Machine learning to predict overall short-term mortality in cutaneous melanoma

P103

Claudia Cozzolino[1,*], Alessandra Buja[2], Massimo Rugge[3,4], Andrea Miatton[2], Manuel Zorzi[3], Antonella Vecchiato[1], Paolo Del Fiore[1], Saveria Tropea[1], Alessandra Brazzale[5], Giovanni Damiani[6], Luigi dall'Olmo[1,7], Carlo Riccardo Rossi[7], Simone Mocellin[1,7]

* Correspondence: claudia.cozzolino@iov.veneto.it

Objective

Cutaneous malignant melanoma (CMM) is one of the deadliest skin cancers due to its aggressiveness and high probability of misdiagnosis. Melanoma accounted for 5.6% of all new cancer cases in the U.S. in 2021, and its incidence has been on a steady global increase over the past few decades. In Italy, the estimated total number of CMM new cases was 14,900 (8,100 in males, 6,700 in females) in 2020, while 169,900 people are estimated to be alive following a melanoma diagnosis. CMM is the third most common malignancy among Italians aged 50 years or less. These numbers raise concerns in optimizing the efficacy and the efficiency of management, as well as over the economic impact of this disease on healthcare systems.

Despite advances in early detection and treatment, CMM continues to be a disease with highly variable outcomes. Developments in systemic adjuvant medications for stage III and stage IV melanomas are contributing to improved outcomes even for high-risk patients, but there are still gaps in our ability to correctly stage melanomas. Internationally, the prognostic assessment of CMM outcome is based on the American Joint Committee on Cancer (AJCC) melanoma staging system.

Recent theoretical applications of various artificial intelligence (AI) algorithms for the development of updated staging systems for CMM have produced promising results in oncological research, optimizing cancer care, and, ultimately, in personalized cancer therapy Unlike traditional computer programming, AI is not reliant on a pre-determined algorithm to produce an output, but analyzes input data with its associated output to process a model that can then be used to infer on similar datasets. The main advantage of AI lies in its ability to analyze multiple measures in complex and large data sets, combining information, and weighing the relative impact in relation to the target outcomes.

Based on the CMM clinico-pathological profile as recorded by the regional population-based Veneto cancer registry, this study aims to explore the consistency of the AI in predicting short term overall mortality in CMM patients and then to provide a useful tool for clinical practice.

Materials and Methods

The data were sourced from the Veneto Cancer Registry (RTV), a high-resolution, population-based dataset covering the regional population (approximately 4.9 million residents), and the regional health service records. Cancer registration procedures were based on information collected from various sources (e.g., pathology reports, death certificates, and the health service's administrative records).

All incident cases of invasive CMM registered by the RTV in 2015 (1,279 cases) and 2017 (1,368 cases) were included. The following variables were considered for this study: demographics (age and sex); histological subtypes of CMM (malignant (NOS), superficial spreading, nodular, lentigo maligna, acral-lentiginous, desmoplastic, spitzoid, and those arising from the blue nevus variant); tumor site (lower limbs, upper limbs, head, hands and feet, and trunk); CMM growth phase (radial versus vertical); ulceration (absent versus present); Breslow thickness (≤ 0.75, 0.76-1.50, 1.51-3.99, ≥ 4.00 mm); CMM regression (absent versus present); tumor-infiltrating lymphocytes (TIL) (absent versus present); mitotic count (number of mitoses per mm^2); T, N, and M AJCC stages at diagnosis (8th edition); sentinel lymph node biopsy (SLNB); SLNB maximum metastasis diameter (in mm); number of positive lymph nodes (after SLNB or lymphadenectomy); and, evaluation of overall survival (OS) time truncated at 3 years from diagnosis.

Descriptive statistics were obtained, representing categorical variables as frequencies and proportions, while summarising continuous numerical variables with means, medians, and minimum–maximum intervals. A univariate analysis was performed with Cox Regression to verify the strength and direction of each independent variable on CMM mortality. The correlation matrix was also calculated to inspect interdependencies and redundancies in the data.

Some features were not available for all subjects. Being the missing values evenly distributed across the cohort, it was considered preferable to proceed with imputation strategies as opposed to discarding incomplete records so as not to excessively reduce the sample and lose information. Simple feature imputation and multivariable regression were adopted to fill in missing data.

The prognostic algorithms were based on supervised learning techniques. Several machine learning (ML) models were trained to predict mortality risk expressed as a binary label: survived versus deceased within 3 years of diagnosis.

Support Vector Machine (SVM), Random Forest (RF), Gradient Boosting (GB), and k-Nearest Neighbors (kNN) were implemented as shallow classifiers with varying degrees of complexity and interpretability.

Ten-fold cross-validation (CV) was performed on the training set in order to understand the best preprocessing procedures (feature selection, scaling, etc.) and optimize model hyperparameters.

Given that the task of predicting CMM mortality risk is naturally defined as an imbalanced classification problem, the fitting and test evaluations were measured in terms of balanced accuracy and F1-score.

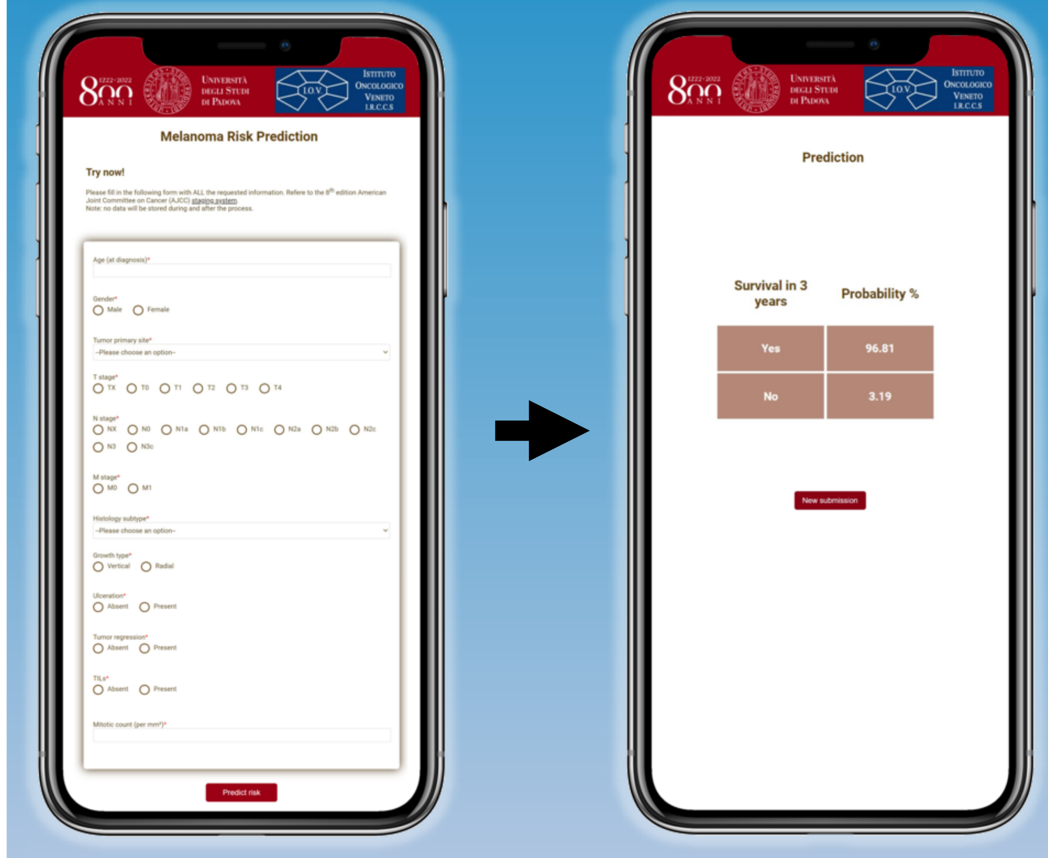

Figure 1. Usage example of the web application implemented on top of the best machine learning model. Given a set of clinico-pathological features, our app predicts the 3-year survival probability.

Results

The overall mortality was 10.4% at 3 years after diagnosis, with a mean follow-up of 1,032.8 days. The univariate analysis revealed that older age, male sex, a vertical growth pattern, a thicker Breslow depth, ulceration presence, TILs absence, a higher mitotic count, SLNB positivity with a wider SLNB max diameter, and a greater number of positive lymph nodes are all statistically associated with short-term CMM mortality.

The primary site is relevant when the tumor is located on the hands, feet, or head. Nodular and malignant (NOS) histological subtypes had the highest hazard ratios (HR 2.55 and 2.30, respectively); acral-lentiginous subtype as the reference category), while superficial spreading corresponds to better outcomes (HR 0.40).

As expected, the correlation analysis revealed interdependence between T stage values and Breslow thickness, as well as between N stages and SLNB positivity. For this reason, the ML models were trained and evaluated on two different variable subsets: once excluding Breslow, number of positive lymph nodes, SLNB positivity and maximum diameter, once excluding T and N stages.

Considering the classifiers prognostic performances, the GB and RF ensemble models reasonably outperformed the others having a more complex structure. In general, the use of T and N stages in risk prediction resulted in slightly higher evaluation scores than Breslow thickness and lymph node status. In terms of F1 score (0.67), the SVM yielded the best result. With a balanced accuracy of 89%, the RF model proved to be the best option. The model's accuracy is shown as an ROC survival curve in Figure 3, achieving an area under the curve (AUC) of 0.91.

As RF is a tree-based model, it was possible to extract each feature's Gini importance score. Figure 2 represents the most important CMM risk prediction variables, as determined by the optimal model. The patient's age, mitotic rate, T4 staging, the presence of ulceration, and metastasis appear to have the greatest influence on the classification of short-term mortality.

Finally, a web application was built on top of the best model developed (an example in Figure 1).

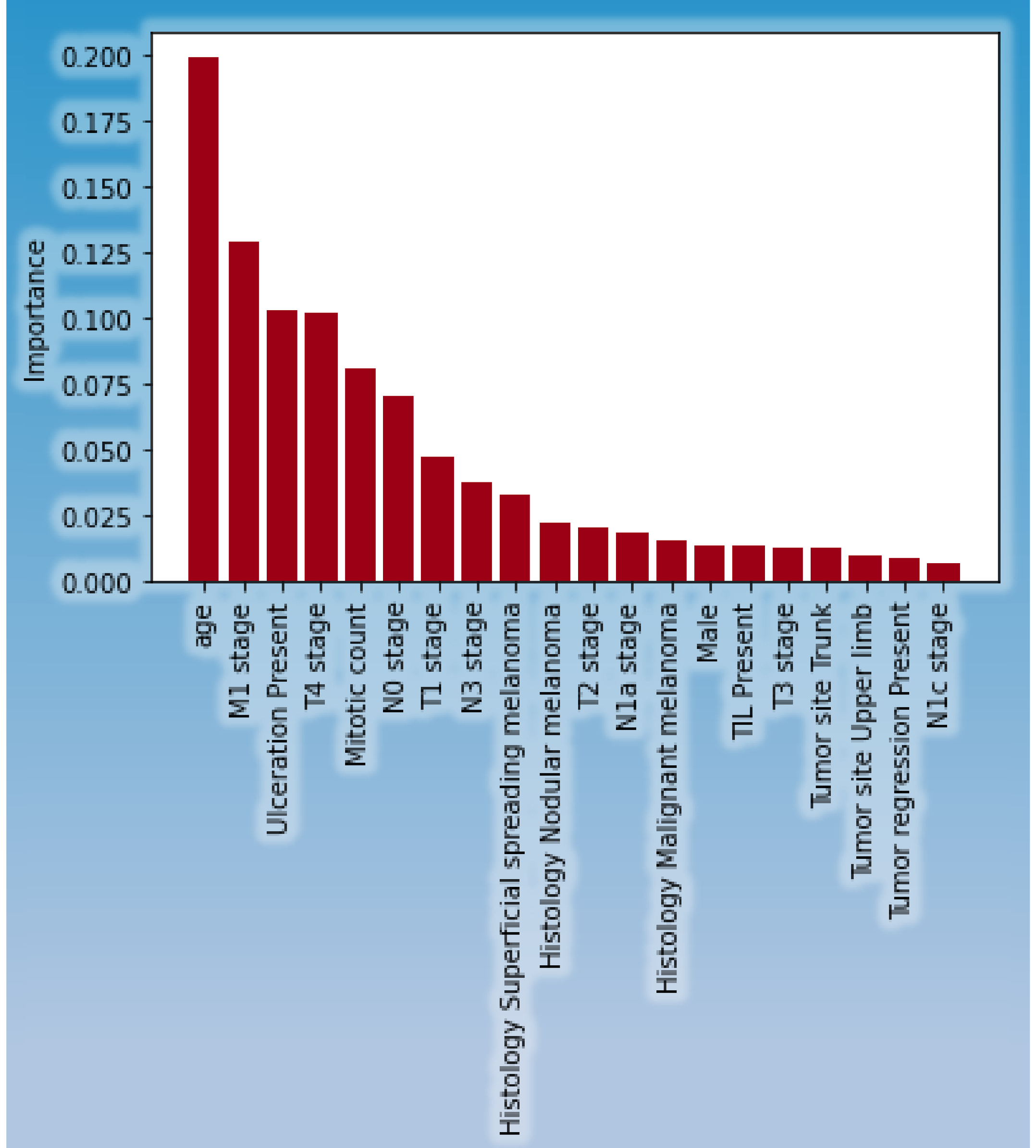

Figure 2. Most important features according to the RF model in predicting CMM risk (based on the Gini impurity criterion).

Conclusion

In recent years, machine learning has been applied extensively to improve melanoma risk stratification and prognosis prediction. Most research has focused on finding new clinical and pathological markers. Nevertheless, none of the new, promising, prognostic variables have yet been added to the AJCC system, which is currently the gold standard staging system. A more accurate prognostic tool is needed to increase the survival of melanoma patients by preventing recurrence and providing the most appropriate follow-up regimens.

We decided to focus on the development of an algorithm based on known and validated prognostic factors, with the aim of using machine learning to improve prediction capabilities and facilitate the application of this novel melanoma risk stratification tool. The results of an initial univariate analysis of available data were consistent with those of earlier scientific literature.

While histologic characteristics, including thicker Breslow depth, the presence of ulceration, SLNB positivity, and the absence of TILs, are widely accepted, our data also suggest that primary site location, histology subtype, and N stage, may have different relevance depending on the specific class considered in the prognosis. These findings suggest that a better classification of existing prognostic factors is possible.

The results of training a new model through machine learning are promising. Using only routinely collected information, our algorithm was able to attain an accuracy of 89% and an AUC value of 0.91. Comparatively, one previous study on the prognostic accuracy of the AJCC staging system, 8th edition, reported an AUC of only 0.74 (on a cohort of 1,462 patients).

In addition, we developed a web-based application to make the research results accessible and applicable with minimal effort in the current clinical setting.

Applications based on machine learning will undoubtedly shape the future of medicine, however, the real-world validation of the results attained is a necessary step to understand the actual effectiveness of the tool and promote this technology's integration into everyday clinical practice.

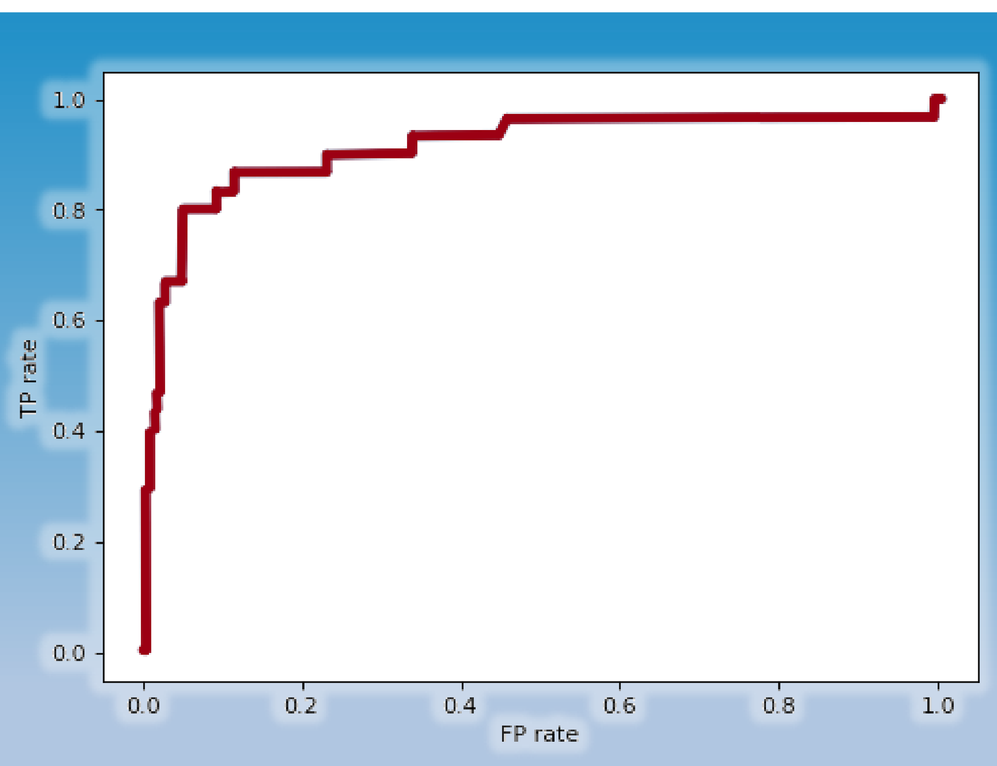

Figure 3. RF model 3-year survival/mortality prediction ROC curve (AUC = 0.911).

References

1. Abbas O, Miller DD, Bhawan J. Cutaneous Malignant Melanoma: Update on Diagnostic and Prognostic Biomarkers. The American Journal of Dermatopathology 2014;36:363–379.
2. Davis LE, Shalin SC, Tackett AJ. Current state of melanoma diagnosis and treatment. Cancer Biology & Therapy 2019;20:1366–1379.
3. Melanoma of the Skin - Cancer Stat Facts. SEER. Available at https://seer.cancer.gov/statfacts/html/melan.html. Accessed April 1, 2022.
4. Rastrelli M, Tropea S, Rossi CR et al. Melanoma: Epidemiology, Risk Factors, Pathogenesis, Diagnosis and Classification. in vivo 2014;:7.
5. Pavri SN, Clune J, Ariyan S et al. Malignant Melanoma: Beyond the Basics. Plastic and Reconstructive Surgery 2016;138:330e–340e.
6. Associazione Italiana di Oncologia Medica, Gruppo di Lavoro Registri Tumori Italiani, SIAPEC-IAP et al. I Numeri del Cancro in Italia 2021. Available at https://www.aiom.it/wp-content/uploads/2021/10/2021_NumeriCancro_web.pdf. Accessed April 12, 2022.
7. Buja A, Rugge M, De Luca G et al. Cutaneous Melanoma in Alpine Population: Incidence Trends and Clinicopathological Profile. Current Oncology 2022;29:2165–2173.
8. Keung EZ, Gershenwald JE. The eighth edition American Joint Committee on Cancer (AJCC) melanoma staging system: implications for melanoma treatment and care. Expert Review of Anticancer Therapy 2018;18:775–784.
9. Ma E, Hoegler K, Zhou A. Bioinformatic and Machine Learning Applications in Melanoma Risk Assessment and Prognosis: A Literature Review. Genes 2021;12:1751.
10. Dlamini Z, Francies FZ, Hull R, Marima R. Artificial intelligence (AI) and big data in cancer and precision oncology. Comput Struct Biotechnol J. 2020 Aug 28;18:2300-2311. doi: 10.1016/j.csbj.2020.08.019
11. Farina E, Nabhen JJ, Dacoregio MI, Batalini F, Moraes FY. An overview of artificial intelligence in oncology. Future Sci OA. 2022 Feb 10;8(4):FSO787. doi: 10.2144/fsoa-2021-0074.
12. Gareau DS, Browning J, Correa Da Rosa J et al. Deep learning-level melanoma detection by interpretable machine learning and imaging biomarker cues. J Biomed Opt 2020;25. doi:10.1117/1.JBO.25.11.112906.
13.Ali Z, Yousaf N, Larkin J. Melanoma epidemiology, biology and prognosis. European Journal of Cancer Supplements 2013;11:81–91.
14. Howard J. Artificial intelligence: Implications for the future of work. Am J Ind Med 2019;62:917–926.
15. Hamet P, Tremblay J. Artificial intelligence in medicine. Metabolism 2017;69:S36–S40.
16. Mintz Y, Brodie R. Introduction to artificial intelligence in medicine. Minimally Invasive Therapy & Allied Technologies 2019;28:73–81.
17. Chen M, Decary M. Artificial intelligence in healthcare: An essential guide for health leaders. Healthc Manage Forum 2020;33:10–18.
18. Choi RY, Coyner AS, Kalpathy-Cramer J et al. Introduction to Machine Learning, Neural Networks, and Deep Learning. Neural Networks;:12.
19. Jiang T, Gradus JL, Rosellini AJ. Supervised Machine Learning: A Brief Primer. Behavior Therapy 2020;51:675–687.
20. Sultan AS, Elgharib MA, Tavares T et al. The use of artificial intelligence, machine learning and deep learning in oncologic histopathology. J Oral Pathol Med 2020;49:849–856.
21. Jiang Y, Yang M, Wang S et al. Emerging role of deep learning-based artificial intelligence in tumor pathology. Cancer Communications 2020;40:154–166.
22. Ferre F, de Belvis AG, Valerio L et al. Italy: health system review. Health Syst Transit 2014;16:1–168.
23. Portale Sanità Regione del Veneto - PDTA Melanoma. Available at https://salute.regione.veneto.it/web/rov/pdta-melanoma. Accessed April 20, 2020.
24. AIOM. LINEE GUIDA MELANOMA EDIZIONE 2020. Available at https://www.aiom.it/wp-content/uploads/2020/10/2020_LG_AIOM_Melanoma.pdf.
25. National Comprehensive Cancer Network Cutaneous melanoma guidelines. 2021. NCCN. Available at https://www.nccn.org/guidelines/guidelines-detail. Accessed March 22, 2022.
26. Overview | Melanoma: assessment and management | Guidance | NICE. Available at https://www.nice.org.uk/guidance/ng14. Accessed March 22, 2022.
27. Pedregosa, F. et al. Scikit-learn: Machine Learning in Python, Journal of Machine Learning Research 2011, 12, 2825—2830.
28. Perrier A. Feature Importance in Random Forests, 2015. Available at https://alexisperrier.com/datascience/2015/08/27/feature-importance-random-forests-gini-accuracy.html
29. Arora C, Kaur D, Lathwal A et al. Risk prediction in cutaneous melanoma patients from their clinico-pathological features: superiority of clinical data over gene expression data. Heliyon 2020;6:e04811.
30. Wu P-C, Chen Y-C, Chen H-M et al. Prognostic factors and population-based analysis of melanoma with sentinel lymph node biopsy. Sci Rep 2021;11:20524.
31. Bobos M. Histopathologic classification and prognostic factors of melanoma: a 2021 update. Ital J Dermatol Venereol 2021;156. doi:10.23736/S2784-8671.21.06958-3.
32. Kanaki T, Stang A, Gutzmer R et al. Impact of American Joint Committee on Cancer 8th edition classification on staging and survival of patients with melanoma. European Journal of Cancer 2019;119:18–29.
33. Kycler W, Grodecka-Gazdecka S, Bręborowicz J et al. Prognostic factors in melanoma. Reports of Practical Oncology & Radiotherapy 2006;11:39–48.
34. Bartlett E, Karakousis G. Current Staging and Prognostic Factors in Melanoma. Surgical Oncology Clinics of North America 2015;24. doi:10.1016/j.soc.2014.12.001.
35. Triantafyllidis AK, Tsanas A. Applications of Machine Learning in Real-Life Digital Health Interventions: Review of the Literature. J Med Internet Res 2019;21:e12286.
36. Wong SL, Kattan MW, McMasters KM et al. A Nomogram That Predicts the Presence of Sentinel Node Metastasis in Melanoma With Better Discrimination Than the American Joint Committee on CancerStaging System. Ann Surg Oncol 2005;12:282–288.
37. Safran T, Viezel-Mathieu A, Corban J et al. Machine learning and melanoma: The future of screening. Journal of the American Academy of Dermatology 2018;78:620–621.