Published in final edited form as:
Ann Surg Oncol. 2012 January ; 19(1): 287–293. doi:10.1245/s10434-011-1979-6.

Validation of Statistical Predictive Models Meant to Select Melanoma Patients for Sentinel Lymph Node Biopsy

Michael S. Sabel, MD, **John D. Rice, MSPH**, **Kent A. Griffith, MS, MPH**, **Lori Lowe, MD**, **Sandra L. Wong, MD**, **Alfred E. Chang, MD**, **Timothy M. Johnson, MD**, and **Jeremy M.G. Taylor, PhD**
Departments of Surgery (MSS, SLW, AEC, TMJ), Dermatology (TMJ), Pathology (LL), University of Michigan Health System and the Biostatistics Core of the University of Michigan Comprehensive Cancer Center (JDR, KAG, JMGT), Ann Arbor, Michigan

Abstract

Introduction—To identify melanoma patients at sufficiently low risk of nodal metastases who could avoid SLN biopsy (SLNB). Several statistical models have been proposed based upon patient/tumor characteristics, including logistic regression, classification trees, random forests and support vector machines. We sought to validate recently published models meant to predict sentinel node status.

Methods—We queried our comprehensive, prospectively-collected melanoma database for consecutive melanoma patients undergoing SLNB. Prediction values were estimated based upon 4 published models, calculating the same reported metrics: negative predictive value (NPV), rate of negative predictions (RNP), and false negative rate (FNR).

Results—Logistic regression performed comparably with our data when considering NPV (89.4% vs. 93.6%); however the model's specificity was not high enough to significantly reduce the rate of biopsies (SLN reduction rate of 2.9%). When applied to our data, the classification tree produced NPV and reduction in biopsies rates that were lower 87.7% vs. 94.1% and 29.8% vs. 14.3%, respectively. Two published models could not be applied to our data due to model complexity and the use of proprietary software.

Conclusions—Published models meant to reduce the SLNB rate among patients with melanoma either underperformed when applied to our larger dataset, or could not be validated. Differences in selection criteria and histopathologic interpretation likely resulted in underperformance. Development of statistical predictive models must be created in a clinically applicable manner to allow for both validation and ultimately clinical utility.

Introduction

Lymphatic mapping and sentinel lymph node (SLN) biopsy provides critical staging information for clinically node negative melanoma patients. In addition, the early removal of micrometastatic disease from the regional nodes may have a therapeutic benefit.[1] For these reasons, SLN biopsy has been widely accepted as a component of the surgical management of melanoma and is considered by some to be the standard of care for a subset of melanoma patients.[2] However, the procedure is not without controversy.[3, 4] Although the Multicenter Selective Lymphadenectomy Trial-I (MSLT-I) trial, which randomized clinically node negative patients to wide excision alone versus wide excision with SLN biopsy, suggested a

*Corresponding author: Michael S. Sabel, MD, 3304 Cancer Center, 1500 East Medical Center Drive, Ann Arbor, MI 48109-0932; Fax: (734) 647-9647; msabel@umich.edu.

potential therapeutic benefit for SLN biopsy among the node-positive patients, overall the study demonstrated no survival benefit to the procedure.[1]

Many surmise that one reason the MSLT-I trial failed to demonstrate any impact of SLN biopsy on overall survival is that any therapeutic benefit would be limited to those patients harboring micrometastases. With the present indications for SLN biopsy, only around 20% of patients are found to be node positive. This means 4 out of 5 patients undergoing SLN biopsy derive no direct benefit from the procedure, other than reassurance. Improved selection of patients for SLN biopsy, by increasing the fraction of patients found to have metastatic disease, would not only improve the relative therapeutic benefit of the procedure, but minimize the cost and morbidity. To that end, multiple groups, including ours, have retrospectively examined their SLN outcomes in an attempt to identify clinical and histopathologic features that may correlate with positive SLN.[5–8] Some of these studies attempted to create statistical models with a high accuracy for predicting the SLN status.[8, 9] To potentially identify those patients at sufficiently low risk of nodal metastases whom could avoid the procedure, Mocellin et al.[10] examined four different predictive statistical models, focusing on negative predictive value (NPV), SLN biopsy reduction rate (rate of negative predictions, RNP), and the false negative rate (FNR), or how many patients with positive biopsies were predicted to be negative by the model. They reported predictive models that could identify approximately a quarter of patients who may be spared SLN biopsy, with a false negative error rate of 1–2%. In their conclusion, the authors ask institutions involved in the management of melanoma to independently validate their models. In this study, we attempted to do just that.

Methods

In the Mocellin paper, a sample of 1132 melanoma patients, from two institutions were used to construct 4 statistical models using-logistic regression, classification trees, random forests, and support vector machines (SVM), with the intention of selecting models that allowed for a significant reduction in the number of biopsies performed while incurring a minimum of false negatives. To validate these findings, after obtaining approval from the University of Michigan Institutional Review Board (IRB), we queried our comprehensive prospectively-collected melanoma database to identify consecutive melanoma patients 16 years of age and older who underwent SLN biopsy at the University of Michigan.

Patients undergo same-day preoperative lymphoscintigraphy in which 99mTc sulfur colloid is injected intradermally around the primary lesion or biopsy site 2–4 hours prior to surgery. Either isosulfan blue dye (Lympazurin 1%) or diluted methylene blue dye is injected intradermally around the primary lesion or biopsy site after the induction of general anesthesia. Any blue, hot (defined as the hottest node and any other nodes greater than 10% of the counts per minute of the hottest node) or suspicious lymph nodes are considered SLNs. SLN specimens are fixed in 10% neutral-buffered formalin, serially sectioned longitudinally at 2–3mm intervals, and entirely embedded in paraffin. Hematoxylin and eosin (H&E) staining was performed on two 5-μm thick sections from each tissue block. In addition, four unstained serial sections are prepared for possible future immunohistochemical analysis. If SLN(s) were negative by routine light microscopy, immunohistochemical staining is routinely performed using antibodies to the melanoma markers Melan-A (1:12.5 dilution: DAKO Corporation, Carpineteria, CA) and S-100 protein (1:500 dilution: DAKO Corporation, Carpineteria, CA).

Using information available from the Mocellin paper itself and/or the web address cited for supplementary material, we attempted to duplicate their statistical models, using our data with respect to outcome of interest (PSLN) and independent covariates, for purposes of

predicting the occurrence of PSLN. The outcome of interest was the discovery of any positive sentinel lymph nodes. The covariates of interest included the following patient/tumor characteristics: patient age and gender, tumor thickness (Breslow depth), Clark level, regression, ulceration, histologic subtype, and mitotic index. In order to evaluate similarities and differences in the predictive ability of Mocellin's statistical models when applied to our data as compared to when applied their data, we calculated the same metrics (NPV, RNP, and FNR) they had previously published.

Sensitivity, Specificity, Negative Predictive Value (NPV), and False Negative Rate (FNR) are each calculated per their accepted definitions as follows; Sensitivity = True Positives (TP) / {TP + False Negatives (FN)}, Specificity = True Negatives (TN) / {TN + False Positives (FP)}, NPV = TN / {TN + FN}, and FNR is equal to 1 − Sensitivity and defined as FN / FN + TP. SNB reduction rate = {TN + FN} / {Total = TN + FN + TP + FP} and the error rate = FN / Total. All metrics were calculated as in Mocellin et.al.

If data was missing, this was handled both by using only cases on which we had complete data or employing multiple imputation using Imputation and Variance Estimation Software (IVEware). This free, downloadable software, maintained by the researchers at the Survey Methodology Program, Survey Research Center at the University of Michigan, uses a multivariate sequential regression approach to imputing missing values in analytical datasets. Multiple imputed datasets can be created, such that missing data is replaced by ‘sensible’ guesses based upon the remainder of the patient’s data. Multiple imputation methods can then be used to aggregate the estimates from the regression models from each of the imputed datasets. [Raghunathan et. Al., Survey Methodology, June 2001). Full details of the programs and its uses can be found at http://www.isr.umich.edu/src/smp/ive/.

Results

We identified 2313 patients who underwent SLN biopsy for melanoma between 1998 and 2009. The clinical and demographic characteristics are summarized in Table 1, and are compared to the Mocellin data. There were several differences between our patient populations. Gender and age distribution were essentially the same. Although the median Breslow thickness was similar, the data used by Mocellin et al had a larger fraction of melanomas less than 1.0mm in Breslow thickness than our dataset (23.8% vs. 10.5%). This likely reflects differences in the application of sentinel lymph node biopsy. Whereas they performed SLN biopsy on patients with melanoma <1mm if Clark level was IV or V or ulceration was present, we do not use Clark level as an indication for SLN biopsy. Instead, we perform SLN biopsy on melanoma <1mm only in the presence of other adverse histologic features such as angiolymphatic invasion, ulceration, high mitotic rate and young age, resulting in fewer patients with thin melanoma being mapped. The Mocellin population also has a higher fraction of Clark level II and III patients, whereas for patients in whom data were available, we had far fewer Clark level II and III melanomas. Complicating this was the undocumented Clark level for a fraction of our patients. We stopped routinely recording Clark level as it was not predictive of SLN status,[5, 9] however this variable appears in all of Mocellin et al.’s models. We therefore approached this in two ways: analyzing only cases on which we had complete data (1667 subjects, or 72% of the original data set), and employing multiple imputation using IVEware (Version 0.2, Institute for Social Research, University of Michigan, 2010) in order to generate full data sets for further analysis. Other differences between our two datasets included a much higher rate of regression in the Mocellin data (nearly half of the patients compared to only 13.4%) and a higher mitotic rate among our patients, the latter likely reflecting our use of mitotic rate as a selection criteria for SLN biopsy.

All patient/tumor characteristics selected for modeling by Mocellin et al. seemed to be of especially strong predictive value for their data. Of their four models, three (logistic, tree, and SVM) were, on the whole, largely similar both in the set of variables they identified as significant as well as in their performance as measured by NPV (93.6%, 94% and 93%, respectively), RNP (27.5%, 29.8% and 30.1%, respectively), and FNR (1.8%, 1.8% and 2.1%, respectively). However, the random forest seemed to perform somewhat better, with a 97% NPV and a false negative rate (FNR) of 0.5%. This was achieved, however, at the cost of a lesser reduction in biopsies (RNP), at only 18%.

Logistic Regression

We began with what Mocellin et al. note is the simplest and most commonly applied model, logistic regression, for which the coefficients in the prediction model were given in their paper. We applied their model to our data twice, first using only observations on which we had the full set of observations, then on data imputed using IVEware; there were only negligible differences in the results.

Using a cutoff value of 0.1 for the predicted probability to define cases as having a positive SNB at the individual level (not to be confused or aligned with the proportion of positivity at the population level), we found that the logistic model's performance with our data produced worse NPV yet better FNR. These results are shown in Table 2. However, it seems that the model’s specificity is not high enough to be efficient at reducing the rate of biopsies. Raising the probability cutoff to 0.2 lowers NPV by a negligible amount, while lowering the cutoff to 0.05 has minimal impact on the already low false-negative rate and error rate, while lowering the number of SLN biopsy cases potentially omitted.

Classification Tree

When evaluating their classification tree when applied to our data (Table 3) compared to applied to their data, the performance was reduced for 2 of 3 metrics. Comparing our metrics to theirs, the NPV was slightly lower and the RNP was substantially lower; however, the error rate was slightly lower for our data.

We also present in Tables 2 and 3 sensitivity and specificity of the logistic and tree models applied to the Mocellin data and applied to our data. While the sensitivity and specificity of these models applied to our data was quite good, it was not quite as good as for these models applied to the Mocellin data.

Random Forest and SVM

Unfortunately, we were unable to validate the remaining two models, the random forest and SVM. These are sophisticated models, which if carefully constructed have the potential for superior performance; however by their nature they cannot be summarized by a simple graph or equation. In the end, despite multiple correspondences with the authors, because of the complexity of the models in question, complicated by the use of proprietary software (DTREG), we could not apply these models to our data. This was especially disheartening because the random forest model was potentially the best of the four by Mocellin et al. metrics.

Discussion

Breslow thickness is a validated and highly reproducible factor predictive of SLN status.[11] Based on these findings, SLN biopsy has generally been recommended in any melanoma patient with a Breslow thickness ≥ 1.0mm, although some institutions use a threshold of 0.75mm. Using a cutoff of 1.0mm, the SLN positivity rate ranges from approximately 15%

to 25%. Several groups have attempted to identify patient and primary tumor characteristics that may correlate with SLN positivity (Table 4). The primary goal of many of these studies has been to identify adverse factors that may select patients with thin melanomas who may benefit from the procedure. Balch et al,[2] in arguing that SLN biopsy should be the standard of care, state that the risk of nodal metastases should be sufficient to justify the procedure, and approximate this rate as ≥10%. It follows logically, that if there are adverse features that can identify patients with thin melanoma who should have the procedure (risk >10%) then it would also be ideal to identify a subset of patients with melanoma ≥1mm, lacking those features and having a risk <10%, who do not need the procedure. A better selection of patients would not only significantly decrease the costs and morbidity associated with the surgical therapy of melanoma, but potentially increase the relative impact of SLN biopsy on outcome.

In 2006, we proposed this idea, based on a multivariate model based on logistic regression techniques showing the interactions between Breslow thickness, age and mitotic rate on SLN positivity.[9] As an example, our model would predict patients >65 years of age with melanomas <2mm on the head and neck or upper extremity (and lacking angiolymphatic invasion) would not fit the criteria put forth by Balch et al. However, as Mocellin et al correctly point out, our model was based on maximizing predictive accuracy, which may not be the ideal clinical criteria. Differentiating a patient with a 30% risk from a 50% risk may be biologically interesting, but is clinically irrelevant as both would be recommended for SLN biopsy. In contrast, models maximizing the negative predictive value would better select patients who may avoid the procedure. Using clinicopathologic data from two centers, one in Italy and one in Australia, they conclude these models can spare a high percentage of patients from SLN biopsy with a small error rate, and sought validation from other institutions. Given our interest in this approach, we were highly motivated to do so.

We were able to apply the logistic regression and classification tree models to our data; however, we found them not as useful as reported. For the logistic model, the cutpoint for the predicted probability of PLSN was not stated in the Mocellin et.al. paper making exact comparison difficult. However, using a cutpoint of 0.1, the logistic regression model reduced the SLN biopsy rate by 2.9% for UM data, far below the 27.5% reported by the Mocellin group. The prediction metrics were similar for the classification tree applied to UM data; however, the decrease in the fraction of patients recommended to undergo SLN was not as large. Although the reduction in the number of SLN procedures would not have been as large if the classification tree was applied prospectively to our data, there would still have been a 14% reduction in SLNB, with a simultaneous reduction in FNR and no increase in error rate. This highlights the strong potential to these predictive models. It is highly possible that the differences seen may be related to differences in the selection criteria used to recommend SLN biopsy, which thereby creates the sample data on which their models are constructed. Mocellin et. al. include patients with thin melanomas having a Clark level of IV or V. This represents approximately a fifth of their analyzed sample. Their models suggest the potential for a significant reduction in the fraction of these patients recommended for SLN biopsy. However, at our institution, SLN biopsy would not have been recommended for these patients in the first place, without additional worrisome characteristics. Therefore, the levels of reduction in SLN biopsy reported by Mocellin et. al. may be less significant using data from institutions that already have a higher recommendation threshold for SLN biopsy.

In attempting to validate these results, several additional issues related to predictive models in oncology became apparent. One important lesson in moving forward is to remember that in order for these models to be clinically relevant, it is necessary for them to be available in a readily usable form, first for institutions to validate the results, and ultimately for

clinicians. Despite multiple attempts, only 1 out of the 4 models (classification tree) could be exactly validated, with a second method, logistic regression, duplicated but with the probability cutpoint used for classification only estimated, given that it was not provided in the original published manuscript. The random forest and SVM methods could not be validated. Given that their models are sophisticated, requiring specialized proprietary software, the models are not readily generalizable. As the goal of this avenue of research is to provide models that are applicable to a wide variety of data sources, consideration must be given to the availability of the software. There are a number of R packages, for example, that allow for the creation and storage of such models. R is a freely available language and environment for statistical computing. Like commercial products such as SAS or STATA, R is a statistical analysis platform that is often used in the academic statistical and biostatistical communities. 'R packages' are add-on features that expand the statistical application of the R program and can be freely downloaded at http://cran.r-project.org. Alternatively, a website can be established in which the user would enter the patient characteristics (Breslow depth, mitotic index, etc) and the website would return the prediction from the models, which is a trivial exercise for logistic models and classification trees; yet would require additional efforts for the complexity of the classification routines employed by random forests and support vector machine techniques.

Finally, another challenge to the creation and application of these predictive models is the variability in the interpretation of histopathologic features. The histolopathologic interpretation of melanoma with evaluation of the prognostic parameters is the single most important factor in the selection criteria for SLN biopsy. Accordingly, differences in data sets can often be traced to differences in histologic interpretation from institution to institution. Regression emerged as one factor in selecting patients likely to be node negative, and has been previously reported. However, regression was present in almost half of the patients in the Mocellin dataset, as compared with only 13% of our dataset. Among similar series, regression was only present in a small percentage of patients (between 7% and 18%),[5, 6, 12–15] with one outlier (28%).[7] The reported presence of regression in a larger proportion of cases in the Mocellin dataset compared to others is likely a reflection of a lower histologic threshold and/or looser definition of regression. Early histologic evidence of regression can be difficult to identify and is often times equivocal. The earliest stage of regression is lymphocytic inflammation, yet it is difficult, if not impossible, to distinguish early lymphocytic regression from conventional host response of tumor infiltrating lymphocytes. At our institution, we refrain from reporting regression unless unequivocal, well-developed regressive fibrosis is present. This difference in histologic thresholds may, in part, explain the difference in reported rate of regression between the 2 datasets.

Similarly, there is a striking difference in percent of melanomas with a low (<1/mm^2) mitotic rate between the 2 datasets. Mocellin et al. reported that 37.4% of their melanomas had a low mitotic rate, whereas only 15.7% of our cases had less than 1 mitosis per mm^2. While this may represent our use of mitotic rate in selecting patients for SLN biopsy, to some degree it may also represent differences in how mitotic rate is measured. While there may be some inter-institution variability in the histologic application of counting mitotic figures per mm^2, it is likely that the noted differences largely stem from a difference in intensity and vigilance in the search for mitoses. Histologically, one mm^2 corresponds to approximately 4 to 6 high power fields. If 4 high power fields of melanoma are just randomly evaluated, it would not be unusual for not a single mitoses to be identified. At our institution, the dermatopathologists specifically seek out mitotically active areas from which to obtain their count. Historically, our data has shown that even a single mitosis identified in the dermal invasive component of melanoma can stratify patients into a higher risk group for SLN positivity.[6] Thus, our dermatopathologists are sensitized to the prognostic and therapeutic implications between identifying no mitoses, which corrresponds to the "low"

category versus finding even a solitary mitotic figure which would then stratify to the "medium" mitotic index category.

In conclusion, the logistic regression and classification tree models reported by Mocellin et al. provided high sensitivity but low specificity when applied to our data, and had less impact on the reduction in the number of SLN performed than reported. While the logistic regression model has negligible impact on reducing SLN biopsy, the classification tree model performed better. Differences in both selection criteria and the interpretation of histopathologic features likely contribute to variation in model performance. The random forest and SVM models were unable to be clinically validated.

References

1. Morton DL, Thompson JF, Cochran AJ, et al. Sentinel-node biopsy or nodal observation in melanoma. NEJM. 2006; 355(13):1307–1317. [PubMed: 17005948]
2. Balch CM, Morton DL, Gershenwald JE, et al. Sentinel node biopsy and standard of care for melanoma. J Am Acad Dermatol. 2009; 60(5):872–875. [PubMed: 19389531]
3. Baldwin BT, Cherpelis BS, Sondak VK, et al. Sentinel lymph node biopsy in melanoma: Facts and controversies. Clinics in Dermatology. 2010; 28:319–323. [PubMed: 20541686]
4. Zitelli JA. Sentinel lymph node biopsy: an alternate view. Dermatol Surg. 2008; 34:544–549. [PubMed: 18261108]
5. Sondak VK, Taylor JM, Sabel MS, et al. Mitotic rate and younger age are predictors of sentinel lymph node positivity: lessons learned from the generation of a probabilistic model. Ann Surg Oncol. 2004; 11(3):247–258. [PubMed: 14993019]
6. Kruper LL, Spitz FR, Czerniecki BJ, et al. Predicting sentinel node status in AJCC stage I/II primary cutaneous melanoma. Cancer. 2006; 107:2436–2445. [PubMed: 17058288]
7. Mandala M, Imberti GL, Piazzalunga D, et al. Clinical and histopathological risk factors to predict sentinel lymph node positivity, disease-free and overall survival in clinical stages I–II JCC skin melanoma. European Journal of Cancer. 2009; 45:2537–2545. [PubMed: 19553103]
8. Wong SL, Kattan MW, McMasters KM, et al. A nomogram that predicts the presence of sentinel node metastasis in melanoma with better discrimination than the American Joint Committee on Cancer Staging System. annals of Surgical Oncology. 2005; 12(4):282–288. [PubMed: 15827679]
9. Paek SC, Griffith KA, Johnson TM, et al. The impact of factors beyond Breslow depth on predicting sentinel lymph node positivity in melanoma. Cancer. 2007; 109(1):100–108. [PubMed: 17146784]
10. Mocellin S, Thompson JF, Pasquali S, et al. Sentinel node status prediction by four statistical models: Results from a large bi-institutional series (n=1132). Annals of Surgery. 2009; 250:964–969. [PubMed: 19953714]
11. Lens MB, Dawes M, Newton-Bishop JA, et al. Tumour thickness as a predictor of occult lymph node metastases in patients with stage I and II melanoma undergoing sentinel lymph node biopsy. British Journal of Surgery. 2002; 89(10):1223–1227. [PubMed: 12296887]
12. Dadras SS, Lange-Asschenfeldt B, Velasco P, et al. Tumor lymphangiogenesis predicts melanoma metastasis to sentinel lymph nodes. Modern Pathology. 2005; 18(9):1232–1242. [PubMed: 15803182]
13. McMasters KM, Wong SL, Edwards MJ, et al. Factors that predict the presence of sentinel lymph node metastasis in patients with melanoma. Surgery. 2001; 130:151–156. [PubMed: 11490343]
14. Nguyen CL, McClay EF, Cole DJ, et al. Melanoma thickness and histology predict sentinel lymph node status. american Journal of Surgery. 2001; 181:8–11. [PubMed: 11248167]
15. Wagner JD, Gordon MS, Chuang T-Y, et al. Predicting sentinel and residual lymph node basin disease after sentinel lymph node biopsy for melanoma. Cancer. 2000; 89(2):453–462. [PubMed: 10918179]
16. Mraz-Gernhard S, Sagebiel RW, Kashani-Sabet M, et al. Prediction of sentinel lymph node micrometastasis by histological features in primary cutaneous malignant melanoma. arch Dermatol. 1998; 134:983–987. [PubMed: 9722728]

17. Rousseau DL, Ross MI, Johnson MM, et al. Revised American Joint Committee on Cancer staging criteria accurately predict sentinel lymph node positivity in clinically node-negative melanoma patients. Annals of Surgical Oncology. 2003; 10(5):569–574. [PubMed: 12794025]

18. Sassen S, Shaw HM, Colman MH, et al. The complex relationship between sentinel node positivity, patient age and primary tumor desmoplasia: Analysis of 2303 melanoma patients treated at a single center. Annals of Surgical Oncology. 2007; 15(2):630–637. [PubMed: 18080717]

19. Taylor RC, Patel A, Panageas KS, et al. Tumor-infiltrating lymphocytes predict sentinel lymph node positivity in patients with cutaneous melanoma. Journal of Clinical Oncology. 2007; 25(7): 869–875. [PubMed: 17327608]

20. Niakosuri F, Kahn HJ, McCready DR, et al. Lymphatic invasion identified by monoclonal antibody D2-40, younger age and ulceration: Predictors of sentinel lymph node involvement in primary cutaneous melanoma. arch Dermatol. 2008; 144(4):462–467. [PubMed: 18427039]

21. Mitra A, Conway C, Walker C, et al. Melanoma sentinel node biopsy and prediction models for relapse and overall survival. British Journal of Cancer. 2010

Table 1

Demographic and Clinical Summary

Variables	Mocellin (n = 1132)	Michigan data (n = 2313)
Sex		
Female	44.5%	37.4%
Male	55.5%	62.6%
Age (yr)		
Median	54	54
Range	14–85	16–95
Breslow (mm)		
Median	1.61	1.7
Range	0.14–35	0.35–28
T1 (< 1 mm)	23.8%	10.5%
T2 (1–2 mm)	38.9%	48.4%
T3 (2–4 mm)	24.6%	28.2%
T4 (> 4 mm)	12.7%	12.4%
Unknown	-	0.5%
Clark		
II	5.1%	0.6%
III	33.1%	7.9%
IV	54.9%	65.9%
V	6.9%	2.7%
Unknown	-	23.0%
Ulceration		
Present	29.7%	23.7%
Absent	70.3%	73.7%
Unknown	-	2.6%
Regression		
Present	47.1%	13.4%
Absent	52.9%	84.5%
Unknown	-	2.2%
Histological subtype		
SSM	55.0%	50.7%
NM	34.0%	20.6%
Others	11.0%	28.8%
Mitotic index		
High (> 5 mm^2)	30.8%	20.5%
Medium (1–5 mm^2)	31.8%	57.8%
Low (<1 mm^2)	37.4%	15.7%
Unknown	-	6.0%

Table 2

Comparison of the performance of the Logistic Regression Model on Mocellin data and Michigan Data.

	Mocellin	UM (0.1 cutoff)	UM (0.2 cutoff)
Negative Predictive Value	93.6%	89.4%	86.1%
SLN Reduction Rate	27.5%	2.9%	26.0%
False Negative Rate	8.6%	1.2%	14.1%
Error Rate	1.8%	0.3%	3.6%
Sensitivity	91.4%	98.8%	85.9%
Specificity	32.3%	3.4%	30%

Table 3

Performance of the Classification Tree Model using University of Michigan Data

	Mocellin	UM
Negative Predictive Value	94.1%	87.7%
SLN Reduction Rate	29.8%	14.3%
False Negative Rate	8.6%	6.7%
Error Rate	1.8%	1.8%
Sensitivity	91.4%	93.3%
Specificity	35.2%	16.9%

Table 4

Studies of clinical and histopathologic factors predictive of SLN involvement.

Author	Year	N	Institution	Factors Associated with SLN Positivity on Multivariate Analysis
Mraz-Gernhard[16]	1998	215	Univ. of California	Breslow thickness, ulceration (present), mitotic rate (≥6), ulceration (present), angiolymphatic invasion (present), microsatellites (present)
Nguyen[14]	2000	112	Medical Univ. of South Carolina	Breslow thickness (>1.5mm), ulceration (present), angiolymphatic invasion (present)
Wagner[15]	2000	275	Indiana Univ.	Breslow thickness (≥1.25), ulceratioin (present), Mitotic rate (>5)
McMasters[13]	2001	1058	Sunbelt Melanoma Trial	Breslow thickness, Clark level, ulceration (present), age (≤60)
Rousseau[17]	2003	1375	MDACC	Breslow thickness, ulceration (present), age (≤50), tumor site (axial)
Sondak[5]	2004	419	Univ. of Michigan	Breslow thickness, age, mitotic rate (continuous variables)
Wong[8]	2005	979	MSKCC & Sunbelt Melanoma Trial	Breslow thickness, age, site, Clark level, ulceration
Kruper[6]	2006	682	Univ. of Pennsylvania	Breslow thickness, mitotic rate (0.1 to 5, >5), TIL (absent)
Mocellin	2006	246	Univ. of Padova, Italy	Breslow thickness, mitotic rate, satellitosis
Sassen[18]	2007	2303	Sydney Melanoma Unit	Breslow thickness (>2.0mm), age (<40), histology (desmoplastic), ulceration (present).
Taylor[19]	2007	875	MSKCC	Breslow thickness, ulceration (present), gender (male), TIL (absent)
Paek[9]	2007	1130	Univ. of Michigan	Breslow thickness, mitotic rate, age (all continuous), site (trunk, leg), angiolymphatic invasion (present)
Niakosari[20]	2008	96	University of Toronto	Lymphatic invasion (present), age (<50), ulceration (present)
Mandala[7]	2009	394	Riuniti Hospital, Italy	Breslow thickness, site (axial), TIL (absent)
Mitra[21]	2010	561	Multiple locations in UK	Breslow thickness, mitotic rate (0.1 to 6, >6), site (H&N, leg, trunk)