SKIN

IN-DEPTH REVIEWS

Development of a metastatic risk model for cutaneous squamous cell carcinoma

Scott W. Fosko MD[a], Melinda B. Chu MD[b], Brandon T. Beal MD[c], Maulik Dhandha MD[d], Eric S. Armbrecht PhD[e]

[a]Mayo Clinic, Department of Dermatology, Jacksonville, FL [b]OncoDerm Associates, St. Louis, MO [c]Cleveland Clinic, Department of Dermatology, Cleveland, OH [d]Saint Louis University, Department of Dermatology, St. Louis, MO [e]Saint Louis University, Center for Outcomes Research, St. Louis, MO

ABSTRACT

Background: Cutaneous squamous cell carcinoma (cSCC) is the most common cancer capable of metastasis. Due to its high incidence and lack of inclusion in national databases it has been difficult to identify high-risk factors associate with metastasis. The development of a cSCC metastatic risk model would help physicians identify patients who are at risk for metastasis, and would allow for the initiation of early aggressive management to improve outcomes.

Aim: Explore different statistical approaches to develop a model to predict cSCC metastasis that is accurate and reflects routine clinical practice.

Methods: All cSCCs diagnosed and treated at Saint Louis University from January 2010 to March 2012 were included. Three statistical approaches were studied: multivariable logistic regression (MLR), pattern classification (PC) and sum score method (SSM). Two models using the SSM were created with a different number of factors used to merit assignment to the metastatic cohort: 2 factors (S2) or >2 factors (S2+). For each model, sensitivity (SN), specificity (SP) and positive predictive value (PPV) were calculated.

Results: SN, SP, and PPV for each model were: MLR: SN 4.3%, SP 97.4%, PPV 16.0%; S2: SN 78.3%, SP 83.7%, PPV 12.5%; S2+: SN 60.9%, SP 96.5% PPV 34.1%; PC: SN 73.9%, SP 95.9%, PPV 34.7%.

Conclusions: The PC model was the most accurate. The S2+ model had a lower SN, but would be easier to implement as clinicians would only have to sum high-risk factors.

INTRODUCTION

Regional and/or nodal metastasis is reported to occur in ~3-6% of cutaneous squamous cell carcinomas (cSCCs).[1-3] The task of identifying the rare metastatic case of cSCC has been likened to looking for "nodal needles in the cSCC haystack."[4] Despite the low percentage of metastatic cSCCs, cSCC metastasis is associated with significant

morbidity and mortality, and represents a significant public health burden due to the prevalence of cSCC, with greater than 700,000 new cases diagnosed annually in the United States.[1,5]

While imaging and sentinel lymph node biopsy (SLNB) may assist in the detection of nodal disease, the clinical situations in which they are best utilized are not well defined.[4,6] The development of a cSCC metastatic risk index that physicians can use in clinic to guide decision making would help to initiate early aggressive management and improve outcomes for high-risk cSCC. However, the development of cSCC metastatic risk index presents a number of unique challenges:

1) cSCCs are extremely common, but no centralized cancer registries exist. Data collection for cSCCs is not mandated by the National Cancer Institute's Survival, Epidemiology, End Results Program (SEER) database or the American Cancer Society as it is for other skin cancers (i.e. melanoma).
2) The high incidence of cSCC and low rate of metastasis make it difficult to obtain the detailed data necessary to build a cancer registry.
3) There is currently no standardized definition of high-risk cSCC.[7-12]
4) cSCCs often have multiple high-risk factors.
5) Both tumor and host factors influence metastatic risk, but it has been difficult to determine which combination of high-risk factors contribute most.[7-12]

The aim of this study is to explore different statistical approaches to develop a model to predict cSCC metastasis that is accurate and reflects routine clinical practice.

MATERIALS AND METHODS

Study cohort: This study received Institutional Review Board approval and was conducted at the Saint Louis University Department of Dermatology from January 2010 to March 2012. The dataset comprised all cSCCs diagnosed, managed, or treated by any specialty (e.g., dermatology, otolaryngology, plastic surgery, surgical oncology, radiation oncology, and oncology).The cSCC cases were first identified using International Classification of Diseases 9th Revision codes and then verified by medical and pathology chart review. Patient and tumor characteristics including the presence of the American Joint Committee on Cancer (AJCC) and National Comprehensive Cancer Network (NCCN) high-risk cSCC factors† were recorded.[11,12]

Model development: Patient and tumor characteristic variables derived from the NCCN high-risk factors were evaluated for inclusion in the statistical models. The NCCN high-risk factors are inclusive of the criteria the AJCC uses to stage tumors. Five variables were excluded due to rarity (sample prevalence <5%) or absence from routine clinical charting: (1) site of prior radiation or chronic inflammatory process; (2) Breslow Depth (BD) >2mm, (3) Clark level (CL) IV or V; (4) neurological symptoms; and (5) lymphatic or vascular involvement. One variable was excluded due to its subjective definition which increased misclassification bias: (6) poorly defined border. After reviewing the results, 2

† NCCN high-risk factors: Location/size; poorly defined boarders; recurrent; immunosuppression; sight of prior radiation therapy or chronic inflammatory process; rapidly growing tumor;neurologic symptoms; poorly differentiated; acantholytic,adenosquamous, desmoplastic, or metaplastic subtypes;Breslow depth =>2 mm or Clark level IV or V; andperineural, lymphatic, or vascular involvement.

additional variables were excluded due to the lack of a significant association (α ≥ 0.1) with the outcome: (7) immunosuppression and (8) acantholytic, adenosquamous, desmoplastic, or metaplastic histologic subtypes. The net result of this process was 5 variables: (1) size by anatomic location; (2) recurrent; (3) rapidly growing; (4) moderate or poorly differentiated histology; (5) perineural invasion (PNInv).

Since it was observed that some anatomic areas, regardless of tumor size, were associated with metastases, a binary summary variable for anatomic location was created. This variable described whether the tumor fulfilled high-risk location by combining NCCN areas M and H‡. To further improve model fit the NCCN variable for location/size (high-risk areas H, M, & L) was refined using a binary variable with site-specific cut-points determined by an analysis of tumor size stratified by location and metastatic status. To create this variable, the mean diameter of tumors in the metastatic vs. nonmetastatic groups was compared at each anatomic location. After evaluation, it was decided the mean diameter of tumors in the nonmetastatic group plus 2 standard deviations (encompassing 95% of cases) should be used as the benchmark to determine if tumor size in the context of location would be high or low risk. Said in another way, size was considered to be a significant factor at a specific anatomic location if the mean size of the tumors in the metastatic cohort was more than 2 standard deviations greater than the mean size of tumors in the nonmetastatic group at that site. See Table 1.

Final model variables: After analyses and variable reduction, 6 variables significantly associated with metastasis were included in the models: (1) anatomic location; (2) moderately or poorly differentiated histology; (3) perineurial invasion; (4) rapidly growing; (5) recurrent; (6) size in context of location.

Model analysis: Four models using 3 statistical approaches were studied to determine their ability to accurately predict metastatic status: (1) multivariable logistic regression (MLR); (2) pattern classification (PC); (3) and sum score method (SSM). Metastasis was defined as pathologic identification of cSCC in a lymph node, the parotid gland, or distant metastasis. Sensitivity (SN), specificity (SP) and positive predictive value (PPV) were calculated for each model.

Multivariable logistic regression: In this model, multiple dependent variables (i.e. risk factors) were assessed and compared to the outcome variable, which is dichotomous (i.e. metastasis or no metastasis).[13]

Sum score method: In this method, a cutoff sum score for group assignment is selected prior to the analysis. Two models using the SSM were created with a different number of factors used to merit assignment to the metastatic cohort: 2 factors only (S2) or ≥2 factors (S2+).[14] The sum score for each case is calculated by adding up the total number of risk factors present. There was no weighting of the factors§.

‡ NCCN defines high-risk location/size as any size area H, "mask areas of the face" (central face, eyelids, eyebrows, periorbital, nose, lips [cutaneous and vermilion], chin, mandible, preauricular and postauricular skin/sulci, temple, ear), genitalia, hands, and feet; area M ≥ 10 mm (scalp, forehead, cheeks, neck, and pretibia); and area L ≥ 20 mm (trunk and extremities [excluding pretibial, hands, feet, nail units, and ankles]).

§ For example, a recurrent cSCC (1 point - recurrence) on the ear (1 point – high-risk location regardless of size) that is rapidly growing (1 point – rapidly growing) is assigned a sum score of 3 points. This tumor would be categorized as *nonmetastatic* in the S2 model where tumors with only 2 risk factors are designated to be metastatic. In the S2+ model, this tumor would be assigned to the predicted metastatic group.

Pattern classification method: The PC method has been described as a "20 questions" approach where one can intuitively classify a pattern through a sequence of questions. These patterns can be depicted as a decision tree.[15] The model attempts to accurately classify the outcome of each case with the least number of decisions and make the simplest decision tree.

RESULTS

The study cohort included data on 800 cSCCs from 585 patients. Dermatology diagnosed, managed, and/or treated 93.4% of cSCCs. There were 23 cases of metastasis (2.9%). Most patients (93.7%) contributed 1 or 2 tumors to the data set. Eleven patients (1.9%) had greater than 5 tumors and none of these were metastatic.

There were 225 tumors located in area L. Almost all metastatic cases (95.7%, n=22/23) were located in areas M and H. The metastatic rate for head and neck cSCC was 4.2% (n=22/519). Metastatic cases were observed on 11 distinct anatomic sites with cheek (n=7) and preauricular area (n=3) being most common. There were 2 metastatic tumors on each of the following sites: scalp, temple, lip and neck. Forehead, ear, nose, chin, and arm each had 1 metastatic case.

Odds ratios and p-values were calculated for variables significant for metastasis: poor or moderate differentiation, anatomic location, size in context of location, rapidly growing, recurrent, and PNInv. See Table 2.

Sensitivity, Specificity, and Positive Predictive Value: The SN, SP, and PPV for each model are depicted in Table 3. The lowest sensitivity was observed in MLR analysis (4.3%); the highest, S2 method (78.3%). The lowest specificity was seen in the S2 method, 83.7%; all other models had specificities > 95%. PPVs ranged from 12.5%-34.7%. MLR and S2 method had similar low PPV (16.0% and 12.5% respectively). The highest PPV was observed in the PC analysis (34.7%), which was closely followed by S2+ (34.1%).

TABLES

Anatomic location	Tumor diameter defined as high-risk by anatomic location
Face	≥ 2 cm
Lips	> 2.5 cm
Scalp	≥ 4cm
Neck	> 3 cm
Extremities	> 3 cm

Table 1. Anatomic location and tumor diameter cutoffs defining high-risk

Risk factor	Odds ratio	*P*-value
Poorly or moderately differentiated histology	5.88	.001
Anatomic location[a]	4.11	.18
Size in context of location (see Table 1)	4.01	.10
Rapidly growing	3.03	.07
Recurrent	2.71	.09
Perineural invasion	2.03	.28

Table 2. Final model variables independently associated with metastasis

[a] Tumors in areas M and H as defined by the National Comprehensive Cancer Network.

	Multivariate logistic regression	Sum score (S2)[a]	Sum score (S2+)[b]	Pattern classification
Sensitivity	4.3%	78.3%	60.9%	73.9%
Specificity	97.4%	83.7%	96.5%	95.9%
Positive predictive value	16.0%	12.5%	34.1%	34.7%

Table 3. Sensitivity, specificity, and positive predictive value for each statistical model

[a]Sum score (2): sum score method performed where tumors with any 2 risk factors were assigned to metastatic cohort.

[b]Sum score (2+): sum score method performed where tumors with 2 or more risk factors were assigned to metastatic cohort.

DISCUSSION

In this analysis, we sought to explore different statistical models, MLR, SSM, and PC, and their SN, SP, and PPV in detecting metastasis. The PC model had the highest accuracy: SN 73.9%, SP 95.9%, and PPV 34.7%. Also of note was the finding that 95.7% (n=22/23) of metastatic cSCC were located on the head and neck. That the PC model had the highest values for SN, SP, and PPV when examined in aggregate suggests both the combination and the additive effect of risk factors contribute to cSCC metastasis.

While statistically the pattern classification method might appear complex and inaccessible, the approach is fundamentally similar to the way physicians are trained to diagnosis and manage diseases. All available data is assimilated and considered with certain data weighted more or less. For example, a 1 cm moderately-differentiated cSCC with PNInv on the forearm is considered to have a lower metastatic risk than a 2.5 cm well-differentiated tumor on the ear. Both tumors exhibit 2 risk factors in our models, but the tumor on the ear would prompt higher concern than the cSCC on the arm.

When developing a prognostic model, ease of use in clinical practice must be considered. While the SN of S2+ is lower than PC, the simplistic S2+ approach may be more feasible to implement (the user just has to calculate a score based on predefined variables). In this model, any 3 combinations of risk factors in Table 2 would be concerning for metastasis. The most well-known use of the SSM might be the CHADS2 score which measures stroke risk in patients with atrial fibrillation.[16]

In this study we identified several variables independently associated with cSCC metastasis which are relevant to clinical practice: poorly or moderately differentiated histology, anatomic location (areas H and M), size in context of location (Table 1), rapidly growing, recurrent, and PNInv. Poorly or moderately differentiated histology, anatomic location (areas H and M), and size in context of location (Table 1) all had odds ratios greater than 4 and are also commonly noted in other studies as significant high-risk factors for metastasis. [8-12,17]

The NCCN guidelines represent the broadest and most comprehensive definition of high-risk cSCC factors currently available.[12] And while the AJCC has attempted to refine cSCC high-risk factors for staging purposes and prognostic value, they have not been successful.[7-9] Jambusaria-Pahlajani et al[8] developed an alternative staging system for cSCC which provides greater ability to stratify high-risk cSCCs and improves prognostic accuracy.[9] Their alternative staging system uses 4 high-risk factors: poor differentiation,

PNInv, tumor diameter ≥ 2 cm, and invasion beyond subcutaneous fat. In another recent study, Thompson et al.[10] performed a meta-analysis of 23,421 cSCCs and found invasion beyond subcutaneous fat, BD > 2mm, diameter > 2 cm, poor differentiation, PNInv, immunosuppression, and location on the temple, ear, and lip to be significant risk factors for metastasis. Lastly, Peat et al.[17] performed a risk stratification analysis for metastasis from cSCC and found poor and moderate differentiation, PNInv and lymphovascular invasion, diameter ≥ 2 cm, and CL V to be important variables for metastasis. Thought the aims of these studies and ours were all slightly different, there was significant overlap between these studies high-risk factors for metastasis and ours - poor differentiation, PNInv, and location/size.

While the prevalence rate for cSCC metastasis of 2.9% in this study is consistent with published studies[1-3], the low prevalence negatively impacted the PPV in all models, and did not allow any model to near or even exceed a PPV of 50%. The PC model had the highest PPV at 34.7%. This is a universal factor that makes identifying high-risk and metastatic cSCC so difficult, the extremely high incidence of cSCC and the extremely low prevalence of cSCC metastasis. In the southern half of the United States metastatic cSCC is estimated to have a higher mortality than melanoma due to the high incidence of cSCC and despite the low prevalence of cSCC metastasis[1]; thus, demonstrating the importance of identifying these high-risk tumors early.

While our results are informative, the study has several limitations: 1) This is a single institution study, 2) we were not able to evaluate the clinical importance of BD or CL due to the absence of the routine collection of this data in clinical practice. In our experience, BD and CL are not routinely recorded as it would be impractical to have dermatopathologists report BD or CL for every cSCC, thus while this is a limitation, it is a reflection of clinical practice and 3) there were no standard criteria for which patients received SLNBs which could impact rates of metastasis within this population.

Conflict of Interest Disclosures: none.

Funding: none.

Corresponding Author:
Brandon T. Beal, MD
Department of Dermatology
Dermatology & Plastic Surgery Institute
Cleveland Clinic
9500 Euclid Ave / A61
Cleveland, OH 44195
Phone: (216) 444-5772
Fax: (216) 636-0435
e-mail: bealb@ccf.org

References:

1. Karia PS, Han J, Schmults CD. Cutaneous squamous cell carcinoma: estimated incidence of disease, nodal metastasis, and deaths from disease in the United States, 2012. J Am Acad Dermatol. 2013;68(6):957-966.

2. Breuninger H, Eigentler T, Bootz F, Hauschild A, et al. Brief S2k guidelines-Cutaneous squamous cell carcinoma. J Dtsch Dermatol Ges. 2013;11 Suppl 3:37-45, 39-47.

3. Czarnecki D, Staples M, Mar A, Giles G, Meehan C. Metastases from squamous cell carcinoma of the skin in southern Australia. Dermatology. 1994;189(1):52-54.

4. Karia PS, Schmults CD. Screening for nodal metastasis and its challenges: nodal needles in the SCC haystack. JAMA Dermatol. 2014;150(1):16-17.

5. Donaldson MR, Coldiron BM. No end in sight: the skin cancer epidemic continues. Semin Cutan Med Surg. 2011;30(1):3-5.

6. Ruiz ES, Karia PS, Morgan FC, Schmults CD. The positive impact of radiologic imaging on high-stage cutaneous squamous cell carcinoma management. J Am Acad Dermatol. 2017 Feb;76(2):217-225.

7. Chu MD, Slutsky JB, Dhandha MM, Beal BT et al. Evauation of the definitions of "high-risk" cutaneous squamous cell carcinoma using the America Joint Committee on Cancer staging criteria and National Comprehenisve Cancer Network guidelines. J Skin Cancer. 2014;154340.

8. Jambusaria-Pahlajani A, Kanetsky PA, Karia PS, Hwang WT et al. Evaluation of AJCC tumor staging for cutaneous squamous cell carcinoma and a proposed aternative tumor staing system. JAMA Dermatol. 2013 Apr;149(4):402-10.

9. Karia PS, Jambusaria-Pahlajani A, Harrington DP, Murphy GF, et al. Evaluation of American Joint Committee on Cancer, International Union Agaist Cancer, and Brigham and Women's Hospital tumor staging for cutaneous squamous cell carcinoma. J Clin Oncol. 2014 Feb 1;32(4):327-334.

10. Thompson AK, Kelley BF, Prokop LJ, Murad MH, Baum CL. Risk Factors for Cutaneous Squamous Cell Carcinoma Recurrence, Metastasis, and Disease-Specific Death: A Systematic Review and Meta-analysis. JAMA Dermatol. 2016;152(4):419-428.

11. Edge SB, Byrd DR, Compton CC, et al., editors. AJCC Cancer Staging Manual. 7th edition. New York, NY, USA: Springer; 2010. Cutaneous squamous cell carcinoma and other cutaneous carcinomas; pp. 301–314.

12. National Comprehensive Cancer Network. Squamous cell skin cancer (version 1.2017). Available from: URL: http://www.nccn.org/professionals/physician_gls/pdf/ squamous.pdf. Accessed April 2017.

13. Hidalgo B, Goodman M. Multivariate or multivariable regression? Am J Public Health. 2013;103(1):39-40.

14. DiStefano C ZM, Mindrila D. Understanding and Using Factor Scores: Considerations for the Applied Researcher Practical Assessment Research & Evaluation. 2014.

15. RO. D. Decision Trees. Pattern Classification New York, NY: Wiley & Sons; 2004:394-396.

16. Gage BF, Waterman AD, Shannon W, Boechler M, et al. Validation of clinical classification schemes for predicting stroke: results from the National Registry of Atrial Fibrillation. JAMA. 2001;285(22):2864-2870.

17. Peat B, Insull P, Ayers R. Risk stratification for metastasis from cutaneous squamous cell carcinoma of the head and neck. ANZ J Surg. 2012;82(4):230-233.