction. We were unable to confirm these results.

Results of our first study would suggest the period of survival to be associated with COX-2 over-expression and absent Caspase 3 expression, as described in chapter 2. In the internal validation study, described in chapter 4, the association between COX-2 over-expression and period of survival was confirmed. We could not confirm the association between absent Caspase 3 expression and period of survival. In the external validation study, described in chapter 5, an association between the period of survival and COX-2 over-expression could not be demonstrated in a similar and independent study group of vulva cancer patients.

At first sight, this disappointing outcome contrasts with the promising results of other studies. In the analysis of this apparent discrepancy we addressed two questions. First, did methodological or technical errors in our study contribute to this result? Second, is it possible that results of other researchers are less promising than they appeared to be?

One of the methodological imperfections in our study is the limited statistical power, particularly in the first study. This is a direct consequence of the low incidence of vulva cancer. The first study group comprised 50 patients, treated between 1995 and 1998. Immunohistochemical staining was assessed using a three-point scale to define intensity and distribution. Cut-off levels were chosen depending on the discriminative power of the categories in the first study group. Another population is likely to have a best cut-off point at a different level, and thus worse test results, when the initial cut-off point is re-used. (1)

The relatively small sample size also constitutes a threat to the proper use of the TMA. The TMA approach has been designed to survey tumor populations and not to examine individual tumors.(2) It cannot be avoided that some cellular abnormalities are missed when the analysis of potentially heterogeneous tumors is restricted to a small number of samples per patient measuring only 0.6 mm in diameter. Although the agreement
between expression-patterns in triplicate 0.6 mm core biopsies on TMA and on full sections was generally good, as described in chapter 3, it was highly dependent on the antigen tested and the scoring system used. For Caspase 3 a kappa of 0.4 (slight agreement) was found. This must have had an adverse effect on the reproducibility of the association between absent Caspase 3 expression and survival, as found in the first study. Furthermore a small sample size will be of influence particularly in the assessment of antigens that come to expression focally and in a small number of cells, as is the case of Caspase 3.

The small number of patients available for the study limits the type of statistical tests that can be applied. Multivariate analysis was performed with a proportional hazard model. Another, possibly more promising approach would have been a hierarchical cluster analysis. Cluster analysis organizes cases according to the similarity or dissimilarity of immunostaining profiles, placing the cases with similar immuno-profiles together as neighboring rows in the clustergram. It has been previously observed that the potential of combinations of prognostic markers is superior to that of a single marker. The size of our data set unfortunately did not allow a hierarchical cluster analysis.

Many factors will influence the technical preparation of the immunoreactions and a wide variety of causes could result in incorrect interpretations. Variables affecting immunoassaying may involve different steps of the routine technical tissue processing. The fixation, dehydration and embedding processes, or external factors, such as the influence of storage time and temperature, causing degradation of antigen may be of influence on the final result.

The TMA technology allows the simultaneous analysis of a large number of tumors under strictly standardized conditions. Most of the limitations of immunohistochemical tests previously mentioned have been overcome by this technique. The TMA used for external validation was constructed in the laboratory of the department of pathology of the University Medical Centre Groningen. As external validation of the association between COX-2 over-expression and survival failed, an analysis was made of the causative factors possibly related to the technical process. A difference in time-interval between the construction of the TMA and the immunostaining was established. Several other studies demonstrated a deleterious effect of storing pre-cut slides at room temperature. This phenomenon may, to some extent, have contributed to the lack of reproducibility.

Although the prognostic value of COX-2 expression is not extensively tested in vulva cancer, it is studied in squamous cell cancer of the esophagus. As discussed in chapter 5, the association between prognosis of patients with esophagus cancer and COX-2 expression is described in thirteen different papers. There is a considerable difference
in the cut-off levels used, resulting in variability of outcome. As far as we can ascertain, these results have not been internally or externally validated. The need for validation of diagnostic tests is widely recognized.(8-10) A review article on the validation of diagnostic accuracy, reported most studies to have a worse accuracy than the original one.(1)
Positive results of immunohistochemical tests in particular have to be judged with great caution before completion of the required validation studies. It may well be that the prognostic significance of COX-2 over-expression, in patients with esophagus cancer will not stand after external validation of test results.

Although we did not succeed in demonstrating additional predictive value of any protein marker it is too early to conclude that further efforts in this direction have to be abandoned. The use of TMA is reliable, provided that the study group is large enough. Testing with a large number of protein markers is preferred to testing with single markers as it allows the application of hierarchical clustering. This could well be a promising approach as tumors with different biological behavior are more likely characterized by a group of markers than by a single one.

As long as no additional markers with discriminative potential are found, prognostic classification has to be based on long established clinical pathologic markers.
The current FIGO classification discriminates 4 stages depending on the size of the tumor, the existence of uni-or bilateral groin lymph node metastases and the local extension of tumor.

Tumors with a diameter of 2 cm or less are classified as stage I, tumors with a diameter of more than 2 cm are classified as stage II, as along as there are no lymph node metastases and no local extension into surrounding anatomical structures as urethra, vagina or anus. As stated in the introduction this classification has its limitations as 30% of the patients without lymph node metastases will get a recurrence in 10 years of which 50% will finally die from the disease. It is good to recall that a scientific basis for this cut-off level of 2 cm is lacking. In the largest prospective study on prognostic factors in vulva cancer, including 588 patients, the 5-years survival of patients with lesions up to 8 cm in diameter was not significantly shorter than the survival of those with lesions smaller than 2 cm. (11) This study proposed a cut-off level of 8 cm.
As lymph node status is the most powerful predictor of survival it is not very likely that other independent clinical pathologic risk factors can be identified in study groups comprising patients with and without lymph node metastases. There is no information on useful prognostic factors for patients without lymph node metastases. It is possible
that protein markers will eventually contribute to a better prognostic classification in this particular group of patients.

FIGO stage III includes tumors that involves the lower urethra, vagina or anus (T3) and/or have metastasized unilaterally to groin lymph nodes. Stage IVA entails both patients with bilateral lymph node metastases (irrespective of local tumor extension) and patients without lymph node metastases but with extension into the upper urethra, vagina, bladder mucosa, rectum or pelvic bones (T4). It is not clear from the literature whether the variable ‘bilateral lymph node metastases’ has independent prognostic value or just reflects the presence of multiple lymph node metastases. The existence of bilateral lymph node metastases would appear to lose its prognostic significance when correction is made for the number of positive nodes, as is shown in chapter 6. We therefore suggest removing single or dual laterality of lymph node metastases from the FIGO staging system.

When the limited data on prognostic significance of lymph node associated factors are reviewed extra-capsular spread turns out to be strongly associated with impaired survival. This is in accordance with our conclusion in the study on prognostic significance of lymph node associated factors as described in chapter 6. The importance of extra-capsular spread for the period of survival justifies its inclusion in the staging system. However, this introduction is dependent on acceptance of an internationally recognized definition of extra-capsular spread.

Limited data are available on the independent prognostic significance of T-status. A recent study on patients with stage III vulva cancer convincingly showed that patients with T3 tumors have a significantly better prognosis than patients with stage III disease based on lymph node metastases.(12) These authors did not address the question of a possible influence of T-status on the prognosis. This answer has to be provided by comparing the prognosis of patients with a different T-status but without lymph node metastases. Answering this question is even more difficult as surgery is generally not the standard treatment for patients with T4 tumors. Consequently detailed exploration does not take place. This limits the evaluation of the data of these patients and it is an argument to assign patients with T4 tumors to a separate category.

It is recognized that patients with one lymph node metastasis have a worse prognosis than patients without lymph node metastases. It is generally accepted to treat all patients with more than one positive node with adjuvant radiotherapy. At this moment, consensus is lacking on the treatment of choice for those patients with only one lymph node metastasis. To contribute to the solution of this question, we analyzed the data of patients from three different hospitals with vulva cancer and just one lymph node
metastasis and no extra-capsular spread. As stated in the introduction, previous studies recommending adjuvant radiotherapy did not take extra-capsular spread (the most important lymph node associated prognostic factor) sufficiently into consideration. (13) We selected for our study, described in chapter 7, patients with only one lymph node, without extra-capsular spread, to rule out its predominant effect. We could not demonstrate a beneficial effect of adjuvant radiotherapy for this selected group.

In summary, sufficient data are produced to necessitate partial revision of the current FIGO system. Particularly, the suggested revision of stage III and IV will result in improved terms of discrimination. As clinical pathologic factors with enough discriminative power are lacking in the group of patients without lymph node metastases, changes in the classification of stage I and II can not be expected to contribute to an improvement of discrimination between patients with a high and low risk for recurrence and death of disease. Protein markers, which reflect the biological behavior of the tumor, may have additional value in this group of patients. When a group of patients with a specific risk for an adverse outcome has been identified, therapy aimed at reducing this risk must not be implemented before conclusive evidence of the beneficial effect has been provided.
References

(1) van den Bruel A, Aertgeerts B, Buntinx F. Results of diagnostic accuracy studies are not always validated. J Clin Epidemiol 2006; 59(6):559-566.


