Defining optimal oncolytic virus treatment and diagnostics in high risk melanoma patients

Stahlie, E.H.A.

CHAPTER 3

External validation of a Dutch predictive nomogram for complete response to T-VEC in an independent American patient cohort

Emma H.A. Stahlie*
Michael J. Carr*
Jonathan S. Zager
Alexander C. J. van Akkooi
** Authors contributed equally to this study*

Annals of Surgical Oncology 2022 Mar;29(3):1637-1644.

ABSTRACT

Purpose
Talimogene Laherparepvec (T-VEC) is a modified herpes simplex virus type-1 used as intralesional immunotherapy in stage IIIB-IVM1a melanoma patients. Recently, Stahlie et al. published a predictive model for complete response (CR) to T-VEC. This study aimed to externally validate this model in an independent, American patient cohort.

Methods
In total 71 stage IIIB-IVM1a melanoma patients treated with T-VEC at Moffitt Cancer Center were included. A second nomogram was built incorporating the same predictive factors: tumor size (diameter of largest metastasis), type of metastases (cutaneous, subcutaneous and nodal) and number of metastases (cut-off:<20 & >20). Predictive accuracy was assessed through calculation of overall performance, discriminative ability and calibration.

Results
The two cohorts were similar in many clinicopathologic factors and only differing in tumor mutational status and use of systemic therapy prior to T-VEC. In the validation cohort, 37 (52%) patients showed CR, 22 (31%) partial response (PR), 2 (5.6%) stable disease (SD) and 10 (15%) progressive disease (PD). Of those who demonstrated a CR, 16 (43%) recurred. Overall performance was good (0.164) and discriminative power resulted in fair discriminative ability (0.827). The calibration curve showed slight underestimation for predicted probabilities >0.15 and slight overestimation <0.15.

Conclusion
The original model as well as the validation model show comparable and good predictive accuracy. The validation model reinforces the conclusion that for the best response to T-VEC, it should be used early on in the course of the disease, when the tumor burden is cutaneous with smaller diameter and fewer of metastases.

INTRODUCTION

Talimogene Laherparepvec (T-VEC) is an oncolytic immunotherapy used for the treatment of stage IIIB-IVM1a melanoma patients[1]. It is injected directly into metastatic lesions; therefore, in order to be eligible for treatment with T-VEC, patients must have readily identifiable cutaneous, subcutaneous, and/or nodal lesions. T-VEC is based on the herpes simplex virus type 1 that has undergone several genetic modifications which enhance its clinical activity against tumor cells but also limit replication in normal cells[2].

The phase III OPTiM trial, of which the first results were published in 2015, was the first to present T-VEC as a novel potential new treatment option, demonstrating long-lasting complete responses (CR) and a longer median overall survival (OS) when compared to treatment with GM-CSF[3]. Shortly after, T-VEC was approved by the U.S. Food and Drug Administration (FDA), followed by approval by the European Medicines Agency (EMA) with the slight difference of a limited registration for stage IIIB-D, M1a unresectable melanoma with injectable cutaneous, subcutaneous and/or nodal metastases. T-VEC has since been increasingly used across different centers around the world.

Hence, multiple studies have investigated T-VEC in real-world practice, and complete response- as well as overall response rates surpass those of the OPTiM trial (CR 11% and ORR 26%), the highest reported being 62% and 88.5%, respectively[4, 5]. Although these response rates seem high, they still vary greatly per study (e.g. studies by Mohr et al. and Perez et al., reporting CR rates of 21% and 44%, respectively) and alongside differences in patient selection, seem highly dependent on certain patient- and/or tumor characteristics[6, 7]. A recently published study by Stahlie et al. set out to identify predictive factors for a CR, and concluded that the best moment to treat melanoma patients with T-VEC is when their tumor burden is still low, suggesting use earlier in the course of the disease[5]. Furthermore, their results showed that patients with cutaneous metastases have the best chance of achieving a CR compared to those with subcutaneous and/or nodal disease. These conclusions are based on a prediction model which estimates the probability of achieving a CR, which the authors developed to optimize patient selection for treatment with T-VEC. The present study set out to externally validate this nomogram model in a comparable, but independent, American patient cohort. By validating the nomogram, we hope that it will help clinicians in daily clinical practice to assist in their selection of patients who might benefit the most from T-VEC treatment.

PATIENTS AND METHODS

Patients and treatment

The study was approved by the institutional review board of Moffitt Cancer Center. All patients older than 18 years of age with stage IIIB-IVM1a melanoma, treated with T-VEC monotherapy at Moffitt Cancer Center between November 2014 and August 2020, were included. All patients were staged according to the American Joint Committee on Cancer (AJCC) 8th edition. The eligibility criteria were the same as those utilized in the study by Stahlie et al.: a follow-up time beyond 6 months from the start of treatment and patients had to have injectable cutaneous, subcutaneous or nodal metastases. Prior treatment for metastases was not reason for patient exclusion. A database was prospectively maintained with patient- and tumor characteristics and treatment outcomes. Treatment with T-VEC was done according to the guidelines and recommendations of

the manufacturer (Initial dose 10^6 pfu/mL, subsequent doses 10^8 pfu/mL. AMGEN Inc., Thousand Oaks, CA).[8, 9] Evaluation of complete clinical response to T-VEC was confirmed through biopsy of the treated lesion.

Variables and statistical analyses

Baseline and clinicopathologic characteristics were collected for all patients. Response rates were assessed according to the World Health Organization criteria and in case of pathological evaluation as described by Tetzlaff et al.[10, 11] Overall response rate (ORR) was defined as all patients with CR or partial response (PR) as best response (defined as the best response a patient achieves during the study treatment).

In order to externally validate the prediction model published by Stahlie et al. on the patient cohort included in this study, a comparable model was built incorporating the same variables: type of metastases (categorical), number of metastases (categorical) and diameter of the largest metastasis (continuous). The type of metastases were assessed as follows: patients with cutaneous metastases only were categorized as cutaneous, patients with subcutaneous as well as cutaneous metastases were categorized as subcutaneous metastases, and patients with cutaneous and/or subcutaneous as well as nodal metastases were classified as lymph node metastases. The number of metastases were defined as a categorical variable by cases with ≤ 20 metastases in one group and those with > 20 metastases in a second group.

The predictive accuracy of this second model was also assessed through calculation of overall performance, discriminative ability and calibration. Overall performance was calculated with the Brier score. Discriminative ability

of the model was assessed by calculating the area under receiver-operating characteristics (ROC) curve. This value ranges between 0.5 and 1.0, the first indicating no discrimination above chance and the latter indicating perfect discrimination unrelated to chance. To assess calibration, a calibration plot was generated in which the observed and predicted probability of achieving a CR were plotted. When the prediction is perfectly calibrated, it corresponds to the 45 degree line: the intercept and calibration slope are 0 and 1, respectively. All points below and above this line, reflect over- and underestimation. 1000 bootstrap resamples were used to reduce overfit bias. With bootstrapping the performance of the nomogram is simulated, as if it were applied to future patients.

The model in this study was used for external validation of the prediction model published by Stahlie et al., therefore these models were compared. Comparisons of clinicopathologic characteristics between cohorts were performed using GraphPad Prism version 9.1.2. P-values < 0.05 were considered statistically significant. Descriptive statistics, Fisher's exact test, and Chi square test were used for comparison of categorical and continuous data. All nomogram analyses were performed using R version 3.6.1.

RESULTS

A total of 71 patients treated at Moffitt Cancer Center met eligibility criteria and were evaluated as part of the external validation cohort. The median age at initial T-VEC injection was 77 years (range 45 – 94 years). Most patients were men (53%, n=38), with injections administered on lesions of the head and neck (25%, n=18), trunk (5.6%, n=4), upper extremity (18%, n=13), and lower extremity (51%, n=36). Concurrent nodal disease was injected in 5 (7.0%) patients, with AJCC 8 staging groups represented by 29 (41%), 37 (52%), and 5 (7.0%) having stage IIIB, IIIC, IIID+IVM1a disease, respectively. (Table 1)

Comparison of model and validation cohorts

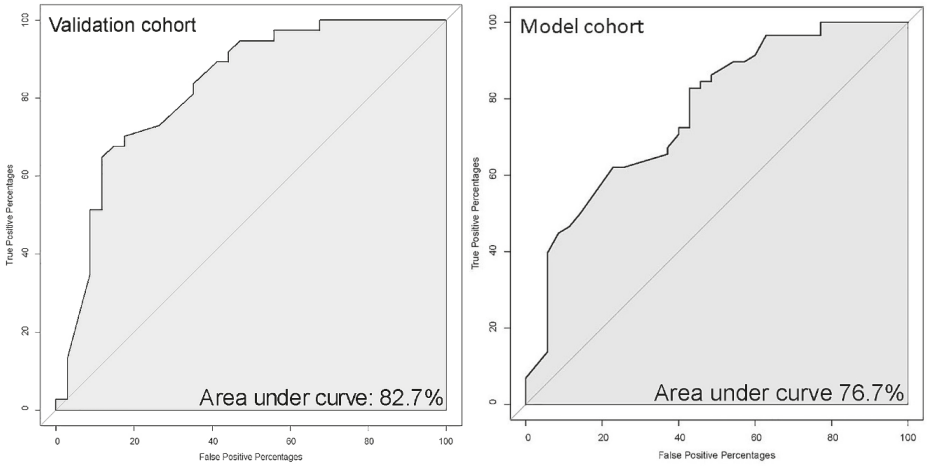
Variable	Level	Model Cohort[5]	Validation Cohort	p-value
Sex (%)	Female Male	53 (57%) 40 (43%)	33 (46%) 38 (54%)	0.21
Median age in years (range)		72 (30 - 97)	77 (45 - 94)	**0.002**
Ethnicity	Caucasian	93 (100%)	71 (100%)	1.00
Substage (AJCC 8)	IIIB IIIC IIID + IVM1a	30 (32%) 56 (60%) 7 (8%)	29 (41%) 37 (52%) 5 (7%)	0.52
Mutation status	Wildtype BRAF mut. NRAS mut. Other/Unknown	15 (16%) 40 (43%) 20 (22%) 18 (19%)	39 (55%) 12 (17%) 0 (0%) 20 (28%)	**<0.001**
Location metastases	Extremity Trunk Head/neck	68 (73%) 13 (14%) 12 (13%)	49 (69%) 4 (5.6%) 18 (25%)	0.05
Number of metastases	≤ 20 lesions > 20 lesions	73 (78%) 20 (22%)	63 (89%) 8 (11%)	0.10
Mean diameter of the largest metastases in mm (range)		15 (2 – 100)	10 (4 – 86)	0.22
Type of metastases	Cutaneous only Subcutaneous Lymph nodes	32 (34%) 53 (57%) 8 (8.6%)	28 (39%) 38 (54%) 5 (7.0%)	0.78
Pre-T-VEC treatment				
Radiotherapy	No Yes	89 (96%) 4 (4%)	66 (93%) 5 (7.0%)	0.50
Systemic therapy	No Yes	77 (83%) 16 (17%)	42 (59%) 29 (41%)	**0.001**
Perfusion	No Yes	70 (75%) 23 (25%)	61 (86%) 10 (14%)	0.12
Surgery	No Yes	5 (5.4%) 88 (95%)	0 (0%) 71 (100%)	0.08
Median number T-VEC cycles (range)		8 (3 – 34)	6 (2 - 27)	**0.004**
Median time to clinical response, weeks (range)		5 (3 - 19)	7 (2 – 23)	**0.001**
Response to T-VEC	CR PR SD+PD	58 (62%) 16 (17%) 19 (20%)	37 (52%) 22 (31%) 12 (17%)	0.12

Table 1. *Clinicopathologic features of patients treated with T-VEC. Data are expressed as n (%) unless otherwise specified. Bold values indicate statistical significance (p < 0.05). Abbreviations: AJCC: American Joint Committee on Cancer, CR: complete response, PR: partial response, SD: stable disease, PD: progressive disease.*

Statistically significant differences in clinicopathologic characteristics between model and validation cohorts were found in several variables. Model cohort patients were found to have younger median age (72 vs. 77 years, respectively, p=0.002), more often had BRAF-mutant tumors (43% vs. 17%, p<0.001), less often received systemic therapy prior to T-VEC (17% vs. 41%, p=0.001), received more cycles of T-VEC injections (median 8 vs. 6 cycles, p=0.004), and had shorter time to clinical response (5 vs. 7 weeks, p=0.001) when compared to validation cohort patients. (Table 1) No significant differences between cohorts were found among melanoma stage at first T-VEC injection, location of T-VEC injections, number of metastases, size of largest metastases, type of metastases, or clinical response to T-VEC injections.

Validation cohort patients received a median 6 (range 2-27) treatment cycles before cessation of T-VEC. Median time to clinical response was 7 (range 2-23) weeks. Over a median follow-up of 14 (range 0-70) months from final T-VEC injection, treatment with T-VEC resulted in a pathologic CR in 37 (52%) of patients, in 22 (31%) a PR, in 2 (5.6%) stable disease, and in 10 (14%) progression of disease (PD). Prior to T-VEC, 29 (41%) patients received systemic therapy with checkpoint inhibition, and/or regional therapy with isolated limb infusion (n=10, 14%) or radiotherapy (n= 5, 7%). Of those who demonstrated a best response of CR, 16 (43%) recurred, 3 (19%) with local recurrence only, 5 (31%) with regional and distant recurrence, and 8 (50%) local, regional, and distant recurrence. These patients went on to receive a secondary series of T-VEC, surgery, or systemic treatments with checkpoint or targeted therapy. Patients with best response of PR, SD, or PD (n=34) were mostly treated with checkpoint inhibition (n=20) or targeted therapy (n=4) after T-VEC. Adverse effects to T-VEC were seen in 9 patients and included fever, chill, fatigue, and muscle ache. No patients stopped T-VEC secondary to adverse effects, no patients required hospitalization, and there were no deaths associated with treatment. The prediction model developed by Stahlie et al. was applied to this validation cohort and compared to that of the original cohort used to create the nomogram. The Brier score of the validation cohort was 0.164, similar to that of 0.182 in the model cohort, aligning with good overall performance. The ROC curve demonstrated an area under the curve (AUC) of 0.827 in the validation cohort, again similar to 0.767 of the model cohort, supporting a fair discriminatory capability between predictors used to set up the model. (Figure 1a) The calibration plot of the validation model showed underestimation for predicated probabilities >0.15 and overestimation <0.15, again similar to that of the model cohort underestimation of <0.55 and overestimation of >0.55 (Figure 1b). The nomogram for each model is provided (Figure 2).

1a: ROC curve



1b: Calibration plot

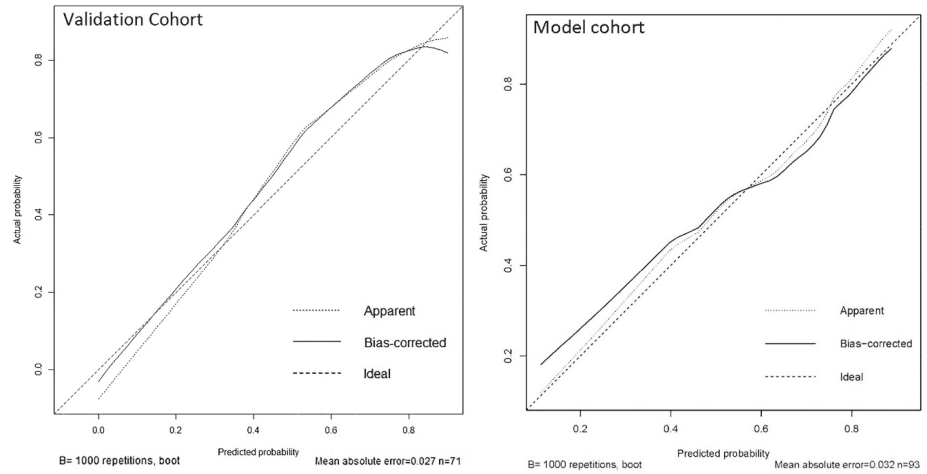


Figure 1. *Comparison of (a) ROC curve and (b) Calibration Plot for the Model and Validation Cohort*

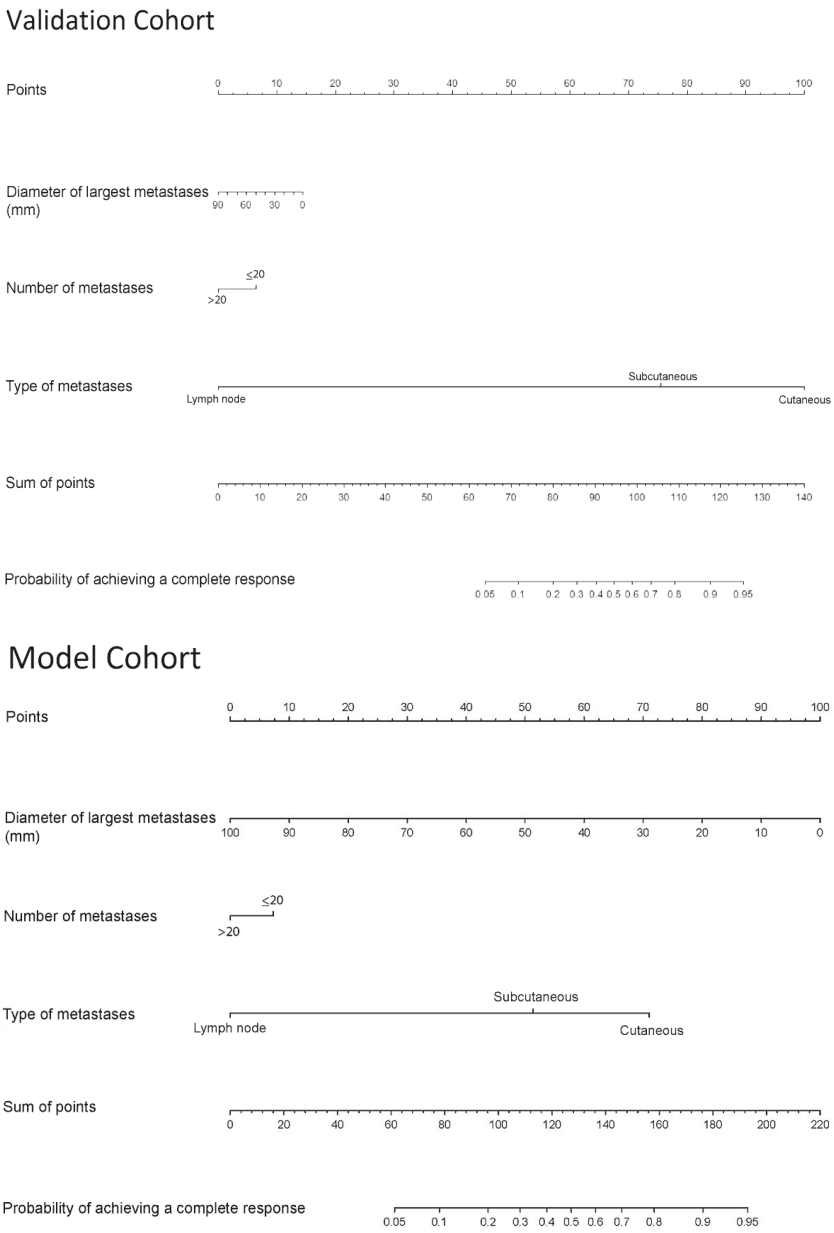


Figure 2: *Side-by-side Comparison of the Model and Validation Cohort Nomogram. The probability of CR is calculated by drawing a line from each predictor to the 'points' axis at the top. The points corresponding to each predictor should then be summed and this number should be located on the 'sum of points' axis. A line from the 'sum of points' axis can then be drawn onto the probability scale at the bottom. Example in the validation cohort nomogram: patient with 2 cutaneous metastases of which the largest has a diameter of 30mm. Probability of achieving a CR to T-VEC is 90%.*

DISCUSSION

The prediction model that was originally developed by Stahlie et al., combined known and statistically significant variables to predict a CR to T-VEC in patients with stage IIIB-IV1a melanoma. The current study provides validation of this nomogram through direct comparison of the Brier score, area under the ROC curve, and calibration plot of a secondary externally developed nomogram. In each of these testing mechanisms, the model and validation nomograms were similar in value. Descriptively, the two cohorts were similar in many clinicopathologic factors and only differing in tumor mutational status and use of systemic therapy prior to T-VEC.

Nomograms are tools that use biologic or clinical variables to depict a statistical prognostic model which can be used to determine diagnosis and predict prognosis or responses[12]. Use of nomograms is popular, as they provide superior disease- and treatment-related risk estimations that benefit can patients as well as clinicians. As T-VEC is a relatively new treatment option for melanoma patients and real-world data has only recently become increasingly available, this is the first time predictive factors were used to set up a prediction model for estimating the response. Stahlie et al. chose to focus only on tumor characteristics contributing to tumor burden and therefore only included the maximum lesion diameter, type of metastases, and the number of metastases in the prediction model. Previous studies also reported an association between the size of the treated lesion(s) and the response to T-VEC, though significance was only achieved for the ORR and the durable response rate[13, 14]. Stahlie et al. was the first to show a significant association between the type of metastases (cutaneous, subcutaneous, and nodal metastases) and the response to T-VEC: the more superficial the better the response. It is still unclear what causes this difference, but several causes have been hypothesized: biologically heterogeneous tumor sites might lead to different antitumor responses, metastasis from subcutaneous and nodal metastases to other anatomic sites might be easier due to the deep(er) location and finally the challenges that come with intratumoral injection might play a role[15, 16]. Although the number of metastases is not reported as predictive variable in previous literature, it was included in the model as it contributes to the patient's tumor burden.

The validation model incorporated the same variables as the original model and it is striking that the developed nomograms look very alike. When examining the validation nomogram and comparing it to the original, the most notable differences can be found in (1) the length of the 'diameter of largest metastasis'-axis and (2) the shorter probability scale. It seems that in contrast to the diameter being the most significant predictive factor in the original cohort, in the validation

cohort the type of metastases was the most significant. Yet, from both models one can conclude that the patients with a low tumor burden and cutaneous metastases have the highest probability of achieving a CR. These results reinforce the conclusion that T-VEC should be used early in the course of the disease for the best results.

In most centers, surgical excision is still the first-line treatment for limited recurrent in-transit disease. However, these patients often develop additional recurrences shortly after: a study by Dong et al. reported that over a median follow-up time of 40 months, only 19% of 648 patients did not develop a recurrence after resection of in-transit metastases[17]. Although these patients would nowadays be eligible for adjuvant systemic therapy, still 40% and 46% of stage IIIB-C patients (7th AJCC) treated with adjuvant pembrolizumab or nivolumab (respectively), will develop a recurrence within 3 to 4 years of follow-up. Unfortunately, none of the studies investigating T-VEC have reached a similar follow-up time, but a durable response rate of 51% in stage IIIB-IVM1a patients does look promising[5]. Other studies have also shown a positive association between treatment-naïve patients and the response to T-VEC[3, 5, 18]. Therefore, it might be more arguable to treat these patients with T-VEC rather than to surgically resect their in-transit metastases, as it seems newly diagnosed stage IIIB-IVM1a patients with a small tumor burden have the best chance of achieving a CR.

Although the current study reports a favorable response rate, a significant portion of patients have later progression of disease. Incorporating systemic immunotherapy with T-VEC seems a logical next step to treating this group of patients. Several studies have evaluated the efficacy T-VEC in combination with systemic immunotherapy, including T-VEC in combination with ipilimumab (n=19; ORR 50%, CR 22%)[19], T-VEC in combination with pembrolizumab (n=21; ORR 67%, CR 43%)[20] and three ongoing phase II trials evaluating T-VEC in combination with pembrolizumab or nivolumab (NCT02965716, NCT04068181 and NCT04330430). One recently published retrospective review evaluated the use of T-VEC in patients who progress on systemic immunotherapy, reporting that initiation of T-VEC sequentially to or concurrently with systemic immunotherapy did not significantly affect in-field response and in-field complete response improved survival outcomes[21].

The validation model shows good overall performance and discrimination is robust and comparable to that observed in the Stahlie publication with AUC of 0.826 versus AUC of 0.767, respectively. Moreover, the calibration curve indicates mostly underestimation >0.15, but the difference with the ideal line is almost negligible. Therefore, we believe the nomogram can confidently be used by

clinicians when selecting patients with the highest chance to develop a CR for treatment with T-VEC.

This study has its limitations. It was consciously chosen to incorporate only three variables; all easily obtainable tumor characteristics contributing to the patient's tumor burden. However, in the future, analysis of larger sample sizes might lead to the discovery of more valuable significant predictors for response. Incorporating these in the nomogram, could make it even more accurate and meaningful for future clinical use. Also interesting would be to focus more on the durable response (CR + PR lasting continuously for a minimum of 6 months), since a long-term CR is more valuable than a short-term CR[3]. Although the same in- and exclusion were used for the validation cohort as for the model cohort, a selection bias might play a role due to the retrospective setup of the study: a different inclusion time period was used and the selection of patients for treatment with T-VEC might differ per center.

Finally, our nomogram was developed and validated only in Western patients, therefore further validation is needed for application of our nomogram in non-Western patients.

CONCLUSION

This study successfully externally validated a predictive nomogram for CR on T-VEC monotherapy in stage IIIB-D, IVM1a melanoma. From both models can be concluded that for the best response to T-VEC, it should be used early on in the course of the disease, when the patient's tumor burden is cutaneous with smaller diameter and fewer number of metastases.

REFERENCES

1. Kaufman HL, Bines SD. OPTIM trial: a Phase III trial of an oncolytic herpes virus encoding GM-CSF for unresectable stage III or IV melanoma. *Future Oncol* 2010; 6(6):941-9.
2. Rehman H, Silk AW, Kane MP, et al. Into the clinic: Talimogene laherparepvec (T-VEC), a first-in-class intratumoral oncolytic viral therapy. *J Immunother Cancer* 2016; 4:53.
3. Andtbacka RH, Kaufman HL, Collichio F, et al. Talimogene Laherparepvec Improves Durable Response Rate in Patients With Advanced Melanoma. *J Clin Oncol* 2015; 33(25):2780-8.
4. Franke V, Berger DMS, Klop WMC, et al. High response rates for T-VEC in early metastatic melanoma (stage IIIB/C-IVM1a). *Int J Cancer* 2019; 145(4):974-978.
5. Stahlie EHA, Franke V, Zuur CL, et al. T-VEC for stage IIIB-IVM1a melanoma achieves high rates of complete and durable responses and is associated with tumor load: a clinical prediction model. *Cancer Immunol Immunother* 2021.
6. Perez MC, Miura JT, Naqvi SMH, et al. Talimogene Laherparepvec (TVEC) for the Treatment of Advanced Melanoma: A Single-Institution Experience. *Ann Surg Oncol* 2018; 25(13):3960-3965.
7. Mohr P, Haferkamp S, Pinter A, et al. Real-World Use of Talimogene Laherparepvec in German Patients with Stage IIIB to IVM1a Melanoma: A Retrospective Chart Review and Physician Survey. *Adv Ther* 2019; 36(1):101-117.
8. Amgen. Imlygic (talimogene laherperepvec) suspension for intralesional injection: US prescribing information. 2015.
9. Harrington KJ, Michielin O, Malvehy J, et al. A practical guide to the handling and administration of talimogene laherparepvec in Europe. *Onco Targets Ther* 2017; 10:3867-3880.
10. Tetzlaff MT, Messina JL, Stein JE, et al. Pathological assessment of resection specimens after neoadjuvant therapy for metastatic melanoma. *Ann Oncol* 2018; 29(8):1861-1868.
11. Geneva W, Organization WH, Organization WH. WHO Handbook for reporting results of cancer treatment. *WHO offset publication* 1979(48).
12. Balachandran VP, Gonen M, Smith JJ, et al. Nomograms in oncology: more than meets the eye. *Lancet Oncol* 2015; 16(4):e173-80.
13. Zhou AY, Wang DY, McKee S, et al. Correlates of response and outcomes with talimogene laherperpvec. *J Surg Oncol* 2019; 120(3):558-564.
14. Masoud SJ, Hu JB, Beasley GM, et al. Efficacy of Talimogene Laherparepvec (T-VEC) Therapy in Patients with In-Transit Melanoma Metastasis Decreases with Increasing Lesion Size. *Ann Surg Oncol* 2019; 26(13):4633-4641.
15. Chang CJ, Tai KF, Roffler S, et al. The immunization site of cytokine-secreting tumor cell vaccines influences the trafficking of tumor-specific T lymphocytes and antitumor efficacy against regional tumors. *J Immunol* 2004; 173(10):6025-32.
16. Li H, van der Leun AM, Yofe I, et al. Dysfunctional CD8 T Cells Form a Proliferative, Dynamically Regulated Compartment within Human Melanoma. *Cell* 2020; 181(3):747.
17. Dong XD, Tyler D, Johnson JL, et al. Analysis of prognosis and disease progression after local recurrence of melanoma. *Cancer* 2000; 88(5):1063-71.
18. Ressler JM, Karasek M, Koch L, et al. Real-life use of talimogene laherparepvec (T-VEC) in melanoma patients in centers in Austria, Switzerland and Germany. *J Immunother Cancer* 2021; 9(2).

19. Puzanov I, Milhem MM, Minor D, et al. Talimogene Laherparepvec in Combination With Ipilimumab in Previously Untreated, Unresectable Stage IIIB-IV Melanoma. *J Clin Oncol* 2016; 34(22):2619-26.
20. Long G, Dummer R, Andtbacka R, et al. Follow-up analysis of MASTERKEY-265 phase 1b (ph1b) study of talimogene laherparepvec (T-VEC) in combination (combo) with pembrolizumab (pembro) in patients (pts) with unresectable stage IIIB–IVM1c melanoma (MEL). Society for Melanoma Research Fifteenth International Congress, 2018.
21. Carr MJ SJ, DePalo D, Rothermel LD, Song Y, Straker RJ, Baecher K, Louie RJ, Stahlie EHA, Wright GP, Naqvi SMH, Kim Y, Sarnaik AA, Karakousis GC, Lowe MC, Delman KA, van Akkooi ACJ, Ollila DW, Collichio F, Zager JS. Talimogene Laherparepvec (TVEC) for the Treatment of Advanced Locoregional Melanoma after Failure of Immunotherapy: A Multi-Institutional Experience. *Annals of Surgical Oncology* 2021.