Evaluation of the budding and depth of invasion (BD) model in oral tongue cancer biopsies

Alhadi Almangush¹,²,³, Ilmo Leivo⁴, Maria Siponen⁵,⁶,⁷, Elias Sundquist⁵,⁸, Rayan Mroueh⁹, Antti A. Mäkitie⁹, Ylermi Soini¹⁰, Caj Haglund¹¹,¹², Pentti Nieminen¹³, Tuula Salo²,⁵,⁸,¹⁴

¹Department of Pathology, University of Helsinki, Helsinki, Finland
²Department of Oral and Maxillofacial Diseases, University of Helsinki, Helsinki, Finland
³Institute of Dentistry, University of Misurata, Misurata, Libya
⁴Department of Pathology, University of Turku, Turku, Finland
⁵Cancer and Translational Medicine Research Unit, University of Oulu, Oulu, Finland
⁶Department of Oral and Maxillofacial Diseases, Kuopio University Hospital, Kuopio, Finland.
⁷Institute of Dentistry, University of Eastern Finland, Kuopio, Finland
⁸Medical Research Center, Oulu University Hospital, Oulu, Finland.
⁹Department of Otorhinolaryngology– Head and Neck Surgery, Helsinki University Hospital and University of Helsinki, Helsinki, Finland.
¹⁰Department of Pathology and Forensic Medicine, University of Eastern Finland, Kuopio and Cancer Center of Eastern Finland, Kuopio, Finland.
¹¹Department of Surgery, University of Helsinki and Helsinki University Hospital, Helsinki, Finland.
¹²Research Programs Unit, Translational Cancer Biology, University of Helsinki, Helsinki, Finland
¹³Medical Informatics and Statistics Research Group, University of Oulu, Oulu, Finland.
¹⁴HUSLAB, Department of Pathology, Helsinki University Central Hospital, University of Helsinki, Helsinki, Finland.

Word count: 2238

Corresponding author: Alhadi Almangush DDS, PhD
Department of Pathology, University of Helsinki, Helsinki, Finland
E-mail: alhadi.almangush@helsinki.fi
Telephone number: 00358 45 2044668
Abstract

It is of great clinical importance to identify simple prognostic markers from preoperative biopsies that could guide treatment planning. Here, we compared tumor budding (B), depth of invasion (D) and the combined scores (i.e. BD histopathologic model) in preoperative biopsies and the corresponding postoperative specimens of oral tongue squamous cell carcinoma (OTSCC). Tumor budding and depth of invasion were evaluated in the pre- and postoperative samples from 100 patients treated for OTSCC. Sensitivity and specificity statistics were used. Our results showed statistically significant ($P < 0.001$) relationship between pre- and postoperative BD scores. There was an agreement between the pre- and postoperative BD model scores in 83 cases (83%) with 57.1% sensitivity (95% CI: 39.4% to 73.7%) and 96.9% specificity (95% CI: 89.3% to 99.6%). Our findings suggest that the BD model, analyzed from representative biopsies, could be used for the treatment planning of OTSCC.

Key words: Oral tongue cancer, Tumor budding, Tumor depth, BD model, Prognosis
Introduction

Oral or mobile tongue squamous cell carcinoma (OTSCC) has shown increased incidence in several countries [1]. Aggressive behavior and poor prognosis are reported even at early stages of the tumor [2, 3]. A preoperative biopsy is routinely obtained for histopathologic diagnosis of suspicious tongue lesions. Although several prognostic markers for OTSCC have been published, there is still a lack of validated markers that could easily be evaluated in preoperative OTSCC histological sections. Therefore, identification of prognostic marker(s) in biopsy specimens would be a valuable tool for treatment planning (local resection with or without the neck dissection).

We have previously introduced the budding and depth (BD) histopathologic model as a prognostic tool in OTSCC [4]. The prognostic value of this model has been validated in cohorts of oral squamous cell carcinomas (OSCC) [5, 6]. In these studies, the BD model was shown to have superior prognostic power when compared to the other previously introduced histopathologic grading systems, such as WHO grading [7], malignancy grading of the deep invasive margins [8], and histological risk score [9]. Additionally, tumor budding is associated with the progression and prognosis of several epithelial cancers, such as head and neck [10], esophageal [11], colorectal [12], pancreatic [13], lung [14] and breast [15]. Specifically in OTSCC, budding correlates with occult cervical lymph node metastasis and poor prognosis [16, 17]. Similarly, the depth of invasion is a prognostic marker for OTSCC [17]. Recently, pre- and postoperative samples were compared in a study of 91 OSCC cases, and it was shown that both budding and tumor depth correlated significantly with relapse-free survival [18]. To our knowledge, however, there is no sizeable cohort where the BD model has been tested in OTSCC biopsies and compared to
the corresponding postoperative OTSCC samples. The aim of this study was to analyze the
sensitivity and specificity of preoperative BD scores of hematoxylin and eosin stained
OTSCC biopsies compared to the postoperative BD scores of the corresponding cases.

Material and Methods

Hematoxylin and eosin (HE) stained slides from pre- and postoperative samples of 145
patients diagnosed with OTSCC at the University Hospitals of Helsinki, Kuopio and Oulu
between the years 1981 and 2016 were retrieved for this study. The use of pre- and
postoperative samples and the data inquiry was approved by the ethics committees of
Helsinki, Kuopio and Oulu University Hospitals. All patients were diagnosed by incisional
biopsy and treated by surgical excision of the tumor. Patients without either pre- or
postoperative counterparts available were excluded. Cases received preoperative therapy
were also excluded. A total of 100 cases were eligible for the comparative analyses.

Tumor budding (B) was defined as the presence of single cancer cell or cluster of
less than five cancer cells. The invasive front (IF) was evaluated under low magnification
(×4) and then the field with highest density of tumor budding was counted under high
magnification (×20). The depth of tumor invasion (D) was measured from the surface of
the tumor to the deepest point of invasion. The scoring was performed by an independent
researcher (AA) and reviewed by an experienced head and neck pathologist (IL). BD scores
were assigned as previously described [4] (Fig. 1). In brief, score 0 refers to <5 buds at the
IF and <4 mm in depth. Score 1 refers to either presence of ≥5 buds at the IF or a deep
tumor of ≥4 mm in depth. Score 2 refers to the presence of ≥5 buds at the IF and a deep
tumor of ≥4 mm in depth.
Statistical analysis: All analyses were performed with IBM SPSS version 20. The statistical significance of the relationship between pre- and postoperative measures was evaluated using chi-square test. For sensitivity and specificity statistics with their 95% confidence intervals (95% CI), BD scores of low and intermediate were combined together (low and intermediate vs. high) to evaluate the predictive power of preoperative score for the postoperative score of the corresponding sample.

Results

Patient Characteristics

One case received preoperative therapy and was therefore excluded from our analysis. A total of 100 patients were enrolled in the statistical analyses of the study. There were 51 males (51.0%). Stage distribution was as follows: 41 cases (41.0%) were assigned as stage I, 40 (40.0%) as stage II, 9 (9.0%) as stage III, and 10 (10.0%) as stage IV. The mean age at diagnosis was 60.8 years (range 27 to 91 years). All tumors were located on the oral tongue (OTSCC).

Histopathologic correlation between biopsy specimens and surgical resection specimens

Tumor budding (B)

The number of tumor budding in biopsies ranged from 0 to 13 buds (median 1, mean 3.5); and for corresponding postoperative samples from 0 to 17 buds (median 3, mean 3.9). Of the cases, 82 (82%) had the same B category (low <5 buds or high ≥5 buds) in pre- and postoperative samples. The association between pre- and postoperative B was statistically
significant ($P$ value of chi-square test $< 0.001$). The preoperative scores showed a good sensitivity of 59.1% (95% CI: 43.3% to 73.7%) and a high specificity of 100% (95% CI: 93.6% to 100%) in predicting the postoperative score of the same case (Table 1).

**Depth of invasion (D)**

In biopsy specimens, depth values ranged from 0.5 to 10 mm (mean 4.1 mm, median 4 mm); and for the corresponding postoperative samples ranged from 0.5 to 23 mm (mean 6.3 mm, median 6 mm). Of the cases, 77 (77.0%) had the same D category (superficial $< 4$ mm or deep $\geq 4$ mm) in pre- and postoperative samples. The relationship between the pre- and postoperative D value was statistically significant ($P$ value of chi-square test $< 0.001$). The preoperative measurement showed a high predictive power of postoperative measurement with 77.1% sensitivity (65.6% to 86.3%) and 76.7% specificity (57.2% to 90.1%) (Table 1).

**BD model**

For preoperative samples, 35 cases (34.7%) had BD score 0, 43 cases (42.6%) had score 1, and 23 cases (22.8%) had score 2. In postoperative samples, 21 cases (20.8%) had score 0, 44 cases (43.6%) had score 1, and 36 cases (35.6%) had score 2. There was a significant association between scores of BD model and cTNM stage (two-sided $P= 0.001$). The BD histological model showed a highly significant relationship between pre- and postoperative measurements ($P$ value of chi-square test $< 0.001$). There was an agreement between the pre- and postoperative scores of the BD model in 83 cases (83.0%) with 57.1% sensitivity (95% CI: 39.4% to 73.7%) and 96.9% specificity (95% CI: 89.3% to 99.6%) (Table 1).
Discussion

During the histopathologic evaluation of OTSCC, pathologists attempt to identify histopathological prognostic markers. Identification of such markers, especially in early stage tumors, could guide clinical treatment decisions. Recently, our group suggested the BD model as a prognostic tool in a large multicenter cohort of OTSCC [4]. This model is now also validated in other cohorts of OSCC [5, 6]. In the multivariate analysis of these previous studies, BD model showed superior prognostic power compared to the other parameters. In this study, we demonstrated a significant relationship between the BD scores in pre- and postoperative OTSCC samples. This finding is particularly useful for making treatment decisions at an early stage, but occasionally highly aggressively behaving cases.

The use of BD model in daily practice might provide a reliable additional prognostic tool that could overcome the shortcoming of currently used preoperative tumor size staging (T) and histopathological biopsy grading. Both of these commonly used preoperative analyses, tumor clinical size measurement and cancer cells differentiation grading, have been criticized for their low prognostic power of the cancer [2, 19].

Cancer cell can invade individually or in a collection of cell clusters [20]. Different patterns for head and neck cancers invasion has been suggested, including worst pattern of invasion (WPOI) and tumor budding. WPOI was introduced as a part of histologic risk model [21], and it was shown as a useful prognostic marker in early OTSCC [2]. However, tumor satellites, which represent type 5 WPOI and are defined as tumor island/s located one mm or more away from the main tumor or next closest satellite, require the evaluation of all tumor sections [21] and thus this score remains inapplicable for biopsy specimen
analyses. On the other hand, tumor budding is a recently introduced histopathologic pattern which has been reported as a promising prognosticator in several carcinomas [10-13, 15] and has been successfully evaluated in preoperative biopsy [18, 22]. A five-bud cutoff point has widely been used in OSCC [17, 23-25] and other cancers [26, 27] to stratify the tumors into low risk (<5 buds) or high risk groups (≥5 buds).

Depth of invasion has been reported as a significant prognosticator in OTSCC [17, 28, 29]. The cutoff point of 4mm depth is widely accepted and has been validated in recent OTSCC studies [2, 30, 31]. Of note, a meta-analysis has also concluded that 4mm would be an optimal cutoff point [32]. In this study, we found a good correlation between the depth in pre-operative and postoperative samples when we stratified the cases into two categories (superficial <4mm vs. deep ≥4mm). However, when the exact measurements (i.e. quantitative) of pre- and postoperative depth were compared, the correlations were low. This was expected as in postoperative samples the measurement could be taken at several sites of the cancer sections, while in the preoperative biopsies the measurement is possible only from a limited tumor area. More importantly, in the present series both measurements were in the same category (superficial <4mm or deep ≥4mm) in 77% of the cases. For the remaining cases, low quality biopsies (e.g. superficial samples missing the deepest part of the tumor) did not allow accurate measurement of the invasion depth. The validity of preoperative tumor depth evaluation by ultrasonography or magnetic resonance imaging (or both) has been confirmed in many studies [33-36]. This should be considered as a surrogate method in case the entire tumor thickness is unclear in the preoperative biopsy. Additionally, the depth of invasion evaluation from fresh-frozen intraoperative
sections has shown a strong association with the postoperative measurement [37]. Such procedures could also reduce the inaccuracy in the preoperative biopsy measurement.

Similar to our multicenter Finnish study of 100 OTSCC patients, a Japanese group has recently published a study of 91 OSCC cases from tongue and floor of mouth [18]. They concluded that the budding scores in particular showed a significant correlation between biopsies and corresponding resected specimens. Such a correlation was also observed with the depth measurements, but with less accuracy. These two separate cohorts both highlight the usefulness of preoperative evaluation of the budding, and in cases of representative, sufficiently deep biopsies, also the depth of invasion. Of note, the results of our current study are based on SCC cases from the mobile tongue only, a subsite of oral cavity, in which SCCs are mostly human papilloma virus (HPV) negative [38, 39]. In contrast, HPV-positive head and neck SCCs most commonly occur in the oropharynx (including base of the tongue) and are reported to have a favorable prognosis [40].

All our cases that had different scores (about 17%) in the BD model in preoperative compared to postoperative samples had non-representative biopsies. These biopsies were often badly fragmented, too superficial, or had some technical artifacts, such as tangential cutting. Therefore, we strongly recommend that clinicians carefully take a large (at least 4 mm wide and 4 mm deep) biopsy (or several biopsies from different parts of the tumor) that includes the deepest part of the tumor. A high quality biopsy would allow the pathologist to evaluate the BD model accurately. However, if the quality of the biopsy is low (too shallow, fragmented, or not in the deepest area), the BD model evaluation would be inadequate.
Several previous findings have shown that the BD model is a simple and predictive histopathological grading system for OSCC patients [4-6, 18]. Here, we demonstrated that in satisfactory biopsies, the BD model can be evaluated from HE-stained slides, and the BD scores significantly corresponded to the scores of postoperative tumor resection samples.

Compliance with ethical standards: Institutional Review Board approval was obtained from the ethics committees of Helsinki, Kuopio and Oulu University Hospitals.

Funding: This study was supported by Finnish Dental Society (Alhadi Almangush), Maritza and Reino Salonen Foundation (Ilmo Leivo) and the Finnish Cancer Society (Tuula Salo).

Conflict of interest: The authors declare that they have no conflict of interest.

References


expression are useful predictors of nodal involvement in T1 esophageal squamous cell carcinoma. Anticancer Res 33:5023-5029


### Table 1. Distribution of cases according to preoperative and postoperative scores

<table>
<thead>
<tr>
<th>Preoperative budding</th>
<th>Postoperative budding</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low (%)</td>
<td>High (%)</td>
</tr>
<tr>
<td>Low</td>
<td>56 (75.7)</td>
<td>18 (24.3)</td>
</tr>
<tr>
<td>High</td>
<td>0</td>
<td>26 (100)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Preoperative depth</th>
<th>Postoperative depth</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Superficial (%)</td>
</tr>
<tr>
<td></td>
<td>n (%)</td>
</tr>
<tr>
<td>Superficial</td>
<td>23 (59.0)</td>
</tr>
<tr>
<td>Deep</td>
<td>7 (11.5)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Preoperative BD score</th>
<th>Postoperative BD score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low or intermediate (%)</td>
</tr>
<tr>
<td></td>
<td>n (%)</td>
</tr>
<tr>
<td>Low or intermediate</td>
<td>63 (80.8)</td>
</tr>
<tr>
<td>High</td>
<td>2 (9.1%)</td>
</tr>
</tbody>
</table>
Figure 1 legend

Score 0 (A-D): Small magnification of preoperative biopsy (A) of superficial tumor without tumor budding in the higher magnification (B) of the IF. Small magnification for the corresponding resection specimen (C), and higher magnification of the IF (D).

Score 1 (E-H): Small magnification of preoperative biopsy of very deep tumor (E), without tumor budding at the invasive front (F). Small magnification of the corresponding resection specimen (G) and higher magnification of the IF (H) which shows no tumor budding.

Score 2 (I-L): Small magnification of preoperative biopsy of very deep tumor (I), with the presence of tumor budding at the IF in the higher magnification (J). Small magnification of the corresponding resection specimen (K) and higher magnification of the IF (L) which shows tumor budding.

* IF: Invasive front.
Small magnification: ×20
Higher magnification: ×100