New models to predict depth of infiltration in endometrial carcinoma based on transvaginal sonography

Frank De Smet[1], Ignace Vergote[3,*], Jos De Brabanter[1,2], Maja L. Konstantinovic[3], Thierry Van den Bosch[3], Philippe Moerman[4], Bart De Moor[1], Frederic Amant[3], Dirk Timmerman[3]

[1]Department of Electrical Engineering ESAT-SCD, Katholieke Universiteit Leuven, Kasteelpark Arenberg 10, B-3001 Leuven-Heverlee, Belgium.

[2]Sint-Lieven hogeschool, Departement industrieel ingenieur (Associatie K.U.Leuven), Gebroeders Desmetstraat 1, B-9000 Gent, Belgium.

[3]Department of Obstetrics and Gynecology, University Hospitals, Katholieke Universiteit Leuven, Herestraat 49, B-3000 Leuven, Belgium.

[4]Department of Pathology, University Hospitals, Katholieke Universiteit Leuven, Herestraat 49, B-3000 Leuven, Belgium.

[*]Presenting author

Objective: We assessed histopathological and ultrasound parameters obtained at transvaginal sonography with color Doppler imaging (CDI) for the prediction of the depth of myometrial invasion in 175 patients with endometrial cancer.

Methods: We collected data from 97 consecutive women and divided them into two groups (group I: stage Ia and Ib – group II: stage Ic and higher). With stepwise logistic regression we examined which variables significantly contributed in a logistic regression model. We used these variables to train a logistic regression model and Least Squares Support Vector Machines (LS-SVM) with linear and RBF (Radial Basis Function) kernels. Finally, these models were prospectively validated on a second set of 78 new patients.

05-140

De Smet F., Vergote I., De Brabanter J., Konstantinovic M.L., Van den Bosch T., Moerman P., De Moor B., Amant F., Timmerman D., ``New models to predict depth of infiltration in endometrial carcinoma based on transvaginal sonography", in *International Journal of Gynecological Cancer (14th International meeting of the ESGO) (ESGO 14)*, Istanbul, Turkey, Sep. 2005, 2 p., Lirias number: 181559.

Results: The ratio of the endometrial and uterine volume had the largest AUC (78%) from all ultrasound parameters, which was smaller than the AUC (79%) of the subjective assessment of an experienced gynaecological ultrasonographer. The AUCs of the parameters obtained after CDI were low. Stepwise logistic regression selected degree of differentiation, number of fibroids, endometrial thickness and volume. Compared with the AUC (72%) of the subjective assessment, prospective evaluation of the models only resulted in a higher AUC (77%) for the LS-SVM with an RBF kernel.

Conclusions: CDI does not contribute in the prediction of the surgical stage of endometrial cancer and single morphological parameters perform insufficiently. An LS-SVM model with an RBF kernel gives the best predictions and seems more reliable than the subjective assessment of an experienced ultrasonographer.

New models to predict depth of infiltration in endometrial carcinoma based on transvaginal sonography

Frank De Smet, Jos De Brabanter, Thierry Van den Bosch, Nathalie Pochet, Frédéric Amant, Caroline Van Holsbeke, Philippe Moerman, Bart De Moor, Ignace Vergote, Dirk Timmerman

Department of Electrical Engineering ESAT-SCD, Katholieke Universiteit Leuven, Kasteelpark Arenberg 10, B-3001 Leuven-Heverlee, Belgium (F De Smet, MD, MscEng, PhD; J De Brabanter, MscEng, PhD; N Pochet, Msc; B De Moor, MscEng, PhD), Medical Direction, Alliance of Christian Sickness Funds, Haachtsesteenweg 579, B-1031 Brussel, Belgium (F De Smet, MD, MscEng, PhD), Sint-Lieven hogeschool, Departement industrieel ingenieur (Associatie K.U.Leuven), Gebroeders Desmetstraat 1, B-9000 Gent, Belgium (J De Brabanter, MscEng, PhD), and Departments of Obstetrics and Gynecology (T Van den Bosch, MD; Frédéric Amant, MD, PhD; Caroline Van Holsbeke, MD; I. Vergote, MD, PhD; D Timmerman, MD, PhD) and Pathology (Ph Moerman, MD, PhD), University Hospitals, Katholieke Universiteit Leuven, Herestraat 49, B-3000 Leuven, Belgium.

Corresponding author:

Prof. Dirk Timmerman

Department of Obstetrics and Gynaecology

University Hospital Gasthuisberg, Katholieke Universiteit Leuven

Herestraat 49, B-3000 Leuven, Belgium

Tel 32-16-344204 – Fax 32-16-344205 – E-mail: dirk.timmerman@uz.kuleuven.ac.be

Running head: Depth of infiltration in endometrial carcinoma

Abstract

Objectives: Preoperative knowledge of the depth of myometrial infiltration is important in patients with endometrial carcinoma. This study aimed at assessing the value of histopathological parameters obtained from an endometrial biopsy (Pipelle® de Cornier – results available preoperatively) and ultrasound measurements obtained after transvaginal sonography (TVS) with color Doppler imaging (CDI) in the preoperative prediction of the depth of myometrial invasion as determined by the final histopathological examination of the hysterectomy specimen (golden standard).

Methods: We first collected ultrasound and histopathological data from 97 consecutive women with endometrial carcinoma and divided them into two groups according to the surgical stage (stage Ia and Ib versus stage Ic and higher). The area (AUC) under the Receiver Operating Characteristic (ROC) curve of the subjective assessment of depth of invasion by an experienced gynecologist and of the individual ultrasound parameters were calculated. Subsequently we used these variables to train a logistic regression model and Least Squares Support Vector Machines (LS-SVM) with linear and RBF (Radial Basis Function) kernels. Finally, these models were prospectively validated on data from 78 new patients in order to make a preoperative prediction of the depth of invasion.

Results: The ratio of the endometrial and uterine volume (EV/UV) had the largest AUC (78%) from all ultrasound parameters, while the AUC of the subjective assessment was 79%. The AUCs of the blood flow indices were low (range 51%-64%). Stepwise logistic regression selected the degree of differentiation, the number of fibroids, the endometrial thickness and volume of the tumor. Compared with the AUC (72%) of the subjective assessment, prospective evaluation of the mathematical models resulted in a higher AUC

(77%) for the LS-SVM model with an RBF kernel, but this difference was not significant. *Conclusions:* Single morphological parameters do not improve the predictive power when compared to the subjective assessment of depth of myometrial invasion of endometrial cancer and blood flow indices do not contribute to the prediction of the stage. In this study an LS-SVM model with an RBF kernel gives the best prediction and might be more reliable than the subjective assessment but this has to be confirmed by larger prospective studies.

Key words: endometrial carcinoma, transvaginal sonography, color Doppler imaging, staging, logistic regression, least squares support vector machines.

Introduction

Carcinoma of the endometrium is the most common female pelvic malignancy [1]. Initial preoperative evaluation of patients suspected of having a carcinoma of the endometrium includes a transvaginal sonography (TVS) with or without color Doppler imaging (CDI) and endometrial biopsy.

The distinction between FIGO surgical stage Ib and Ic endometrial carcinoma (assessed postoperatively) is determined by the degree of myometrial invasion (Ib is less and Ic is more than 50% invasion).[2] This is an important prognostic factor,[3] that determines the treatment protocol in many institutions. Accurate preoperative distinction between patients with stage Ia or Ib disease and patients with stage Ic or higher would allow identification of high-risk patients who might need pelvic lymphadenectomy. This might be important as in many countries, patients who will need lymphadenectomy would be referred to a gynecologic oncologist, while patients not requiring lymphadenectomy would be operated by a general gynecologist or surgeon.

Several techniques are used to estimate the depth of myometrial invasion, but all have specific limitations. Intraoperative gross visual inspection or frozen section do not allow preoperative planning of the surgical procedure. Franchi et al.[4] reported an accuracy of 85.3% in predicting the degree of myometrial invasion in a series of 403 patients using intraoperative gross visual inspection, whereas Kucera et al.[5] found an accuracy of 88% at frozen section in a combined set of 624 patients. The contrast-enhanced magnetic resonance imaging (MRI) being the most reliable method. In a meta-analysis, Kinkel et al.[6] reported an area under the Receiver Operating Characteristic (ROC) curve (AUC) of 91% with respect to the prediction of myometrial invasion.

However, MRI is costly, has a limited availability and is not appropriate for all patients (e.g., claustrophobia, obesity, and contrast allergies). Different groups[7-17] have studied the value of TVS and CDI using different morphological or CDI parameters with a considerable variation in the results. Arko et al[18] published one of the largest series that investigated the use of TVS to estimate the depth of myometrial invasion in 120 patients. This study reported an accuracy of 73% in predicting myometrial invasion.

In the present study on patients with endometrial carcinoma, we analyzed ultrasound measurements obtained from TVS with CDI and histopathological data, obtained from preoperative endometrial biopsy (Pipelle® de Cornier)). We then explored whether they contribute in predicting myometrial invasion as assessed postoperatively by the final histopathological examination (golden standard). Moreover, we aimed to construct models that predict the presence of deep myometrial invasion, which could help the clinician to preoperatively identify patients that might need more extensive surgery.

Patients and Methods

We first collected data from 97 consecutive patients with endometrial carcinoma, who were scanned between September 1994 and February 2000 by a single operator (DT) [19]. Further on, we will refer to these patients as the training set. The mean age of these patients was 65.9 years (range 45-83) with 88 women being postmenopausal. The distribution of the different surgical FIGO stages was as follows: 24 stage Ia, 35 Ib, 12 Ic, 8 II, 13 III and 5 IV. The histopathological subtypes were: 76 endometrioid adenocarcinoma, 3 serous papillary, and 18 mixed type (5 with a clear cell and 3 with a serous papillary component). Fifty-four tumors were highly, 18 moderately and 25 were poorly differentiated. Tumors with a serous papillary or a clear cell component were considered to be poorly differentiated. All patients gave informed consent and underwent preoperative ultrasound examination with TVS and CDI in the department of Obstetrics and Gynecology (University Hospitals Leuven) using the same protocol. The uterus was assessed both in sagittal and coronal planes with an Acuson Sequoia ultrasound system (Siemens-Acuson Inc., Mountain View, CA, USA), equipped with highly sensitive color Doppler imaging, and a MultiHertz intravaginal probe with a field of view of 140°. The color Doppler imaging examination always included measurements of flow indices from both uterine arteries and subendometral blood vessels. High-quality transparent color copies (Agfa Drystar, Agfa Gevaert, Mortsel, Belgium) and schematic hand made drawings of the ultrasonographic findings were obtained for every patient. Histopathology was assessed preoperatively using an endometrial biopsy by Pipelle® de Cornier which was shown to accurately reflect the histopathological parameters[20-22]. The patients were divided into two groups as determined by the final histopathologic examination of the hysterectomy specimen: patients with surgical stage Ia or Ib and patients with surgical stage Ic or higher.

Several morphological parameters visualized by gray scale TVS are available for univariate analysis (endometrial (ET) and myometrial (MT) thickness; endometrial (EV) and uterine (UV) volume; ET/AP (AP = uterine anteroposterior diameter); EV/UV; MT/AP; endometrial echogenicity (EE: homogeneous or heterogeneous); endometrial lining (EL: regular or irregular)). Endometrial and uterine volume (expressed in mL) were calculated from three measurements of the endometrium or the uterus in two perpendicular planes and the volume was calculated according to the formula for a prolate ellipsoid: $\pi/6 \times D1 \times D2 \times D3$ (where D1, D2, and D3 represent the three diameters of the structure). Blood flow indices included intratumoral peak systolic velocity (PSV), time-averaged maximum mean velocity (TAMXV), resistance index (RI) and pulsatility index (PI). Furthermore uterine artery PSV, TAMXV (maximum of the values measured at both the left and right uterine artery, i.e. the worst case), RI and PI (minimum of the values measured at both the left and right uterine artery) were measured. The subjective assessment by the gynecologist of the depth of myometrial invasion (using a 4-value scoring system - 0: stage Ia; 1: Ib; 2: Ic; 3: II or higher) was also recorded. The gynecologist was not blinded to the histological results and tumor grading but he mainly based his assessment on the volume of the tumor and the remaining myometrium between tumor and serosa.

Univariate analysis

Univariate analysis was performed using the SAS software package (Release 8.01). We used the Wilcoxon rank-sum test (for continuous data) and the Fisher's exact test (for categorical data) to calculate p-values that reflect whether there is a significant difference for a certain variable between patients with surgical stage Ia or Ib and patients with surgical stage Ic or higher [23]. In addition, the ROC curves and the AUCs were estimated[24] and compared[25] for the individual parameters using custom scripts written in MATLAB (Version 6.5 Release 13 – also see Epstein et al.[26] where the same scripts were applied). The optimal cut-off point on the ROC curve was defined as the point that obtained the best trade off between sensitivity and specificity (point with a tangent line with slope 1, for which it can be proven that it maximizes the sum of the sensitivity and specificity). The resulting sensitivity and specificity values were also calculated. For all hypothesis tests two-sided tests were used and $p < 0.05$ was used as the level of significance.

Multivariate analysis

We trained three models (training = using the patients of the training set to determine the coefficients of a model in order to optimize its ability to differentiate between patients with and without deep myometrial invasion) based on a set of variables selected after stepwise logistic regression analysis. Subsequently, these models were prospectively validated on a new and independent set of new patients. A schematic overview of the multivariate analysis procedure is given in Figure 1.

VARIABLE SELECTION:

With multivariate stepwise logistic regression analysis (using stepwise selection in the LOGISTIC procedure from SAS) we aimed to select the variables that significantly contribute in a standard logistic regression model that predicts deep myometrial invasion. We considered the following variables for inclusion in the model: the ultrasound parameters discussed above, the number of fibroids detected during ultrasound examination (NF; range 0-2; this parameter was previously reported to be a potential factor disturbing sonographic prediction (overestimation of invasion)[27]), the degree of differentiation of the cancer, presence of a clear cell component, and presence of a serous papillary component. Note that the latter three (histopathological) variables were assessed on an endometrial biopsy (Pipelle® de Cornier) that is available preoperatively. In the model, obtained at the end of the stepwise logistic regression analysis, only variables having a coefficient significantly different from zero (p-value < 0.05 - Wald Chi-Square statistic) were allowed[28]. Note that only 74 of the 97 patients from the training set could be used for the stepwise logistic regression analysis because of missing values in some of the considered variables.

MODEL BUILDING:

The variables selected after the stepwise logistic regression analysis were subsequently used to fit a standard logistic regression model and Least Squares Support Vector Machine (LS-SVM) models[29] with linear and RBF (Radial Basis Function) kernels to the training set.

Support vector machines (SVM) are a relatively new method to solve classification problems and have already been extensively used for various applications, including medical ones[30] (for more details, see the Opinion published in the same issue of this journal).

Since the models in this section are only based on a subset of the variables used during variable selection and since only the patients without missing values in any of the variables can be taken into account, the number of patients in the model building step

(94) was larger than the number of patients (74) used in the variable selection step. As described above, the single valued output of the models can also be analysed and compared using the Wilcoxon rank-sum test and ROC curves, and can also be used to estimate an optimal cut-off point or threshold for these models. Patients with a model output larger than this cut-off are then predicted to have deep myometrial invasion.

The standard logistic regression model was fitted with the LOGISTIC procedure from SAS. The class labels for patients with stage Ia or Ib were 0, and 1 for patients with stage Ic or higher. The Wald Chi-Square statistic was used to assess the significance of the coefficient of a certain variable in the fitted model.

Using LS-SVMlab version 1.5 [29,31] for MATLAB we trained two LS-SVM models using a linear and an RBF kernel. It is possible to write an LS-SVM with a linear kernel as a simple linear equation in its variables. An LS-SVM with an RBF kernel has a more complex form, explaining why it is not explicitly stated in this manuscript.

PROSPECTIVE VALIDATION:

In the previous section, the AUCs of the mathematical models were estimated using the same collection of patients that was used to fit or train these models. This can possibly lead to results that are too optimistic. Therefore, we prospectively validated our models using independent data from 78 consecutive new patients. We will refer to these patients as the "independent test set"), that became available *after* the derivation of these models (prospectively collected). The mean age of the patients in the test set was 64.1 years and 72 of them were postmenopausal. They were assessed using the same protocol as used for the patients of the training set. The distribution of the FIGO stages was: 14 stage Ia, 36 Ib, 16 Ic, 1 II, 9 III and 2 IV. The following histopathological subtypes were present: 59 endometrioid adenocarcinoma, 1 mucinous, 2 serous papillary, 15 mixed type (9 with a serous papillary and 4 with a clear cell component), and 1 endometrial tumor with unspecified histopathologic subtype. Forty tumors were highly, 14 moderately and 24 poorly differentiated. Using this independent test data, we calculated the AUCs of the three models discussed above and compared them with the AUC of the subjective assessment of the expert. We also evaluated the performance of our models at the optimal cut-off points obtained after the ROC-analysis of the training set. We used the method described by Hanley and McNeil [24,25] to estimate the sample size needed to reach statistical significance.

Results

The results – based on the training set – of the univariate analysis of the ultrasound parameters and the subjective assessment are presented in Table 1. EV/UV had the largest AUC (78%) from all the ultrasound parameters, which was comparable to the AUC of the subjective assessment (79% - difference not statistically significant). Also, there was no significant difference between the AUC of EV/UV and the AUCs of ET, MT, EV, ET/AP and MT/AP. Compared to the latter, the AUCs of the blood flow indices were low. Only the values of the uterine artery RI and PI were (borderline) significant at the 5% level between stage Ia and Ib versus stage Ic and higher.

Multivariate stepwise logistic regression selected the degree of differentiation, the number of fibroids (NF), ET and EV as variables that significantly contributed in a standard logistic regression model aiming to discriminate between patients with and without deep myometrial invasion on the final histopathological assessment. None of the blood flow indices were included.

The resulting logistic regression model fitted to the training data is given by (note that since we only had to take the missing variables in the four selected variables into account, 94 patients could be used to fit the three models, which is more than the number of patients (74) that was used for variable selection):

$$y = \frac{\exp(\beta_0 + \beta_1.\mathrm{DD1} + \beta_2.\mathrm{DD2} + \beta_3.\mathrm{NF} + \beta_4.\mathrm{ET} + \beta_5.\mathrm{EV})}{1 + \exp(\beta_0 + \beta_1.\mathrm{DD1} + \beta_2.\mathrm{DD2} + \beta_3.\mathrm{NF} + \beta_4.\mathrm{ET} + \beta_5.\mathrm{EV})}$$

where DD1 and DD2 equal 1 if, respectively, the tumor is moderately and poorly differentiated and 0 in other cases and where y is the model output which is a number on a continuous scale between 0 and 1. A patient is predicted to have a tumor stage Ia or Ib if y is smaller or equal than a certain cut-off level and predicted to have a tumor stage Ic or higher if y is larger than this cut-off level. The coefficients are (rounded off to two decimal places): $\beta_0 = -3.70$ (95% CI [-5.53, -1.86], $p < 0.0001$), $\beta_1 = 2.36$ ([0.82, 3.91], $p = 0.0027$), $\beta_2 = 2.42$ ([1.00, 3.84], $p = 0.0008$), $\beta_3 = -2.45$ ([-4.23, -0.67], $p = 0.0070$), $\beta_4 = 0.20$ ([0.07, 0.32], $p = 0.0021$), and $\beta_5 = -0.11$ ([-0.19, -0.03], $p = 0.0054$). These coefficients indicate that the predicted probability of deep myometrial invasion increases when the degree of differentiation and the endometrial thickness increase and that the predicted probability of deep myometrial invasion decreases when the number of fibroids and the endometrial volume increase. The negative influence of the endometrial volume seems unexpected, but can be seen as a non-linear effect of the endometrial thickness (since EV ~ ET^3). The performance of the standard logistic regression model on the training data and the optimal cut-off level are also summarized in Table 1.

The resulting LS-SVM model with a linear kernel fitted to the training data is given by:

$$y = \beta_0 + \beta_1.\mathrm{DD} + \beta_2.\mathrm{NF} + \beta_3.\mathrm{ET} + \beta_4.\mathrm{EV}$$

where DD equals 1, 2 and 3 if the degree of differentiation is highly, moderately and poorly differentiated, respectively and where y is the model output which is a number on a continuous scale. Again, a patient is predicted to have a tumor stage Ia or Ib if y is smaller or equal than a certain cut-off level and predicted to have a tumor stage Ic or higher if y is larger than this cut-off level. The coefficients are (rounded off to two decimal places): $\beta_0 = -1.44$, $\beta_1 = 0.37$, $\beta_2 = -0.37$, $\beta_3 = 0.05$, and $\beta_4 = -0.03$. According to the sign of these coefficients, the influence of the different variables is qualitatively the same as in the logistic regression model.

As mentioned previously, the LS-SVM model with an RBF kernel cannot be written in a simplified form (it is a sum with 95 terms) and is therefore not explicitly stated here. However, it could easily be implemented in, for example, Microsoft Excel. The model output is a single and continuous number that needs to be compared with a certain cut-off level. The performance of the LS-SVM models with a linear and RBF kernel on the training data and the optimal cut-off levels are also described in Table 1.

Evaluated on the training set, the standard logistic regression and the LS-SVM models with a linear and RBF kernel had a larger AUC than the subjective assessment. This difference was only significant for the LS-SVM with an RBF kernel ($p < 0.0001$) and borderline significant for the standard logistic regression model ($p = 0.0595$).

The results of the prospective validation (the evaluation of the models on the independent test set was only possible in 76 (out of 78) patients because of missing values in EV) are presented in Table 2 and Figure 2. From these we can conclude that prospective evaluation on the independent test set only resulted in a better AUC for the LS-SVM model with a RBF kernel (difference not significant) and in an equally good AUC for the LS-SVM model with a linear kernel when compared to the AUC of the subjective assessment. The performance of the standard logistic regression model was poor here. For the optimal cut-off value, the likelihood ratio for a positive result (LR+) of the subjective assessment was better than the LR+ of the LS-SVM models. The opposite was true for the likelihood ratio for a negative result (LR-). This means that - at the chosen cut-off level - the LS-SVM models are better at ruling out deep myometrial invasion than ruling it in when compared to the subjective assessment.

Discussion

Our study indicates that single morphological parameters do not improve the predictive power when compared to the subjective assessment and that spectral Doppler analysis does not contribute to the prediction of the degree of myometrial invasion in endometrial cancer. Combining the degree of differentiation, the endometrial thickness, endometrial volume and the number of fibroids in an LS-SVM model with a linear or RBF kernel might deliver predictions that are as reliable as the subjective impression by an experienced sonologist. Assuming that a real difference exists between the true AUC of the LS-SVM model with an RBF kernel and the true AUC of the subjective assessment, the number of patients in the independent test set, however, was not sufficient to reach statistical significance in a prospective evaluation. If the values in Table 2 represent the true AUCs (i.e. those that would be achieved by infinite populations), one would need a sample size of approximately 919 patients to be able to detect – with 80% power – the difference between these AUCs as statistically significant[32]. Confirmation of the performance of LS-SVM models with an RBF kernel in larger prospective studies is therefore necessary.

As could be expected and as explained in the Opinion of this issue, the performance on the test set or level of generalization of the LS-SVM model with a linear kernel was better than the performance of the standard logistic regression model. Evaluation on the training set (Table 1) gave the opposite order of performance, although the difference was small. The LS-SVM model with an RBF kernel had the best overall performance, both on the training as on the independent test set. This is an indication that non-linear effects might play a role in the distinction between patients with and without deep myometrial invasion. The better LR- of the LS-SVM model could be helpful to detect patients who might benefit from a pelvic lymphadenectomy by an experienced surgeon.

It is important to emphasize that the models described in this study might not be ready to be implemented in routine clinical practice. First of all, the measurements that were considered in our study all originated from the same sonologist. Because of differences that might exist between different centers, or even individual sonologists (who might for example use different ultrasound equipment), the models discussed here should be tested using multicenter prospective data using a stringent and detailed protocol. We have planned this multicenter prospective study. Moreover, the techniques used by the same expert might undergo subtle changes throughout time, causing a drop in model performance when the model is applied on new patients. These comments also apply to the evaluation of the degree of differentiation, a variable that was also included in our models. This parameter is, at least partially, a subjective measure that can also differ between centers, between pathologists and in time. There is also the possibility of change in characteristics of the population of patients, causing new patients to be drawn from a different distribution than the one that was used to derive the models. This again might cause a drop in model performance when applied to the new data.

Despite the possible limitations, we believe that the proposed models could represent a simple and inexpensive method that might contribute to the preoperative distinction between low- and high-risk patients, allowing for better preoperative allocation of patients with endometrial carcinoma. Therefore, further research is needed in this area.

Acknowledgments

This research is supported by Research Council KUL: GOA-Mefisto 666, GOA AMBioRICS, IDO (IOTA Oncology, Genetic networks), several PhD/postdoc & fellow grants; Flemish Government: FWO: PhD/postdoc grants, projects G.0115.01 (microarrays/oncology), G.0240.99 (multilinear algebra), G.0407.02 (support vector machines), G.0413.03 (inference in bioi), G.0388.03 (microarrays for clinical use), G.0229.03 (ontologies in bioi), G.0241.04 (Functional Genomics), G.0499.04 (Statistics), research communities (ICCoS, ANMMM, MLDM); IWT: PhD Grants, STWW-Genprom (gene promotor prediction), GBOU-McKnow (Knowledge management algorithms), GBOU-SQUAD (quorum sensing), GBOU-ANA (biosensors); Belgian Federal Science Policy Office: IUAP P5/22 (`Dynamical Systems and Control: Computation, Identification and Modelling, 2002-2006) ; EU-RTD: FP5-CAGE (Compendium of Arabidopsis Gene Expression); ERNSI: European Research Network on System Identification; FP6-NoE Biopattern; FP6-IP e-Tumours.

References

1. Young RC. Gynecologic malignancies. In *Harrison's principles of internal medicine - 14th edition*, Fauci AC, Braunwald E, Isselbacher KJ, Wilson JD, Martin JB, Kasper DL, Hauser SL, Longo DL (eds). McGraw-Hill: New York, 1998; 605-611.
2. Levine DA, Hoskins WJ. Update in the management of endometrial cancer. *Cancer J* 2002; **8 Suppl 1**: S31-40.
3. Ludwig H. Prognostic factors in endometrial cancer. *Int J Gynecol Obstet* 1995; **49 Suppl**: S1-7.
4. Franchi M, Ghezzi F, Melpignano M, Cherchi PL, Scarabelli C, Apolloni C, Zanaboni F. Clinical value of intraoperative gross examination in endometrial cancer. *Gynecol Oncol* 2000; **76**: 357-361.
5. Kucera E, Kainz C, Reinthaller A, Sliutz G, Leodolter S, Kucera H, Breitenecker G. Accuracy of intraoperative frozen-section diagnosis in stage I endometrial adenocarcinoma. *Gynecol Obstet Invest* 2000; **49**: 62-66.
6. Kinkel K, Kaji Y, Yu KK, Segal MR, Lu Y, Powell CB, Hricak H. Radiologic staging in patients with endometrial cancer: a meta-analysis. *Radiology* 1999; **212**: 711-718.
7. Artner A, Bosze P, Gonda G. The value of ultrasound in preoperative assessment of the myometrial and cervical invasion in endometrial carcinoma. *Gynecol Oncol* 1994; **54**: 147-151.

8. Cacciatore B, Lehtovirta P, Wahlstrom T, Ylostalo P. Preoperative sonographic evaluation of endometrial cancer. *Am J Obstet Gynecol* 1989; **160**: 133-137.

9. Develioglu OH, Bilgin T, Yalcin OT, Ozalp S, Ozan H. Adjunctive use of the uterine artery resistance index in the preoperative prediction of myometrial invasion in endometrial carcinoma. *Gynecol Oncol* 1999; **72**: 26-31.

10. Fishman A, Altaras M, Bernheim J, Cohen I, Beyth Y, Tepper R. The value of transvaginal sonography in the preoperative assessment of myometrial invasion in high and low grade endometrial cancer and in comparison to frozen section in grade 1 disease. *Eur J Gynaecol Oncol* 2000; **21**: 128-130.

11. Gordon AN, Fleischer AC, Reed GW. Depth of myometrial invasion in endometrial cancer: preoperative assessment by transvaginal ultrasonography. *Gynecol Oncol* 1990; **39**: 321-327.

12. Karlsson B, Norstrom A, Granberg S, Wikland M. The use of endovaginal ultrasound to diagnose invasion of endometrial carcinoma. *Ultrasound Obstet Gynecol* 1992; **2**: 35-39.

13. Olaya FJ, Dualde D, Garcia E, Vidal P, Labrador T, Martinez F, Gordo G. Transvaginal sonography in endometrial carcinoma: preoperative assessment of the depth of myometrial invasion in 50 cases. *Eur J Radiol* 1998; **26**: 274-279.

14. Prompeler HJ, Madjar H, du Bois A, Lattermann U, Wilhelm C, Kommoss F, Pfleiderer A. Transvaginal sonography of myometrial invasion depth in endometrial cancer. *Acta Obstet Gynecol Scand* 1994; **73**: 343-346.

15. Weber G, Merz E, Bahlmann F, Mitze M, Weikel W, Knapstein PG. Assessment of myometrial infiltration and preoperative staging by transvaginal ultrasound in

patients with endomettial carcinoma. *Ultrasound Obstet Gynecol* 1995; **6**: 362-367.

16. Sahakian V, Syrop C, Turner D. Endometrial carcinoma: transvaginal ultrasonography prediction of depth of myometrial invasion. *Gynecol Oncol* 1991; **43**: 217-219.
17. Szantho A, Szabo I, Csapo ZS, Balega J, Demeter A, Papp Z. Assessment of myometrial and cervical invasion of endometrial cancer by transvaginal sonography. *Eur J Gynaecol Oncol* 2001; **22**: 209-212.
18. Arko D, Takac I. High frequency transvaginal ultrasonography in preoperative assessment of myometrial invasion in endometrial cancer. *J Ultrasound Med* 2000; **19**: 639-643.
19. Timmerman D, De Smet F, De Brabanter J, Thijs I, Moreau Y, De Moor B, Vergote I. Pre-operative prediction of depth of myometrial invasion in patients with endometrial cancer: evaluation of ultrasound parameters and development of a new logistic regression model. *Int J Gynecol Cancer 2003*; **13 suppl 1**: S26.
20. Dijkhuizen FP, Mol BW, Brolmann HA, Heintz AP. The accuracy of endometrial sampling in the diagnosis of patients with endometrial carcinoma and hyperplasia: a meta-analysis. *Cancer* 2000; **89**: 1765-1772.
21. Clark TJ, Mann CH, Shah N, Khan KS, Song F, Gupta JK. Accuracy of outpatient endometrial biopsy in the diagnosis of endometrial cancer: a systematic quantitative review. *BJOG* 2002; **109**: 313-321.
22. Amant F, Moerman P, Neven P, Timmerman D, Van Limbergen E, Vergote I. Endometrial cancer. *Lancet* 2005; **366**:491-505.

23. Dawson-Saunders B, Trapp RG. *Basic & Clinical Biostatistics, 2nd edition.* Appleton & Lange: Connecticut, 1994.

24. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982; **143**: 29-36.

25. Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology* 1983; **148**: 839-843.

26. Epstein E, Skoog L, Isberg PE, De Smet F, De Moor B, Olofsson PA, Gudmundsson S, Valentin L. An algorithm including results of gray-scale and power Doppler ultrasound examination to predict endometrial malignancy in women with postmenopausal bleeding. *Ultrasound Obstet Gynecol* 2002; **20**: 370-376.

27. Weber G, Merz E, Bahlmann F, Mitze M, Weikel W, Knapstein PG. Assessment of myometrial infiltration and preoperative staging by transvaginal ultrasound in patients with endometrial carcinoma. *Ultrasound Obstet Gynecol* 1995; **6**: 362-367.

28. Hosmer DW, Lemeshow S. *Applied Logistic regression.* John Wiley & Sons: New York, 1989.

29. Suykens JAK, Van Gestel T, De Brabanter J, De Moor B, Vandewalle J. *Least Squares Support Vector Machines.* World Scientific: Singapore, 2002.

30. Van Gestel T, Suykens J, Baesens B, Viaene S, Vanthienen J, Dedene G, De Moor B, Vandewalle J. Benchmarking Least Squares Support Vector Machine Classifiers. *Machine Learning* 2004; **54**: 5-32.

31. LS-SVMlab version1.5. http://www.esat.kuleuven.ac.be/sista/lssvmlab/

32. Jaeschke R, Guyatt G, Sackett DL. Users' guides to the medical literature. III. How to use an article about a diagnostic test. A. Are the results of the study valid? B. What are the results and will they help me in caring for my patients? Evidence-Based Medicine Working Group. *JAMA* 1994; **271**: 389-391 and 703-707.

Tables

Table 1: Univariate analysis of the ultrasound parameters, the subjective assessment, the standard logistic regression model and the LS-SVM models with a linear and RBF kernel (training set, $N = 97$).

		Range	AUC [95% CI]	Optimal cut -off value*	Sensitivity (%)	Specificity (%)	Mean or proportion in stage Ia or stage Ib	Mean or proportion in stage Ic or higher	p-value
ET (mm)		2-65	0.76 [0.66, 0.86]	14	81	64	15	25	<0.0001
MT (mm)		2-18	0.71 [0.59, 0.82]	8	74	61	8.8	6.4	0.001
EV (ml)		0-84	0.76 [0.66, 0.86]	4.9	71	69	8.2	18	<0.0001
UV (ml)		16-1075	0.61 [0.49, 0.72]	89	58	69	91	147	0.08
ET/AP		0.07-1.5	0.75 [0.65, 0.86]	0.43	72	71	0.37	0.54	<0.0001
EV/UV		<0.0001-0.75	0.78 [0.68, 0.87]	0.09	69	80	0.07	0.15	<0.0001
MT/AP		0.04-0.44	0.75 [0.64, 0.85]	0.17	74	75	0.24	0.15	<0.0001
EE (% heterogeneous)		-	0.60 [0.49, 0.72]	-	65	56	44 %	65 %	0.06
EL (% irregular)		-	0.61 [0.50, 0.73]	-	78	44	56 %	78 %	0.03
Intra-tumoral	PSV (cm/sec)	0-0.96	0.61 [0.49, 0.73]	0.13	59	64	0.14	0.21	0.09
	TAMXV (cm/sec)	0-0.77	0.61 [0.49, 0.73]	0.06	82	46	0.09	0.14	0.09
	RI	0.05-1	0.62 [0.48, 0.75]	0.5	50	78	0.62	0.54	0.08
	PI	0.23-6.0	0.61 [0.48, 0.74]	0.61	38	88	1.4	1.1	0.10
Uterine artery	PSV (cm/sec)	0.09-2.1	0.51 [0.39, 0.65]	0.62	31	84	0.49	0.53	0.81
	TAMXV (cm/sec)	0.04-0.75	0.57 [0.45, 0.70]	0.25	37	80	0.20	0.24	0.27
	RI	0.41-1.2	0.64 [0.52, 0.76]	0.71	49	78	0.78	0.71	0.03
	PI	0.16-6.0	0.64 [0.52, 0.76]	1.3	49	78	1.9	1.5	0.04
Subjective assessment (stage Ia: O; stage Ib: 1; stage Ic: 2; Stage II or higher: 3)		0-3	0.79 [0.69, 0.88]	1	61	86	0: 51 % 1: 36 % 2: 12 % 3: 2 %	0: 13 % 1: 26 % 2: 39 % 3: 21 %	<0.0001
Standard logistic regression		0-1	0.89 [0.83, 0.96]	0.45	77	86	0.21	0.65	<0.0001
LS-SVM with linear kernel		-1.5 – 1.4	0.88 [0.81, 0.95]	-0.31	91	73	-0.52	0.20	<0.0001
LS-SVM with RBF kernel		-1.2 – 0.93	0.99 [0.97, 1]	-0.30	97	100	-0.74	0.56	<0.0001

ET=Endometrial Thickness; MT=Myometrial Thickness; EV=Endometrial Volume; UV=Uterine Volume; AP=uterine anteroposterior diameter; EE=Endometrial Echogenicity; EL=Endometrial Lining; PSV=Peak Systolic Velocity; TAMXV= Time-Averaged Maximum Mean Velocity; RI=Resistance Index; PI=Pulsatility Index. *The optimal cut-off point was defined as the point that obtained the best trade-off between sensitivity and specificity. P-values show the statistical significance of any differences between results shown in the previous two columns.

Table 2: Prospective validation: performance of the standard logistic regression model and the LS-SVM models with linear and RBF kernels for the patients of the independent test set ($N = 78$ for the subjective assessment and $N = 76$ for EV/UV and the mathematical models). Comparison with the ultrasound parameter (EV/UV) from Table 1 with the best discriminatory potential and the subjective assessment.

	AUC [95% CI]	Optimal cut -off value*	Sensitivity (%)	Specificity (%)	LR+⁺	LR-⁻
EV/UV	0.70 [0.58, 0.82]	0.085	57	72	2.1	0.59
Subjective assessment	0.72 [0.59, 0.84]	1	61	80	3.0	0.49
Standard logistic regression	0.66 [0.53, 0.79]	0.45	50	75	2.0	0.67
LS-SVM with linear kernel	0.72 [0.59, 0.84]	-0.31	75	69	2.4	0.36
LS-SVM with RBF kernel	0.77 [0.66, 0.87]	-0.30	79	67	2.4	0.32

*The optimal cut-off values were taken from Table 1 as evaluated on the training set.
LR+⁺ = Likelihood Ratio for a positive result; LR-⁻ = Likelihood Ratio for a negative result;

Figure legends

Figure 1. Schematic overview of multivariate analysis and model building. (1) Variable selection step: using stepwise logistic regression analysis and the training set, the variables that significantly contributed in a standard logistic regression model (that aims to predict the degree – more or less than 50% – of myometrial invasion as assessed by the final histopathological examination) were selected. Note that the values for all variables that were considered for inclusion in the logistic regression model are known preoperatively and could therefore be used to make a (preoperative) prediction of the result of the final (and postoperative) histopathological examination of the degree of myometrial invasion. (2) Model training (determination of the coefficients of a model in order to optimize its classification performance using the training set): a standard logistic regression model and LS-SVM models with linear and RBF kernels (also aiming to predict the result of the final histopathological assessment) were fit to the training data. The variables used in these models were restricted to the variables selected in step 1. Model training also involved the determination of an optimal cut-off level with the best trade off between sensitivity and specificity as assessed on an ROC curve. Patients with a model output larger than the cut-off are predicted to have an endometrial cancer stage Ic or higher. Because the calculations in step 1 and 2 were based on the patients without missing values in any of the variables, the latter implies that the number of patients used in step 2 (where only a subset of the variables was taken into account) could be larger than in step 1. (3) Prospective validation: the models trained in step 2 were subsequently applied on an independent set of new patients that had not been used in model training. Receiver Operating Characteristic (ROC) curves (and the associated areas under the

curve (AUC)) were constructed by comparing the model output with the final histopathological assessment of the degree of myometrial invasion. (4) Finally, the model AUCs were compared with the AUC of the expert subjective assessment of the same independent test set patients.

Figure 2. Comparison of the ROC curves for the subjective assessment, the standard logistic regression model, and the LS-SVM models with a linear and RBF kernel for the patients of the independent test set ($N = 78$ for the subjective assessment and $N = 76$ for EV/UV and the mathematical models).

Figure 1

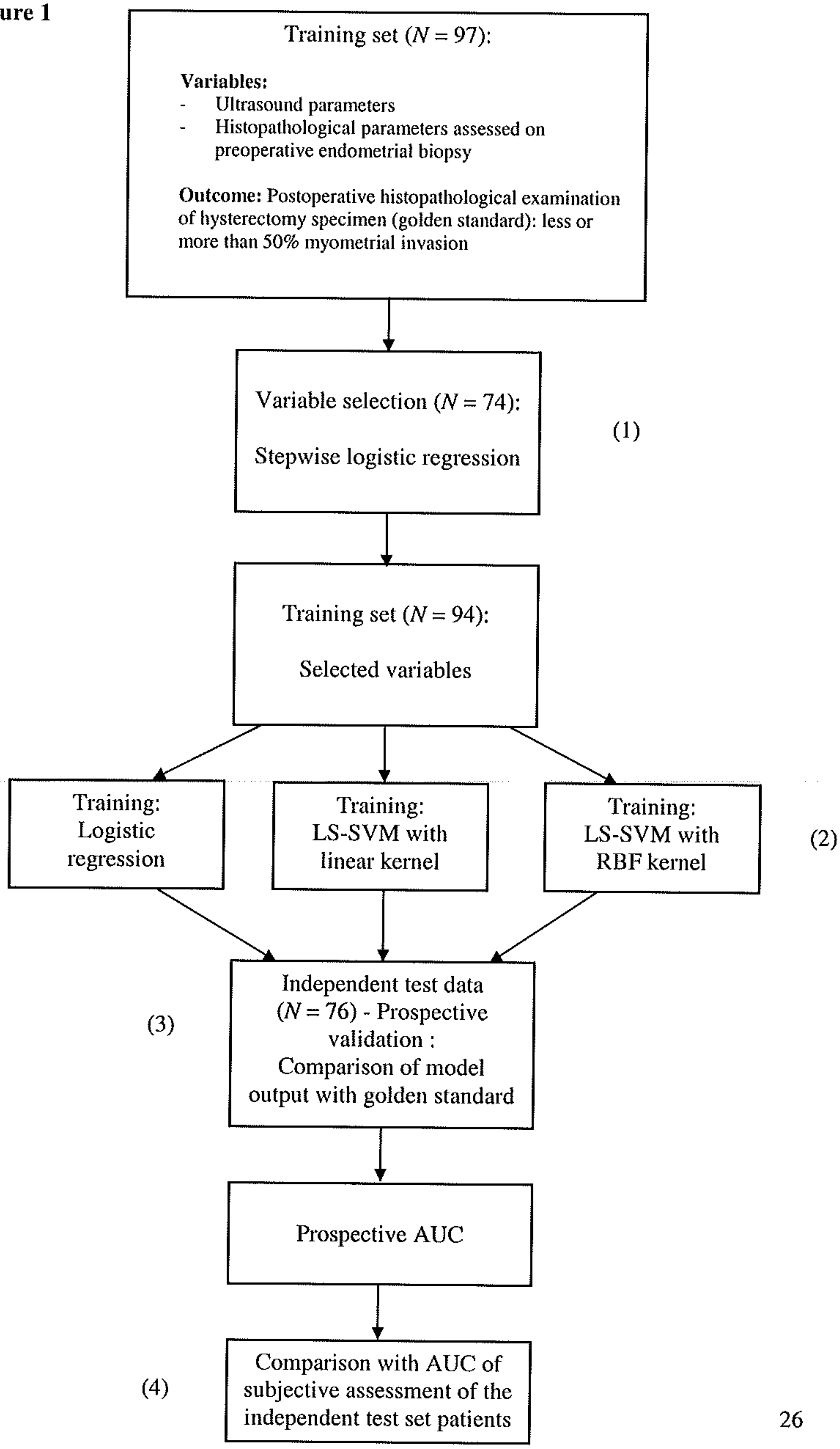
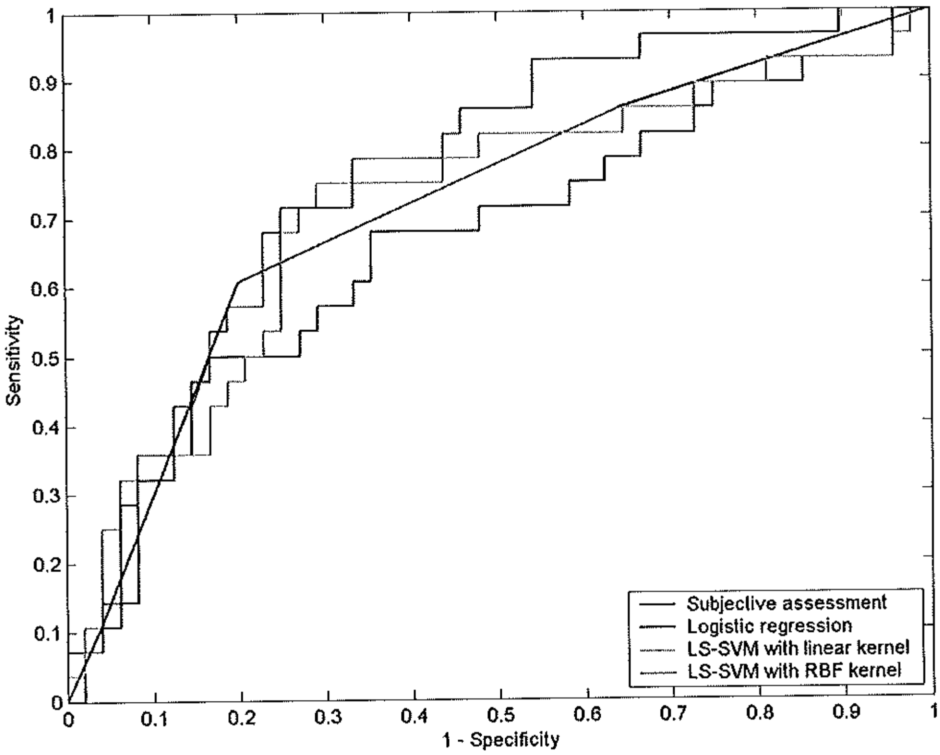

Training set (N = 97):
Variables:
- Ultrasound parameters
- Histopathological parameters assessed on preoperative endometrial biopsy
Outcome: Postoperative histopathological examination of hysterectomy specimen (golden standard): less or more than 50% myometrial invasion
Variable selection (N = 74):
Stepwise logistic regression
(1)
Training set (N = 94):
Selected variables
Training: Logistic regression
Training: LS-SVM with linear kernel
Training: LS-SVM with RBF kernel
(2)
Independent test data (N = 76) - Prospective validation : Comparison of model output with golden standard
(3)
Prospective AUC
Comparison with AUC of subjective assessment of the independent test set patients
(4)

Figure 2



Sensitivity
1 - Specificity
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
Subjective assessment
Logistic regression
LS-SVM with linear kernel
LS-SVM with RBF kernel