



Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.ejcancer.com



Original Research

The EORTC-DeCOG nomogram adequately predicts outcomes of patients with sentinel node–positive melanoma without the need for completion lymph node dissection[☆]



Daniëlle Verver^{a,*,3,1}, A. Rekkas^{b,3,1}, Claus Garbe^{c,4},
David van Klaveren^{d,3}, Alexander C.J. van Akkooi^{e,3},
Piotr Rutkowski^{f,3}, Barry W.E.M. Powell^{g,3}, Caroline Robert^{h,3},
Alessandro Testori^{i,3}, Barbara L. van Leeuwen^{j,3},
Astrid A.M. van der Veldt^{k,3}, Ulrich Keilholz^{l,3}, Rudolf Stadler^{m,4},
Alexander M.M. Eggermont^{n,3}, Cornelis Verhoef^{a,3}, Ulrike Leiter^{c,4,2},
Dirk J. Grünhagen^{a,3,2}

^a Department of Surgical Oncology, Erasmus MC Cancer Institute, University Medical Centre Rotterdam, Rotterdam, the Netherlands

^b Department of Medical Statistics, Leiden University Medical Center, Leiden, the Netherlands

^c Department of Dermatology, University Hospital Tuebingen, Tuebingen, Germany

^d Department of Medical Statistics, Erasmus Medical Centre, Rotterdam, the Netherlands

^e Department of Surgery, The Netherlands Cancer Institute – Antoni van Leeuwenhoek Amsterdam, the Netherlands

^f Department of Soft Tissue/Bone Sarcoma and Melanoma, Maria Skłodowska-Curie Institute – Oncological Center, Warsaw, Poland

^g Melanoma Unit, St George's Foundation University Hospital, London, UK

^h Department of Dermatology and Allergology, Cancer Institute Gustave Roussy, Villejuif, France

ⁱ Fondazione IRCCS policlinico san Matteo, Pavia, Italy

^j Department of Surgical Oncology, Groningen University, University Medical Centre Groningen, Groningen, the Netherlands

^k Department of Medical Oncology, Erasmus MC Cancer Institute, University Medical Centre Rotterdam, Rotterdam, the Netherlands

^l Director of the Charité Comprehensive Cancer Center, Charité - University of Medicine Berlin, Berlin, Germany

^m Department of Dermatology, University Hospital Johannes Wesling Minden, Minden, Germany

ⁿ Princess Máxima Center for Pediatric Oncology, Utrecht, the Netherlands

Received 23 December 2019; received in revised form 14 April 2020; accepted 18 April 2020

Available online 23 May 2020

[☆] **Presentations:** The content of this manuscript has not been published elsewhere. However, parts of this content have been presented during the EORTC Melanoma Group Fall Meeting 5th of October 2018 in Mallorca, Spain.

* **Corresponding author:** Department of Surgical Oncology, Erasmus MC Cancer Institute, dr. Molewaterplein 40, 3015 GD Rotterdam, the Netherlands.

E-mail address: d.verver@erasmusmc.nl (D. Verver).

¹ Shared first author. ² Shared last author. ³ On behalf of the European Organization for Research and Treatment of Cancer (EORTC) Melanoma Group (MG). ⁴ On behalf of the German Dermatologic Cooperative Oncology Group study (DeCOG-SLT).

<https://doi.org/10.1016/j.ejca.2020.04.022>

0959-8049/© 2020 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

KEYWORDS

Melanoma;
Sentinel lymph node;
Nomogram;
Prognosis;
Adjuvant therapy

Abstract Purpose: Based on recent advances in the management of patients with sentinel node (SN)–positive melanoma, we aimed to develop prediction models for recurrence, distant metastasis (DM) and overall mortality (OM).

Methods: The derivation cohort consisted of 1080 patients with SN-positive melanoma from nine European Organization for Research and Treatment of Cancer (EORTC) centres. Prognostic factors for recurrence, DM and OM were studied with Cox regression analysis. Significant factors were incorporated in the models. Performance was assessed by discrimination (c-index) and calibration in cross-validation across centres. The models were externally validated using a prospective cohort consisting of 705 German patients with SN-positive: 473 trial participants of the German Dermatologic Cooperative Oncology Group study (DeCOG-SLT) and 232 screened patients. A nomogram was developed for graphical presentation.

Results: The final model for recurrence and the calibrated models for DM and OM included ulceration, age, SN tumour burden and Breslow thickness. The models showed reasonable calibration. The c-index for the recurrence, DM and OM model was 0.68, 0.70 and 0.70, respectively, and 0.70, 0.72 and 0.74, respectively, in external validation. The EORTC-DeCOG model identified a robust low-risk group, with all identified low-risk patients (approximately 4% of the entire population) having a 5-year recurrence probability of <25% and an overall 5-year recurrence rate of 13%. A model including information on completion lymph node dissection (CLND) showed only marginal improvement in model performance.

Conclusions: The EORTC-DeCOG nomogram provides an adequate prognostic tool for patients with SN-positive melanoma, without the need for CLND. It showed consistent results across validation. The nomogram could be used for patient counselling and might aid in adjuvant therapy decision-making.

© 2020 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The American Joint Committee on Cancer (AJCC) staging system is the most widely accepted approach to melanoma staging [1,2]. Patients are classified into distinct stages based on the tumour node metastasis criteria where nodal status is based on number of positive lymph nodes after completion lymph node dissection (CLND) in case of a positive sentinel node (SN) or after a therapeutic lymph node dissection in case of clinically apparent nodal disease. Recently there have been many advances in the care of patients with SN-positive melanoma that also affect staging, namely CLND is no longer routine practice as the Multicenter Selective Lymphadenectomy Trial-II (MSLT-II) and the German Dermatologic Cooperative Oncology Group study (DeCOG-SLT) demonstrated no survival benefit for CLND [3–6] and as immune checkpoint inhibition and targeted therapy have been introduced in the adjuvant setting with highly encouraging results [7–10]. Consequently the AJCC staging system is likely to be less appropriate for patients with SN-positive melanoma not undergoing CLND because of decreased discriminatory ability [11] as the number of positive nodes after sentinel lymph node biopsy (SLNB) is not an independent prognostic factor [3,4] (in contrast to involved non-SNs retrieved after CLND [3]). As a result, omitting CLND could result in poorer risk stratification and impaired selection for adjuvant

therapy. On the other hand, SN tumour burden has been shown to be an independent predictor of involved non-SNs [12–14], and therefore SN tumour burden may serve as a surrogate.

The objective of the present study was to identify independent prognostic factors in a large European SN-positive melanoma population, using solely information from the primary melanoma and the SLNB, to develop a prediction model for recurrence, distant metastasis (DM) and overall mortality (OM), presented in the form of a nomogram. The resulting model could aid in adjuvant therapy decision-making. The prediction models were externally validated using a large prospective German cohort.

2. Patients and methods

2.1. Cohort characteristics

2.1.1. Derivation cohort

The retrospective derivation cohort consisted of 1080 patients with SN-positive melanoma who underwent SLNB between 1993 and 2008 in one of nine EORTC Melanoma Group centres that have been previously collected and described [11,15–17]. The current study only excluded duplicate cases ($n = 2$), leading to a total of 1078 eligible SN-positive patients. The two duplicate cases concerned an error in that database. The applied procedures have been described previously [11].

2.1.2. Validation cohort

The prospective German validation cohort involved two sets of patients. The first set consisted of 473 patients who were included in the DeCOG-SLT multicentre randomised phase-3 trial comparing survival between patients with SN-positive melanoma who did or did not undergo CLND [4]. The second set consisted of an additional 219 patients from a single centre (University Hospital, Tuebingen) who were initially screened for inclusion in the DeCOG-SLT trial but were not included because of meeting the trial's exclusion criteria (e.g. head and neck melanoma, age >75 years), unwillingness to participate, or no known reason. They also did or did not undergo CLND and were followed and prospectively registered in accordance with similar protocols. All patients had a tumour thickness of at least 1 mm and underwent surgery between 2006 and 2014. The study design, applied procedures and follow-up protocols have been described in detail elsewhere [4]. There was no overlap between the derivation cohort and validation cohort.

2.2. Outcomes

Outcomes of interest were first recurrence, first DM and OM. Time to recurrence was calculated from date of SLNB to date of first recurrence or date of death by any cause. Time to first DM was calculated from date of SLNB to date of first DM or date of death by any cause. Time to OM was calculated from date of SLNB to date of death by any cause.

2.3. Statistical analysis

The checklist proposed by the AJCC was used for guidance in building a high-quality prediction model [18]. Associations between possible prognostic factors and recurrence were studied with Cox regression analysis. The following eight variables were identified as possible prognostic factors based on clinical experience, literature review and availability of sufficient data: sex, age, ulceration, location, histology, Breslow thickness, total number of SNs removed and total number of positive SNs. To make efficient use of available data, an advanced multiple imputation of missing values strategy (5 imputations) was applied [19]. This was done separately for each derivation centre to avoid using information of missings in cross-validation. The possible non-linearity of continuous variables was modelled by logarithmic transformation. Independent prognostic factors were selected with multivariable backwards selection. Linear predictor values (the sum of truncated predictor values times their predictor effects) were scaled and rounded to a risk score with integer values between 0 and 100. Because recurrence, DM and OM are strongly related, the final recurrence prediction model based on data from all nine EORTC centres was used as

a basis for predicting DM and OM, where the baseline hazard and the slope of the recurrence prediction model were calibrated to DM and OM [20]. This approach is beneficial as it provides a unique risk score for each individual that translates into probabilities of all outcomes of interest, instead of developing three independent prediction models. To test the validity of our approach, we did develop these independent models and compared them with the calibrated models. The absolute risk prediction of each outcome was plotted against the risk score. To reduce overestimation of events occurring in patients with extremely high scores, scores were truncated at an integer of 23, corresponding to the 99th percentile of score distribution. Model performance was assessed by examining discrimination and calibration. Discrimination was measured using the concordance index (c-index); the closer to 1, the better the discrimination, and a value of 0.5 indicates that the model is no better than a chance [21]. Calibration was assessed visually by plotting the predicted probability against the actual observed frequency in quintiles of predicted outcomes. A 45° line indicates perfect calibration (when the predictive value of the model perfectly matches the patient's actual risk). Any deviation above or below the 45° line indicates underprediction or overprediction, respectively. A nomogram was developed for graphical presentation of the models. To evaluate generalisability of the models across different centres, an internal–external cross-validation was performed in which the model was fitted using data from eight centres and validated in the centre that was left out [22]. In addition we performed external validation using the prospective German cohort. We first needed to develop a model for recurrence where we replaced the continuous variable SN tumour burden with the categorical substitute used in the prospective German cohort (single cells, <0.5 mm, 0.5–1.0 mm, >1.0–2.0 mm, >2.0–5.0 mm and >5.0 mm). For the derivation cohort, single cells were defined as <0.1 mm according to the Rotterdam criteria [23]. Single cells in the validation cohort were not specifically defined, but as the Rotterdam criteria were used for measuring SN tumour burden, definitions are likely to correlate. The performance of this altered model was compared with the final recurrence model used for the nomogram. Subsequently the altered model was externally validated with the 692 patients from the prospective German cohort. To test how much the information on additional positive nodes retrieved after CLND would add to the discrimination of the prediction model, we also developed a prediction model in which the variable, additional positive nodes after CLND, was added. This model was based on 1015 patients that underwent CLND in the derivation cohort.

Furthermore we calculated the model performance for recurrence, DM and OM of the AJCC 7th edition classification, AJCC 8th edition classification and the

simple classification that was published previously (i.e. absent/present ulceration and low/high SN tumour burden) was tested [11]. Lastly the observed outcomes per group for all classifications were estimated using the Kaplan Meier analysis. All statistical tests were two-sided, with a $P < 0.05$ considered statistically significant. All statistical analyses were performed using SPSS version 22.0 (IBM, Armonk, New York, USA) and R (version 2.15, R Foundation for Statistical Computing, Vienna, Austria, 2011).

Table 1
Baseline characteristics of derivation and validation cohort.

| Characteristic | Derivation cohort (n = 1078) | Validation cohort (n = 692) | P value ^b |
|-------------------------------|---------------------------------|--------------------------------|----------------------|
| Age, years ^a | n = 1077 51 (40–62) | 57 (46–68) | <0.001 ^c |
| Gender | | | <0.001 |
| Female | 509 (47.2) | 267 (38.6) | |
| Male | 569 (52.8) | 425 (61.4) | |
| Breslow, mm ^a | n = 1076 3.0 (1.9–4.8) | 2.4 (1.6–4.0) | <0.001 ^c |
| Ulceration | n = 1015 | n = 596 | 0.570 |
| Absent | 536 (52.8) | 306 (51.3) | |
| Present | 479 (47.2) | 290 (48.7) | |
| Location | | | <0.001 |
| Extremity | 614 (57.0) | 335 (47.0) | |
| Trunk | 426 (39.5) | 355 (51.3) | |
| Head and neck | 38 (3.5) | 12 (1.7) | |
| Positive SNs | n = 984 | n = 690 | <0.001 |
| 1 node | 775 (78.8) | 623 (90.3) | |
| 2 nodes | 164 (16.7) | 60 (8.7) | |
| >2 nodes | 45 (4.6) | 7 (1.0) | |
| SN tumour burden, mm | 0.9 (0.4–2.5) | - | - |
| SN tumour burden, extended | | n = 626 | <0.001 |
| Single cells ^d | 113 (10.5) | 187 (29.9) | |
| <0.5 mm | 221 (20.5) | 57 (9.1) | |
| 0.5–1.0 mm | 235 (21.8) | 208 (33.2) | |
| >1.0–2.0 mm | 200 (18.6) | 114 (18.2) | |
| >2.0–5.0 mm | 195 (18.1) | 36 (5.8) | |
| >5.0 mm | 114 (10.6) | 24 (3.8) | |
| SN tumour burden, simple | | n = 626 | <0.001 |
| ≤ 1.0 mm | 569 (52.8) | 452 (72.2) | |
| > 1.0 mm | 509 (47.2) | 174 (27.8) | |
| CLND | | | <0.001 |
| No | 63 (5.8) | 384 (55.5) | |
| Yes | 1015 (94.2) | 308 (44.5) | |
| Positive non-SNs ^e | n = 1007 | n = 302 | 0.088 |
| None | 804 (79.8) | 229 (75.8) | |
| 1 node | 127 (12.6) | 53 (17.5) | |
| >1 node | 76 (7.5) | 20 (6.6) | |

CLND, completion lymph node dissection; IQR, interquartile range; SN, sentinel node. Values in parentheses are percentages unless indicated otherwise.

^a Values are median (IQR).

^b Chi-square test.

^c Except Mann–Whitney U test.

^d For the derivation cohort, single cells were defined as metastasis <0.1 mm.

^e Information retrieved after CLND.

3. Results

The retrospective derivation cohort consisted of 1078 and the prospective validation cohort of 692 patients with SN-positive. Patients in the validation cohort had less extensive disease in terms of Breslow thickness, number of positive SNs and tumour burden in the SN compared with those in the derivation cohort (Table 1).

In the derivation cohort, recurrence at five-years occurred in 496 patients (46.0%), DM in 437 patients (40.5%) and OM in 364 patients (33.8%). Median follow-up time for all survivors was 106 months (interquartile range [IQR] 61–130 months). In the prospective validation cohort, recurrence at five-years occurred in 267 patients (38.6%), DM in 223 patients (32.2%) and OM in 174 patients (25.1%). Median follow-up time for all survivors was 66 months (IQR: 48–94 months).

3.1. Models for recurrence, distant metastasis and overall mortality

The final multivariable Cox model for recurrence after backwards selection included four independent prognostic factors: ulceration, age, Breslow thickness and SN tumour burden (Table 2). Logarithmic transformation of the continuous variables adequately represented their effects. The c-index for the final recurrence model was 0.68 (95% confidence interval [CI]: 0.65–0.70). In cross-validation, the recurrence model was reasonably calibrated across nine centres in general, only in smaller centres there was substantial underestimation of the risk (Fig. S1).

The association between linear predictors of recurrence and DM was of the same size (calibration slope: 1.01, 95% CI: 0.87–1.16). The c-index for the calibrated model for DM was 0.70 (95% CI: 0.67–0.72) and was reasonably calibrated across nine centres in cross-validation (Fig. S2). The performance of this calibrated model, based on the baseline hazard and the slope of the recurrence model, was similar to that of the independently developed prediction model for DM (c-index: 0.70, 95% CI: 0.68–0.73).

Table 2

Final model for 5-year recurrence (hazard ratio with 95% confidence interval)^a.

| | Hazard ratio | Lower 95 | Upper 95 |
|------------------|--------------|----------|----------|
| Age | 1.28 | 1.12 | 1.45 |
| Breslow | 1.41 | 1.23 | 1.61 |
| SN tumour burden | 1.59 | 1.39 | 1.81 |
| Ulceration | | | |
| Absent | Reference | | |
| Present | 1.41 | 1.16 | 1.73 |

SN, sentinel node.

^a Includes the final independent prognostic factors selected with multivariable backwards selection.

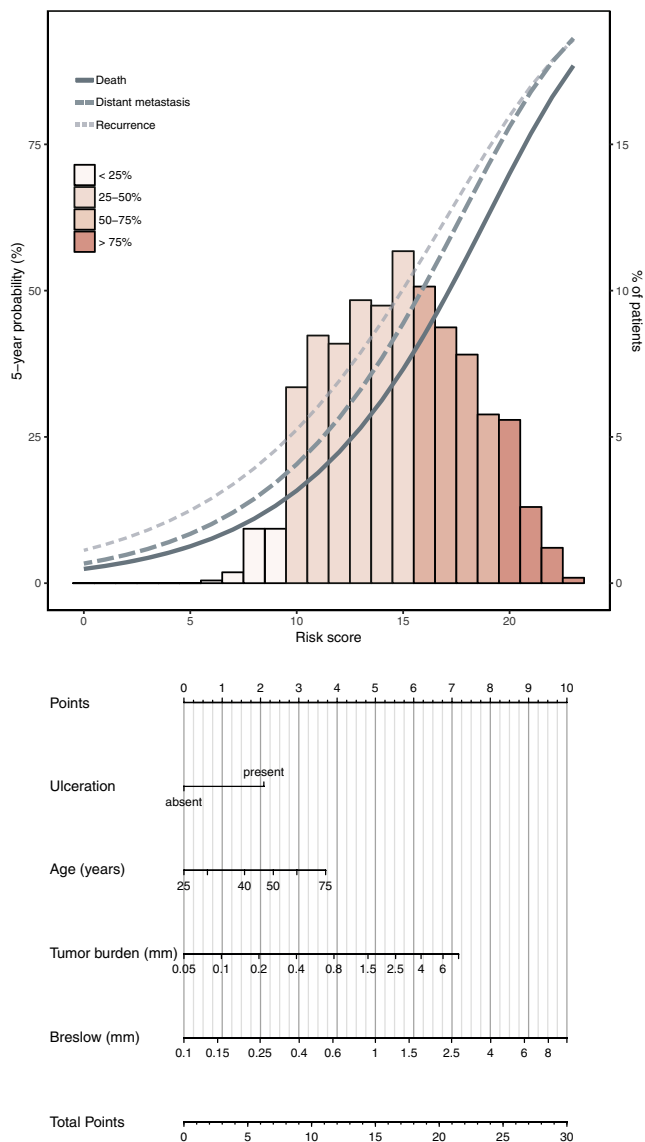


Fig. 1. Nomogram and risk distribution. The curves refer to predicted recurrence, distant metastasis or overall mortality at 5 years. The histogram refers to the risk score distribution in the cohort; each bar represents the proportion of patients in the cohort that was assigned that specific score. The histogram was divided into four risk groups based on the risk of recurrence: low risk: <25%, intermediate risk: 25–50%, high risk: 50–75% and very high risk: >75%. The nomogram incorporates four factors: ulceration, age, SN tumour burden and Breslow thickness. To calculate an individual's probability of 5-year recurrence, distant metastasis and overall mortality, values for the prognostic factors must be determined first (for example: absent ulceration, 35 years, SN tumour burden 0.8 mm and Breslow thickness 1.0 mm). Second, for each value the corresponding points can be obtained by drawing a line from each value towards the point axis (in example: 0, 1, 4 and 5 points, respectively). Third, the points must be added up to obtain the total risk score (in example: risk score of 10). Finally, the 5-year recurrence, distant metastasis and overall mortality probability can be read by moving vertically from the x-axis (total risk score) to the predicted risk curves and corresponding probabilities on the left y-axis (in example: 26% for recurrence, 20% for distant metastasis and 16% for overall

The association between linear predictors of recurrence and OM was of the same size (calibration slope: 1.04, 95% CI: 0.88–1.20). The c-index for the calibrated model for OM was 0.70 (95% CI: 0.67–0.73), and was reasonably calibrated across nine centres in cross-validation (Fig. S3). The performance of this calibrated model was similar to that of the independently developed prediction model for OM (c-index: 0.70, 95% CI: 0.68–0.73).

A four-item risk score was developed, assigning points to each prognostic factor based on the magnitude of association with recurrence. A nomogram to calculate the score and the risk of recurrence, DM and OM is presented in Fig. 1. The scores were divided into four risk groups based on the 5-year probability of recurrence: <25% (low risk; score 6–9; 4.1% of the population); 25–50% (intermediate risk; score 10–15; 52.9% of the population); 50–75% (high risk; score 16–19; 33.2% of the population); and >75% (very high risk; score 20–23; 10.0% of the population). The observed outcomes for recurrence, DM and OM per risk group are shown in Table 3.

3.2. External validation

For external validation purposes, an altered recurrence model was developed using the categorised SN tumour burden variable used in the prospective German cohort (Table S1). This altered model showed similar performance compared with the final recurrence model (c-index 0.68, 95% CI: 0.65–0.70). In external validation, the c-index for the altered recurrence model was 0.70 (95% CI: 0.67–0.74), for DM 0.72 (95% CI: 0.68–0.75) and for OM 0.74 (95% CI: 0.71–0.78). The calibration plots indicate good calibration, though there may be slight underestimation for higher-risk patients in the recurrence and OM models (Fig. S4).

3.3. Additional prognostic value of CLND

An extended model for recurrence was created by adding the variable, number of additional positive nodes after CLND, to the final recurrence model. This extended model for recurrence had a c-index of 0.69 (95% CI: 0.67–0.72). The calibrated extended models for DM and OM showed c-indices of 0.72 (95% CI: 0.69–0.74) and 0.72 (95% CI: 0.69–0.75), respectively.

3.4. Simple classification

A simplified version of the model stratifies patients into four groups based on ulceration and SN tumour burden:

mortality). The percentage of patients in the entire population (1078) that also had a total risk score of 10 can be determined from the histogram, as well as the corresponding percentage of patients on the right y-axis (in example: 7%).

Table 3

Observed outcomes per classification in the derivation cohort.

| Risk groups | Recurrence | Distant metastasis | Overall mortality |
|---------------------------------------|------------------|--------------------|-------------------|
| EORTC-DeCOG model | | | |
| Low risk (<25% recurrence) | 0.13 (0.06–0.20) | 0.10 (0.04–0.16) | 0.07 (0.02–0.13) |
| Intermediate risk (25–50% recurrence) | 0.38 (0.33–0.43) | 0.31 (0.26–0.36) | 0.25 (0.21–0.30) |
| High risk (50–75% recurrence) | 0.61 (0.56–0.66) | 0.55 (0.49–0.60) | 0.49 (0.43–0.54) |
| Very high risk (>75% recurrence) | 0.82 (0.73–0.88) | 0.78 (0.69–0.84) | 0.70 (0.61–0.77) |
| Simple classification | | | |
| Group 1 | 0.32 (0.26–0.36) | 0.26 (0.21–0.30) | 0.21 (0.16–0.25) |
| Group 2 | 0.52 (0.44–0.58) | 0.48 (0.40–0.54) | 0.41 (0.33–0.47) |
| Group 3 | 0.49 (0.41–0.55) | 0.42 (0.34–0.48) | 0.35 (0.28–0.42) |
| Group 4 | 0.73 (0.67–0.77) | 0.69 (0.63–0.74) | 0.60 (0.53–0.66) |
| AJCC 7th edition | | | |
| IIIA ≤ 1.0 mm | 0.32 (0.26–0.37) | 0.25 (0.20–0.30) | 0.20 (0.15–0.25) |
| IIIA >1.0 mm | 0.50 (0.42–0.57) | 0.46 (0.38–0.53) | 0.40 (0.32–0.46) |
| IIIB | 0.63 (0.58–0.67) | 0.57 (0.52–0.62) | 0.49 (0.44–0.53) |
| IIIC | 0.60 (0.02–0.84) | 0.62 (0.02–0.85) | 0.63 (0.02–0.86) |
| AJCC 8th edition | | | |
| IIIA ≤ 1.0 mm | 0.27 (0.20–0.34) | 0.21 (0.15–0.28) | 0.15 (0.09–0.21) |
| IIIA >1.0 mm | 0.37 (0.23–0.49) | 0.34 (0.20–0.46) | 0.27 (0.14–0.38) |
| IIIB | 0.43 (0.36–0.48) | 0.35 (0.29–0.41) | 0.30 (0.24–0.36) |
| IIIC | 0.64 (0.59–0.68) | 0.48 (0.53–0.63) | 0.50 (0.45–0.55) |
| IIID | 0.66 (0.00–0.90) | 0.68 (0.00–0.91) | 0.70 (0.00–0.92) |

AJCC, American Joint Committee on Cancer; CI, confidence interval. Values in parentheses are 95% confidence intervals. This table provides observed outcomes for 5-year recurrence, distant metastasis and overall mortality per risk group, e.g. when classified as low risk according to the EORTC-DeCOG model the observed 5-year recurrence was 0.13 with a 95% confidence interval of 0.06–0.20.

1) absent ulceration and ≤1.0 mm; 2) absent ulceration and >1.0 mm; 3) present ulceration and ≤1.0 mm and 4) present ulceration and >1.0 mm¹¹. The c-indices for this classification in predicting recurrence, DM and OM were 0.63 (95% CI: 0.61–0.65), 0.64 (95% CI: 0.62–0.67) and 0.64 (95% CI: 0.61–0.67), respectively. The observed outcomes for recurrence, DM and OM per risk group are shown in Table 3.

3.5. The American Joint Committee on Cancer (AJCC) classifications

Patients were classified based on the 7th AJCC classification into IIIA ≤ 1.0 mm, IIIA >1.0 mm, IIIB and

IIIC and based on the 8th edition into IIIA ≤ 1.0 mm, IIIA >1.0 mm, IIIB, IIIC and IIID. The c-indices for predicting recurrence, DM and OM for the 7th AJCC edition were 0.61 (95% CI: 0.59–0.63), 0.62 (95% CI: 0.60–0.65) and 0.62 (95% CI: 0.59–0.65), respectively, and for the 8th AJCC edition 0.62 (95% CI: 0.59–0.64), 0.63 (95% CI: 0.60–0.65) and 0.63 (95% CI: 0.61–0.66), respectively. The observed outcomes for recurrence, DM and OM for both AJCC classifications are shown in Table 3. A cross-table comparing the patients staged in accordance with the AJCC classifications and the risk groups based on the EORTC-DeCOG model is illustrated in Table 4. An overview of c-indices for all the different models is presented in Table 5.

Table 4

Cross-table comparing EORTC-DeCOG risk groups with the 7th and 8th AJCC classification, based on 937 complete cases.

| EORTC-DeCOG classification | AJCC 7th classification | | | | Total |
|---------------------------------|-------------------------|--------------|------|------|-------|
| | IIIA ≤1.0 mm | IIIA >1.0 mm | IIIB | IIIC | |
| Low risk (score 6–9) | 82 | 2 | 5 | 0 | 89 |
| Intermediate risk (score 10–15) | 207 | 83 | 93 | 2 | 385 |
| High risk (score 16–19) | 15 | 89 | 230 | 1 | 335 |
| Very high risk (score 20–23) | 0 | 11 | 113 | 4 | 128 |
| Total | 304 | 185 | 441 | 7 | 937 |
| EORTC-DeCOG classification | AJCC 8th classification | | | | Total |
| | IIIA ≤1.0 mm | IIIA >1.0 mm | IIIB | IIIC | |
| Low risk (score 6–9) | 62 | 3 | 22 | 2 | 89 |
| Intermediate risk (score 10–15) | 96 | 43 | 157 | 89 | 385 |
| High risk (score 16–19) | 0 | 10 | 71 | 253 | 335 |
| Very high risk (score 20–23) | 0 | 0 | 2 | 122 | 128 |
| Total | 158 | 56 | 252 | 466 | 937 |

AJCC, American Joint Committee on Cancer.

Table 5
C-indices with 95% confidence intervals for the different prediction models.

| | Recurrence | Distant metastasis ^a | Overall mortality ^a |
|-----------------------------------|------------------|---------------------------------|--------------------------------|
| EORTC-DeCOG prediction model | 0.68 (0.65–0.70) | 0.70 (0.67–0.72) | 0.70 (0.67–0.73) |
| EORTC-DeCOG altered model | | | |
| Derivation cohort | 0.68 (0.65–0.70) | 0.70 (0.67–0.72) | 0.70 (0.67–0.73) |
| External validation | 0.70 (0.67–0.74) | 0.72 (0.68–0.75) | 0.74 (0.71–0.78) |
| EORTC-DeCOG extended model | 0.69 (0.67–0.72) | 0.72 (0.69–0.74) | 0.72 (0.69–0.75) |
| EORTC-DeCOG simple classification | 0.63 (0.61–0.65) | 0.64 (0.62–0.67) | 0.64 (0.61–0.67) |
| AJCC 7th edition ^b | 0.61 (0.59–0.63) | 0.62 (0.60–0.65) | 0.62 (0.59–0.65) |
| AJCC 8th edition ^c | 0.62 (0.59–0.64) | 0.63 (0.60–0.65) | 0.63 (0.61–0.66) |

^a Calibrated models.

^b For IIIA ≤ 1.0 mm, IIIA > 1.0 mm, IIIB, IIIC.

^c For IIIA ≤ 1.0 mm, IIIA > 1.0 mm, IIIB, IIIC, IIID. AJCC, American Joint Committee on Cancer; SN, sentinel node.

4. Discussion

The present study developed and validated a nomogram to predict five-year recurrence, DM and OM in patients with SN-positive melanoma, by solely using information from the primary melanoma and SLNB. The resulting patient-specific probabilities could be used to tailor adjuvant therapeutic strategies for patients with SN-positive melanoma, without the prerequisite to undergo CLND and thereby avoiding potential significant morbidity. The greatest contemporary value of our prognostic nomogram is the possibility of identifying patients at sufficiently low risk for recurrence, DM and OM in whom adjuvant therapy could be omitted.

Although the FDA and EMA pragmatically approved adjuvant therapy for all stage-III patients, it is still under debate which patients should not be considered candidates. Patients with stage IIIA ≤ 1.0 mm (AJCC 7th edition) were considered low risk in most adjuvant therapy trials and were therefore not included (one even excluded all IIIA patients) [7–9,24,25]. The current study indicates that when the AJCC 8th edition criteria are used for defining IIIA ≤ 1.0 mm instead of the 7th edition, it results in improved selection of low-risk patients in terms of predicted prognosis (e.g. 5-year recurrence probability of 27% versus 32%, respectively). A recent study also showed that including SN tumour burden to the 8th AJCC staging system has crucial prognostic relevance [26]. Of note our EORTC-DeCOG model is able to identify an even more robust low-risk group, as all identified low-risk patients (which approximately concerned 4% of the entire population after imputation) had a 5-year recurrence probability of $< 25\%$ and an overall 5-year observed recurrence rate of 13%. However, identifying more robust low-risk groups comes at the cost of fewer patients being assigned low risk (see Table 4). Nonetheless a major advantage of our EORTC-DeCOG model is that it provides a more continuous type of predicted probabilities. As a result it is possible to derive risk groups based on outcome probabilities and/or risk scores (e.g. low risk; scores

6–9; recurrence probability of $< 25\%$) which is in contrast to the AJCC classifications where exact patient/tumour characteristics define the risk groups (e.g. IIIA ≤ 1.0 mm: T1a/b-T2a + N1a-N2a with ≤ 1.0 mm SN tumour burden). In the current study we choose to derive risk groups based on the recurrence probability, as this seems the most relevant outcome in the context of selecting patients for adjuvant therapy; other cut-off values and/or outcomes are possible. In conclusion, the EORTC-DeCOG model not only outperforms the AJCC classifications in terms of overall model discrimination (see Table 5), but also seems to be able to identify a more robust low-risk group in whom it may be justified to forego adjuvant therapy.

The previously published simplified model, based on ulceration and SN tumour burden, harboured the least performance, though still reasonable, and showed similar predicted prognosis for the low-risk group as the 7th AJCC edition. Whether to implement a more complex model versus a less robust model is a balance between performance and simplicity. In our opinion, the simple model could serve as an easy user-friendly prognostic tool for daily clinical practice and to generally inform patients, but for more adequate risk estimates and decisions upon (adjuvant) treatment, we advocate using the comprehensive EORTC-DeCOG model. Noteworthy, besides the common prognostic factors (i.e. ulceration, Breslow thickness and SN tumour burden), the current study also identified increasing age as an independent prognostic factor for recurrence, DM and OM. This finding is supported by other studies reporting on the significance of the patient's age [27].

Stratifying for ulceration and SN tumour burden only was previously demonstrated to yield similar discriminatory ability for melanoma-specific mortality as stratifying for AJCC substages which included information on nodal status after CLND [11]. The additional value of non-SN status retrieved after CLND was also tested in the current study, by developing an extended model. This model showed only marginal

improvement in performance (e.g. c-index for the recurrence model increased from 0.68 to 0.69), thereby indicating that omitting CLND has very limited consequences for prognostication if SN tumour burden is taken into account.

This study has several limitations. First is the retrospective design of the derivation cohort, which has inherent biases. However, the models proved to be successful in external validation. Performance was comparable between the derivation and prospective validation cohort, even though the latter cohort included patients with relatively better prognosis (e.g. less extensive disease) and largely represents a clinical trial population. Adjuvant interferon- α therapy was intended in approximately 60% of the patients included in the DeCOG-SLT trial, which is another possible limitation [4]. It could have potentially influenced outcomes, especially in patients with ulcerated melanomas as ulceration seems to be a predictive factor for IFN sensitivity [28,29]. Furthermore, it is unknown how many patients in the validation cohort received effective novel therapy after recurrence. Because patients were included from 2006 through 2014, it is likely some patients did. As patients in the derivation cohort were included from 1993 through 2008, novel therapies probably had limited effect. To date, no novel biomarker has been validated that suffices to predict long-term clinical benefits and subsequently could be incorporated in the models, despite efforts in this direction (e.g. PD-L1) [30]. In addition, other prognostic factors such as mitotic rate or microsatellites could not be incorporated in the present models because of insufficient data. Another limitation is the inadequate representation of patients with SN-positive with a head and neck melanoma in both cohorts. For the validation cohort this is largely explained as it was an exclusion criterion in the DeCOG-SLT trial, and for the derivation cohort this might be partially explained by the historical concerns of poor safety, accuracy and prognostication. Similar numbers (~5%) have been reported in other European cohorts [31,32], while particularly American cohorts have reported higher numbers (>10%) [3,33]. With the introduction of adjuvant therapies, the number of performed SLNBs in head and neck melanomas is likely to increase.

Considering the advances in the management of patients with SN-positive melanoma, it becomes highly relevant to have a prediction model that provides precise patient-specific probabilities based on solely factors from the primary melanoma and the SLNB. The EORTC-DeCOG nomogram is the first that meets these demands, and as a result it could be used for patient counselling and assist in trial design. In addition it might aid in adjuvant therapy decision-making. To facilitate its use, an online calculator has been developed and can be accessed at <https://www.evidencio.com/models/show/2010>.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial or not-for-profit sectors.

Conflict of interest statement

D.V. has served in a consultancy/advisory role for Amgen, not related to this work. U.L. has served in a consultancy/advisory role for Pierre Fabre, Roche and Novartis and has received travel/accommodation expenses from Pierre Fabre and Novartis (all not related to this work). A.C.J.v.A. has served in a consultancy/advisory role for Amgen, Bristol-Meyers Squibb, Novartis, MSD-Merck, Merck-Pfizer and Roche and has received funding from Amgen and Novartis (all not related to this work). P.R. has received honoraria from BMS, Novartis, Roche, MSD and Amgen; has served in advisory role for Novartis, MSD, BMS, Pierre Fabre and Blueprint Medicines; has been a member of the speakers bureau of Pfizer and Pierre Fabre and has received research funding by the government (all not related to this work). C.R. has received honoraria from BMS, MSD, Novartis and Roche and Pierre Fabre and has served in advisory/consultancy role for BMS, MSD, Novartis, Roche and Pierre Fabre (all not related to this work). A.T. has received travel/accommodations expenses from Agenus (not related to this work). A.A.M.v.d.V. has received honoraria from MSD, BMS, Pfizer, Eisai, Roche, Novartis and Ipsen and has served as a consultant for MSD, BMS, Pfizer, Eisai, Roche, Novartis and Ipsen (all not related to this work). U.K. has received honoraria from AstraZeneca, BMS, Merck, Novartis and Pfizer; has served in consulting/advisory role for AstraZeneca, BMS and Merck; has been a member of the speakers bureau of AstraZeneca, BMS, Merck, Novartis and Pfizer; has received research funding from AstraZeneca and Pfizer and has received travel/accommodation expenses from AstraZeneca, BMS, Ipsen and Pfizer (all not related to this work). R.S. has received honoraria from Takeda, 4SC, Novartis, Roche and MSD; has served in advisory role for Takeda, 4SC, Novartis and Roche and has received research funding from the government (all not related to this work). C.G. has received honoraria from Amgen, BMS, MSD, Novartis, Philogen, Roche and Sanofi; has served in consulting/advisory role for Amgen, BMS, MSD, Novartis, Philogen, Roche and Sanofi; has received research funding from BMS, Novartis and Roche and has received travel/accommodation expenses from Amgen, BMS, MSD, NeraCare, Novartis, Philogen, Roche and Sanofi (all not related to this work). A.M.M.E. has received honoraria from Biocad, BMS, GSK, IO Biotech, MSD, Nektar, Novartis and Pfizer; has served in advisory/consultancy role for BMS, GSK,

MSD, Novartis and Pfizer and has received travel/accommodation expenses from BMS and MSD (all not related to this work). D.J.G. has received honoraria from Amgen, BMS and Abbvie (all not related to this work). All other authors declare no competing interests.

Acknowledgements

None.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.ejca.2020.04.022>.

References

- [1] Gershenwald JE, Scolyer RA, Hess KR, Sondak VK, Long GV, Ross MI, et al. Melanoma staging: evidence-based changes in the American Joint Committee on Cancer eighth edition cancer staging manual. *CA Cancer J Clin* 2017;67:472–92.
- [2] Balch CM, Gershenwald JE, Soong SJ, Thompson JF, Atkins MB, Byrd DR, et al. Final version of 2009 AJCC melanoma staging and classification. *J Clin Oncol* 2009;27:6199–206.
- [3] Faries MB, Thompson JF, Cochran AJ, Andtbacka RH, Mozzillo N, Zager JS, et al. Completion dissection or observation for sentinel-node metastasis in melanoma. *N Engl J Med* 2017;376:2211–22.
- [4] Leiter U, Stadler R, Mauch C, Hohenberger W, Brockmeyer N, Berking C, et al. Complete lymph node dissection versus no dissection in patients with sentinel lymph node biopsy positive melanoma (DeCOG-SLT): a multicentre, randomised, phase 3 trial. *Lancet Oncol* 2016;17:757–67.
- [5] Leiter U, Stadler R, Mauch C, Hohenberger W, Brockmeyer N, Berking C, et al. Final analysis of DeCOG-SLT trial: survival outcomes of complete lymph node dissection in melanoma patients with positive sentinel node. *J Clin Oncol* 2018;36 (suppl; abstr 9501).
- [6] Leiter U, Stadler R, Mauch C, Hohenberger W, Brockmeyer N, Berking C, et al. Final analysis of DeCOG-SLT trial: No survival benefit for complete lymph node dissection in patients with melanoma with positive sentinel node. *J Clin Oncol* 2019; JCO1802306.
- [7] Eggermont AMM, Blank CU, Mandala M, Long GV, Atkinson V, Dalle S, et al. Adjuvant pembrolizumab versus placebo in resected stage III melanoma. *N Engl J Med* 2018;378:1789–801.
- [8] Long GV, Hauschild A, Santinami M, Atkinson V, Mandala M, Chiarion-Sileni V, et al. Adjuvant dabrafenib plus trametinib in stage III BRAF-mutated melanoma. *N Engl J Med* 2017;377:1813–23.
- [9] Weber J, Mandala M, Del Vecchio M, Gogas HJ, Arance AM, Cowey CL, et al. Adjuvant nivolumab versus ipilimumab in resected stage III or IV melanoma. *N Engl J Med* 2017;377:1824–35.
- [10] Eggermont AM, Chiarion-Sileni V, Grob JJ, Dummer R, Wolchok JD, Schmidt H, et al. Prolonged survival in stage III melanoma with ipilimumab adjuvant therapy. *N Engl J Med* 2016;375:1845–55.
- [11] Verver D, van Klaveren D, van Akkooi ACJ, Rutkowski P, Powell B, Robert C, et al. Risk stratification of sentinel node-positive melanoma patients defines surgical management and adjuvant therapy treatment considerations. *Eur J Cancer* 2018;96:25–33.
- [12] Gershenwald JE, Andtbacka RH, Prieto VG, Johnson MM, Diwan AH, Lee JE, et al. Microscopic tumor burden in sentinel lymph nodes predicts synchronous nonsentinel lymph node involvement in patients with melanoma. *J Clin Oncol* 2008;26:4296–303.
- [13] Lee JH, Essner R, Torisu-Itakura H, Wanek L, Wang H, Morton DL, et al. Factors predictive of tumor-positive non-sentinel lymph nodes after tumor-positive sentinel lymph node dissection for melanoma. *J Clin Oncol* 2004;22:3677–84.
- [14] Murali R, Desilva C, Thompson JF, Scolyer RA. Non-Sentinel Node Risk Score (N-SNORE): a scoring system for accurately stratifying risk of non-sentinel node positivity in patients with cutaneous melanoma with positive sentinel lymph nodes. *J Clin Oncol* 2010;28:4441–9.
- [15] van der Ploeg AP, van Akkooi AC, Rutkowski P, Nowecki ZI, Michej W, Mitra A, et al. Prognosis in patients with sentinel node-positive melanoma is accurately defined by the combined Rotterdam tumor load and Dewar topography criteria. *J Clin Oncol* 2011;29:2206–14.
- [16] Oude Ophuis CM, van Akkooi AC, Rutkowski P, Powell WE, Robert C, Testori A, et al. Timing of completion lymphadenectomy after positive sentinel node biopsy in patients with melanoma. *Br J Surg* 2017;104:726–33.
- [17] Oude Ophuis CM, Verhoef C, Rutkowski P, Powell BW, van der Hage JA, van Leeuwen PA, et al. The interval between primary melanoma excision and sentinel node biopsy is not associated with survival in sentinel node positive patients - an EORTC Melanoma Group study. *Eur J Surg Oncol* 2016;42:1906–13.
- [18] Kattan MW, Hess KR, Amin MB, Lu Y, Moons KG, Gershenwald JE, et al. American Joint Committee on Cancer acceptance criteria for inclusion of risk models for individualized prognosis in the practice of precision medicine. *CA Cancer J Clin* 2016;66:370–4.
- [19] van Buuren S, Groothuis-Oudshoorn K. MICE: multivariate imputation by chained equations in R. *J Stat Software* 2011;45:1–67.
- [20] Steyerberg EW. Clinical prediction models. Springer; 2009.
- [21] Harrell Jr FE, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med* 1996;15:361–87.
- [22] Steyerberg EW, Harrell Jr FE. Prediction models need appropriate internal, internal-external, and external validation. *J Clin Epidemiol* 2016;69:245–7.
- [23] van Akkooi AC, Nowecki ZI, Voit C, Schafer-Hesterberg G, Michej W, de Wilt JH, et al. Sentinel node tumor burden according to the Rotterdam criteria is the most important prognostic factor for survival in melanoma patients: a multicenter study in 388 patients with positive sentinel nodes. *Ann Surg* 2008;248:949–55.
- [24] Jeffrey SW, Mario M, Michele Del V, Helen G, Ana MA, Charles Lance C, et al. Adjuvant therapy with nivolumab (NIVO) versus ipilimumab (IPI) after complete resection of stage III/IV melanoma: updated results from a phase III trial (CheckMate 238). *J Clin Oncol* 2018;36. 9502-9502.
- [25] Hauschild A, Dummer R, Schadendorf D, Santinami M, Atkinson V, Mandala M, et al. Longer follow-up confirms relapse-free survival benefit with adjuvant dabrafenib plus trametinib in patients with resected BRAF V600-mutant stage III melanoma. *J Clin Oncol* 2018;36(35):3441–9.
- [26] Satzger I, Leiter U, Grager N, Keim U, Garbe C, Gutzmer R. Melanoma-specific survival in patients with positive sentinel lymph nodes: relevance of sentinel tumor burden. *Eur J Cancer* 2019;123:83–91.
- [27] Kretschmer L, Starz H, Thoms KM, Satzger I, Volker B, Jung K, et al. Age as a key factor influencing metastasizing patterns and disease-specific survival after sentinel lymph node biopsy for cutaneous melanoma. *Int J Cancer* 2011;129:1435–42.

- [28] Eggermont AM, Suci S, Testori A, Santinami M, Kruit WH, Marsden J, et al. Long-term results of the randomized phase III trial EORTC 18991 of adjuvant therapy with pegylated interferon alfa-2b versus observation in resected stage III melanoma. *J Clin Oncol* 2012;30:3810–8.
- [29] Eggermont AM, Suci S, Rutkowski P, Kruit WH, Punt CJ, Dummer R, et al. Long term follow up of the EORTC 18952 trial of adjuvant therapy in resected stage IIB-III cutaneous melanoma patients comparing intermediate doses of interferon-alpha-2b (IFN) with observation: ulceration of primary is key determinant for IFN-sensitivity. *Eur J Cancer* 2016;55:111–21.
- [30] Nishino M, Ramaiya NH, Hatabu H, Hodi FS. Monitoring immune-checkpoint blockade: response evaluation and biomarker development. *Nat Rev Clin Oncol* 2017;14:655–68.
- [31] Roulin D, Matter M, Bady P, Lienard D, Gugerli O, Boubaker A, et al. Prognostic value of sentinel node biopsy in 327 prospective melanoma patients from a single institution. *Eur J Surg Oncol* 2008;34:673–9.
- [32] Richtig G, Richtig E, Neiss AN, Quehenberger F, Gmainer DG, Kamolz LP, et al. Does the time interval between sentinel lymph node biopsy and completion lymph node dissection affect outcome in malignant melanoma? A retrospective cohort study. *Int J Surg* 2020;75:160–4.
- [33] Klemen ND, Han G, Leong SP, Kashani-Sabet M, Vetto J, White R, et al. Completion lymphadenectomy for a positive sentinel node biopsy in melanoma patients is not associated with a survival benefit. *J Surg Oncol* 2019;119:1053–9.