EPCAM IS ASSOCIATED WITH LOCOREGIONAL RECURRENCE OF CERVICAL CANCER AFTER RADIOTHERAPY

A Thesis Submitted to
the Department of Biomedical Science
in Partial Fulfillment of the Requirements
for the Master's Degree of Science

Le Trong Dat

Jan 2020

National Cancer Center
Graduate School of Cancer Science and Policy

Professor Kim, Joo-Young

We hereby approve the M.S. thesis of
Le Trong Dat.

Jan 2020

Mi Kyung, Kim sign
Chairman of Thesis Committee

Kim, Joo-Young sign
Thesis Committee Member

Chang-ok Goo sign
Thesis Committee Member

NATIONAL CANCER CENTER
GRADUATE SCHOOL OF CANCER SCIENCE AND POLICY

ABSTRACT

EpCAM is Associated with Locoregional Recurrence of Cervical Cancer after Radiotherapy

Purpose:

Our study aimed to investigate the association between the expression of epithelial cell adhesion molecule (EpCAM) and prognosis of cervical cancer after radiotherapy. We also examined the correlation between pathologist's immunohistochemical (IHC) scoring process of EpCAM and H-score analyzed by an image analysis software.

Methods:

Two hundred two pre-radiotherapy tumor samples were obtained from cervical cancer patients who received radiotherapy between September 2003 and August 2010 at National Cancer Center, Korea. EpCAM expression was graded by one pathologist based on the staining intensity and the area of positive staining: 0 = positive in less than 5% of tumor area, 1+ = focal weak positive 5-20%, 2+ = positive 20-50%, 3+ = diffuse strong positive more than 50%. H-score was calculated using InForm® Image Analysis Software 2.2 and TIBCO Spotfire® (Perkin Elmer) in a subset of 75 patients. The Cox proportional hazards model and Kaplan-Meier curve were used to investigate the EpCAM expression associated

with locoregional recurrence-free survival (LRFS), disease-free survival (DFS), and overall survival rate (OS).

Results:

Between two groups of EpCAM developed from intensity score for analysis (IHC grade 0 vs. 1+2+3), the Log-rank test showed a significant difference in LRFS (p-value=0.041). The 5-year LRFS rate was 95.8% in the IHC grade 0 group and 80.1% in the IHC grade 1+2+3 group. However, EpCAM expression was only marginally associated with LRFS (P value=0.074) for univariable and DFS (P value=0.086) for multivariable in cox regression model. H-score was shown to increase with the pathologist's score with the Spearman's correlation coefficient of 0.914 (P value <0.001). Kappa statistic also revealed substantial agreement between H-score and pathologist's score with a coefficient of 0.75 (0.63-0.87).

Conclusion:

The data revealed that EpCAM showed a tendency being associated with locoregional recurrence in cervical cancer after radiotherapy. H-score and pathologist's score showed a strong correlation suggesting this new scoring process can be used as an assistive tool or to replace the manual scoring process because this method reduces the workload burden for pathologists as well as the bias of interobserver.

CONTENT

LIST OF TABLE

LIST OF FIGURES

LIST OF ABBREVIATION

EpCAM	Epithelial cell adhesion molecule
IHC	Immunohistochemistry
LRFS	Locoregional recurrence-free survival
DFS	Disease-free survival
OS	Overall survival
SCC	Squamous cell carcinoma
AD	Adenocarcinoma
ADS	Adenosquamous carcinoma
FIGO	International Federation of Gynecology and Obstetrics stage for cervical carcinoma
CCRT	Concurrent chemoradiotherapy
WD	Well differentiated
MD	Moderately differentiated
PD	Poorly differentiated
WBC	White blood cells
SCC-Ag	Squamous cell carcinoma antigen
CEA	Carcinoembryonic antigen
HR	Hazard ratio
CI	Confidence interval

1. INTRODUCTION

1.1 Cervical cancer

Cervical cancer is a disease in which malignant cells form in the tissues of the cervix. Among women worldwide, cervical cancer ranks fourth for both incidence and mortality. It remains an enormous burden of global health, causing 569,847 new cases and 311,365 deaths every year and these numbers are predicted to increase in the future [1]. Recently, thanks to the effectiveness of the screening test for cervical cancer, the incidence and mortality rates are significantly declining in developed countries. However, the incidence rate remains high in developing countries where it accounts for 85% of all cervical cancer cases worldwide [2]. In South Korea, according to the cancer statistics from 1999 to 2015, cervical cancer is the seventh most common cancer and the third leading cause of cancer-related deaths among women. In 2015, 3,582 new cases were diagnosed, accounting for 3.5% of all cancer with the crude incidence rate and age-standardized rate were 14.1 and 9.1, respectively. In general, over the period, the number of newly diagnosed cases gradually decreased from 16.3 in 1999 to 9.1 in 2015 [3]. In contrast, the incidence of cervical cancer with carcinoma in situ increased in all ages [4] because of the early diagnosis and treatment at a precancerous stage of cancer [2]. Similar to the downward trend of incidence, the age-standardized mortality rate of cervix cancer has decreased since 2003, with an estimated annual percent change of 4.9% during 2003-2015. The number of deaths from cervical

cancer reached the highest point in 2003 (n=1,111), and in 2016, there were 897 deaths reported, making up 2.9% of all cancer-related deaths [5].

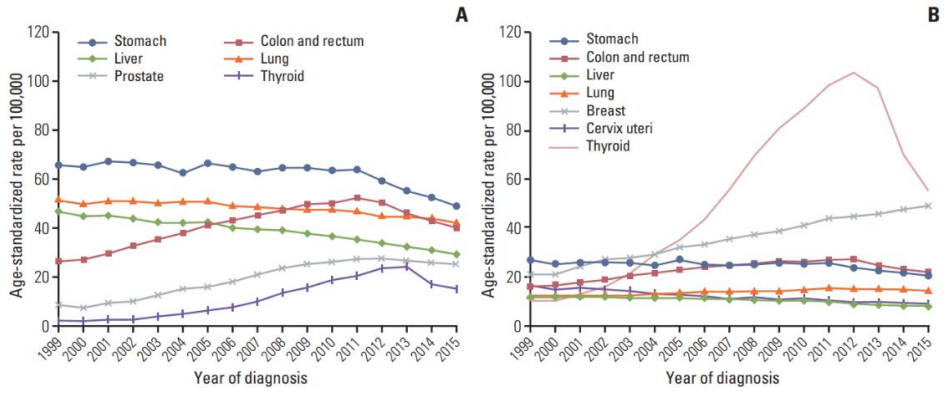


Figure 1. Trends in age-standardized incidences of selected cancers by sex from 1999 to 2015 in Korea. (A) Men. (B) Women [3].

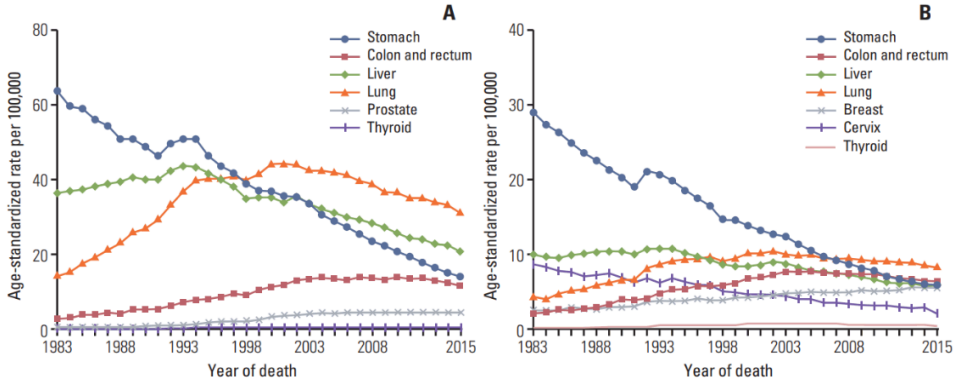


Figure 2. Annual age-standardized cancer mortalities of selected cancers by sex from 1983 to 2015 in Korea. (A) Men. (B) Women [3].

There are two main types of cervical cancer depending on the location in which cancer develops: squamous cell carcinoma (SCC) and adenocarcinoma (AD). Squamous cell carcinomas is the most common type of cervical cancer

which makes up for approximately 80 to 90 percent of total cervical cancer cases. This cancer subtype develops from squamous epithelial cells that line the bottom of the cervix. The second most popular subtype of cervix cancer is adenocarcinomas which develop in the glandular cells that line the upper portion of the cervix. Cervical adenocarcinomas account for about 10 to 20 percent of cervical cancers [6]. Besides these two main subtypes, some other types of cervical cancer have been found in the cervix but they only account for a small percentage of all cancer cases. One of them is known to be the combination of SCC and AD which has been called adenosquamous carcinoma (ADS) [7].

Currently, there are several options of treatment available for patients with cervical cancer, including surgery, radiation therapy, chemotherapy, targeted agent therapy, and immunotherapy. However, according to FIGO cancer report 2018 on cancer of the cervix uteri, in the management of cervical cancer, surgery and radiotherapy are considered primary treatments after clinical staging work-up, with chemotherapy a valuable adjunct [8]. Besides, practice guidelines for the management of cervical cancer in Korea 2017 also suggest that for early-stage disease, the primary treatment is either surgery or radiation therapy [9]. The way that external radiation therapy or internal radiation therapy is given depends on the type and stage of the cancer being treated. Although surgery is preferred for the disease at the early stages, if there are contraindications for surgery or anesthesia, radiotherapy has been shown to be able to replace surgery because of the ability to offer comparable results in terms of local control and survival. This is confirmed

in a randomized study carried out by Landoni et al. that compared surgery and radiotherapy by randomly assigning 343 eligible women into surgery or external radiotherapy arm. The final result did not exhibit a significant difference in twenty-year OS in the radiotherapy group compared with the surgery group (77% vs 72%, P=0.280) [10]. At present, definitive radiotherapy or concurrent chemoradiation (CCRT) is preferred in patients who might need postoperative radiotherapy because these patients are at a high risk of suffering from treatment-related illnesses from both surgery and radiotherapy. The cure rate of cervix uteri cancer in patients receiving surgery or concurrent chemoradiation therapy (CCRT) is up to 80% – 90% for the early stages (stages I–II), and this number decreases to 60% for stage III disease [8].

Recurrence in cervical cancer refers to the cases where cancer returns months or years after the completion of primary treatment. The cancer recurrence can be either a local recurrence or metastatic recurrence, or there may be a combination of these two recurrence types [8]. It is known that cervical cancer spreads progressively and predictably through regional lymphatics, suggesting cancer recurrence is an outcome of inadequate initial therapy. Even though the development and widespread adoption of screening as well as early detection technique has lower overall mortality from this disease, there is no critical improvement has been made for the prognosis of patients with advanced cancer. As stated by a large number of studies, clinicopathological features of patients with cervical carcinomas, such as clinical stage [11, 12], tumor size [13] are presented

to be predictors for the survival of cervical cancer patients. However, these characteristics by themselves are insufficient to forecast the prognosis of cervical cancer patients. Therefore, many scholars have put a lot of effort into finding novel molecular biomarkers that can help clinical practitioners to obtain more precise anticipation about the cervical cancer patient's prognosis.

1.2 EpCAM expression in cervical cancer

EpCAM is a 40 KD transmembrane glycoprotein, which is discovered for the first time by Herlyn, M., et al. in a study carried out on colon cancer [14]. This protein is also known with other names depending on the monoclonal antibodies used to identify its presence, e.g. CD326, MK-1, and TACSTD-1. Its molecular structure consists of three domains: extracellular domain (265 amino acid), single transmembrane domain (23 amino acid) and intracellular domain (26 amino acid) [14]. EpCAM is expressed in many kinds of epithelium in healthy persons including fetal lung, liver, skin, and germ cells [15] and its biological functions in cell signaling, proliferation, differentiation, formation and maintenance of organ morphology have been described [16]. In epithelial tumor tissues, EpCAM is detected at high concentrations, which has been revealed to promote the proliferation of tumors [17, 18]. Recent studies carried out in patients of breast cancer, ovarian cancer as well as pancreatic, urothelial, and gallbladder carcinoma proved that higher EpCAM expression levels are related to poor survival outcomes [19-24]. In cervical cancer, a study conducted on patients with SCC stated that an increased level of EpCAM is found in the majority of cervical SCC, but not in the

normal epithelial area adjacent to SCC. Moreover, EpCAM expression could be used to distinguish between high-grade squamous intraepithelial lesions (HSIL) and low grade squamous intraepithelial lesions (LSIL) [25]. However, in this type of cancer, the relation of EpCAM to survival outcomes of patients after primary treatment, including radiation therapy is not fully understood.

1.3 Immunohistochemistry in cervical cancer

Currently, one of the most common techniques used to identify the presence and the extent of EpCAM expression is IHC. The basic underlying principle of this technique is the detection of the molecular biomarker of interest in the cell of the biopsy tissue samples by the use of antibodies binding specifically to that biomarker. In this method, visualization of the reaction between antibody and antigen is accomplished by using a secondary antibody in combination with an enzyme, such as alkaline phosphatase, which catalyzes a color-producing reaction. Traditionally, EpCAM expression is judged by a pathologist under a light microscope to assign a score ranging from 0 to 3 based on the visual parameters set. This method is still considered as a gold standard to evaluate IHC until now. However, this visual scoring is highly possible to cause some problems due to subjectivity in interpretation in addition to the time consumption which makes these methods low throughput to meet the growing need of large cancer hospitals. Therefore, there is another approach had been developed with the hope that it can assist investigators to overcome the limitations of pathologist's manual evaluation, which is called H-score. This method was first developed in 2003 by Hirsch et al.

and following the principle, the EpCAM expression will be examined and IHC score will be calculated by a software [26]. Since its inception, H-score has been widely applied in many kinds of researches by scientists to assess the expression level of biomarkers on different types of cancer [27, 28]. This method was even proven to be more accurate than the traditional evaluation methods used earlier [29] and it has also been recommended by reputable organizations worldwide [30]. However, in Korea, the application of this method has not been well studied.

1.4 Purpose of study

With the background mentioned above, the present study aims to investigate the association between EpCAM expression and prognosis of cervical cancer after radiotherapy as well as examine the correlation between pathologist's score of EpCAM and H-score.

2. MATERIALS AND METHODS

2.1 Subjects and follow-up data

Our study included samples collected from original tumors of 215 patients before undergoing radiotherapy treatment at the National Cancer Center, Korea between September 2003 and August 2010. The study was performed according to the approval of our Institutional Review Board, and informed consent was obtained from all patients before enrolment to collect and use the tumor samples.

FIGO criteria were employed to determine the tumor stage and the workup of staging identification consisted of a bimanual physical examination, simple chest radiography, cystoscopy and rectosigmoidoscopy in all patients.

Locoregional Free Survival (LRFS), Disease-Free Survival (DFS), and Overall Survival (OS) were the primary endpoints for radiotherapy outcome. LRFS and DFS were calculated from the date that radiotherapy was delivered to the date of local relapse and relapse in any site, respectively. Local recurrence included recurrent diseases at the cervix and parametrial tissues. Locally persistent diseases at 3 months after radiotherapy completed were considered as local relapses. Patients were censored at the time of death as well as at their last follow-up visit.

2.2 Immunohistochemistry

Tissue samples were composed of 2-4 pieces measuring approximately 3 x 3 mm each obtained by multiple punch biopsies. All pieces of tumors were

formalin-fixed and paraffin-embedded into a single block. The FFPE (Formalin-fixed, paraffin-embedded) tissue blocks were cut into 3µm thick sections and placed on silane-coated slide glass (Muto Pure Chemicals, Tokyo, Japan).

2.2.1 Immunohistochemistry evaluation by a pathologist.

EpCAM expression was graded by one pathologist under a light microscope based on two criteria: the staining intensity and the area of positive staining. The final score ranged from 0 to 3:

- 0 = positive < 5% of tumor area.
- 1+ = focal weak positive 5-20%.
- 2+ = positive 21-50%.
- 3+ = diffuse strong positive > 50% in examined tumor area.

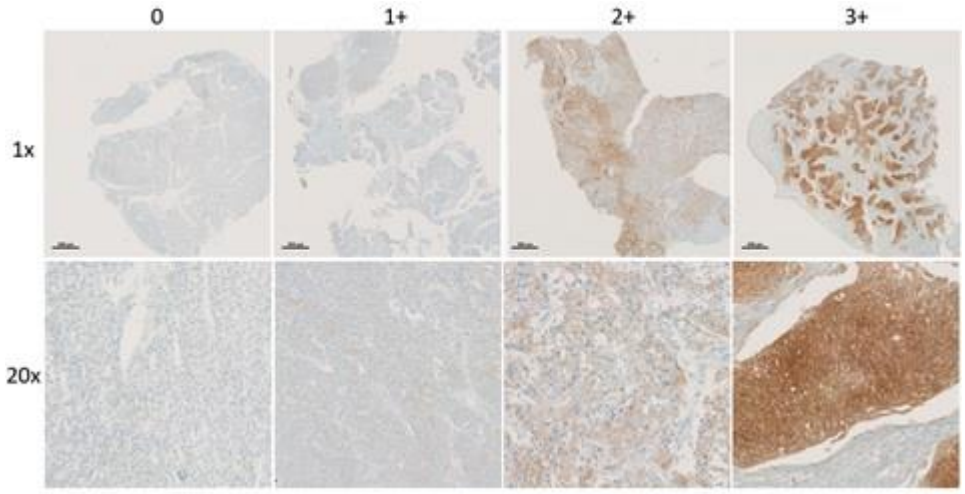


Figure 3. Pathologist's score of EpCAM in cervical cancer tissues.

2.2.2 Immunohistochemistry evaluation by software using H-score.

In parallel, H-score was calculated using InForm® Image Analysis Software 2.2 and TIBCO Spotfire® (Perkin Elmer) in a subset of 75 patients. Following this method, H-score was calculated in two steps: Firstly, each cell in the patient tissue is assessed for staining intensity with a score of 0, +1, +2, and +3. After that, the proportion of cells at each staining intensity level is measured, and finally, an H-score is calculated using the formula as below:

$$[1 \times (\%\ \text{cells } 1+) + 2 \times (\%\ \text{cells } 2+) + 3 \times (\%\ \text{cells } 3+)]$$

The H-score might be either simply based on a controlling staining intensity, or more complexly, could be the sum of H-scores individually counted for each intensity level seen. The final H-score which ranges from 0 to 300 provides more relative weight to stronger intensity in a given tumor sample [28].

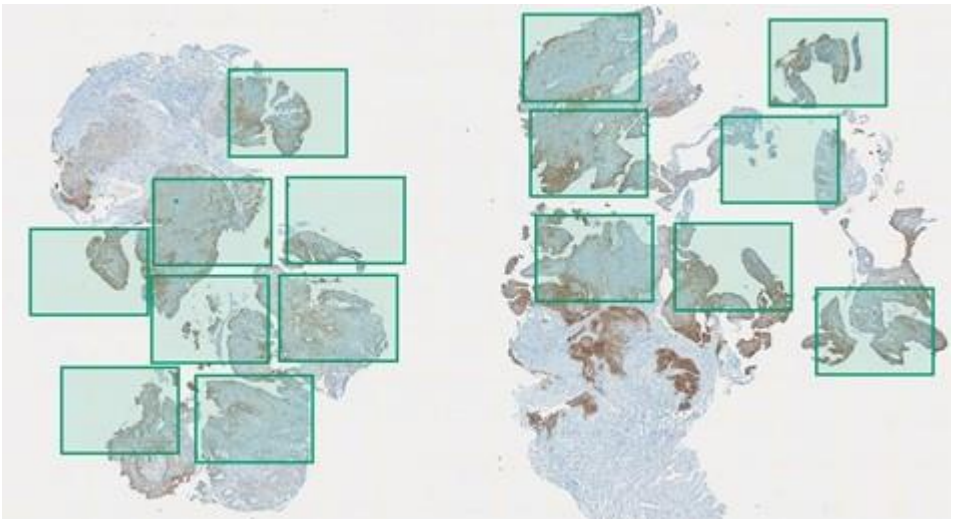


Figure 4. Step of putting Region of Interest (ROI) on the EpCAM-stained tissue slides with minimum required box size so that the computer program can interpret the scanned images of the ROI.

2.3 Statistical analyses

The baseline characteristics were summarized in Table 1. To assess the association between EpCAM expression and clinicopathological characteristics, the continuous variables were compared using the Wilcoxon rank-sum test and categorical variables were compared using the Chi-squared test or Fisher's exact test. LRFS, DFS, and OS were estimated using Kaplan-Meier method and the comparison between survival curves was performed by Log-rank test. The association between EpCAM expression and prognosis was analyzed using the Cox proportional hazards model. EpCAM expression and chosen clinicopathological variables based on univariable P value ≤ 0.2 were entered the multivariable analysis, the final model was determined using the backward selection method with an elimination criterion P value > 0.05.

To explore the correlation between pathologist's score of EpCAM and H-score analyzed by an image analysis software, the distribution of H-score was summarized according to the pathologist's score.

The statistical test was performed two-sided, the P value <0.05 was considered statistically significant. All statistical analyses were performed using the R project software version 3.5.2.

3. RESULTS

3.1 Radiotherapy Outcome

From the date of receiving radiation therapy to the date of last follow up, 72 patients had disease progression, including 29 local recurrences, 19 regional recurrence, and 49 distant metastases. There were 38 patients developed both local and regional recurrences. The patients were followed up for a median period of 100.96 months (range, 4.83 to 162.18 months).

3.1 Clinical Characteristics of Patients

Table 1 summarized information on the baseline characteristics of all patients in this study. The median age at the time of diagnosis was 54 years (range 21-81). The majority of patients was at the locally advanced stage (N = 174, 80.9%), a minority was at late stage, stage III and IVA (N = 41, 19.1%). In contrast, the number of patients with tumor size smaller than 4 cm (N = 123, 57.2%) was similar to that of patients with tumor size equal or greater than 4 cm (N = 92, 42.8%). In terms of smoking status, most of the patients were non-smoker (N = 170, 79.1%). CCRT was administered to 196 patients (91.2%). Furthermore, a small number of patients was AD (N = 19, 8.8%) while there were 196 patients (91.2%) with SCC. A similarity was witnessed in tumor pathological differentiation with the number of patients in poorly differentiation group versus well to moderate differentiation group was 38 (N=38, 18%) and 173 (82%) respectively.

3.2 EpCAM expression in relation to patient characteristics.

The patients were divided into 2 groups according to the EpCAM expression based on the pathologist's score: group 1 (IHC score=0) and group 2 (IHC score = 1+2+3), and then the associations between EpCAM expression and the clinicopathological characteristics of cervical cancer patients were investigated. There were 26 patients were assigned in group one, accounting for 12.9% of total patients. Of the 176 patients in group two, 67 cases (38%) show diffuse strong positive more than 50% of the tumor.

The results showed that EpCAM expression was associated with tumor size, fibrinogen and SCC-Ag level with differences that were statistically significant (Chi-square test, P = 0.02, Wilcoxon rank-sum test, P = 0.045 and 0.008, respectively). EpCAM expression was not associated with any of the other clinicopathological characteristics.

Table 1. Baseline Characteristic Table

Characteristics	N*(%) or median (min-max)
EpCAM	
0	26 (12.9)
1	53 (26.2)
2	56 (27.7)
3	67 (33.2)
Age group	54 (21-81)
age≥40	179 (83.3)
age<40	36 (16.7)
FIGO (2009) stage	
~IIB	174 (80.9)
III/ IVA	41 (19.1)
Tumor size	
< 4 cm	92 (42.8)
≥ 4 cm	123 (57.2)
Smoking	
present	26 (12.1)
ex-smoker	19 (8.8)
non-smoker	170 (79.1)
CCRT	
yes	196 (91.2)
no	19 (8.8)
Histology	
SCC	196 (91.2)
EpCAM 0	26
EpCAM 1	50
EpCAM 2	53
EpCAM 3	54
AD+ADS	19 (8.8)
EpCAM 1	3
EpCAM 2	3
EpCAM 3	13
Differentiation	
WD +MD	173 (82.0)
PD	38 (18.0)
HB	12.2 (5.5 -15.5)
WBC	6920 (3220 -27840)
FIBRINOGEN	338 (167 -749)
SCC-Ag	5.55 (1 -185)
CEA	2.95 (0.5 -1630.9)

Table 2. Association between clinical characteristics and EpCAM group

Characteristics	EpCAM group (N*=202)		
	0	**1+2+3**	**p-value**
Age group			0.774
age≥40	23	147	
age<40	3	29	
FIGO (2009) stage			0.592
~IIB	20	144	
III/ IVA	6	32	
Tumor size			**0.024**
< 4 cm	6	82	
≥ 4 cm	20	94	
Smoking			0.174
present	2	23	
ex-smoker	0	18	
non-smoker	24	135	
CCRT			1.000
yes	24	160	
no	2	16	
Histology			0.141
SCC	26	157	
AD+ADS	0	19	
Differentiation			0.427
WD +MD	23	140	
PD	3	33	
HB	12.1 (6.6 - 14.2)	12.2 (5.5 - 15.5)	0.497
WBC	6620 (3980 - 12650)	6945 (3220 - 27840)	0.425
FIBRINOGEN	290.5 (189 - 516)	345.5 (188 - 749)	**0.045**
SCC	8.6 (1 - 63.1)	4.9 (1 - 185)	**0.008**
CEA	2.3 (0.6 - 32.3)	3.1 (0.5 - 1630.9)	0.130

EpCAM expression was significantly associated with tumor size (Chi-square test, P = 0.02), fibrinogen and SCC-Ag level (Wilcoxon rank-sum test, P = 0.045 and 0.008, respectively).

3.3 EpCAM expression in relation to cancer survival outcomes.

In the univariable analyses, a Cox proportional hazards model was used to investigate factors predictive of cervical cancer recurrence after radiation therapy. Of the investigated variables, age group (HR=2.97, 95%CI = 1.52-5.81, P=0.001) and tumor differentiation (HR=3.23, 95%CI = 1.64-6.35, P<0.001) was significantly associated with LRFS. In multivariable analysis, only these two variables were significant.

In relation to OS, tumor stage (HR=4.20, 95%CI = 2.67-6.63, P<0.001, CCRT status (HR= 2.08, 95%CI = 1.12-3.85, P=0.020), tumor differentiation (HR= 1.87, 95%CI = 1.11-3.14, P=0.018) fibrinogen (HR= 1.002, 95%CI = 1.000-1.005, P=0.015) and SCC-Ag level (HR= 1.014, 95%CI = 1.006-1.021, P<0.001) were significant in univariable model while only tumor stage and tumor differentiation were still statistically significant in multivariable model with both P value were smaller than 0.001.

In addition, tumor stage (HR= 4.07, 95%CI = 2.52-6.58, P<0.001), tumor size (HR= 2.09, 95%CI = 1.26-3.47, P=0.05), tumor histology (HR= 2.34, 95%CI = 1.23-4.46, P=0.009), tumor differentiation (HR= 3.03, 95%CI = 1.84-4.99, P=<0.001), WBC (HR= 1.00, 95%CI = 1.00-1.00, P=0.043) and SCC-Ag (HR= 1.012, 95%CI = 1.004-1.019, P=0.004) level were identified as the significant predictors of DFS. The results were not much different in the multivariable model when tumor stage, tumor size, tumor histology, tumor differentiation were still significant.

In both univariable and multivariable analysis, no significant association was observed between EpCAM expression and survival outcome. EpCAM expression was only marginally associated with LRFS (P value=0.074) for univariable and DFS (P value=0.086) for multivariable in cox regression model.

Kaplan-Meier analysis for LRFS, OS, and DFS of cervical cancer patients according to EpCAM expression was shown in figure 3. Between two groups of EpCAM developed from pathologist's score for analysis (IHC score 0 vs. 1+2+3), worse survival outcome was seen in the group 2 (IHC score 1+2+3) and the Log-rank test showed the significant difference in LRFS (P=0.041). The 5-year LRFS rate was 95.8% in IHC grade 0 group and 80.1% in IHC grade 1+2+3 group. EpCAM expression was not significantly associated with DFS and OS.

3.4 Correlation between pathologist's score and H-score

H-score was calculated for a subset of 75 patients and then was examined in the relationship with the pathologist's score. Spearman's correlation analysis demonstrated that H-score was strongly positively correlated with the pathologist's score (r=0.914, P value <0.001). The level of agreement between pathologist's score and H-score was checked using the Kappa statistic and the result revealed substantial agreement between these two scoring methods (k=0.75 (0.63-0.87)). Median H-score was calculated according to each pathologist's score and Kruskal-Wallis test proved that median H-score is shown to increase with the pathologist's score (P value <0.001).

Table 3. Univariable analyses for survival outcomes with clinicopathological characteristic factors

Characteristics	LRFS			OS			DFS		
	N (event)	HR (95% CI)	p-value	N (event)	HR (95% CI)	p-value	N (event)	HR (95% CI)	p-value
EpCAM									
0	26 (1)	1		26 (9)	1		26 (7)	1	
1+2+3 (≥5%)	176 (35)	6.14 (0.84-44.85)	0.074	176 (67)	1.36 (0.68-2.73)	0.387	176 (63)	1.56 (0.71-3.40)	0.267
Age									
age≥40	179 (25)	1		179 (65)	1		179 (53)	1	
age<40	36 (13)	2.97 (1.52-5.81)	**0.001**	36 (14)	1.13 (0.63-2.01)	0.681	36 (19)	2.08 (1.23-3.51)	**0.006**
FIGO stage									
~IIB	174 (28)			174 (48)	1		174 (44)	1	
III/ IVA	41 (10)	1.89 (0.91-3.91)	0.086	41 (31)	4.20 (2.67-6.63)	**<.001**	41 (28)	4.07 (2.52-6.58)	**<.001**
Tumor size									
< 4 cm	92 (12)			92 (30)	1		92 (21)	1	
≥ 4 cm	123 (26)	1.79 (0.90-3.56)	0.094	123 (49)	1.35 (0.86-2.13)	0.192	123 (51)	2.09 (1.26-3.47)	**0.005**
Smoking									
present	26 (4)	1		26 (8)	1		26 (8)	1	
ex-smoker	19 (7)	2.98 (0.87-10.19)	0.082	19 (5)	0.90 (0.29-2.75)	0.850	19 (10)	2.23 (0.88-5.66)	0.091
non-smoker	170 (27)	1.07 (0.37-3.06)	0.901	170 (66)	1.43 (0.69-2.99)	0.336	170 (54)	1.10 (0.52-2.32)	0.796
Ccrt									
yes	196 (36)	1		196 (67)	1		196 (64)	1	
no	19 (2)	0.61 (0.15-2.56)	0.502	19 (12)	2.08 (1.12-3.85)	**0.020**	19 (8)	1.35 (0.65-2.82)	0.427
Histology									
SCC	196 (32)	1		196 (69)	1		196 (61)	1	
AD+ADS	19 (6)	2.29 (0.96-5.48)	0.063	19 (10)	1.61 (0.83-3.13)	0.160	19 (11)	2.34 (1.23-4.46)	**0.009**
Differentiation									
WD +MD	173 (24)	1		173 (57)	1		173 (47)	1	
PD	38 (13)	3.23 (1.64-6.35)	**<.001**	38 (19)	1.87 (1.11-3.14)	**0.018**	38 (23)	3.03 (1.84-4.99)	**<.001**
HB	215 (38)	0.95 (0.79-1.13)	0.541	215 (79)	0.95 (0.84-1.08)	0.463	215 (72)	0.90 (0.80-1.02)	0.114
WBC	215 (38)	1.00 (1.00-1.00)	0.174	215 (79)	1.00 (1.00-1.00)	0.176	215 (72)	1.00 (1.00-1.00)	**0.043**
FIBRINOGEN	188 (31)	1.002 (0.999-1.005)	0.189	188 (68)	1.002 (1.000-1.005)	**0.015**	188 (60)	1.002 (1.000-1.004)	0.106
SCC-Ag	212 (38)	1.001 (0.985-1.016)	0.931	212 (78)	1.014 (1.006-1.021)	**<.001**	212 (72)	1.012 (1.004-1.019)	**0.004**
CEA	206 (36)	0.9995 (0.995-1.004)	0.834	206 (75)	1.001 (0.9998-1.002)	0.111	206 (70)	1.001 (1.000-1.002)	0.082

Table 4. Multivariable analyses for survival outcomes with clinicopathological characteristic factors

Characteristics	LRFS			OS			DFS		
	N (event)	HR (95% CI)	p-value	N (event)	HR (95% CI)	p-value	N (event)	HR (95% CI)	p-value
EpCAM									
0	26 (1)	1		26 (9)	1		26 (7)	1	
1+2+3 (≥5%)	176 (35)	4.31 (0.59-31.73)	0.152	176 (67)	1.36 (0.66-2.79)	0.41	176 (63)	2.30 (0.89-5.96)	0.086
Age									
age≥40	179 (25)			179 (65)			179 (53)	1	
age<40	36 (13)			36 (14)			36 (19)	2.76 (1.37-5.56)	**0.005**
FIGO stage									
~IIB	174 (28)			174 (48)	1		174 (44)	1	
III/ IVA	41 (10)			41 (31)	5.00(2.92-8.56)	**<0.0001**	41 (28)	5.26 (2.85-9.72)	**<.0001**
Tumor size									
< 4 cm	92 (12)			92 (30)			92 (21)	1	
≥ 4 cm	123 (26)			123 (49)			123 (51)	2.55 (1.31-4.98)	**0.006**
Smoking									
present	26 (4)			26 (8)			26 (8)		
ex-smoker	19 (7)			19 (5)			19 (10)		
non-smoker	170 (27)			170 (66)			170 (54)		
Ccrt									
yes	196 (36)			196 (67)			196 (64)		
no	19 (2)			19 (12)			19 (8)		
Histology									
SCC	196 (32)			196 (69)			196 (61)	1	
AD+ADS	19 (6)			19 (10)			19 (11)	3.43 (1.45-8.13)	**0.005**
Differentiation									
WD +MD	173 (24)	1		173 (57)	1		173 (47)	1	
PD	38 (13)	4.33 (2.04-9.16)	**0.0001**	38 (19)	2.83(1.57-5.09)	**0.0005**	38 (23)	2.63 (1.39-4.97)	**0.003**
HB	215 (38)			215 (79)			215 (72)		
WBC	215 (38)			215 (79)			215 (72)		
FIBRINOGEN	188 (31)			188 (68)			188 (60)		
SCC-Ag	212 (38)			212 (78)			212 (72)		
CEA	206 (36)			206 (75)			206 (70)		

Figure 5. Kaplan-Meier for Survival Probability by pathologist's score. Loco-regional recurrence-free survival (LRFS). The P-value for the log-rank test.

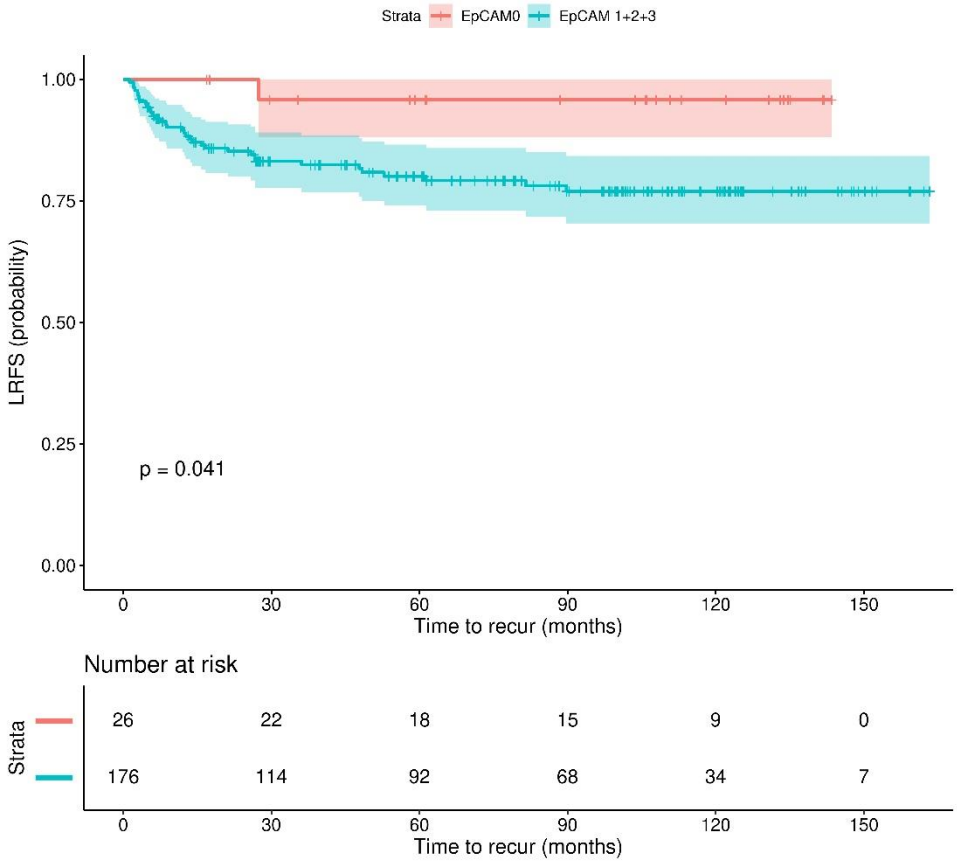


Figure 6. Kaplan-Meier for Survival Probability by pathologist's score. Disease-free survival (DFS). The P-value for the log-rank test.

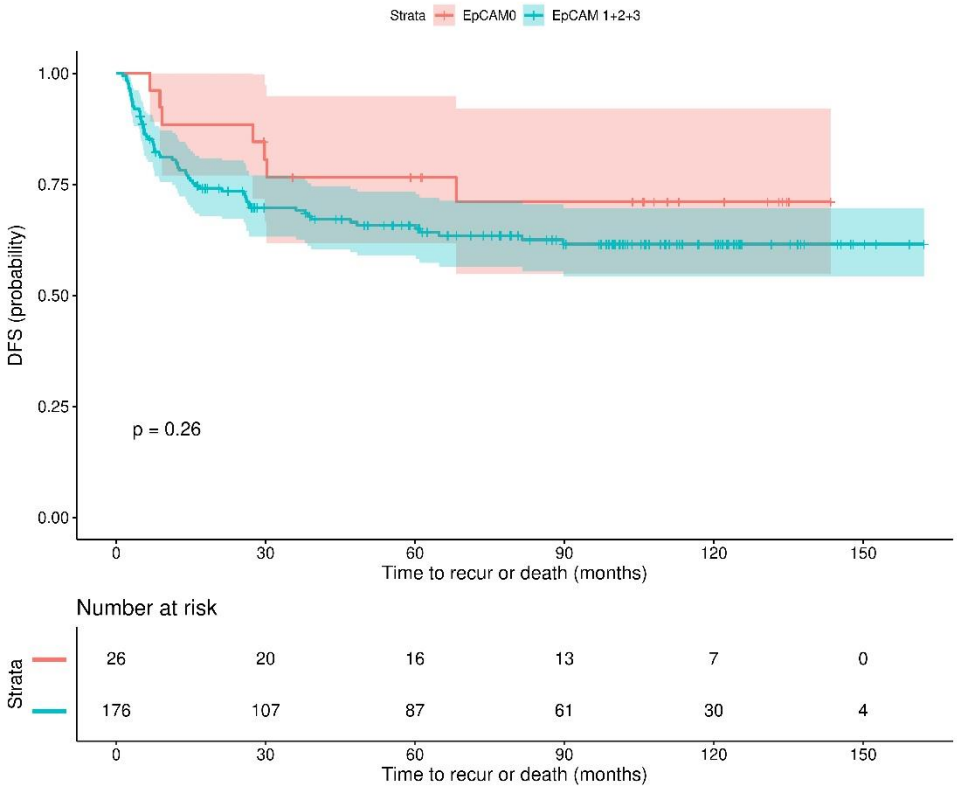


Figure 7. Kaplan-Meier for Survival Probability by pathologist's score. Overall survival rate (OS). The P-value for the log-rank test.

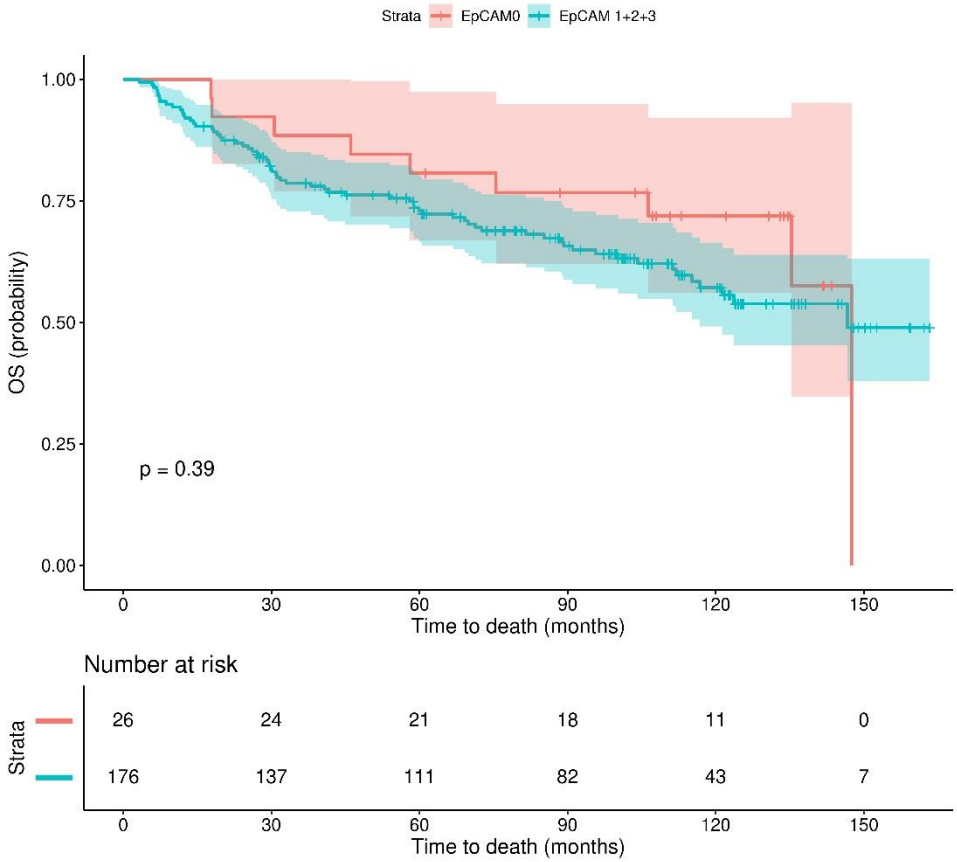
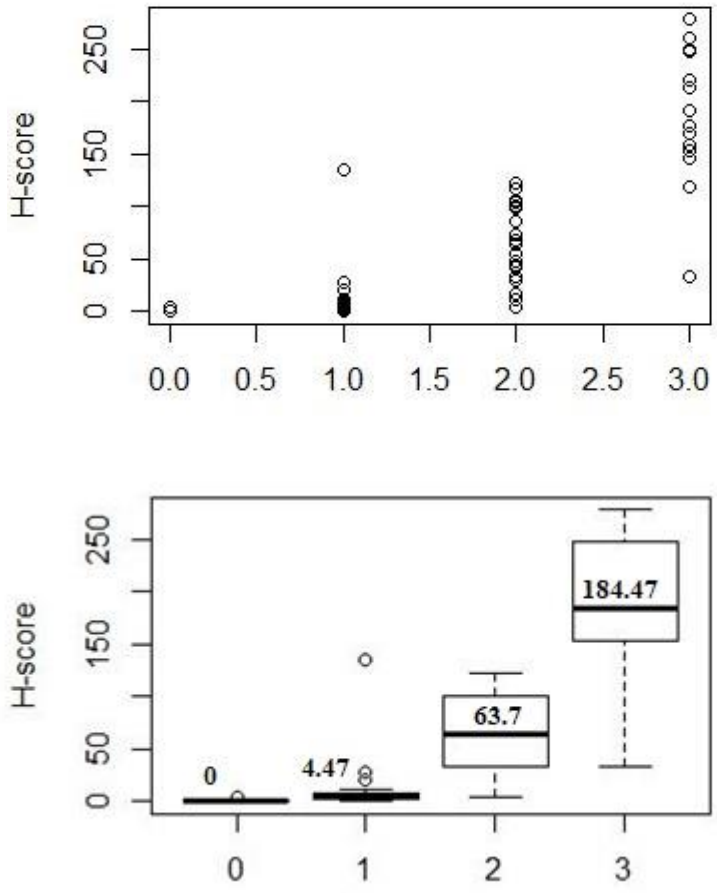


Figure 8. Exploratory Analysis for Correlation between pathologist's score and H-score (n=75)



4. DISCUSSION

In recent years, along with the rapid development of science and technology, the discovery of biomarkers of cancer has, in general, draw a great deal of interest from scientists as well as clinical practitioners because not only are they helpful for diagnosis and detection of life-threatening illnesses at an early stage, monitoring disease progression, but also for predicting disease recurrence and therapeutic efficacy. An ideal tumor marker would be a protein or a fragment of the protein that can be detected easily in the blood or urine of cancer patients but not detected in a healthy person. In this study, we investigated the association between the expression of EpCAM and prognosis of cervical cancer after radiation treatment. Between two groups of EpCAM developed from the pathologist's score for analysis, there were 176 cases observed in EpCAM positive group (IHC score 1+2+3) compared to 26 cases in the negative group (IHC score 0). Our data revealed that EpCAM showed a tendency being associated with locoregional recurrence of cervical cancer after radiotherapy in the log-rank test (P value 0.041). In contrast, when this relation was investigated in the cox regression model, the result did not reach significance in both univariable and multivariable tests. Hence, the predictive value of EpCAM for the prognosis of cervical cancer patients after radiation therapy could not be strongly determined in our study. The log-rank test generated statistically significant result but univariable model did not provide the same while usually these two approaches often produce results that are not much different from each other because they merely examine the relationship of one

main variable of interest with dependent variable without the participation of other factors. This might be explained by the limitation in the number of events that occurred for LRFS which plays an important role in cox regression examination, resulting in inaccurate test outcome and bias in data interpretation. For this reason, further researches should be conducted on samples with a higher number of events for each survival endpoint so that a more definitive conclusion could be made. Since EpCAM was first discovered in 1979, it has been widely studied to assess the function as a biomarker in a variety of cancer types. Many investigations so far have confirmed that high expression levels of EpCAM usually correlate with poor prognosis, e.g. in breast cancer and ovarian cancer as well as in pancreatic, urothelial, and gallbladder carcinoma [19-24] although there are exceptional cases that are renal and thyroid carcinoma, where high levels of EpCAM have been shown to concern with increased survival [31, 32]. In several carcinomas types such as gastric cancer or colorectal cancer, EpCAM has even reported to plays a dual role, either promoting or reducing cancer progression [33]. In cervical cancer, a higher concentration than normal level of this protein was found [25, 34], leading to an assumption that EpCAM may have diagnostic and prognostic values. We have therefore tried to conduct this analysis to verify the hypothesis. Although it was not possible to ascertain the relationship between EpCAM and the patient's prognosis based on our 215 subject cohort, the results of our investigation at least showed an association to some extent of the relationship between EpCAM expression and locoregional recurrence. Additionally, one of the remarkable points in our study was the disproportional distribution of the patient’s number with AD

and ADS according to EpCAM expression classification. As can be seen in Table 2, all patients with AD and ADS were in the EpCAM positive group and there were no cases in EpCAM negative group. AD and ADS, in many comparative studies with SCC, have shown to obtain a worse prognosis [13, 35, 36]. One of the examples for this trend was a study conducted in a 3678 patient cohort; when comparing 5-year survival rate between SCC versus AD and ADS by tumor stage, Chen et al. indicated that the more the stage increases, the difference in 5-year survival rate between SCC versus AD and ADS becomes even more pronounced [36]. Therefore, the relationship between EpCAM positive cases and this histological type of cervical cancer suggests that EpCAM positive expression might be related to poor prognosis in patients with uterine cervix cancer. However, because the number of patients with AD and ADS is extremely small compared to that of SCC due to the high prevalence of SCC in cervical cancer (approximately 80%); thus, it is recommended that the link should be further considered carefully in future studies those posse larger number of AD and ADS or those are solely carried out on patients with this subtype of cancer.

In addition, in the present study, our data also indicated a statistically significant association between EpCAM with some other biological and clinicopathological characteristics. They are tumor size, fibrinogen, and SCC-Ag level. Size of tumor is a well-known traditional clinicopathological factor that has been shown to be a consistent diagnostic and prognostic marker in cervical cancer. Results of both investigations of Berek, J.S., et al.[13] and Eifel, P.J., et al.[37]

showed that the 5-year survival rate of patients significantly dropped from 88%–97% for tumors with a size smaller than 2–3 cm to 50%–62% for those greater than 4–5 cm. Towards fibrinogen, this protein was investigated as a biomarker in cervical cancer due to the assumption that their increased concentration in plasma was derived from tumor cells and their tumor-promoting effects were shown through microenvironment instead of just a response to the tumor. In their research, Polterauer, S., et al. proved this assumption and concluded that fibrinogen plasma levels can serve as an additional independent prognostic parameter in patients with cervical cancer [38]. In a recent study conducted by Zhao, K., et al., fibrinogen was combined with platelet to examine the correlation with clinicopathological characteristics in cervical cancer. The researcher found that fibrinogen levels, but not platelet levels, were an independent prognostic factor for poor survival in early-stage patients [39]. As a result, the finding of the positive association between EpCAM expression and these two factors (P value = 0.024 and 0.045 for tumor size and fibrinogen, respectively) in our investigation once again strengthened the hypothesis about EpCAM's prognostic value in patients with cervical cancer. However, in contrast with the finding above, EpCAM expression exhibited a reverse relationship with SCC-Ag since higher concentration of SCC-Ag level was witnessed in the negative group of EpCAM (P value = 0.008). This reduces the certainty of our hypothesis about EpCAM because SCC-Ag is an antigen found in normal cervix epithelium with its elevated serum level can be detected in approximately 60% of patients with SCC [40]. In several studies, SCC-Ag proved a predictive value for the prognosis of cervical carcinomas and it is also concluded

to be useful for the need to monitor patients after primary treatments [41-43]. Our data also indicated that SCC-Ag was significantly related to poor OS (P<0.001) and DFS (P=0.004) in univariable analysis. Therefore, later researches should also further clarify this issue.

Our study also examined the correlation between pathologist's immunohistochemical (IHC) grading of EpCAM and H-score analyzed by an image analysis software in a subset of 75 patients. H-score was shown to increase with the increasing intensity score with the Spearman's correlation coefficient of 0.914 (P value <0.001) and Kappa coefficient of 0.75 (0.63-0.87). To our best knowledge, H-score still has not been used commonly to evaluate EpCAM IHC for cervical cancer patients in Korea although this updated scoring system was developed for the first time in 2003 by Hirsch et al. [26]. Therefore, more researches in the future are necessary to be able to strongly assert the use in a common way of H-score in clinical practice. Even until now, IHC data scored by pathologists is still considered as a gold standard in clinical practice. In this method, the expert pathologist investigates qualitatively the stained tissue slide under a microscope to provide score; which is clinically used for therapeutic decision making. Such qualitative judgment is time-consuming and more often suffers from interobserver bias because the decision of a pathologist may vary from one pathologist to another depending on the experience. For this reason, automated IHC measurements promise to overcome these limitations and there has been many researches conducted to compare the effectiveness of an automated IHC scoring

with traditional pathologist's scoring. For example, to verify the utility of the H-score, Jung, Y., et al. compared it with manually constructed numerical values graded by a pathologist and finally concluded that using H-score was reasonable for profiling protein expressions [44]. Besides, In a multicenter cohort study of Debaugnies, F., et al. conducted in 2016, the performances of the H-score was compared versus the adapted hemophagocytic lymphohistiocytosis (HLH)–2004 guidelines which are still the most widely used criteria to define and diagnose HLH in clinical practice in Adult and Pediatric Patients. Interestingly, the diagnostic sensitivity and specificity were even higher for H-score compared to HLH-2004 guidelines, 100% and 80% for children and 90% and 79% for adults, respectively [29]. Our study provided more information on this platform with evidence of a high correlation between H-score which is generated by a software and expert's score. With the positive results demonstrated in the above studies, this scoring approach nowadays is also widely accepted and recommended by leading associations and organizations [30] because it is not only precise in ranges of staining that are difficult to be judged with the bare eyes [45] but it is also able to produce continuous data [46].

5. CONCLUSION

The data revealed that EpCAM showed a tendency being associated with locoregional recurrence in cervical cancer after radiotherapy. H-score and pathologist's score showed a strong correlation suggesting this new scoring process can be used as an assistive tool or to replace the manual scoring process because this method reduces the workload burden for pathologists as well as the bias of interobserver.

BIBLIOGRAPHY

1. Bray, F., et al., *Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries.* CA Cancer J Clin, 2018. **68**(6): p. 394-424.
2. Miller, K.D., et al., *Cancer treatment and survivorship statistics, 2016.* CA Cancer J Clin, 2016. **66**(4): p. 271-89.
3. Jung, K.W., et al., *Cancer Statistics in Korea: Incidence, Mortality, Survival, and Prevalence in 2015.* Cancer Res Treat, 2018. **50**(2): p. 303-316.
4. Lim, M.C., et al., *Incidence of cervical, endometrial, and ovarian cancer in Korea, 1999-2010.* J Gynecol Oncol, 2013. **24**(4): p. 298-302.
5. Kweon, S.S., *Updates on Cancer Epidemiology in Korea, 2018.* Chonnam Med J, 2018. **54**(2): p. 90-100.
6. Hasan, H., *cervical cancer: current and emerging trends in detection and treatment.* The Rosen Publishing Group, Inc., 2009.
7. Kurman RJ, C.M., Herrington CS, Young RH. WHO Classification of Tumours of Female Reproductive Organs, 4th edn. Lyon: IARC; 2014., 2014.
8. Bhatla, N., et al., *Cancer of the cervix uteri.* Int J Gynaecol Obstet, 2018. **143 Suppl 2**: p. 22-36.

9. Lim, M.C. and M. Lee, *Practice guidelines for management of cervical cancer in Korea: a Korean Society of Gynecologic Oncology Consensus Statement.* 2017. **28**(3): p. e22.

10. Landoni, F. and A. Colombo, *Randomized study between radical surgery and radiotherapy for the treatment of stage IB-IIA cervical cancer: 20-year update.* 2017. **28**(3): p. e34.

11. Slomovitz, B.M., et al., *Uterine papillary serous carcinoma (UPSC): a single institution review of 129 cases.* Gynecol Oncol, 2003. **91**(3): p. 463-9.

12. Matsuo, K., et al., *Validation of the 2018 FIGO cervical cancer staging system.* Gynecol Oncol, 2019. **152**(1): p. 87-93.

13. Berek, J.S., et al., *Adenocarcinoma of the uterine cervix: histologic variables associated with lymph node metastasis and survival.* Obstet Gynecol, 1985. **65**(1): p. 46-52.

14. Herlyn, M., et al., *Colorectal carcinoma-specific antigen: detection by means of monoclonal antibodies.* Proc Natl Acad Sci U S A, 1979. **76**(3): p. 1438-42.

15. Balzar, M., et al., *The biology of the 17-1A antigen (Ep-CAM).* J Mol Med (Berl), 1999. **77**(10): p. 699-712.

16. Schnell, U., V. Cirulli, and B.N. Giepmans, *EpCAM: structure and function in health and disease.* Biochim Biophys Acta, 2013. **1828**(8): p. 1989-2001.

17. Trzpis, M., et al., *Epithelial cell adhesion molecule: more than a carcinoma marker and adhesion molecule.* Am J Pathol, 2007. **171**(2): p. 386-95.

18. Cirulli, V., et al., *KSA antigen Ep-CAM mediates cell-cell adhesion of pancreatic epithelial cells: morphoregulatory roles in pancreatic islet development.* J Cell Biol, 1998. **140**(6): p. 1519-34.

19. Spizzo, G., et al., *Prognostic significance of Ep-CAM AND Her-2/neu overexpression in invasive breast cancer.* Int J Cancer, 2002. **98**(6): p. 883-8.

20. Spizzo, G., et al., *High Ep-CAM expression is associated with poor prognosis in node-positive breast cancer.* Breast Cancer Res Treat, 2004. **86**(3): p. 207-13.

21. Varga, M., et al., *Overexpression of epithelial cell adhesion molecule antigen in gallbladder carcinoma is an independent marker for poor survival.* Clin Cancer Res, 2004. **10**(9): p. 3131-6.

22. Scheunemann, P., et al., *Occult tumor cells in lymph nodes as a predictor for tumor relapse in pancreatic adenocarcinoma.* Langenbecks Arch Surg, 2008. **393**(3): p. 359-65.

23. Brunner, A., et al., *EpCAM overexpression is associated with high-grade urothelial carcinoma in the renal pelvis.* Anticancer Res, 2008. **28**(1a): p. 125-8.

24. Fong, D., et al., *Ep-CAM expression in pancreatic and ampullary carcinomas: frequency and prognostic relevance.* J Clin Pathol, 2008. **61**(1): p. 31-5.

25. Chantima, W., et al., *EpCAM expression in squamous cell carcinoma of the uterine cervix detected by monoclonal antibody to the membrane-proximal part of EpCAM.* BMC Cancer, 2017. **17**(1): p. 811.

26. Hirsch, F.R., et al., *Epidermal growth factor receptor in non-small-cell lung carcinomas: correlation between gene copy number and protein expression and impact on prognosis.* J Clin Oncol, 2003. **21**(20): p. 3798-807.

27. Mazieres, J., et al., *Evaluation of EGFR protein expression by immunohistochemistry using H-score and the magnification rule: re-analysis of the SATURN study.* Lung Cancer, 2013. **82**(2): p. 231-7.

28. Pirker, R., et al., *EGFR expression as a predictor of survival for first-line chemotherapy plus cetuximab in patients with advanced non-small-cell lung cancer: analysis of data from the phase 3 FLEX study.* Lancet Oncol, 2012. **13**(1): p. 33-42.

29. Debaugnies, F., et al., *Performances of the H-Score for Diagnosis of Hemophagocytic Lymphohistiocytosis in Adult and Pediatric Patients.* Am J Clin Pathol, 2016. **145**(6): p. 862-70.

30. Fitzgibbons, P.L., et al., *Template for reporting results of biomarker testing of specimens from patients with carcinoma of the breast.* Arch Pathol Lab Med, 2014. **138**(5): p. 595-601.

31. Ralhan, R., et al., *EpCAM nuclear localization identifies aggressive thyroid cancer and is a marker for poor prognosis.* BMC Cancer, 2010. **10**: p. 331.

32. Seligson, D.B., et al., *Epithelial cell adhesion molecule (KSA) expression: pathobiology and its role as an independent predictor of survival in renal cell carcinoma.* Clin Cancer Res, 2004. **10**(8): p. 2659-69.

33. van der Gun, B.T., et al., *EpCAM in carcinogenesis: the good, the bad or the ugly.* Carcinogenesis, 2010. **31**(11): p. 1913-21.

34. Richter, C.E., et al., *Primary cervical carcinoma cell lines overexpress epithelial cell adhesion molecule (EpCAM) and are highly sensitive to immunotherapy with MT201, a fully human monoclonal anti-EpCAM antibody.* Int J Gynecol Cancer, 2010. **20**(9): p. 1440-7.

35. Zhou, J., et al., *Comparison of clinical outcomes of squamous cell carcinoma, adenocarcinoma, and adenosquamous carcinoma of the uterine cervix after definitive radiotherapy: a population-based analysis.* J Cancer Res Clin Oncol, 2017. **143**(1): p. 115-122.

36. Chen, R.J., et al., *Influence of histologic type and age on survival rates for invasive cervical carcinoma in Taiwan.* Gynecol Oncol, 1999. **73**(2): p. 184-90.

37. Eifel, P.J., et al., *Adenocarcinoma of the uterine cervix. Prognosis and patterns of failure in 367 cases.* Cancer, 1990. **65**(11): p. 2507-14.

38. Polterauer, S., et al., *Fibrinogen plasma levels are an independent prognostic parameter in patients with cervical cancer.* Am J Obstet Gynecol, 2009. **200**(6): p. 647.e1-7.

39. Zhao, K., et al., *Prognostic significance of pretreatment plasma fibrinogen and platelet levels in patients with early-stage cervical cancer.* Gynecol Obstet Invest, 2015. **79**(1): p. 25-33.

40. Farghaly, S.A., *Tumor markers in gynecologic cancer.* Gynecol Obstet Invest, 1992. **34**(2): p. 65-72.

41. Bolli, J.A., et al., *Squamous cell carcinoma antigen: clinical utility in squamous cell carcinoma of the uterine cervix.* Gynecol Oncol, 1994. **55**(2): p. 169-73.

42. Hong, J.H., et al., *The prognostic significance of pre- and posttreatment SCC levels in patients with squamous cell carcinoma of the cervix treated by radiotherapy.* Int J Radiat Oncol Biol Phys, 1998. **41**(4): p. 823-30.

43. Gaarenstroom, K.N., et al., *Can initial serum cyfra 21-1, SCC antigen, and TPA levels in squamous cell cervical cancer predict lymph node metastases or prognosis?* Gynecol Oncol, 2000. **77**(1): p. 164-70.

44. Jung, Y., et al., *Identification of prognostic biomarkers for glioblastomas using protein expression profiling.* Int J Oncol, 2012. **40**(4): p. 1122-32.

45. Rimm, D.L., *What brown cannot do for you.* Nat Biotechnol, 2006. **24**(8): p. 914-6.

46. Rimm, D.L., et al., *Bimodal population or pathologist artifact?* J Clin Oncol, 2007. **25**(17): p. 2487-8.

ACKNOWLEDGEMENT

First, I would like to express gratitude to my advisor Professor Joo-Young Kim for giving me incredible knowledge and advice throughout my master's degree. Without her guidance and persistent support, this dissertation would not be possible.

I would also like to give sincere appreciation to Professor Chong Woo Yoo and Professor Joon Mee Kim for their help in many ways throughout my master's training. Moreover, it would be an omission if I do not thank the committee member Professor Mi Kyung Kim for her insightful comments in evaluating my thesis, which fulfilled my study from various perspectives.

My special thanks to Miss Eun Young Park, my dear sister Hye-Jin Shin, my little brother Tung Hoang and my friends in GCSP for treating me so kind and for their assistance throughout my school life.

Last but not least, from the bottom of my heart, I would like to send my deepest appreciation to my family in Vietnam, my wife and my daughter for their unconditional support and belief not only in this journey but also in everything of my life.