Title: Machine learning prediction for early-stage melanoma outcomes: recurrence-free survival, disease-specific survival, and overall survival

Guihong Wan, PhD[1*], Hannah Rashdan, MD[1*], Olivia M. Burke, BS[1], Sara Khattab, MD[1], Nga Nguyen, MD, MPH[1], Bonnie W. Leung, MD[1], Emma Beagles, MD[1], Crystal T. Chang, MD[1], Kun-Hsing Yu, MD, PhD[2,3], Mia S. DeSimone, MD, MPH[3**], and Yevgeniy R. Semenov, MD, MA[1**]

[1]Department of Dermatology, Massachusetts General Hospital, Harvard Medical School
[2]Department of Biomedical Informatics, Harvard Medical School
[3]Department of Pathology, Brigham and Women's Hospital, Harvard Medical School

*Designates co-first authors.
**Designates co-senior authors.

Corresponding author:
Yevgeniy R. Semenov, MD, MA
Department of Dermatology
Massachusetts General Hospital
Harvard Medical School
40 Blossom Street
Bartlett Hall 6R, Room 626
Boston, MA 02114
Email: ysemenov@mgh.harvard.edu

Tables: 2
Figures: 1
References: 13
Supplementary figures: 3
Supplementary tables: 1

Keywords: machine learning; early-stage melanoma; melanoma recurrence; melanoma-specific survival; overall survival

ABSTRACT

This study compared machine-learning models for predicting recurrence-free survival (RFS), disease-specific survival (DSS), and overall survival (OS) using clinicopathologic data from 1,621 stage I/II primary cutaneous melanoma patients. Our time-to-event models achieved concordance indices of 0.829 for RFS, 0.812 for DSS, and 0.778 for OS. Tumor thickness and mitotic rate were the most important predictors for RFS. Charlson comorbidity score and insurance type were critical for DSS and OS.

Melanoma is a highly aggressive cancer with a substantial risk of recurrence. Most patients with melanoma recurrence had the disease that was initially early-stage (stage I/II) at diagnosis.[1] Tumor thickness, clinical stage, mitotic rate, age at diagnosis, melanoma location, and ulceration status have been associated with melanoma recurrence.[2,3] However, widely adopted, comprehensive risk stratification approaches for early-stage melanoma patients remain lacking, with the current reliance on AJCC staging alone offering limited predictive accuracy for identifying patients at the highest risk of recurrence.[4]

The recent approval by the US Food and Drug Administration of adjuvant immune checkpoint inhibitor (ICI) therapy for managing stage IIB and IIC melanoma underscores the need for risk stratification approaches to identify early-stage melanoma patients with more aggressive disease for whom the potential benefits of immunotherapy may outweigh risk of toxicities associated with this therapeutic class. The extension of ICI therapy to this population was based on clinical trial results demonstrating improved recurrence-free survival (RFS) among patients treated with adjuvant ICIs compared to those who only underwent surgical excision.[5] While RFS is an important endpoint, it has limitations, particularly in a setting where most recurrences are localized and are, therefore, not expected to impact mortality directly.[2] Considering that a major goal of cancer treatment is to extend survival, disease-specific survival (DSS) and overall survival (OS) are considered gold standard endpoints in oncology,[6] but typically require trials with larger patient populations and longer follow-up. Thus, it is crucial to examine the extent to which early-stage melanoma risk stratification and management using RFS as an endpoint can be reliably extended to predict DSS and OS.

In this study, we aimed to address these gaps by comparing machine-learning models for predicting RFS, DSS, and OS, and investigating whether models trained on RFS data can also reliably estimate DSS and OS in this population.

Patient characteristics

This study identified 1,621 patients diagnosed with stage I/II primary cutaneous melanoma at Massachusetts General Hospital (MGH) and Brigham and Women's Hospital/Dana-Farber Cancer Institute (BWH/DFCI) between 2000 and 2020. We extracted 31 clinicopathologic features from patients' electronic health records. The characteristics of the study population are presented in **Table 1**, with a complete list of clinicopathologic features provided in **Supplementary Table 1**. In the MGH cohort, 217 (19.7%) patients experienced disease recurrence, 106 (9.6%) died from melanoma, and a total of 382 (34.7%) were dead at the last follow-up. In the BWH/DFCI cohort, 91 (17.5%) patients experienced disease recurrence, 33 (6.3%) died from melanoma, and a total of 105 (20.2%) were dead at the last follow-up. The cumulative incidence of melanoma recurrence, melanoma-specific mortality, and overall mortality are presented in **Figure 1**. The melanoma-specific mortality rate was higher among patients with distant recurrence compared to patients with locoregional recurrence (57.9% vs 32.7%, P-value<0.001). Compared to patients with non-recurrent melanoma, patients with recurrent melanoma had a higher risk of mortality (HR: 2.76; P-value<0.001) (**Supplementary Figure 1).**

Machine-learning model performance

To predict RFS, DSS, and OS, we used two time-to-event machine-learning algorithms, Gradient Boosting Survival (GBS)[10] and Random Survival Forest (RSF),[11] the best performers in our previous studies.[2,7] Models were evaluated by concordance index[12] using (1) five-fold

cross-validation of the MGH cohort (internal validation); (2) the MGH cohort for model training and the BWH/DFCI cohort for testing the models independently (external validation). We also examined the transferability of the model trained for recurrence prediction on the MGH cohort to estimate the risk of melanoma-specific and overall mortality on the BWH/DFCI cohort (cross-transfer validation).

Table 2 presents model performance for predicting melanoma outcomes. In the RFS prediction, the best performance was 0.877 (95% CI: 0.871-0.882) and 0.829 (95% CI: 0.827-0.831) in the internal and external validations, respectively. The best performance in the DSS prediction was 0.875 (95% CI: 0.862-0.888) and 0.812 (95% CI: 0.810-0.814) in the internal and external validations, respectively. In the OS prediction, the best performance was 0.826 (95% CI: 0.821-0.831) and 0.778 (95% CI: 0.777-0.781) in the internal and external validations, respectively. In the cross-transfer validation, the model performance deteriorated significantly for DSS (0.789; 95% CI: 0.786-0.792) and, to a greater extent, OS (0.631; 95% CI: 0628-0.634).

We also investigated the critical features in the predictions by conducting permutation importance (**Supplementary Figures 2 and 3**). In both GBS and RFS models, thickness and mitotic rate were the most important features in the RFS prediction, while Charlson comorbidity score (CCS) is dominantly predictive for OS. In the DSS prediction, thickness and mitotic rate were ranked higher compared to the OS prediction in both GBS and RFS models.

Discussion

This large-scale retrospective cohort study compares the predictive performance of machine-learning models for RFS, DSS, and OS in early-stage melanoma population. The models achieved similar performance for RFS (concordance index: 0.829), DSS (concordance index: 0.812), and OS (concordance index: 0.778) in the external validation. These results

demonstrate that machine-learning models using a fixed set of features can independently predict outcomes of patients with early-stage melanoma.

The performance significantly declined when using the same model trained for predicting recurrence (concordance index: 0.829) to estimate the risk of melanoma-specific mortality and overall mortality (concordance index: 0.789 and 0.631, respectively). This raises concerns regarding the inherent assumption that the same factors that contribute to recurrence also lead to increased melanoma-specific mortality and overall mortality. As such, risk stratification utilizing disease recurrence as a prognostic endpoint can be extended to an extent to predict melanoma-specific mortality but not overall mortality in this population. Considering that recurrence-free survival is the outcome used in many clinical studies guiding oncologic algorithms, including in the most recent approval of ICIs for stage IIB/IIC melanoma,[5] clinicians must practice caution when drawing conclusions regarding melanoma-specific mortality and overall mortality using the same features as for recurrence-free survival.

Furthermore, we found tumor thickness and mitotic rate were primary features in recurrence prediction, whereas CCS was the most notable feature in overall mortality prediction. Despite near-universal reliance of clinical algorithms on Breslow thickness and stage for prognostication,[1] our results demonstrate limitations of predicting melanoma recurrence and mortality with current tools. Further risk stratification evaluating additional distinguishing features associated with melanoma recurrence and mortality is warranted.

Methods

This study builds on recent studies from our group, which developed time-to-event machine-learning approaches for predicting disease recurrence in early-stage melanoma.[2,7] We identified a retrospective cohort of 1,621 patients with stage I/II primary cutaneous melanoma

diagnosed between January 2000 and February 2020, among which 1,101 were from Massachusetts General Hospital (MGH), and 520 were from Brigham and Women's Hospital/Dana-Farber Cancer Institute (BWH/DFCI). We extracted 31 clinicopathologic features, including 10 clinical features and 21 tumor characteristics (**Supplementary Table 1**), and patient outcomes, including recurrence (status and date of first recurrence), melanoma-specific mortality, and overall mortality (status, death reason, and death date or last follow-up). Inclusion criteria and data extraction were consistent with our previous studies,[2,4,7] except that the diagnosis date of the first melanoma was retained for patients with multiple stage I/II melanomas.

To compare groups, we used Pearson's chi-squared test and Student's t-test for categorical and continuous variables, respectively. Kaplan-Meier curves examined differences in overall survival stratified by melanoma recurrence type. To account for guarantee-time bias, Hazard Ratios (HRs) of melanoma recurrence on survival were computed using time-varying Cox proportional hazards models,[8,9] adjusted for baseline demographics and tumor characteristics.

To predict RFS, DSS, and OS, we used two time-to-event machine-learning algorithms, GradientBoostingSurvivalAnalysis (GBS)[10] and RandomSurvivalForest (RSF),[11] the best performers in our previous studies.[2,7] Models were evaluated by concordance index[12] using (1) five-fold cross-validation of the MGH cohort (internal); (2) the MGH cohort for model training and the BWH/DFCI cohort for independent validation (external). We also examined the transferability of the model trained on the MGH cohort for recurrence prediction to estimate the risk of melanoma-specific and overall mortality on the BWH/DFCI cohort (cross-transfer validation). Each experiment was repeated 20 times. Mean and 95% confidence interval (CI)

were reported. We further ranked features in the prediction by conducting permutation importance with 20 repeats. Experiments were implemented using the Python module scikit-survival 0.18.0.[13]

DATA AVAILABILITY

All relevant data are available from the corresponding author: Yevgeniy R. Semenov. All summary data supporting the findings of this study are available within the article and/or its supplementary materials. The patient data generated for this study can only be shared per specific institutional review board (IRB) requirements. Upon a request to the corresponding author, a data-sharing agreement can be initiated following institution-specific guidelines.

172 **Table 1.** Characteristics of the study population

	MGH (N=1101)	BWH/DFCI (N=520)	P-value
Sex			
Female	467 (42.4%)	246 (47.3%)	0.072
Male	634 (57.6%)	274 (52.7%)	
Age at diagnosis			
Mean (SD)	58.7 (16.2)	60.4 (14.4)	0.034
Median [Min, Max]	60 [12, 92]	62 [20, 93]	
Race			
White	1092 (99.2%)	512 (98.5%)	0.285
Other/Unknown	9 (0.8%)	8 (1.5%)	
Year at diagnosis			
< 2005	317 (28.8%)	0 (0%)	<0.001
2005 - 2009	424 (38.5%)	0 (0%)	
2010 - 2014	249 (22.6%)	352 (67.7%)	
>= 2015	111 (10.1%)	168 (32.3%)	
Clinical stage			
1A	565 (51.3%)	265 (51.0%)	0.739
1B	311 (28.2%)	145 (27.9%)	
2A	89 (8.1%)	39 (7.5%)	
2B	77 (7.0%)	46 (8.8%)	
2C	59 (5.4%)	25 (4.8%)	
Histology type			
Lentigo maligna melanoma	106 (9.6%)	19 (3.7%)	<0.001
Melanoma, NOS	109 (9.9%)	207 (39.8%)	
Nodular melanoma	150 (13.6%)	46 (8.8%)	
Superficial spreading melanoma	736 (66.8%)	248 (47.7%)	
Thickness			
Mean (SD)	1.38 (2.54)	1.43 (1.99)	0.707
Median [Min, Max]	0.72 [0.05, 60]	0.75 [0.12, 20]	
Ulceration			
Absent	960 (87.2%)	455 (87.5%)	0.829
Present	135 (12.3%)	61 (11.7%)	
Unknown	6 (0.5%)	4 (0.8%)	
Mitotic rate (mitoses/mm^2)			
Mean (SD)	2.89 (5.29)	2.89 (4.89)	0.999
Median [Min, Max]	1 [0, 45]	1 [0, 35]	
Tumor location			
Lower limb and hip	211 (19.2%)	107 (20.6%)	0.111
Scalp and neck	203 (18.4%)	116 (22.3%)	
Trunk	414 (37.6%)	167 (32.1%)	
Upper limb and shoulder	273 (24.8%)	130 (25.0%)	
Sentinel lymph node biopsy (SLNB)			
Not indicated	588 (53.4%)	210 (40.4%)	<0.001
Performed; all nodes negative	397 (36.1%)	282 (54.2%)	
Not performed due to unknown reasons	34 (3.1%)	7 (1.3%)	
Not performed due to age or comorbidity	72 (6.5%)	18 (3.5%)	
Deferred by patient	10 (0.9%)	3 (0.6%)	
Recurrence			
No	884 (80.3%)	429 (82.5%)	0.322
Yes	217 (19.7%)	91 (17.5%)	
Melanoma-specific mortality			
No	995 (90.4%)	487 (93.7%)	0.035
Yes	106 (9.6%)	33 (6.3%)	
Overall mortality			
Alive	719 (65.3%)	415 (79.8%)	
Dead	382 (34.7%)	105 (20.2%)	
Diagnosis to recurrence (years)			
Mean (SD)	3.1 (3.1)	2.6 (2.0)	0.097
Median [Min, Max]	2.0 [0.2, 16.8]	2.0 [0.3, 9.3]	
Duration of follow-up (years)			
Mean (SD)	9.2 (6.1)	6.2 (2.4)	<0.001
Median [Min, Max]	9.0 [0.3, 23.0]	6.6 [0.6, 11.7]	

173

Table 2. Prediction of melanoma outcomes

	Prediction	Internal Validation (mean and 95% CI)		External Validation (mean and 95% CI)	
		RSF	**GBS**	**RSF**	**GBS**
Reference[1]	RFS	0.870 0.863-0.876	0.877 0.871-0.882	0.804 0.802-0.806	**0.829** **0.827-0.831**
Individual[2]	DSS	0.873 0.860-0.886	0.875 0.862-0.888	0.783 0.777-0.789	**0.812** **0.810-0.814**
	OS	0.826 0.823-0.828	0.826 0.821-0.831	0.776 0.774-0.779	**0.778** **0.777-0.781**
Transfer[3]	DSS	N/A	N/A	0.749 0.745-0.752	**0.789** **0.786-0.792**
	OS	N/A	N/A	0.614 0.611-0.619	**0.631** **0.628-0.634**

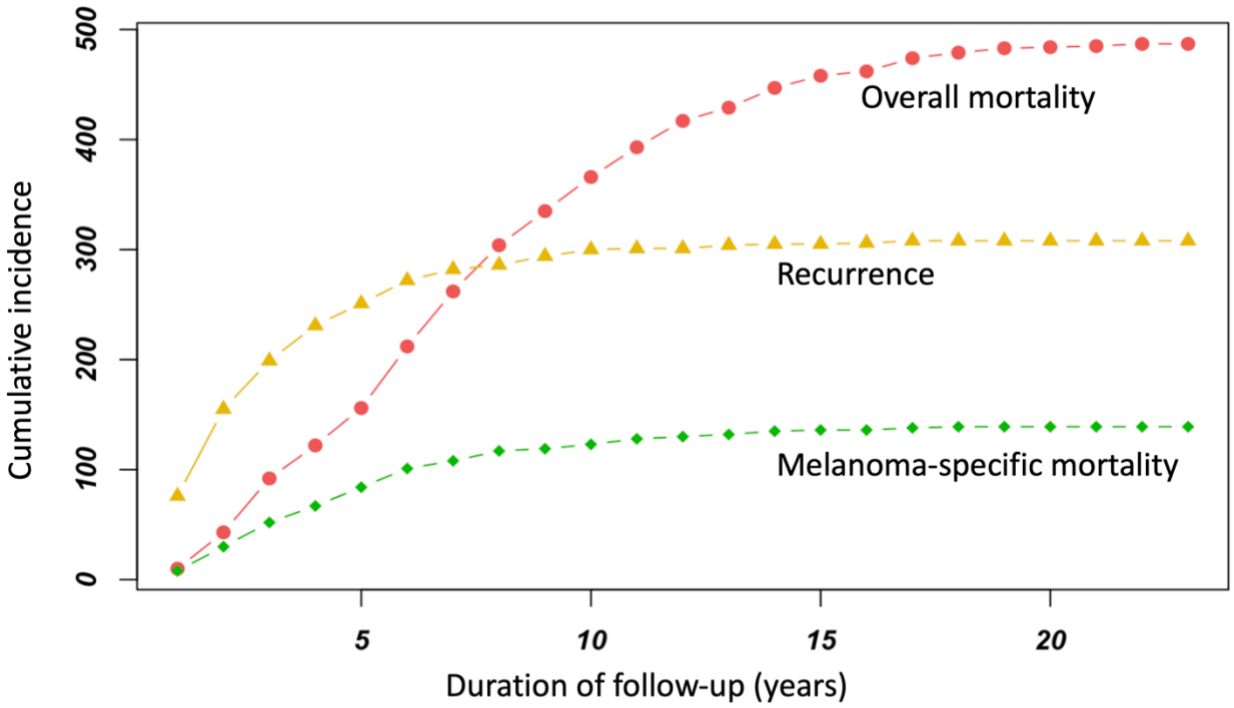
[1]Models were trained and evaluated with RFS as the outcome.

[2]Models were trained and evaluated with DSS and OS as the outcome, respectively.

[3]The model was trained with RFS as the outcome. The same model was evaluated with DSS and OS as the outcome.

N/A: not applicable

Figure 1. Cumulative incidence of melanoma recurrence, melanoma-specific mortality, and overall mortality



Recurrence status	Overall mortality	Melanoma-specific mortality
Distant	97/152 (63.8%)	88/152 (57.9%)
Locoregional	71/156 (45.5%)	51/156 (32.7%)
Non-recurrence	319/1313 (24.3%)	0

Legend: The cumulative incidence pattern of melanoma-specific mortality over time is similar to that of melanoma recurrence, while the cumulative incidence pattern of overall mortality is different compared to that of melanoma recurrence.

Compared to locoregional recurrence, the Hazard Ratios of distant recurrence on overall mortality and melanoma-specific mortality are 2.64 (95% CI: 2.03-3.44; $p<0.001$) and 8.83 (95% CI: 5.87-13.3; $p<0.001$), respectively.

References

1. Bajaj S, Donnelly D, Call M, et al. Melanoma Prognosis: Accuracy of the American Joint Committee on Cancer Staging Manual Eighth Edition. *J Natl Cancer Inst*. 2020;112(9):921-928. doi:10.1093/jnci/djaa008

2. Wan G, Nguyen N, Liu F, et al. Prediction of early-stage melanoma recurrence using clinical and histopathologic features. *npj Precis Onc*. 2022;6(1):1-16. doi:10.1038/s41698-022-00321-4

3. von Schuckmann LA, Hughes MCB, Ghiasvand R, et al. Risk of Melanoma Recurrence After Diagnosis of a High-Risk Primary Tumor. *JAMA Dermatol*. 2019;155(6):688-693. doi:10.1001/jamadermatol.2019.0440

4. Wan G, Leung B, Nguyen N, et al. The Impact of Stage-Related Features in Melanoma Recurrence Prediction: A Machine Learning Approach. *JAAD International*. Published online August 30, 2022. doi:10.1016/j.jdin.2022.08.014

5. Luke JJ, Rutkowski P, Queirolo P, et al. Pembrolizumab versus placebo as adjuvant therapy in completely resected stage IIB or IIC melanoma (KEYNOTE-716): a randomised, double-blind, phase 3 trial. *The Lancet*. 2022;399(10336):1718-1729. doi:10.1016/S0140-6736(22)00562-1

6. Delgado A, Guddati AK. Clinical endpoints in oncology - a primer. *Am J Cancer Res*. 2021;11(4):1121-1131.

7. Wan G, Leung BW, DeSimone MS, et al. Development and validation of time-to-event models to predict metastatic recurrence of localized cutaneous melanoma. *J Am Acad Dermatol*. 2024;90(2):288-298. doi:10.1016/j.jaad.2023.08.105

8. Rashdan H, Zhang S, Wan G, et al. Timing of cutaneous immune-related adverse events impacts survival in patients with cancer treated with immune-checkpoint inhibitors: A multi-institutional cohort study. *J Am Acad Dermatol*. Published online January 28, 2024:S0190-9622(24)00147-6. doi:10.1016/j.jaad.2024.01.038

9. Zhang S, Tang K, Wan G, et al. Cutaneous immune-related adverse events are associated with longer overall survival in advanced cancer patients on immune checkpoint inhibitors: A multi-institutional cohort study. *Journal of the American Academy of Dermatology*. Published online February 2, 2023. doi:10.1016/j.jaad.2022.12.048

10. Survival ensembles | Biostatistics | Oxford Academic. Accessed August 10, 2023. https://academic.oup.com/biostatistics/article/7/3/355/248945?login=false

11. Ishwaran H, Kogalur UB, Blackstone EH, Lauer MS. Random survival forests. *The Annals of Applied Statistics*. 2008;2(3):841-860. doi:10.1214/08-AOAS169

12. Antolini L, Boracchi P, Biganzoli E. A time-dependent discrimination index for survival data. *Stat Med*. 2005;24(24):3927-3944. doi:10.1002/sim.2427

13. Pölsterl S. scikit-survival: A Library for Time-to-Event Analysis Built on Top of scikit-learn. *Journal of Machine Learning Research*. 2020;21(212):1-6.

ACKNOWLEDGMENTS

GW is supported by the National Cancer Institute of the National Institutes of Health under Award Number K99CA286966. YRS is supported in part by the Department of Defense under Award Number W81XWH2110819, the National Institute of Arthritis and Musculoskeletal and Skin Diseases of the National Institutes of Health under Award Number K23AR080791, and the Melanoma Research Alliance Young Investigator Award.

AUTHOR CONTRIBUTIONS

GW, HR, MSD, and YRS had full access to all the data in the study and took responsibility for the integrity of the data and the accuracy of the data analysis. Study concept and design: GW, HR, MSD, and YRS. Data acquisition: HR, OMB, SK, NN, GW, BWL, EB, CTC, BY, MSD, and YRS. Data analysis and interpretation: GW, HR, MSD, and YRS. Drafting of the manuscript: GW, HR, SK, MSD, and YRS. Administrative, technical, or material support: YRS. The study was conducted under the supervision of MSD and YRS. All authors reviewed, edited, and approved the manuscript.

COMPETING INTERESTS:

YRS is an advisory board member/consultant and has received honoraria from Incyte Corporation, Castle Biosciences, Galderma, Pfizer, and Sanofi outside the submitted work.

Supplementary Table 1. Characteristics of the study population for all extracted variables

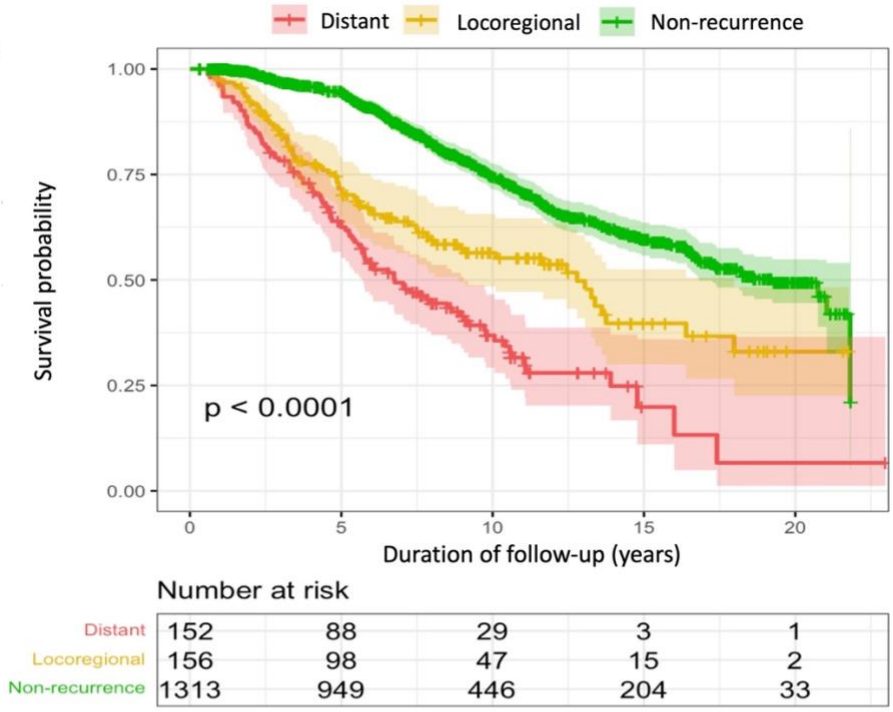
	MGH (N=1101)	BWH/DFCI (N=520)	P-value
1. Sex			
Female	467 (42.4%)	246 (47.3%)	0.072
Male	634 (57.6%)	274 (52.7%)	
2. Age at diagnosis [a]			
Mean (SD)	59 (16)	60 (14)	0.034
Median [Min, Max]	60 [12, 92]	62 [20, 93]	
3. Year at diagnosis [a]			
< 2005	317 (28.8%)	0 (0%)	<0.001
2005 - 2009	424 (38.5%)	0 (0%)	
2010 - 2014	249 (22.6%)	352 (67.7%)	
>= 2015	111 (10.1%)	168 (32.3%)	
4. Race			
White	1092 (99.2%)	512 (98.5%)	0.285
Other/Unknown	9 (0.8%)	8 (1.5%)	
5. Ethnicity			
Hispanic	1 (0.1%)	4 (0.8%)	<0.001
Non-Hispanic	1096 (99.5%)	499 (96.0%)	
Unknown	4 (0.4%)	17 (3.3%)	
6. Median income (thousands of dollars)			
Mean (SD)	105 (36)	102 (34)	0.168
Median [Min, Max]	101 [24, 241]	98 [36, 241]	
7. Insurance type			
Medicaid	17 (1.5%)	8 (1.5%)	<0.001
Medicare	535 (48.6%)	271 (52.1%)	
None or Self-Pay	114 (10.4%)	22 (4.2%)	
Private	435 (39.5%)	219 (42.1%)	
8. Marital status			
Divorced	64 (5.8%)	38 (7.3%)	0.034
Married	746 (67.8%)	376 (72.3%)	
Other	9 (0.8%)	1 (0.2%)	
Single	164 (14.9%)	71 (13.7%)	
Unavailable	33 (3.0%)	7 (1.3%)	
Widowed	85 (7.7%)	27 (5.2%)	
9. First visit due to melanoma (FVM) [b]			
No	280 (25.4%)	129 (24.8%)	0.835
Yes	821 (74.6%)	391 (75.2%)	
10. Charlson comorbidity score (CCS) [b]			
Mean (SD)	2.6 (2.1)	2.6 (1.9)	0.891
Median [Min, Max]	2.0 [0, 12.0]	2.0 [0, 13.0]	
11. Clinical stage			
1A	565 (51.3%)	265 (51.0%)	0.739
1B	311 (28.2%)	145 (27.9%)	
2A	89 (8.1%)	39 (7.5%)	
2B	77 (7.0%)	46 (8.8%)	
2C	59 (5.4%)	25 (4.8%)	
12. Histology type			
Lentigo maligna melanoma	106 (9.6%)	19 (3.7%)	<0.001
Melanoma, NOS	109 (9.9%)	207 (39.8%)	
Nodular melanoma	150 (13.6%)	46 (8.8%)	
Superficial spreading melanoma	736 (66.8%)	248 (47.7%)	
13. Thickness			
Mean (SD)	1.4 (2.5)	1.4 (2.0)	0.707
Median [Min, Max]	0.7 [0.1, 60.0]	0.8 [0.1, 20.0]	
14. Anatomic level			
Mean (SD)	3.3 (0.9)	3.5 (1.0)	<0.001
Median [Min, Max]	4 [2, 6]	4 [2, 6]	
15. Ulceration			
Absent	960 (87.2%)	455 (87.5%)	0.829
Present	135 (12.3%)	61 (11.7%)	
Unknown	6 (0.5%)	4 (0.8%)	
16. Mitotic rate (mitoses/mm^2) [c]			
Mean (SD)	2.9 (5.3)	2.9 (4.9)	0.999
Median [Min, Max]	1.0 [0, 45.0]	1.0 [0, 35.0]	
17. Tumor location			
Lower limb and hip	211 (19.2%)	107 (20.6%)	0.111
Scalp and neck	203 (18.4%)	116 (22.3%)	

Trunk	414 (37.6%)	167 (32.1%)	
Upper limb and shoulder	273 (24.8%)	130 (25.0%)	
18. Laterality			
Left	524 (47.6%)	244 (46.9%)	0.118
Midline	110 (10.0%)	37 (7.1%)	
Right	467 (42.4%)	239 (46.0%)	
19. Sentinel lymph node biopsy (SLNB)			
Not indicated	588 (53.4%)	210 (40.4%)	<0.001
All nodes negative	397 (36.1%)	282 (54.2%)	
Not performed due to unknown reasons	34 (3.1%)	7 (1.3%)	
Not performed due to age or comorbidity	72 (6.5%)	18 (3.5%)	
Deferred by patient	10 (0.9%)	3 (0.6%)	
20. Total surgical margins (cm) [c]			
Mean (SD)	1.3 (0.5)	1.3 (0.5)	0.621
Median [Min, Max]	1.0 [0.2, 3.5]	1.0 [0.1, 3.0]	
21. Check if meeting margin requirements (Margin check)			
Yes	860 (78.1%)	441 (84.8%)	0.006
No	65 (5.9%)	24 (4.6%)	
Unknown	176 (16.0%)	55 (10.6%)	
22. Tumor infiltrating lymphocytes (TIL)			
Absent	247 (22.4%)	82 (15.8%)	<0.001
Present	587 (53.3%)	326 (62.7%)	
Unknown	267 (24.3%)	112 (21.5%)	
23. Tumor infiltrating lymphocytes type (TIL type)			
Absent	247 (22.4%)	82 (15.8%)	<0.001
Brisk	60 (5.4%)	52 (10.0%)	
Non-Brisk	524 (47.6%)	271 (52.1%)	
Unknown	267 (24.3%)	112 (21.5%)	
Unknown type	3 (0.3%)	3 (0.6%)	
24. Radial growth (RG)			
Absent	165 (15.0%)	34 (6.5%)	<0.001
Present	884 (80.3%)	399 (76.7%)	
Unknown	52 (4.7%)	87 (16.7%)	
25. Vertical growth (VG)			
Absent	259 (23.5%)	8 (1.5%)	<0.001
Present	798 (72.5%)	389 (74.8%)	
Unknown	44 (4.0%)	123 (23.7%)	
26. Vertical growth type (VGT)			
Absent	259 (23.5%)	8 (1.5%)	<0.001
Epithelioid	510 (46.3%)	234 (45.0%)	
Epithelioid and nevoid	5 (0.5%)	21 (4.0%)	
Epithelioid and small cell	42 (3.8%)	15 (2.9%)	
Epithelioid and spindled	103 (9.4%)	59 (11.3%)	
Other	10 (0.9%)	13 (2.5%)	
Small cell	11 (1.0%)	3 (0.6%)	
Spindled	49 (4.5%)	5 (1.0%)	
Unknown	44 (4.0%)	123 (23.7%)	
Unknown type	68 (6.2%)	39 (7.5%)	
27. Precursor lesion [d]			
Absent	726 (65.9%)	395 (76.0%)	<0.001
Present	375 (34.1%)	125 (24.0%)	
28. Precursor type [d]			
Absent	726 (65.9%)	395 (76.0%)	<0.001
Benign nevus	164 (14.9%)	75 (14.4%)	
Dermal nevus	1 (0.1%)	1 (0.2%)	
Dysplastic nevus	179 (16.3%)	39 (7.5%)	
Lentigo maligna	25 (2.3%)	1 (0.2%)	
Unknown type	6 (0.5%)	9 (1.7%)	
29. Regression [d]			
Absent	997 (90.6%)	396 (76.2%)	<0.001
Present	104 (9.4%)	124 (23.8%)	
30. Lymphovascular invasion [d]			
Absent	1077 (97.8%)	508 (97.7%)	1
Present	24 (2.2%)	12 (2.3%)	
31. Perineural invasion [d]			
Absent	1078 (97.9%)	514 (98.8%)	0.261
Present	23 (2.1%)	6 (1.2%)	

[a] If a patient has multiple stage I/II primary melanomas, the diagnosis date of the first primary melanoma is used to compute the age and year at diagnosis. Additionally, both variables are considered as continuous features in the prediction models.
[b] International Classification of Diseases (ICD) codes before the primary melanoma diagnosis were used to extract these features. "First visit due to melanoma" indicates that the patient's first visit to the Massachusetts General Hospital and Brigham and Women's Hospital/Dana-Farber Cancer Institute was due to melanoma care: either the evaluation of the lesion that would be diagnosed as melanoma or with an already diagnosed melanoma externally that needs to be treated. In this case, the patient did not visit these institutions previously for treating other diseases. As a result, it is impossible to reliably compute their Charlson Comorbidity score based on their ICD diagnosis codes before the primary melanoma diagnosis. Thus, for this group of patients specifically, we assigned the median value to the Charlson Comorbidity in order to retain the variable in our models.
[c] For continuous features with missing values, we assigned the median values of the samples to the individual, including the Charlson comorbidity score, mitotic rate, and total surgical margins.
[d] Precursor lesion, precursor type, regression, lymphovascular invasion, and perineural invasion were assumed absent if not listed as present in the pathology report. Patients without melanoma pathology reports available were excluded from this study.

Supplementary Figure 1. Association between recurrence-free survival and overall survival

A

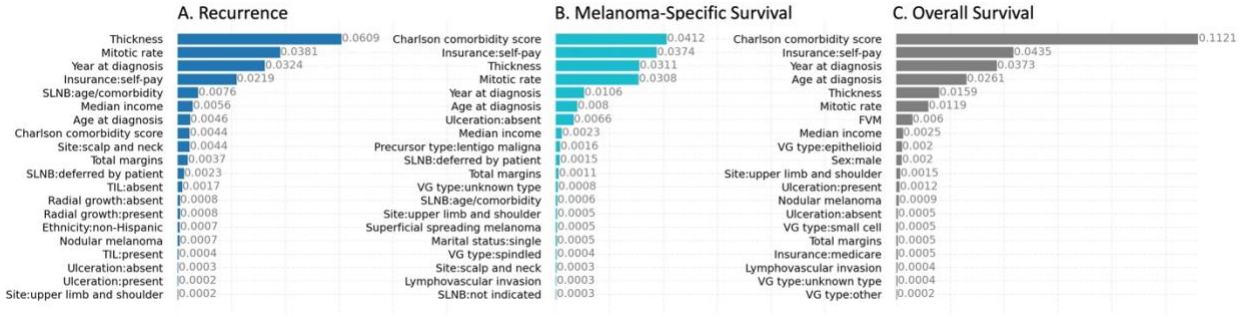


B

Hazard Ratios of recurrence on overall survival compared to non-recurrence

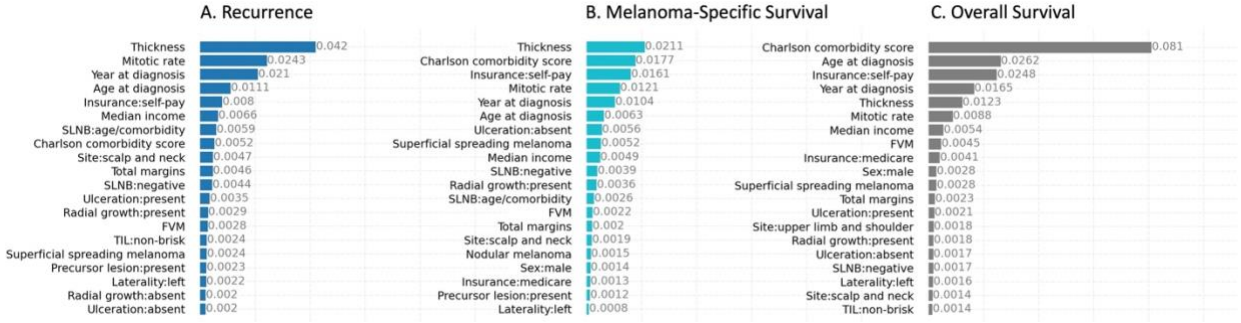
	Hazard Ratio	95% CI	p-value	q-value
Recurrence	2.76	2.21, 3.44	<0.001	<0.001
Distant	5.14	3.83, 6.90	<0.001	<0.001
Locoregional	1.57	1.18, 2.09	0.002	0.008

Supplementary Figure 2. Feature importance in the melanoma recurrence and overall survival prediction models by GBS.



SLNB: sentinel lymph node biopsy; TIL: tumor infiltrating lymphocytes; VG: vertical growth.

Supplementary Figure 3. Feature importance in the melanoma recurrence and overall survival prediction models by RSF.



SLNB: sentinel lymph node biopsy; TIL: tumor infiltrating lymphocytes; VG: vertical growth.