EDITORIAL – LETTER TO THE EDITOR

Prediction of Sentinel Node Positivity Risk in Melanoma Patients: When Appropriate Statistical Methodology was Employed an i31-GEP Test Did Not Outperform a Tool that Uses Clinicopathologic Features Only

Serigne N Lo, PhD [1,2], Alexander HR Varey, PhD [1,2,3], Richard A Scolyer, MD [1,2,4], John F Thompson, MD [1,2,5]

[1] Melanoma Institute Australia, The University of Sydney, North Sydney, NSW, Australia
[2] Faculty of Medicine and Health, The University of Sydney, Sydney, NSW, Australia
[3] Department of Plastic Surgery, Westmead Hospital, Sydney, NSW, Australia
[4] Department of Tissue Oncology and Diagnostic Pathology, Royal Prince Alfred Hospital, Camperdown, NSW, Australia
[5] Department of Melanoma and Surgical Oncology, Royal Prince Alfred Hospital, Camperdown, NSW, Australia

To the editor: In an article published recently by Zakria et al,[1] comparison was made between the performance of two tools that predict the risk of sentinel lymph node (SLN) metastasis for patients with newly-diagnosed melanomas. Using the Melanoma Institute Australia (MIA) nomogram that uses six widely-available clinicopathologic features[2] and an integrated GEP test that combines a proprietary 31-gene expression profile with four clinicopathologic features[3], a cohort of 466 patients with high-risk T1a and T2 melanomas who had undergone SLN-biopsy was analyzed.

The authors concluded that the i31-GEP test outperformed the MIA nomogram in identifying patients who could safely forego SLN-biopsy, but it appears that this conclusion was based on a flawed analysis. Comparison of the precision of the MIA nomogram and the i31-GEP test relied on 95% confidence intervals (CIs), with patients considered low-risk only if they had <5% risk of SLN-positivity and an upper 95% CI ≤10%, and high-risk only if they had >10% SLN-positivity risk and a lower 95% CI ≥5%. However, to compare risk-prediction tools using these criteria, the CIs must be calculated using the same statistical methodology[4], which was not done. The CIs for the MIA online calculator were determined using a parametric method, widely regarded as the 'gold standard' because it incorporates all sources of variability[5]. The CIs for the i31-GEP test, on the other hand, were derived by a locally-estimated scatter-smoothing (LOWESS) spline fitting of the nodal positivity and predicted nodal positivity from the i31-GEP algorithm. This approach, based only on the observed and predicted risks, is data-driven and does not derive any formula for computing the CIs[6].

To demonstrate the importance of calculating CIs using the same methodology when comparing them, we applied the LOWESS approach to compute 95% CIs for the MIA tool development cohort (n=3477). As

expected, LOWESS-based CIs were much narrower than those obtained using the parametric method (see Figures 1a and b). Crucially, when using LOWESS-based CIs, all patients classified by the MIA tool as low-risk or high-risk on point estimates alone also achieved the risk criteria for low-risk and high-risk groups as defined by Zakria et al. (Table 1). A significant difference in CI range using the two approaches is a well-known methodological pitfall[4] and undoubtedly accounts for the apparent difference in precision of the two tools suggested by the authors.

In summary, it appears that the conclusions of the report by Zakria et al were based on the results of an analysis that used inappropriate statistical methodology. An accurate conclusion would be that no difference was observed in the ability of the MIA and i31-GEP tools to identify patients who could safely forego SLN-biopsy. Both predicted SLN-positivity with reasonable precision. When clinicians are deciding which tool is preferable to identify patients most likely to benefit from SLN-biopsy and those unlikely to benefit, they will need to compare the simplicity and immediacy of using the MIA-nomogram, freely available online, with the time delay and substantial cost of using a tool incorporating an i31GEP test.
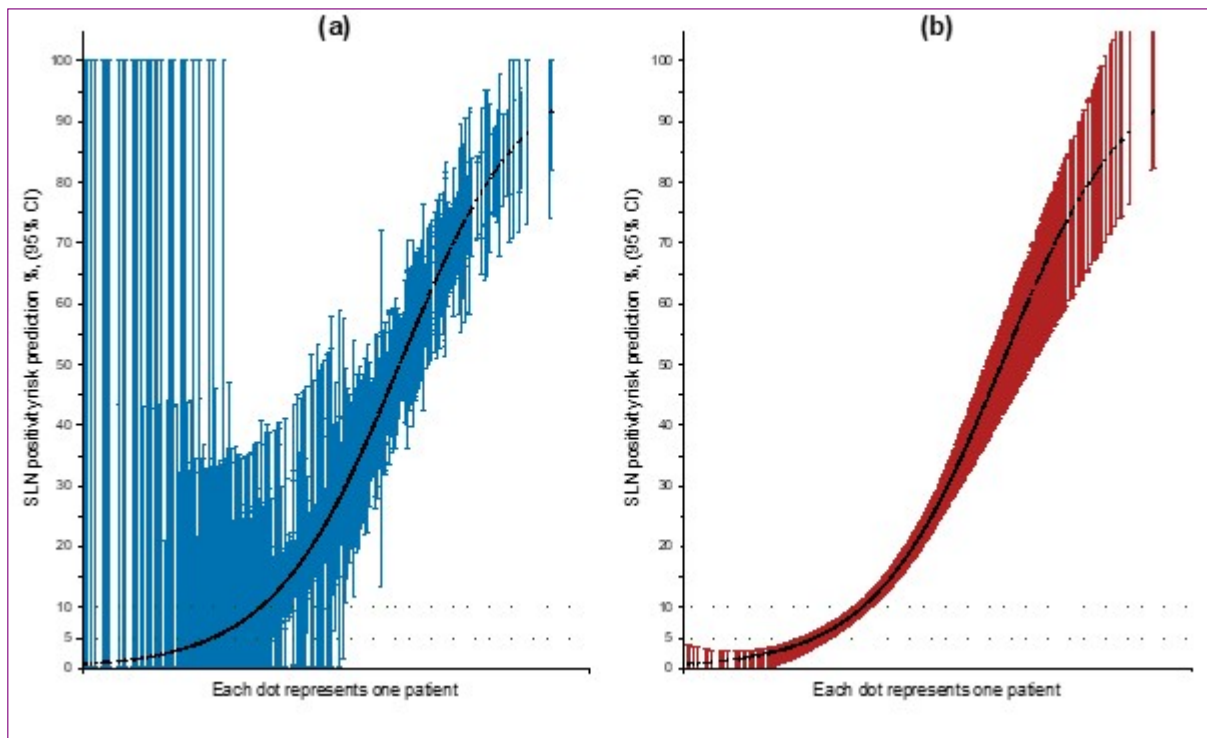
Figure 1. Point estimates and 95% confidence intervals for each of the 3477 patients included in the development cohort for the MIA sentinel node metastasis risk prediction tool. (a) Using the parametric method built into the MIA tool (left panel, blue) and (b) Using the non-parametric LOWESS approach employed by Zakria et al (right panel, red). It can be seen that the LOWESS approach produces much narrower 95% CIs than the "gold standard" parametric approach utilized by the MIA tool.

1128

Table 1. When the MIA development cohort patients they were classified as 'low-risk' or 'high-risk' by point estimates using the MIA tool alone were also classified as such by the point estimates AND 95% CI criteria defined by Zakria et al when CIs were calculated using the LOWESS method.

Low Risk	<5% SLN positive risk	<5% AND upper 95% CL ≤ 10%
	4.4% (153/3477)	100% (153/153)

High Risk	<5% SLN positive risk	<5% AND upper 95% CL ≤ 10%
	78.1% (2717/3477)	100% (2717/2717)

Corresponding Author:
Serigne N Lo
Melanoma Institute Australia, The University of Sydney, North Sydney, NSW, Australia
Email: serigne.lo@melanoma.org.au

References:

1. Zakria D, Brownstone N, Rigel D. The Integrated 31-Gene Expression Profile (i31-GEP) Test for Cutaneous Melanoma Outperforms a Clinicopathologic-only Nomogram at Identifying Patients who can Forego Sentinel Lymph Node Biopsy. SKIN: The Journal of Cutaneous Medicine 2022; 6 (6): 463-473. https://jofskin.org/index.php/skin/article/view/1795/1476

2. Lo SN, Ma J, Scolyer RA et al. Improved Risk Prediction Calculator for Sentinel Node Positivity in Patients With Melanoma: The Melanoma Institute Australia Nomogram. J Clin Oncol 2020; 38 (24): 2719-2727.

3. Whitman ED, Koshenkov VP, Gastman BR et al. Integrating 31-Gene Expression Profiling With Clinicopathologic Features to Optimize Cutaneous Melanoma Sentinel Lymph Node Metastasis Prediction. JCO Precis Oncol 2021; 5.

4. Härdle W, Müller M, Sperlich S et al. 3.5 Confidence Intervals and Confidence Bands. Nonparametric and Semiparametric Models, Berlin: Springer 2004.

5. Hosmer DW, Lemeshow S. Confidence interval estimates of an index of quality performance based on logistic regression models. Stat Med 1995; 14 (19): 2161-2172.

6. Cleveland WS, Devlin SJ. Locally Weighted Regression: An Approach to Regression Analysis by Local Fitting. Journal of the American Statistical Association 1988; 83 (403): 596-610.