ORIGINAL RESEARCH

Establishing an evidence-based decision point for clinical use of the 31-gene expression profile test in cutaneous melanoma

Etan Marks, DO[1], Hillary G. Caruso, PhD[2], Sarah J. Kurley, PhD[2], Sidra Ibad, MD[3], Kristen M. Plasseraud, PhD[2], Federico A. Monzon, MD, FCAP[2], Clay J. Cockerell, MD[4]

[1]UT Southwestern Medical Center, Dallas, TX
[2]Castle Biosciences, Inc., Friendswood, TX
[3]Icahn School of Medicine at Mount Sinai, New York, NY
[4]Cockerell Dermatopathology, Dallas, TX

ABSTRACT

Treatment plans for cutaneous melanoma are based upon individual risk of recurrence. Decisions made post-diagnosis include recommendation for a sentinel lymph node biopsy (SLNB), followed by management decisions such as surveillance, frequency of follow-up, and interdisciplinary consultations including possible adjuvant therapy use. These have traditionally been guided by clinicopathologic factors, but discordance exists, as a substantial number of melanoma deaths occur in patients diagnosed with disease considered to be early stage by such factors, including a negative SLNB. Molecular testing can be used to apply an objective approach that optimizes individualized patient care. The 31-gene expression profile (31-GEP) test has been validated in nearly 1600 patients as an independent predictor of risk of recurrence, distant metastasis and death in Stage I-III melanoma and can guide SLNB decisions in patient subgroups, as demonstrated in 1421 patients. While clinical use of the 31-GEP test has been adopted into routine practice, an evidence-based analysis of a decision point for use in thin, T1 tumors would be clinically useful. To help define an appropriate population for 31-GEP testing, we evaluated changes in patient management, cumulative differential risk across Breslow thicknesses based on a large dataset, and 31-GEP subclass distribution in a clinically tested cohort. Based on this, appropriate use of the 31-GEP test for management decisions was found to be in cutaneous melanoma tumors ≥0.3 mm thick.

INTRODUCTION

The observation that two-thirds of Stage I-III melanoma deaths occur in patients who have no evidence of distant metastatic spread at diagnosis (Stage I-II) suggests that improved prognosis of metastatic risk prognosis is necessary.[1] To address this clinical need, a 31-gene expression profile (31-GEP) test was developed and validated, providing prognostic information based on the biology of the primary melanoma tumor.[2-5] Numerous retrospective and prospective studies have demonstrated that the 31-GEP is a significant and independent predictor of survival, beyond traditional clinicopathologic staging factors and SLN status.[3,4, 6-9] Recognizing the

clinical value of the 31-GEP, physicians have utilized it to inform decisions of intensity of follow-up, including use and frequency of clinical visits, labs, imaging, and referrals.[10-12]

Patients are classified by the 31-GEP test as low-risk (Class 1, 1A lowest risk) or high-risk (Class 2, 2B highest risk) based on differential expression of a panel of 28 discriminating gene targets and 3 control genes.[2,5,13] Previous studies have reported that the 31-GEP identifies high-risk disease in traditionally low-risk subsets of patients, including those with thin (≤1.0 mm, T1) tumors.[3,4,14] This finding allows for the appropriate identification of patients for escalation and de-escalation of interventions. As adoption of the 31-GEP test increases, the natural question arises: what patient population is appropriate for 31-GEP testing? This is a particularly important question for patients with thin melanomas traditionally associated with good outcomes, having a 5-year melanoma-specific survival of 96-99%.[15, 16] A non-invasive prognostic tool such as the 31-GEP can be useful to identify those T1 tumors with increased risk of metastasis and death from melanoma, but the threshold at which the test is expected to have true clinical value has not yet been reported.

In this study, we define a target population for 31-GEP clinical use by i) estimating 5-year recurrence-free (RFS) and distant metastasis-free survival (DMFS) rates associated with T1 tumors and cumulative event rates in T1-T2 tumors using a large multi-center dataset; ii) reviewing 31-GEP clinical reports to evaluate the frequency of when high-risk, Class 2 biology is detected in clinically tested patients with T1 tumors; and iii) analyzing data from previously published clinical utility studies to determine the tumor thickness threshold at which patient management decisions were changed based on the results of the 31-GEP test. The goal of the study is to suggest a more targeted application of the 31-GEP test for patients with thin tumors that can improve resource management and health care outcomes, and to summarize appropriate use relative to the two validated utilities of the 31-GEP test in the context of current melanoma management strategies.

METHODS

Clinical outcomes study cohorts and data collection

31-GEP test results, tumor clinicopathologic data, and clinical outcomes were derived from three non-overlapping published study cohorts, including prospectively-tested patients[6, 7] and archival tumor specimens[3], as well as an independent archival tumor cohort (total n=1479).[17] Median follow-up for the patients in the combined cohort was 3.3 years. All tumor specimens and associated clinical data were collected under institutional (IRB)-approved protocols or were deemed exempt by the IRB.

Cumulative event analysis by Breslow thickness

Total number of cases and recurrence events (events documented as being locoregional or distant recurrences) from the 1479 patients described above were analyzed according to Breslow thickness and 31-GEP class using R.3.6. For tumors between 0.0-2.0 mm Breslow thickness (n=1037) cumulative event rates at each observed case within the 31-GEP subclasses were calculated and plotted. Loess regression was used to fit a smoothed curve to these plotted points. Breslow thickness binning in all analyses was

used as per AJCC v8 guidelines (i.e. rounding to nearest tenth of a mm)[16].

Clinical GEP testing data

De-identified consecutive 31-GEP clinical test results and associated tumor clinicopathologic data (n=579) between 2014 and 2019 from a large dermatopathology practice were used to determine proportions of 31-GEP classes in T1 melanoma tumors by 0.1 mm increments in Breslow thickness. De-identified 31-GEP clinical test results and associated tumor clinicopathologic data between May 2018 to April 2019 were also used to determine proportions of 31-GEP classes in T1 tumors by 0.1 mm increments in Breslow thickness.

Analysis of clinical utility according to Breslow thickness

Raw data from two published, multi-center clinical impact studies[10, 11] (n=403) were used to analyze management changes (including changes in referrals, clinical visits, imaging, laboratory tests, and SLNB recommendations) in consecutively tested patients that were made following 31-GEP test result receipt, according to Breslow thickness. All clinicopathologic data, 31-GEP test results, and clinical use data were collected under IRB-approved protocols.

RESULTS

Outcomes in the T1 population according to Breslow thickness and 31-GEP class

To estimate differences in risks associated with the 31-GEP test results at any given Breslow thickness in T1-T2 tumors, outcomes and clinicopathologic data from 1479 patients with a median follow-up of 3.3 years were used. Cumulative recurrence rates (cumulative recurrences/cumulative cases) in 31-GEP subclasses were calculated at each observed case and plotted. Patient and pathologic characteristics of tumors ≤2.0 mm (n=1037) are shown in Table 1. A smoothed curve was fit using Loess regression to show the estimated cumulative event rate (cumulative recurrences/total cases in 31-GEP subclass) at each observable Breslow thickness. As shown in Figure 1, the predicted cumulative rates of recurrence from the Loess curve for Class 2B separates from the Class 1A tumors between 0.2 and 0.3 mm Breslow thickness. Importantly, the first recurrent/distant metastatic event occurred in a 0.3 mm tumor and thus 5-year RFS and DMFS survival rates are 100% for tumors ≤0.2 mm and significantly different with 31-GEP class in patients with tumors 0.3-1.0 mm thick by log-rank test ($p<0.0001$ for RFS; $p=0.0008$ for DMFS) (Table 2). In patients with tumors 0.3-1.0 mm thick, Class 2B results were associated with RFS and DMFS rates of 75.5% and 82.1%, respectively, compared to rates of 96.8% and 97.4% for Class 1A. As these data suggested a potential target threshold for clinical utility, the threshold of 0.3 mm was explored in the next set of analyses.

Frequency of 31-GEP subclasses in thin melanomas

To determine the prevalence of 31-GEP subclasses in thin melanomas, de-identified test results and clinicopathologic data from a large dermatopathology practice were evaluated for 31-GEP subclass and Breslow thickness (n=437 with Breslow thickness ≤1.0 mm and a 31-GEP subclassification) (Figure 2A). Additionally, de-identified 31-GEP results and clinicopathologic data from tumors 0.1-1.0 mm clinically tested between May 2018 and April 2019 (n=8944) were

evaluated (Figure 2B). In the dermatopathology practice, tumors <0.3 mm Breslow thickness were exclusively Class 1A, while Class 1B, 2A, and 2B results were obtained for tumors ≥0.3 mm Breslow thickness such that 7.1% of tumors ≥0.3 mm and ≤1.0 mm had Class 2 results and 15.7% of tumors ≥0.3 mm had non-Class 1A (e.g. Class 1B, 2A or 2B) results (Figure 2A). In the overall clinically tested population from the last year, tumors <0.3 mm were primarily Class 1A, with few Class 1B, 2A, and 2B tumors. However, 4.3% of tumors ≥0.3 mm were Class 2 and 10.8% were non-Class 1A (Figure 2B).

Physicians' use of 31-GEP test to change management decisions and relationship to Breslow thickness in thin tumors

Four clinical impact studies have shown that the 31-GEP test is associated with a ~50% rate of management changes when incorporated with available clinicopathologic staging.[10-12, 18] Of these, two are multi-center studies of consecutively tested patients that demonstrated changes in clinical visits, SLNB recommendations, frequency and intensity of surveillance imaging, laboratory testing and specialty referrals.[10, 11] Those two studies included a total of 403 patients, of which 160 had primary melanomas ≤1.0 mm thick. To evaluate the impact of 31-GEP on patient management at various Breslow thicknesses, T1 melanomas were categorized by reported Breslow thickness by 0.1 mm increments, and it was determined if a change in management was recommended post-testing. As shown in Figure 3, most of the management changes for thin (T1) melanomas following receipt of the 31-GEP test were recommended for lesions with ≥0.3 mm Breslow thickness. Specifically, 25% of the lesions between 0.3 mm and 1.0 mm had a management change, while 11% of lesions <0.3 mm had a recommended management change.

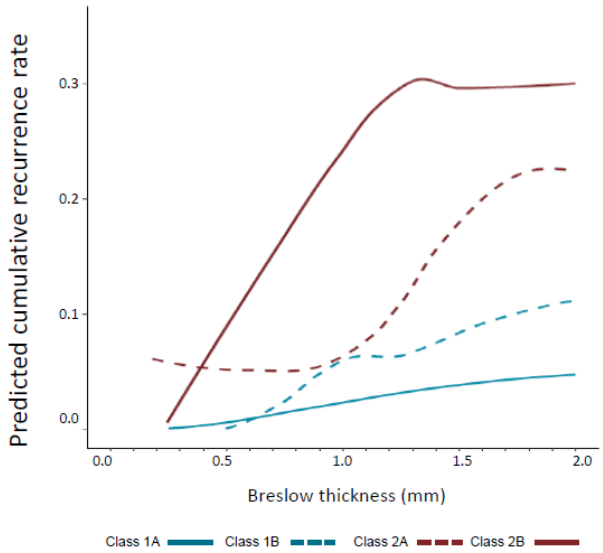

Figure 1. Outcomes associated with 31-GEP subclasses across Breslow depths in T1-T2 tumors. Data from a multi-center cohort of 1479 tumors were used to calculate cumulative recurrence rates (cumulative events/cumulative cases in each 31-GEP subclass) at each observed T1-T2 tumor in the dataset. Loess regression was used to fit a smooth curve to the points and estimate cumulative recurrence rates across Breslow thicknesses between 0.0 and 2.0 mm.

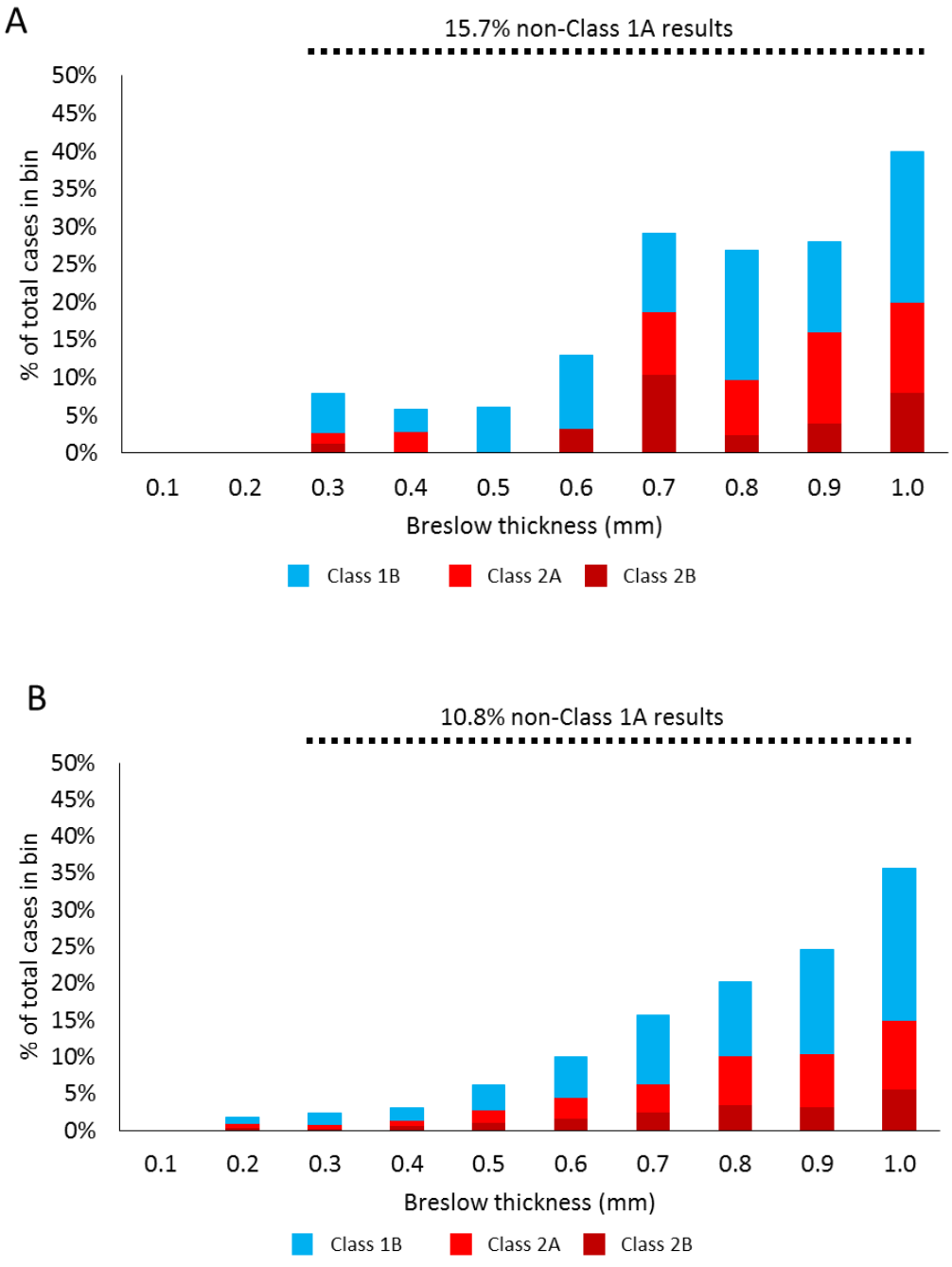

Figure 2. 31-GEP subclass distribution with Breslow thickness in large clinically tested cohorts. A) T1 tumors clinically tested with the 31-GEP test between 2014 to 2019 at a large dermatopathology practice were evaluated for 31-GEP result (n=479). B) T1 tumors clinically tested with the 31-GEP test between May 2018 and April 2019 (n=8944). Percent of tumors in each 31-GEP subclass were plotted according to 0.1 mm Breslow thickness increments. In both panels, non-Class 1A results are shown.

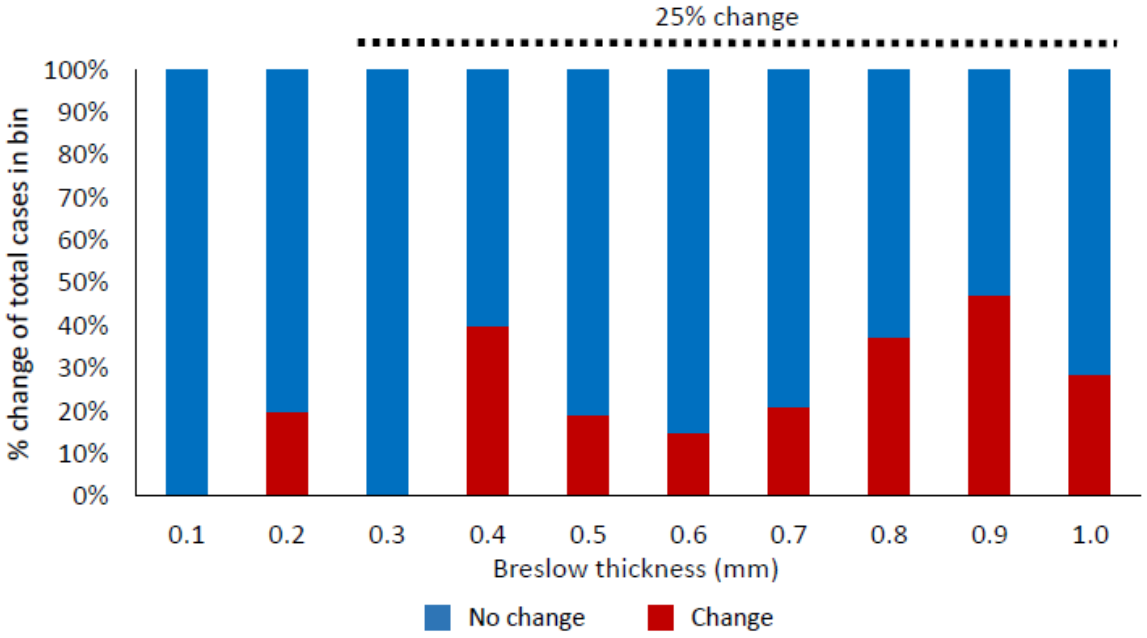

Figure 3. Management changes in 31-GEP-tested T1 tumors by 0.1 mm Breslow thickness increments. Management changes (change in frequency of clinical visits, imaging, referrals, sentinel lymph node biopsy, laboratory tests) after receipt of 31-GEP test results associated with T1 tumors (n=160) from two published studies[10, 11] were compiled and evaluated by 0.1 mm increments in tumor Breslow thickness. Blue bars indicate no change in management and red bars indicate a management change (increase or decrease in intensity).

Table 1. Patient and tumor characteristics of T1-T2 cohort.

Feature	Class 1A (n=718)	Class 1B (n=126)	Class 2A (n=88)	Class 2B (n=105)	Combined (n=1037)
Age (years), p<0.001					
Mean and SD	58.4 (±15.8)	57.8 (±13.5)	61.2 (±14.5)	64.5 (±15.1)	59.2 (±15.4)
Median (Range)	60 (18-91)	57.5 (26-87)	61 (19-88)	67 (20-93)	60 (18-93)
Breslow depth (mm), p<0.0001					
Mean and SD	0.7 (±0.431)	1.1 (±0.436)	1.3 (±0.52)	1.3 (±0.464)	0.9 (±0.495)
Median (Range)	0.6 (0.1-2.0)	1.0 (0.2-2.0)	1.3 (0.2-2)	1.3 (0.1-2)	0.8 (0.1-2)
Ulceration, p<0.0001					
no	605 / 718 (84.26%)	93 / 126 (73.81%)	61 / 88 (69.32%)	50 / 105 (47.62%)	809 / 1037 (78.01%)
unknown	75 / 718 (10.45%)	11 / 126 (8.73%)	10 / 88 (11.36%)	5 / 105 (4.76%)	101 / 1037 (9.74%)
yes	38 / 718 (5.29%)	22 / 126 (17.46%)	17 / 88 (19.32%)	50 / 105 (47.62%)	127 / 1037 (12.25%)

Table 2. 5-year outcomes of patients with T1 tumors above and below 0.3mm.

Group	31-GEP Class	n	RFS (95% CI)	p-value	DMFS (95% CI)	p-value
≤0.2 mm	All	56	100% (100-100%)	N/A	100% (100-100%)	N/A
0.3-1.0 mm	Class 1A	495	96.76% (94.73-98.83%)	p<0.0001	97.35% (95.39-99.34%)	p=0.0008
	Class 1B	62	91.01% (82.97-99.82%)		90.58% (82.21-99.81%)	
	Class 2A	24	90.91% (79.66-100%)		90% (77.77-100%)	
	Class 2B	32	75.46% (60.95-93.42%)		82.14% (67.51-99.95%)	

31-GEP: 31-gene expression profile; RFS: recurrence-free survival; DMFS: distant metastasis-free survival

DISCUSSION

While the risk of mortality is strongly correlated with Breslow thickness, a substantial proportion of melanoma-specific deaths occur in patients initially diagnosed with thin, T1 melanomas,[15, 19-21] because of the relative prevalence of T1 to thick melanomas (T4, Breslow depth >4 mm). Thus, patients with thin melanoma represent the majority of cutaneous melanoma patients and a subset of these patients have poor outcomes, which accounts for a significant percentage of melanoma-related deaths. The 31-GEP test was developed and validated to enhance prognostic accuracy in melanoma,[2] and its overall performance, as well as the performance in thin melanomas, has been previously reported.[3-9, 17]

With any clinical test, it is important to identify a population for whom use of the test is appropriate and has the potential to add value to patient care. This study was aimed at defining an evidence-based target population of tumors for which 31-GEP testing can maximize sensitivity of identifying Class 2 results and where the test result is more likely to result in a clinically appropriate and meaningful change in an individual patient's treatment plan. To do so, we used a combination of outcomes, 31-GEP class frequency and clinical utility to identify a Breslow thickness threshold at and above which the 31-GEP test can be used to estimate differential risk of recurrence and inform management decisions based on that predicted risk. We used a large dataset of patient tumors and clinical outcomes (i.e. recurrence and distant metastatic events) to estimate 5-year RFS and DMFS rates, as well as cumulative event rates at any given Breslow thickness ≤1.0 mm, and found that there was an observable shift in predicted cumulative event rates above 0.3 mm between Class 1A and Class 2B tumors, which have the lowest and highest risks of recurrence, respectively. Of note, estimated cumulative event rates are based on Loess regression fitting to cumulative event rates at each observed case per 31-GEP subclass in the patient cohort. Since the cumulative event rate at each observed case in the data set reflects that case and those with lower Breslow thicknesses, the cumulative event rates derived from the Loess curve may underestimate the risk of recurrence at a given local Breslow thickness. Additionally, there are fewer recurrence events, as expected, in ultra-thin melanomas and this must be considered when interpreting the fitted curves. Thus, the importance of this analysis is the separation of cumulative rates of recurrence between Class 1A and 2B at the 0.3 mm threshold.

We then evaluated 31-GEP class distribution in thin melanoma tumors from a large dermatopathology practice and in all cases clinically tested during the last year. In patients with tumors 0.3 mm to 1.0 mm from the dermatopathology practice, a 7.1% frequency of Class 2 tumors and a 15.7% frequency of non-Class 1A tumors were identified. Distribution of Class results will likely vary practice to practice based on different patient population. Figure 2B shows the distribution across 8944 T1 tumors clinically tested during a one-year period. In this population, very few tumors <0.3 mm were classified as non-Class 1A, indicating a large number of patients would need to be tested in order to identify a high-risk tumor. In tumors 0.3-1.0 mm, 10.8% were non-Class 1A, suggesting a benefit of risk assessment for this portion of the thin tumor population.

Lastly, we identified that most changes in patient management for consecutively tested

melanomas after 31-GEP testing were made for tumors ≥0.3 mm in Breslow thickness. Taken together, results from these three sets of analyses point towards use of the 31-GEP test to guide management decisions for patients with tumors ≥0.3 mm thick for risk of recurrence prediction. Below 0.3 mm in thickness, there are no observed recurrence events in our research patient cohort, suggesting a very low frequency of recurrence/metastasis in this group, and there are very few expected non-Class 1A results. Moreover, physicians have not shown strong impact of test results to change patient management.

In summary, these results suggest a Breslow thickness threshold of 0.3 mm for using the 31-GEP test for guiding management decisions dependent on individual risk of recurrence. In the context of the comprehensive use of the 31-GEP test, the test is also used to guide decisions on using SLNB in patients with T1-T2 tumors, as reported recently[22]. Since the SLNB surgical procedure is currently recommended by NCCN guidelines for T1a tumors with adverse features, there is a population of patients with <0.3 mm tumors with adverse features who could still be considered for this procedure, and in this population there is utility of the 31-GEP test in identifying patients at very low risk of being SLNB positive who can avoid this procedure. The schematic in Figure 4 depicts both ways the test informs management decisions for melanoma patients and recommended thresholds based on published literature and the analysis presented herein.

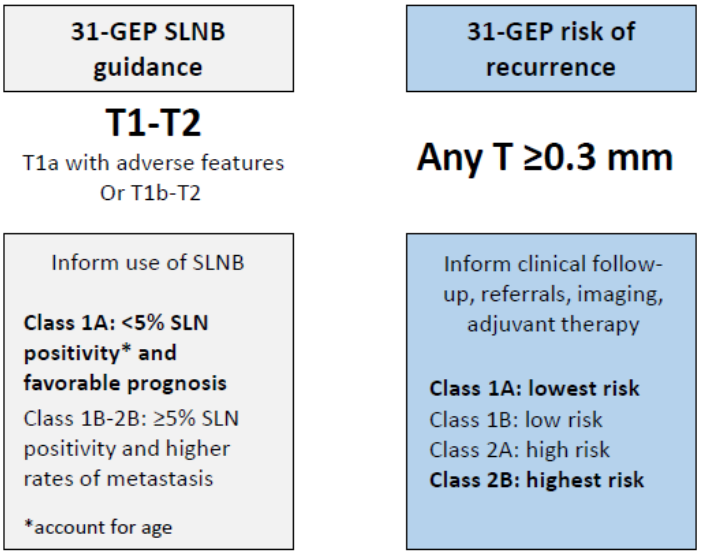

Figure 4. Appropriate use and target populations for 31-GEP test to guide sentinel lymph node biopsy (SLNB) decisions and patient management. Based on published studies and current study data, the 31-GEP test is recommended to be used in patients with T1-T2 tumors (including <0.3 mm tumors with adverse features) for SLNB decision guidance (left) and in tumors with ≥0.3 mm Breslow thickness for risk of recurrence prediction to inform patient management decisions (right).

Conflict of Interest Disclosures: CJC has served as a consultant and speaker for Castle Biosciences, Inc. HGC, SJK, KMP, and FAM are employees and options holders at Castle Biosciences, Inc. EM and SI have no conflicts of interest to declare.

Funding: None.

Corresponding Author:
Etan Marks, DO
UT Southwestern Medical Center
Dallas, TX
etanmarks@gmail.com

References:

1. Morton DL, Thompson JF, Cochran AJ, et al. Final trial report of sentinel-node biopsy versus nodal observation in melanoma. N Engl J Med. 2014;370: 599-609.
2. Gerami P, Cook RW, Wilkinson J, et al. Development of a prognostic genetic signature to predict the metastatic risk associated with cutaneous melanoma. Clin Cancer Res. 2015;21: 175-183.
3. Gastman BR, Gerami P, Kurley SJ, Cook RW, Leachman S, Vetto JT. Identification of patients at risk of metastasis using a prognostic 31-gene expression profile in subpopulations of melanoma patients with favorable outcomes by standard criteria. J Am Acad Dermatol. 2019;80: 149-157.e4.
4. Zager JS, Gastman BR, Leachman S, et al. Performance of a prognostic 31-gene expression profile in an independent cohort of 523 cutaneous melanoma patients. BMC Cancer. 2018;18: 130.
5. Gerami P, Cook RW, Russell MC, et al. Gene expression profiling for molecular staging of cutaneous melanoma in patients undergoing sentinel lymph node biopsy. J Am Acad Dermatol. 2015;72: 780-785.e3.
6. Greenhaw BN, Zitelli JA, Brodland DG. Estimation of prognosis in invasive cutaneous melanoma: An independent study of the accuracy of a gene expression profile test. Dermatol Surg. 2018;44: 1494-1500.
7. Hsueh EC, DeBloom JR, Lee J, et al. Interim analysis of survival in a prospective, multi-center registry cohort of cutaneous melanoma tested with a prognostic 31-gene expression profile test. J Hematol Oncol. 2017;10: 152.
8. Keller J, Schwartz TL, Lizalek JM, et al. Prospective validation of the prognostic 31-gene expression profiling test in primary cutaneous melanoma. Cancer Med. 2019.
9. Podlipnik S, Carrera C, Boada A, et al. Early outcome of a 31-gene expression profile test in 86 AJCC stage IB-II melanoma patients. A prospective multicentre cohort study. J Eur Acad Dermatol Venereol. 2019.
10. Berger AC, Davidson RS, Poitras JK, et al. Clinical impact of a 31-gene expression profile test for cutaneous melanoma in 156 prospectively and consecutively tested patients. Curr Med Res Opin. 2016;32: 1599-1604.
11. Dillon LD, Gadzia JE, Davidson RS, et al. Prospective, multicenter clinical impact evaluation of a 31-gene expression profile test for management of melanoma patients. SKIN. 2018;2: 111-121.
12. Schuitevoerder D, Heath M, Cook RW, et al. Impact of gene expression profiling on decision-making in clinically node negative melanoma patients after surgical staging. J Drugs Dermatol. 2018;17: 196-199.
13. Cook RW, Middlebrook B, Wilkinson J, et al. Analytic validity of DecisionDx-Melanoma, a gene expression profile test for determining metastatic risk in

melanoma patients. Diagn Pathol. 2018;13: 13.
14. Ferris LK, Farberg AS, Middlebrook B, et al. Identification of high-risk cutaneous melanoma tumors is improved when combining the online American Joint Committee on Cancer Individualized Melanoma Patient Outcome Prediction Tool with a 31-gene expression profile-based classification. J Am Acad Dermatol. 2017;76: 818-825.e3.
15. Shaikh WR, Dusza SW, Weinstock MA, Oliveria SA, Geller AC, Halpern AC. Melanoma thickness and survival trends in the United States, 1989 to 2009. J Natl Cancer Inst. 2016;108.
16. Gershenwald JE, Scolyer RA, Hess KR, 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: 472-492.
17. Greenhaw BN, Hsueh EC, Covington KR, Plasseraud KM, Cook RW, Gastman BR. Meta-analysis of the prognostic 31-gene expression profile test in 1261 cutaneous melanoma cases. American Academy of Dermatology. Washington, D. C. , 2019.
18. Farberg AS, Glazer AM, White R, Rigel DS. Impact of a 31-gene expression profiling test for cutaneous melanoma on dermatologists' clinical management decisions. J Drugs Dermatol. 2017;16: 428-431.
19. Gimotty PA, Elder DE, Fraker DL, et al. Identification of high-risk patients among those diagnosed with thin cutaneous melanomas. J Clin Oncol. 2007;25: 1129-1134.
20. Landow SM, Gjelsvik A, Weinstock MA. Mortality burden and prognosis of thin melanomas overall and by subcategory of thickness, SEER registry data, 1992-2013. J Am Acad Dermatol. 2017;76: 258-263.
21. Whiteman DC, Baade PD, Olsen CM. More people die from thin melanomas (1 mm) than from thick melanomas (>4 mm) in Queensland, Australia. J Invest Dermatol. 2015;135: 1190-1193.
22. Vetto JT, Hsueh EC, Gastman BR, et al. Guidance of sentinel lymph node biopsy 316 decisions in patients with T1-T2 melanoma using gene expression profiling. Future Oncol. 2019;15: 1207-1217.