Molecular Biomarkers for Melanoma Screening, Diagnosis and Prognosis: Current State and Future Prospects

Dekker C. Deacon[1]*[†], Eric A. Smith[2†] and Robert L. Judson-Torres[1,3]*

[1] Department of Dermatology, University of Utah, Salt Lake City, UT, United States, [2] Department of Pathology, University of Utah, Salt Lake City, UT, United States, [3] Huntsman Cancer Institute, Salt Lake City, UT, United States

Despite significant progress in the development of treatment options, melanoma remains a leading cause of death due to skin cancer. Advances in our understanding of the genetic, transcriptomic, and morphologic spectrum of benign and malignant melanocytic neoplasia have enabled the field to propose biomarkers with potential diagnostic, prognostic, and predictive value. While these proposed biomarkers have the potential to improve clinical decision making at multiple critical intervention points, most remain unvalidated. Clinical validation of even the most commonly assessed biomarkers will require substantial resources, including limited clinical specimens. It is therefore important to consider the properties that constitute a relevant and clinically-useful biomarker-based test prior to engaging in large validation studies. In this review article we adapt an established framework for determining minimally-useful biomarker test characteristics, and apply this framework to a discussion of currently used and proposed biomarkers designed to aid melanoma detection, staging, prognosis, and choice of treatment.

Keywords: melanoma, biomarkers, diagnostics, prognostics, likelihood ratio

INTRODUCTION

At a fundamental level, clinical tests are designed to aid clinicians in determining a best course of action when there is clinical uncertainty. These tests aim to improve certainty by ruling out a disease state, confirming a disease state, delineating prognosis, and/or identifying the most appropriate treatment regimen. Toward these goals, specific laboratory tests interrogate biomarkers: biological molecules that have clinically-relevant associations with a specific disease state. However, even a biomarker with an established disease state association may not be useful if it does not influence clinical decision-making. There are numerous factors that determine clinical utility: pre-test certainty of disease state, risks associated with making a clinical decision, cost, turn-around time, and the unique circumstances of each patient. Since biomarker development, from discovery to validation, requires substantial resources, investigators should carefully consider what characteristics are necessary for a useful clinical test at each stage of development.

With melanoma and other neoplasms, clinicians seek to balance the efficacy and potential side effects of an intervention with the estimated risk of disease progression. This risk has been traditionally inferred using prognostic tumor staging criteria, which takes into account tumor size, extent, nodal involvement, distant metastases, and histopathological grade and type. In more recent revisions of National Comprehensive Cancer Network (NCCN) and American Joint Committee on Cancer (AJCC) guidelines, select biomarkers have augmented traditional staging algorithms. The level of serum LDH is currently incorporated as a staging (prognostic) factor for malignant

melanoma (1), and the mutational status of the *BRAF* proto-oncogene is predictive of therapeutic response (2–4). However, many critical intervention points in melanoma care remain fraught with uncertainty. Critical shortcomings in diagnosing, staging, and recurrence monitoring for melanoma could be alleviated with appropriate biomarker development. Toward this goal, many biomarker tests have been proposed, some of which are commercially available or are performed in academic referral centers. Presently, none of these tests have been compared to the current standard of care with large scale, randomized, prospective, multi-center, and independently validated studies with long-term clinical follow up.

For this literature review, we summarize the established theoretical framework for determining the minimal test characteristics required to potentially alter clinical decision making at different stages of melanoma care. We then apply this theory to evaluate currently utilized and proposed melanoma biomarkers. A recent review from the Melanoma Prevention Working Group critiqued the rigor in which commercially available prognostic tests have been validated and provided recommendations for care providers (5). Although we touch upon similar analyses here, our goal is to review a larger spectrum of biomarkers for melanoma care—preliminary to practiced—from the perspective of prioritizing further development. Given the substantial resources required for clinical test validation, we hope this discussion will aid the field in deciding which existing tests warrant further validation and which decision points are in most need of further biomarker discovery.

The Case for Biomarkers in Melanoma Diagnosis and Treatment

Despite recent advances in treatment of advanced stage melanoma, particularly with the expansion of the use of immune checkpoint inhibitors (6), melanoma continues to confer significant morbidity and mortality. In 2015, there were 59,782 deaths attributed to melanoma worldwide (7). Nearly 1.6% of all cancer diagnoses are melanoma, and the disease accounts for 0.64% of cancer deaths (8). In the United States specifically, it is expected that there will be over 100,000 new diagnoses of melanoma in 2020 with almost 7,000 deaths attributed to melanoma (9). In Europe, while cost of treatment varies widely based on country and stage, melanoma treatment can range from several thousand to tens of thousands of Euros per patient on average, and these costs are expected to increase with wider adoption of immune checkpoint inhibitors and targeted kinase inhibitors (10).

Varied and unique uncertainties surround clinical decision-making during the detection, diagnosis, and treatment of melanocytic tumors. We have reviewed and analyzed four critical decision points during the identification and treatment of melanoma for which biomarkers that reduce uncertainty have been investigated. These decision points are: (i) deciding whether to biopsy a melanocytic neoplasm, (ii) confirming histopathologic diagnosis, (iii) stratifying risk for lymphatic spread with consideration for SLNB and, (iv) selecting systemic therapy. While this review focuses on cutaneous melanoma,

the considerations discussed could be adapted to potential biomarkers for mucosal and uveal melanoma as well.

Histopathology: the 14-Karat Gold Standard for Diagnosing Melanocytic Lesions

One poignant example of where biomarkers could dramatically enhance melanoma care is with the histopathologic diagnosis of melanoma. Histopathologic diagnosis is the gold standard for melanoma diagnosis, but, despite the advent of immunohistochemistry (IHC) and standardization of diagnostic criteria (11–13), it remains a subjective medical art constrained by significant intra- and interobserver variability. While nuances exist, melanocytic proliferations largely exist on a spectrum from histologically benign to malignant, as described in the MPATH-Dx Classification scheme (**Table 1**) (11). This classification scheme breaks melanocytic lesions into benign (class I), moderately dysplastic nevus and Spitz nevus (class II), severely dysplastic nevus, atypical Spitz nevi (class III), AJCC T1a and T1b invasive melanomas (class IV), and AJCC T1b-T4a invasive melanomas (class V). In the absence of metastasis identified during sentinel lymph node biopsy (SLNB) or further staging, these primary melanomas are considered Stage I-II using the AJCC 8th Edition Pathological Staging Criteria (14).

Although lesions at the extremes of the spectrum (class I and class V lesions) are most likely to be reproducibly diagnosed as such, there remains significant diagnostic discordance among dermatopathologists for each class. In one study evaluating 187 pathologists, MPATH-Dx class I and V have intraobserver concordance rates of 77 and 67%, respectively, and interobserver concordance rates of 71 and 55%, respectively. When compared to consensus diagnosis, the accuracies of individual pathologists were 92% for class I and 72% for class V (15). Separating class II, class III, and class IV is a greater diagnostic challenge. Class II have the poorest concordance (35% intraobserver, 25% interobserver) and consensus accuracy (25%), followed closely by classes III (60% intra- and 45% interobserver consensus with 40% accuracy) and IV (63% intra- and 46% interobserver consensus with 43% accuracy) (15). Similar results have been seen in smaller studies (16, 17). To summarize, these studies concluded that pathologists concurred with their own previous evaluations in only two-thirds to three-quarters of cases, and, for most stages, reached the same diagnosis as their colleagues less than half the time.

Given the poor concordance in distinguishing class II through IV melanocytic lesions, several reports have sought to identify barriers to reproducibility. Unsurprisingly, those pathologists who are trained in dermatopathology, have a significant caseload of melanocytic lesions, and have more than 5 years of post-training experience have better reproducibility and accuracy (18). Dermatopathology training reduced misclassification rates of class II lesions by 10% and class III-IV lesions by 20%. While second and third opinions from trained dermatopathologists can reduce the misclassification rate of a general pathologist, additional peer review has a smaller, but still significant, benefit when a dermatopathologist performs the initial read (19). If

TABLE 1 | Abbreviated MPATH-DX classification and AJCC 8th edition staging criteria for cutaneous melanoma.

MPATH-DX Classification	
Class I	Any benign nevus with or without mild atypia: junctional, compound, intradermal, combined, congenital, Blue, lentiginous nevus, halo nevus
Class II	Dysplastic moderately atypical nevi, conventional Spitz nevi, cellular or epithelioid blue nevi, pigmented or plexiform spindle cell nevi
Class III	Dysplastic severely atypical nevi, melanoma *in situ*, atypical Spitz tumor, atypical pigmented spindle cell tumor, Pagetoid intraepidermal melanocytic proliferation
Class IV	Invasive melanomas ≤1.0 mm thick without ulceration (AJCC 7th edition T1a)
Class V	Invasive melanomas >1.0 mm thick or ≤1.0 mm thick with ulceration (AJCC 7th edition T1b)

AJCC 8th Edition Pathological Staging Criteria	
Stage IA	T1a: <0.8 mm thick without ulceration
	T1b: <0.8 mm thick with ulceration or 0.8-1.0 mm thick without ulceration
Stage IB	T2a: >1-2 mm thick without ulceration
Stage IIA	T2b: >1-2 mm thick with ulceration, T3a: >2-4 mm thick without ulceration
Stage IIB	T3b: >2-4 mm thick with ulceration, T4a: >4 mm thick without ulceration
Stage IIC	T4b: >4 mm thick with ulceration
Stage IIIA	T1a-T2aN1a/2a: ≤2 mm thick with 1-3 positive occult SLN on biopsy
Stage IIIB	T0N1b-1c: No primary tumor but 1 clinically positive SLN or in-transit metastasis, T1a-T3aN1a-2b: Stage IA-IIA primary with 1-3 positive lymph node (LN) and/or in-transit metastasis
Stage IIIC	T3b-4aN1a-3c: Stage IIB primary with any evidence of local metastasis, T4bN1a-2c: Stage IIC primary with 1-3 positive LN or only 1 positive SLN with in-transit metastasis.
Stage IIID	T4bN3a-3c: Stage IIC primary with ≥4 positive LN or ≥2 LN with in-transit metastasis.
Stage IV	T0-4bN0-3cM1: Distant metastasis (M) with any size primary tumor (T) or number of lymph nodes (N)

these findings reflect the pathology community, then three highly trained dermatopathologists reviewing the same class II, III, or IV case, would be expected to have a consensus accuracy of 35, 52, and 62%, respectively (19). These studies indicate that training and experience are not sufficient to reach high levels of accuracy for intermediate lesions. The most beneficial, albeit modest, improvement in concordance was obtained by updating from the AJCC 7th edition diagnostic criteria to the 8th edition (12), suggesting that further refinement of diagnostic criteria could yet improve accuracy. However, diagnostic threshold variability between observers and a tendency to over call diagnoses due to medicolegal concerns (diagnostic drift) are expected to limit the histopathological diagnosis of melanocytic lesions regardless of diagnostic criteria improvements (20–22).

The Promise and Perils of Developing Biomarkers

The above observations highlight the need for objective biomarkers to improve diagnostic accuracy. Biomarkers that bring clarity to the MPATH-Dx classifications would improve patient care by more accurately matching intervention risk to lesion prognosis and save patients and medical systems from expense and potential adverse outcomes related to unnecessary procedures. However, the poor concordance and accuracy of histologic melanoma diagnosis highlight the substantial challenge in developing such biomarkers. The development and validation of clinically useful biomarkers require a gold standard, or a true state, for comparison. If only unambiguous cases are selected for study, then the biomarker, by definition,

provides no clinical value to augment current standard of care. Alternatively, if the gold standards are sufficiently ambiguous (for example, a consensus accuracy of 35%), then a theoretically perfect biomarker is expected to fail validation. Thus, for prospective biomarkers to yield useful test characteristics and address a clinically useful decision-making process, careful selection of the gold standard is required. In addition to validated clinical utility, practical considerations such as cost, invasiveness, tissue requirement, and ease of performing the test may influence biomarker use by clinicians.

The resources required to validate a candidate biomarker are substantial in terms of cost, time, and, most critically, patient volunteers. For example, an international coalition of melanoma experts reported that a randomized clinical trial evaluating the added value of prognostic gene-expression based biomarkers would require cohorts of 1,000 to 9,000 patients, depending on the trial design and hypothesis being tested (5). Therefore, it is critical to carefully consider all of the potential pitfalls of biomarker discovery, development, and implementation prior to committing limited resources to validation of any specific test. Considerations include not only the performance of a proposed biomarker, but also the size and diversity of the cohorts from which it was discovered and validated, and its potential to integrate with current practice and positively influence clinical decision-making. Prior to development, investigators should consider what test characteristics are required for a test to be useful, so that the field can determine how best to distribute limited resources for biomarker discovery and validation. These characteristics are described next.

APPLICATION OF FUNDAMENTAL THEORY OF BIOMARKER UTILITY TO CURRENT MELANOMA BIOMARKERS

Clinical Uncertainty in Melanoma Diagnosis and Treatment

The diversity of clinical circumstances surrounding individual melanomas ensures that there is no "one-size-fits-all" approach to the diagnosis and management of melanoma. Often, the first decision for a clinician in the evaluation of a pigmented skin lesion is to determine whether it is likely benign or has atypical clinical characteristics that warrant biopsy for further evaluation (**Figure 1A**). If a biopsy is performed, a pathologist must evaluate the likelihood of the lesion being malignant or benign (**Figure 1B**). For invasive malignant melanomas, a surgical oncologist must decide whether to offer SLNB to evaluate for early metastatic disease (**Figure 1C**), and selection of systemic therapy becomes necessary with local (e.g., a positive SLNB) or distant metastasis. Each of these decisions represent a junction where the clinician is faced with potential uncertainty and is forced to balance the odds of disease with a significant intervention. Clarification of uncertainty at these decision points might be achieved with objective biomarkers, as discussed in detail below. While these points do not comprise an exhaustive set of decision points in melanoma diagnosis and management that could benefit from biomarker development, tests that aid these decision points would lead to dramatic improvements in patient care. To build a foundation for this discussion, a summary of the fundamental theory of biomarker accuracy and utility is presented.
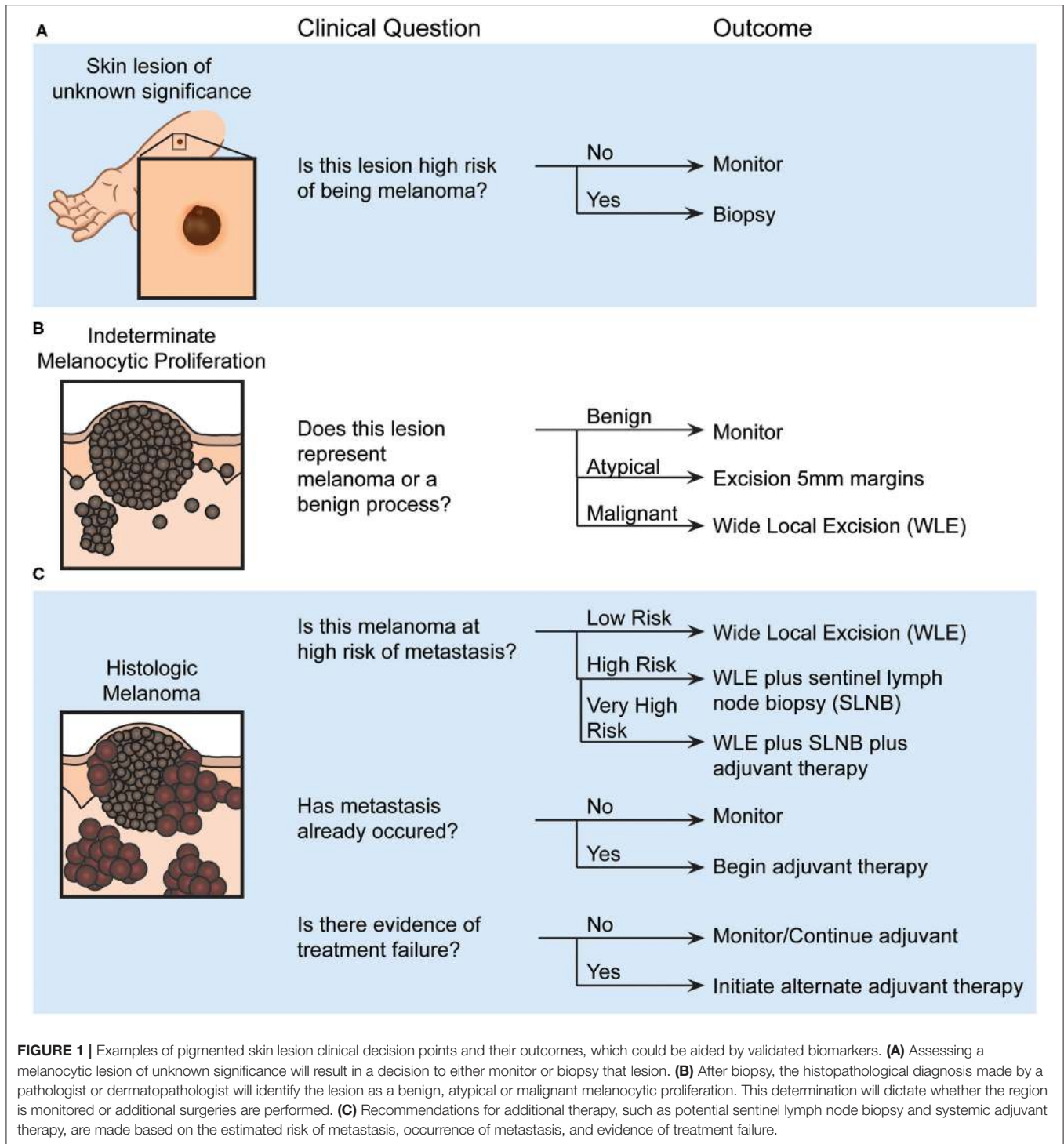
The Theoretical Framework of a Clinically Useful Biomarker

A clinically useful test must be capable of changing clinical decision making. In other words, usefulness is determined by the clinician ordering the test. Clinicians routinely encounter patients presenting with an ambiguous disease state. An illustrative example would be a sore throat that could be due to either *Streptococcus pyogenes* or a viral etiology. In this scenario, a rapid "strep" test or bacterial cultures may be ordered to resolve this ambiguity by increasing the clinician's certainty on whether the intervention of antibiotics would benefit the patient. A clinical test could resolve this uncertainty by decreasing the perceived likelihood of the disease state (exclusion), such as with a negative point-of-care rapid "strep" test. Alternatively, a clinical test could increase the likelihood of a disease state (confirmation), such as with a positive bacterial culture. If the test result shifts the clinician's certainty on the patient's diagnosis above or below their self-determined intervention threshold (e.g., a positive rapid "strep" test resulting in an antibiotic prescription), the test will have usefully influenced a clinical decision.

Similarly, predictive and prognostic tests may aim to determine disease trajectory and can assist in directing further intervention, such as follow up imaging or adjuvant therapy selection.

The potential utility of a biomarker can be estimated by first considering clinicians' certainty of a disease state using current standard of care ("prior odds") and the certainty threshold that must be breached for intervention ("threshold") (**Figure 2A**). It is important to recognize that since both values are derived from individual clinicians' perceptions, they are inherently imprecise and inconsistent. However, ranges for these values can be estimated through analysis of collective clinical decisions or surveying and these ranges are informative when considering the potential utility of a candidate test. When prior odds and intervention thresholds have been estimated, the slope of the line that connects them represents the change in the odds (also called the likelihood ratio or LR) required to change a decision. A clinical test that decreases the perceived likelihood of the disease state results in a negative likelihood ratio or LR- (**Figure 2A**, green line). A clinical test that increases the perceived likelihood of the disease state results in a positive likelihood ratio or LR+ (**Figure 2A**, blue line).

When determining whether a decision point might benefit from biomarker development, it is useful to estimate what values of LR would be required of a theoretical test for posterior odds to surpass the intervention threshold (**Figure 2A**). Likelihood ratios are calculated from the sensitivity and specificity of the test at one reference range or cut-off value [For information on how to calculate these and other established clinical accuracy metrics see (24) and (23)]. Sensitivity describes the ability of the test to discern the presence of a disease and specificity quantifies how often it correctly identifies disease absence. There is often a trade-off between sensitivity and specificity, particularly with test outcomes reported as a continuous variable. Since clinical accuracy is based on assignment of the test cut-off value (a value outside of the "normal" reference range that is flagged as an abnormal result), the sensitivity and specificity of a test can vary wildly based on where a laboratory assigns the cut-off values for normal. To address this issue, test characteristics can be plotted as a receiver operating characteristic (ROC) curve, which plots the sensitivity and specificity at different cut-off points. This also allows one to compare different tests/biomarkers by calculating the area under the curve (AUC). In general, the higher area under the curve reflects a better performing test, with an AUC of 1.0 having perfect sensitivity and specificity and an AUC of 0.5 being useless (**Figure 2B**). Since utility is determined by the clinician ordering and interpreting the test, there is no standardized AUC at which a test becomes clinically useful. Instead, positive and negative likelihood ratios can be derived from the slope between any ROC point and the (0,0) and (1,1) vertices, respectively (**Figure 2B**, point V). If a biomarker test application requires at least a LR+ of 10 and/or LR– of 0.1 to change clinical decision (as in the example presented in **Figure 2A**), then theoretically any point on the ROC curve that lies to the left (for LR+) or above (LR–) of the LR lines could be considered tests worth further development (**Figures 2C,D**). However, ROC curves with higher AUC will provide a larger range of potential test cut-offs that will meet these minimal LRs and will consequently have a higher sensitivity and/or

FIGURE 1 | Examples of pigmented skin lesion clinical decision points and their outcomes, which could be aided by validated biomarkers. **(A)** Assessing a melanocytic lesion of unknown significance will result in a decision to either monitor or biopsy that lesion. **(B)** After biopsy, the histopathological diagnosis made by a pathologist or dermatopathologist will identify the lesion as a benign, atypical or malignant melanocytic proliferation. This determination will dictate whether the region is monitored or additional surgeries are performed. **(C)** Recommendations for additional therapy, such as potential sentinel lymph node biopsy and systemic adjuvant therapy, are made based on the estimated risk of metastasis, occurrence of metastasis, and evidence of treatment failure.

specificity at those cut-offs (**Figures 2C,D**, ROC curves at point W and Y).

In the following sections we apply this theory to estimate the test characteristics required to change clinical decision making in the diagnosis and treatment of melanoma. Where available, we use a combination of published studies and surveys to estimate ranges for the prior odds and intervention thresholds.
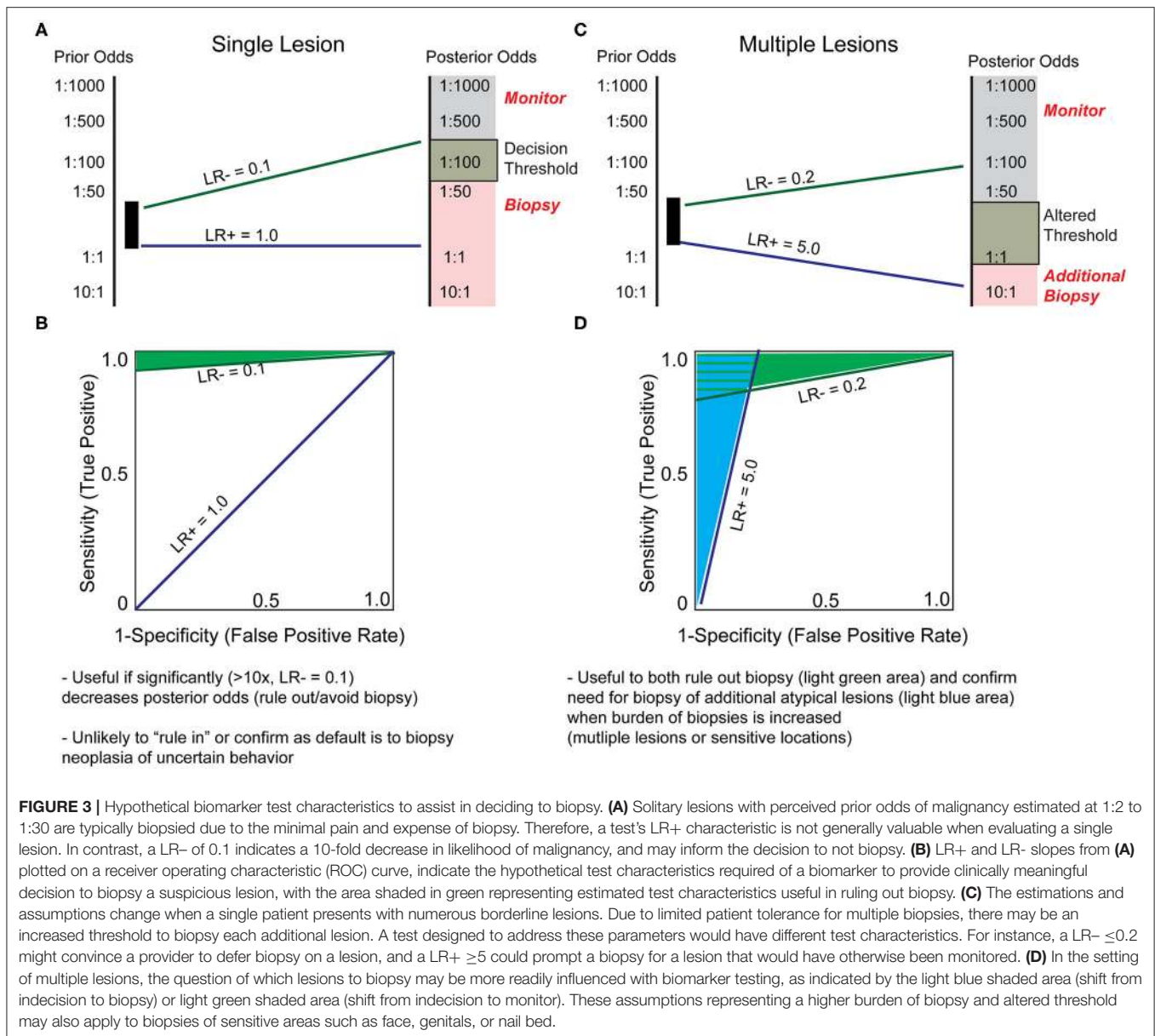
However, every providers' internal threshold of certainty for clinical decision making will vary in the context of patient-specific factors in addition to test characteristics. Thus, these ranges should not be considered exact quantifications but rather estimates to help determine which biomarkers may be worth investing community resources for further discovery and development.

FIGURE 2 | Theoretical framework for biomarker utility analysis. **(A)** A clinician may be challenged with a state of uncertainty, where the perceived probability of the disease (prior odds, purple) does not surpass the clinician's personal thresholds for recommending either major intervention or no intervention (decision thresholds, red dotted lines). A clinician may hope to reduce uncertainty based upon a biomarker test result. A positive result must alter the probability of the disease sufficiently to breach the decision threshold to recommend major intervention (blue line). A negative result must alter the probability of the disease sufficiently to breach the decision threshold to recommend no intervention (green line). In this example, the prior odds are an equal probability of disease or no disease. Based upon this clinician's decision thresholds, an actionable biomarker test result would require either a 10-fold increase or 10-fold decrease in this perceived probability (posterior odds). The extent by which a biomarker test might change the perceived odds can be calculated as a likelihood ratio (LR). A positive LR (LR+, blue line) increases the likelihood of the disease state being present and a negative LR (LR–, green line) decreases the likelihood of the disease state being present. The posterior odds following a biomarker test can be calculated as the prior odds multiplied by the likelihood ratio. **(B)** LR+ and LR– can be plotted against sensitivity and specificity on the axes of a receiver operator characteristic (ROC) curve graph, which represents the sensitivity and specificity of the test at different cut-off values where an "abnormal" result flag would be generated (red curved line). The LR values represent the slope at any given point on the ROC curve, and the LR lines for point V are represented. The area under the ROC curve (AUC) describes the overall sensitivity and specificity of a test at all cut-off points. AUC of 1.0 represents 100% sensitivity and specificity, and a useless test (gray line) has an AUC of 0.5. The red ROC curve has an approximate AUC of >0.9. **(C)** In the example shown in **(A)**, a minimum LR+ of 10 is required to generate posterior odds that breach the clinician's decision threshold for major intervention. The area (shaded in light blue) left of this LR+ slope (dark blue line) represents the sensitivity and specificity characteristics of a test useful for confirming or ruling in a disease state that has an LR+ ≥10. A biomarker test with performance represented by the red ROC curve has a greater range of cut-off values that can produce an LR+ ≥ 10, and has a higher maximal potential sensitivity (point W) compared to the purple ROC curve (point X). This translates to a greater odds of achieving a true test value that has the potential to change the clinician's decision for the red ROC curve, especially with test cut-off values at point W. **(D)** Similar to **(C)**, the area (shaded in green) above the LR– slope (dark green line) represents the sensitivity and specificity characteristics of a test useful for excluding or ruling out the example disease state from **(A)**. The red ROC curve represents a hypothetically useful test with a higher maximal sensitivity capable of achieving a LR– ≤0.1 (point Y), and the purple ROC curve represents a test that is less likely to influence clinical decision making with a lower maximal sensitivity capable of achieving a LR– ≤0.1 (point Z). Figure adapted from (23).

Toward Enhancing Pigmented Lesion Screening With Biomarkers

Many factors are considered when assessing a pigmented lesion in clinic prior to recommending a biopsy. These include but are not 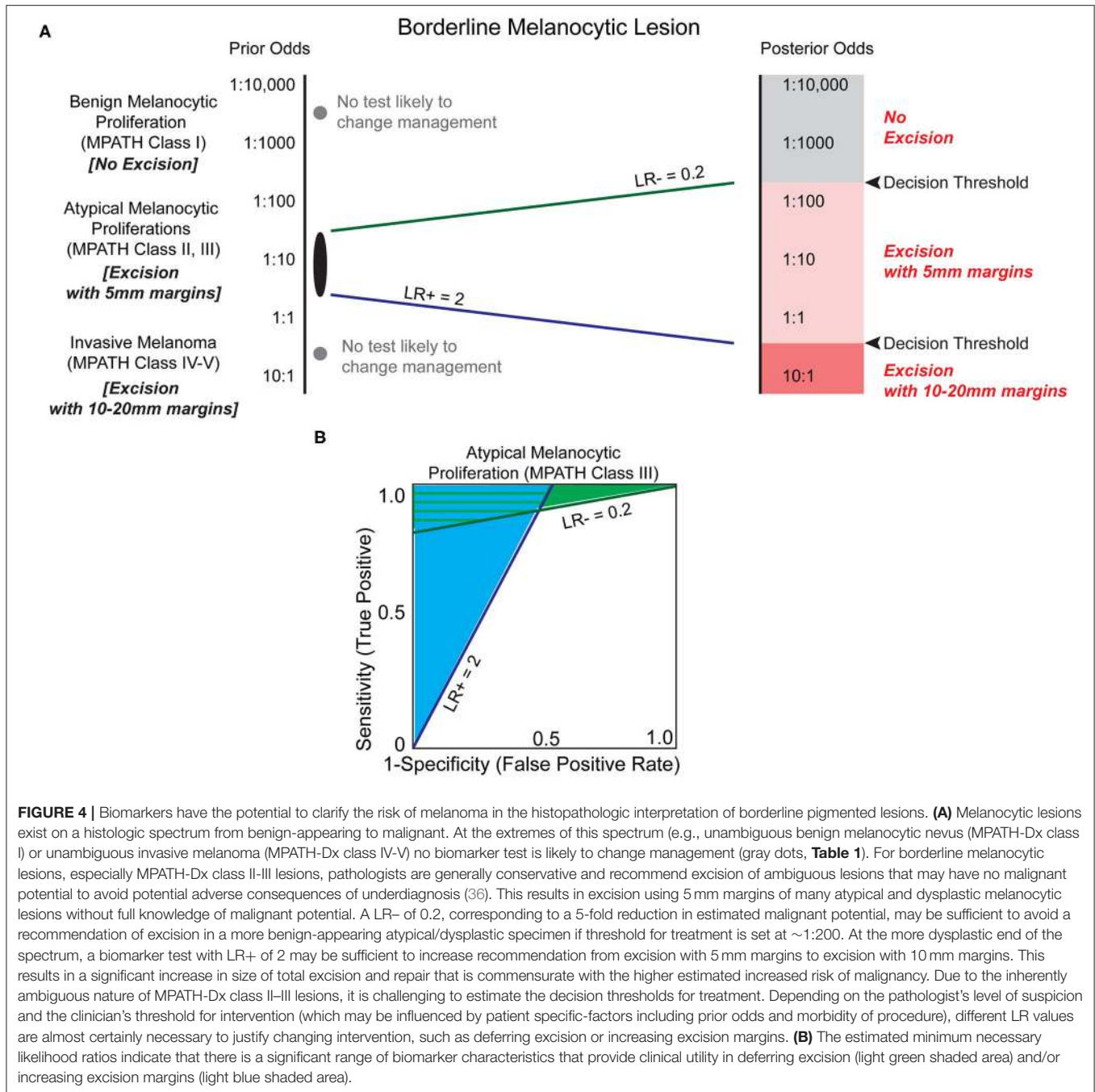limited to clinical impression of the lesion, history of the lesion (either as reported by the patient or through comparison to previous clinical images), the patient's history and risk for melanoma, total number of lesions, and dermatoscopic features. Since the decision to biopsy represents a gestalt of these different factors, it is not surprising that the level of specialty

FIGURE 3 | Hypothetical biomarker test characteristics to assist in deciding to biopsy. **(A)** Solitary lesions with perceived prior odds of malignancy estimated at 1:2 to 1:30 are typically biopsied due to the minimal pain and expense of biopsy. Therefore, a test's LR+ characteristic is not generally valuable when evaluating a single lesion. In contrast, a LR− of 0.1 indicates a 10-fold decrease in likelihood of malignancy, and may inform the decision to not biopsy. **(B)** LR+ and LR- slopes from **(A)** plotted on a receiver operating characteristic (ROC) curve, indicate the hypothetical test characteristics required of a biomarker to provide clinically meaningful decision to biopsy a suspicious lesion, with the area shaded in green representing estimated test characteristics useful in ruling out biopsy. **(C)** The estimations and assumptions change when a single patient presents with numerous borderline lesions. Due to limited patient tolerance for multiple biopsies, there may be an increased threshold to biopsy each additional lesion. A test designed to address these parameters would have different test characteristics. For instance, a LR− ≤0.2 might convince a provider to defer biopsy on a lesion, and a LR+ ≥5 could prompt a biopsy for a lesion that would have otherwise been monitored. **(D)** In the setting of multiple lesions, the question of which lesions to biopsy may be more readily influenced with biomarker testing, as indicated by the light blue shaded area (shift from indecision to biopsy) or light green shaded area (shift from indecision to monitor). These assumptions representing a higher burden of biopsy and altered threshold may also apply to biopsies of sensitive areas such as face, genitals, or nail bed.

training and experience of the physician substantially alter the number of lesions biopsied for each melanoma identified. A recent meta-analysis of 351 studies determined the sensitivity and specificity of identifying melanomas during skin exams was 87.5 and 81.4%, respectively, for dermatologists, but only 79.9 and 70.9% for primary care physicians (25). Unfortunately, access to dermatologists for regular skin exams is neither uniform nor comprehensive, with over half of U.S. citizens having insufficient access to dermatologic care (26). Dermoscopy and reflectance confocal microscopy (RCM) are two non-invasive approaches that increase the number of melanomas biopsied, while reducing the number of benign lesions biopsied (27, 28). However, these techniques also require highly skilled interpretation, and use of RCM is largely limited to dermatologists who specialize in pigmented lesions. Dermoscopy is more broadly used in

clinical practice, though it does not appear to increase diagnostic accuracy in the practice of non-experts (29) and has even been reported to reduce sensitivity of melanoma detection compared to clinical assessment alone (30). An objective biomarker could therefore improve the odds to biopsy a true melanoma, especially in situations where a specialized dermatologist and auxiliary equipment/techniques are not available. In contrast, any increases in certainty might be minimal when used by a highly specialized and experienced pigmented lesion dermatologist.

Ultimately, the cost of a skin biopsy to the individual patient, in terms of both morbidity and finance, is relatively low, but the cost of leaving an early-stage melanoma undiagnosed is exceptionally high. By applying the above theory of biomarker utility, the prior odds of malignancy can be estimated by considering the number of lesions biopsied to diagnose one

FIGURE 4 | Biomarkers have the potential to clarify the risk of melanoma in the histopathologic interpretation of borderline pigmented lesions. (A) Melanocytic lesions exist on a histologic spectrum from benign-appearing to malignant. At the extremes of this spectrum (e.g., unambiguous benign melanocytic nevus (MPATH-Dx class I) or unambiguous invasive melanoma (MPATH-Dx class IV-V) no biomarker test is likely to change management (gray dots, **Table 1**). For borderline melanocytic lesions, especially MPATH-Dx class II-III lesions, pathologists are generally conservative and recommend excision of ambiguous lesions that may have no malignant potential to avoid potential adverse consequences of underdiagnosis (36). This results in excision using 5 mm margins of many atypical and dysplastic melanocytic lesions without full knowledge of malignant potential. A LR− of 0.2, corresponding to a 5-fold reduction in estimated malignant potential, may be sufficient to avoid a recommendation of excision in a more benign-appearing atypical/dysplastic specimen if threshold for treatment is set at ~1:200. At the more dysplastic end of the spectrum, a biomarker test with LR+ of 2 may be sufficient to increase recommendation from excision with 5 mm margins to excision with 10 mm margins. This results in a significant increase in size of total excision and repair that is commensurate with the higher estimated increased risk of malignancy. Due to the inherently ambiguous nature of MPATH-Dx class II–III lesions, it is challenging to estimate the decision thresholds for treatment. Depending on the pathologist's level of suspicion and the clinician's threshold for intervention (which may be influenced by patient specific-factors including prior odds and morbidity of procedure), different LR values are almost certainly necessary to justify changing intervention, such as deferring excision or increasing excision margins. (B) The estimated minimum necessary likelihood ratios indicate that there is a significant range of biomarker characteristics that provide clinical utility in deferring excision (light green shaded area) and/or increasing excision margins (light blue shaded area).

melanoma. This ranges from ∼2 for some dermatologists to ∼30 for primary care physicians (31). Thus, according to current clinical practices, any lesions with prior odds lower than 1/2–1/30 of being malignant are less likely to be biopsied. A biomarker introduced at this stage could therefore provide utility if it conferred a significant degree of certainty that a lesion was not malignant and thus avoided biopsy. For certainty to be sufficiently increased such that the perceived odds of malignancy is reduced from 1/2 and 1/30 prior to the test to <1/30 after the test (posterior odds), an estimated likelihood ratio of 0.1

(LR-) is required (**Figures 3A,B**, green line). Conversely, for a melanocytic tumor eliciting a high suspicion of malignancy from conventional means (high prior odds), a biomarker is unlikely to change decision making, as the clinical threshold to perform the outcome (biopsy) has already been met (LR+ 1.0, **Figures 3A,B**, blue line). Although this analysis suggests that molecular biomarkers are unlikely to change clinical behavior when assessing solitary lesions, there are other clinical scenarios that challenge the assumptions and conclusions of this model. For example, when individual lesions present in sensitive areas,

such as the face or genitals, the morbidity of biopsy and the clinical suspicion threshold required to biopsy are increased. In these cases, a modest LR− may suffice to shift the clinical decision from "biopsy" to "monitor." Alternatively, due to increased morbidity and decreased patient tolerance of multiple skin biopsies, a patient with many suspicious lesions may benefit from a combination "rule in" and "rule out" molecular test (necessary characteristics estimated as LR+ = 5.0 and LR− 0.2, respectively, **Figures 3C,D**) to prioritize lesions to biopsy (32).

The pigmented lesion assay (PLA), developed by DermTech, is the only biomarker currently commercially available for clinical use addressing this decision (33, 34). This test is an adhesive patch-based gene expression test measuring expression of *PRAME* and *LINC00518*, with reported sensitivity of 91% and specificity of 69% (34). The non-invasive nature and high sensitivity of this test has allowed for clinical incorporation in certain specialized scenarios (32). However, a recent study concluded the test is neither widely nor routinely used by pigmented lesion experts, which is consistent with predictions of the above model (35). Beyond the test's performance characteristics, another limitation is the delay between sampling and result, which necessitates a subsequent biopsy visit if the test is positive. If a similarly sensitive point-of-care biomarker test could be developed as a rapid in-office test, it may be more widely adopted.

In summary, because the clinical default in a setting of uncertainty will usually be to biopsy an indeterminate lesion, biomarkers with a large LR− represent the opportunity to change clinical management by avoiding biopsy. For this purpose, a biomarker with maximum specificity, thus minimizing false negatives, is critical since the adverse consequences of missing the diagnosis is high and the morbidity and cost of a skin biopsy is relatively low. There are situations within a dermatology clinic where a biomarker with high sensitivity (LR+) may also prove clinically useful, but these are likely to be more specialized scenarios such as prioritizing lesions for biopsy in a patient with dysplastic nevus syndrome. Combined with practical limitations for such widespread use of non-invasive biomarkers used at this stage, including the ease of use, cost, and the timing of results, we speculate that current technologies for molecular profiling are unlikely to be influential in the routine evaluation of single lesions.

Current Biomarkers for the Histopathological Diagnosis of Pigmented Lesions

The established challenges of histopathologic analysis has motivated the search for biomarkers to assist in melanoma diagnosis. If prior odds based on original histology favor a more benign dysplasia, such as a dysplastic nevus, then it is possible that a moderate LR−, estimated at 0.2, may convince a clinician to monitor a biopsy site rather than pursuing an excision with margins (**Figure 4**). Conversely, if a biomarker applied to the same lesion demonstrated a significant LR+ of 10, then the result may convince a clinician to excise the area with larger margins (10–20 mm instead of 5 mm), despite relatively low prior odds.

For a lesion with higher prior odds of possessing malignant potential but lacking *bona fide* features of melanoma, often termed an atypical melanocytic proliferation, a less compelling LR+ of 2 might be sufficient to upgrade a diagnosis to melanoma, whereas a significantly lower LR- of 0.1–0.01 might be required to decrease excision margins or avoid excision altogether. These analyses demonstrate that there could be great clinical utility of biomarkers examining the malignant potential of intermediate melanocytic lesions, particularly when the biomarker holds a significant likelihood ratio that modifies the prior odds by an order of magnitude or more.

Numerous biomarker tests have been developed to address this issue. **Table 2** reports biomarker tests that are currently commercially available or used in selected referral centers and **Table 3** (discussed below) reports additional candidate biomarker tests reported in the literature. Immunohistochemistry (IHC) biomarkers are usually the most easily adopted test with regards to tissue requirement, cost, and required laboratory equipment and expertise. One promising IHC marker is PRAME (Preferentially expressed Antigen in Melanoma), which has demonstrated an LR+ of 29–62.5 and an LR- of 0.13–0.253 when assessed with an independent validation cohort (39, 40). Considering that it shares 90% concordance with fluorescent *in situ* hybridization (FISH) and 92.7% agreement with the final diagnostic interpretation (40), adopting PRAME IHC could be an economical option for most laboratories to use for routine differentiation of nevi and melanoma. Comparatively, p16 IHC, which is used at some institutions to help distinguish between nevi and melanoma in challenging cases, has no published validation cohorts to identify the clinical characteristics of this test. As a result, the accuracy, reliability, and versatility of p16 is unclear. In addition, the utility of p16 expression depends on the type of lesion since it may help distinguish desmoplastic Spitz nevi from desmoplastic melanoma (37), but fails at distinguishing Spitz nevi from Spitzoid melanoma (38).

FISH is another tissue staining modality that has been used for over a decade to help distinguish nevi from melanoma. It has well-developed scoring criteria, several validation studies, has undergone quality improvement to reduce the number of erroneous Spitzoid classifications, and has been directly compared to both the myPath gene expression panel (GEP) and copy number variation (CNV) assay by comparative genome hybridization (CGH) (42, 43, 49–51). The clinical characteristics of FISH in melanoma diagnosis are highly competitive and cost-effective when compared to the other molecular techniques (**Table 2**) (43, 49–51). However, evaluating FISH staining requires expert interpretation and expensive microscopy equipment, which limits the adoption of the technique outside of reference laboratories.

The CNV by CGH assay has been used clinically for almost two decades and assesses large genomic mutations and deletions across the genome. Melanomas usually possess multiple chromosomal losses and gains (96.2% frequency), Spitzoid nevi frequently have a unique 11p amplification, and non-Spitzoid nevi usually have no detectable chromosomal changes (41).

TABLE 2 | Melanoma biomarkers used by reference centers and dermatopathologists.

Category	Method class	Method specifics	Total study number	Study breakdown	Estimated tissue requirement	Study objective	Number of pathologists confirming cohort	AUC	S/Sp/PPV/NPV, LR+/LR-‡	Independent cohorts assessed**	Qualitative vs. quantitative	Citation
Individual biomarker assays												
Protein expression	IHC	Loss of p16 protein expression	26	11 desmoplastic melanomas, 15 Spitz nevi	Sectioned FFPE slide	Differentiate desmoplastic Spitz nevus and desmoplastic melanoma	NR	NR	82*/100*/NR/NR, >80/0.18	No	Qualitative, requires expert interpretation	(37)
			37	19 Spitzoid melanomas, 18 Spitz nevi	Sectioned FFPE slide	Does not differentiate Spitzoid melanomas from Spitz nevi	Initial diagnosis plus consensus conference with 3+ dermatopathologists	NR	21*/83*/NR/NR, 1.2/0.95			(38)
		PRAME protein expression (4+ staining)	400	255 primary and metastatic melanomas, 145 melanocytic nevi	Sectioned FFPE slide	Support diagnosis of melanoma by PRAME positivity	2+	NR	87*/97*/NR/NR, 29/0.13	Yes	Qualitative, requires expert interpretation	(39)
			110	110 ambiguous cases reviewed by 2+ dermato-pathologists	Sectioned FFPE slide	Validate and compare PRAME IHC to FISH for differentiating ambiguous cases	2+	NR	75.0/98.8/NR/NR, >62.5, 0.253			(40)
Copy Number Variation	CGH	Copy number variation by comparative genomic hybridization	186	132 melanomas, 54 benign nevi (27 Spitz nevi, 19 blue nevi, 7 congenital nevi)	High tissue requirement; remainder of FFPE block	Differentiate melanoma from benign nevi on basis of copy number variation	NR	NR	96*/87*/NR/NR, 7.4/0.046	No	Qualitative, requires expert interpretation	(41)
FISH	FISH	Four probes targeting chromosome 6p23, 6p25, 11q13, and centromere 6	22	12 ambiguous lesions, 10 unequivocal lesions	Medium to High tissue requirement	Validate FISH on histologically ambiguous lesions based on clinical behavior of lesion	Initial diagnosis plus one dermatopathologist review, ambiguous tumors reviewed by two pathologists	NR	60/50/NR/NR, 1.2/0.80	Yes	Qualitative, requires expert interpretation	(42)
		Four probes targeting chromosome 9p21, 6p25, 11q13, and 8q24	424	Training set: 152 melanoma and 170 nevi. Validation set: 51 melanoma and 51 nevi.	Medium to High tissue requirement	Distinguish between melanoma and nevi on basis of chromosomal changes	1	0.94+	94/98/NR/NR, 47/0.061	Yes		(43)
Gene Expression with Myriad myPath	23 GEP qRT-PCR	Weighted 23 gene expression algorithm by qRT-PCR	901	Training set: 254 melanoma and 210 nevi. Validation set: 437 independent lesions	Medium tissue requirement, macro-dissection from sections	Differentiate benign nevi from malignant melanoma	Initial diagnosis, one study dermatopathologist review, if discordance a third dermatopathologist	0.96	90/91/NR/NR, 10/0.11	Yes	Quantitative algorithm, Qualitative output	(44)

(Continued)

TABLE 2 | Continued

Category	Method class	Method specifics	Total study number	Study breakdown	Estimated tissue requirement	Study objective	Number of pathologists confirming cohort	AUC	S/Sp/PPV/NPV, LR+/LR–[‡]	Independent cohorts assessed**	Qualitative vs. quantitative	Citation
			1,400	204 melanoma, 656 nevi		Prospective validation of differentiating between benign and malignant melanocytic lesions	Concordance between 3 experienced dermatopathologists	NR	91.5/92/NR/NR, 11.4/0.0924			(45)
			182	99 melanomas, 83 nevi		Correlate long-term clinical outcomes with gene signature	3	NR	93.8/96.2/NR/NR, 24.7/0.0644			(46)
			50	23 melanomas, 27 nevi		Categorize potential desmoplastic melanomas as likely malignant or likely benign	2 dermatopathologists, independent review	NR	Sensitivity is "about 80%, better than FISH"			(47)
			181	125 diagnostically uncertain lesions, 56 diagnostically certain lesions		Test accuracy of GEP to diagnosis and outcomes (gold standard) in cases with uncertain histo-pathological diagnoses	3 dermatopathologists	NR	90.4/95.5/NR/NR, 20.1/0.101			(48)

(Continued)

TABLE 2 | Continued

Compared techniques	Total study number	Study breakdown	"Gold Standard"	Number of pathologists confirming cohort	Unequivocal histology: S/Sp/PPV/NPV, LR+/LR−‡	Unequivocal Concordance	Challenging histology: S/Sp/PPV/NPV, LR+/LR−‡	Challenging Concordance	Citation
Comparison of biomarker assays									
CGH vs. FISH	30	25 melanomas, 5 nevi	Histology	Confirmation of initial diagnosis by 1 dermatopathologist	–	FISH vs. CGH: 90%	–	–	(49)
Histology vs. FISH vs. myPath	117	Histologically unequivocal: 15 malignant and 24 benign. 78 Histologically challenging cases	Histology	1 pathologist for unequivocal and 3 for challenging cases	FISH: 93/100/NR/NR, >90/0.07 myPath: 62/95/NR/NR, 12.4/0.40	FISH vs. histology: 97%, myPath vs. histology: 83% FISH vs. myPath: 80%	FISH: 56/83/NR/NR, 3.3/0.53 myPath: 52/80/NR/NR, 2.6/0.60	FISH vs. histology: 70%, myPath vs. histology: 64% FISH vs. myPath: 70%	(50)
Histology vs. myPath vs. FISH vs. (discordant cases only) CGH	268	Histologically unequivocal: 198. Histologically challenging: 70	Histology and SNP-array	Challenging cases reviewed by 1 study and 2 independent dermatopathologists	myPath: 67/81/NR/NR, 3.5/0.41	myPath vs. histology: 75%	FISH: 61/100/NR/NR, >60/0.41 myPath: 50/93/NR/NR, 7.1/0.53	FISH vs. histology: 84%, myPATH vs. histology: 74% FISH vs. myPath: 69% Agreement with nine pathologists in discordant cases: CGH vs. histology: 71% FISH vs. histology: 54% MyPath vs. histology: 14%	(51)

AUC, area under (receiver operating characteristic) curve; CGH, comparative genomic hybridization; FFPE, formalin-fixed paraffin embedded; FISH, Fluorescence in situ hybridization; IHC, immunohistochemistry; LR+, positive likelihood radio; LR−, negative likelihood radio; NPV, negative predictive value; NR, not reported; PPV, positive predictive value; qRT-PCR, quantitative real time polymerase chain reaction; S, sensitivity; Sp, specificity.

*Calculated from manuscript data, no validation cohort.

**Tests are considered to have independent cohort validation if an independent clinical cohort was included in the original study or at least one follow-up study used an independent cohort with the same diagnoses or a cohort with ambiguous diagnoses.

‡All LR values are calculated based on reported sensitivity and specificity.

While a powerful technique to distinguish nevi from melanomas, there is no clear CNV-based definition of melanoma. As a result, the CNV cut-off that separates a benign vs. malignant lesion is often a gestalt decision from the dermatopathologists, which makes quantifying the clinical accuracy characteristics challenging. However, the price of CGH has fallen in recent years, the technology can be readily adopted by laboratories with experience in molecular diagnostics, and, it has 90% concordance with FISH in cases with unequivocal histopathology (49). In particularly challenging cases, CGH also outperformed both FISH and myPath in reaching concordance with a consensus histopathological diagnosis (51).

More recently, the Myriad myPath quantitative gene expression panel (GEP) and scoring algorithm was reported to have favorable test characteristics (AUC of 0.96, sensitivity of 90% and specificity of 91%) (44). This platform has subsequently been studied across prospectively submitted samples (45) in the evaluation of desmoplastic melanocytic proliferations (47) and in conjunction with long-term clinical outcomes (46). The most recent study is well-designed to assess function in histopathologically uncertain lesions and exemplifies how to use clinical outcome data as the gold standard instead of histopathology (48). Despite its impressive clinical characteristics, when assessed individually, the myPath methodology is not as competitive in comparison studies to FISH and CGH when histopathology is the gold standard (50, 51). Additional studies of the myPath test compared directly against FISH and CGH with long term follow up of clinical outcomes as the gold standard are warranted to determine the true relative utility of these tests.

In summary, the existing literature most highly supports the use of PRAME IHC, FISH, and CGH to aid dermatopathologists in routinely stratifying melanoma risk. While the clinical characteristics of CGH are uncertain, the presence of amplifications and deletions would likely prompt a conservative practitioner to upgrade the diagnosis to melanoma (presumed LR+ >2). FISH and PRAME IHC easily reach the modeled threshold of utility, even in challenging cases, but would not be sufficient to rule out melanoma (LR− is not <0.1–0.01) (50, 51, 63), and it is unclear if a CGH result revealing no amplifications or deletions would have a sufficient LR- to overturn a conservative melanoma diagnosis. The only available biomarker that could potentially have such a robust LR− is myPath, but such performance would need to be proven in a head-to-head comparison with the other testing modalities. As it stands, there is no currently used biomarker with a robust enough LR+ and LR− to be definitive for including or excluding a melanoma diagnosis with certainty, which leaves room for the further development of novel tests and more rigorous validation of current tests.

Potential Future Biomarkers for the Diagnosis of Pigmented Lesions

Since distinguishing early invasive melanoma from pre-malignant melanocytic lesions by routine histopathology remains a diagnostic challenge (15), numerous research groups have attempted to develop ancillary tests to augment melanoma diagnosis. Presented below and summarized in **Table 3** are a set of studies that examined promising biomarkers. All of these studies utilize methodologies obtainable by academic reference laboratories, all have at least one test characteristic defined, and most have been assessed in at least one independent test cohort.

One study used IHC to develop a semi-quantitative scoring system based on the percent expression of Ki-67 (MIB-1), p16, and HMB45 expression in melanocytes. Of these commonly available stains, Ki-67 has the best individual sensitivity and specificity, but the addition of p16 and HMB45 increased this to 97.4% sensitivity and 97.3% specificity with an AUC of 0.987. Notably, the IHC panel accurately predicted the metastatic evolution of tumors that were initially classified as melanocytic or Spitzoid nevi, and it appropriately downgraded one melanoma to a benign nevus (62). While these results are impressive, this test requires stain scoring which is notoriously subjective. Therefore, maintaining a high interobserver concordance with three IHC stains could be challenging, especially with scoring borderline p16 and HMB45 cases. Another proposed diagnostic marker is 5-hydroxymethylcytosine (5-hmC), which is decreased in melanoma compared to nevi, keratinocytes, and lymphocytes. Staining variation limits the sensitivity and specificity of 5-hmC scoring system to 92.7 and 97.7%, respectively, though there is a characteristic biphasic staining pattern between the melanoma and nevi components of a transformed nevus that could be clinically useful according to a pilot study (61).

Melanomas lose primary cilia after transformation whereas they are retained in benign nevi (64–66). The presence of primary cilia have been shown to differentiate typical and atypical Spitzoid nevi from malignant tumors with a sensitivity of 81% and specificity of 65% (AUC 0.84), but, with the incorporation of cytologic atypia, hyperchromatism, and asymmetry in making a diagnosis, the combined specificity and AUC were increased (sensitivity 72%, specificity 89%, AUC 0.92) (60). This is an improvement over standard diagnosis of Spitzoid tumors, which are classically difficult to distinguish between benign and malignant subtypes. Recent pilot studies have also suggested visualization of primary cilia using IHC may also provide similar utility in other melanoma types (67–69).

Discrimination between nevus and melanoma may also be accomplished using differences in gene expression. The most accessible method of analysis is to perform quantitative real time polymerase chain reaction (qRT-PCR) on a single transcript target from tissue or serum. For instance, silver homolog (also called *SILV*, *PMEL*, *gp100*, or *ME20M*) expression from FFPE tissue is significantly lower in melanomas compared to nevi, and it has favorable characteristics (AUC 0.94) as a single biomarker (57). Similarly, a qRT-PCR panel for 6 microRNAs (miRNA) in serum demonstrated positive and negative likelihood ratios of 20 and 0.07, respectively, in distinguishing uveal melanoma and nevi (59). Additional work remains to be completed to confirm clinical utility of these studies as neither contained validation cohorts and both had small sample sizes. The importance of a validation cohort is highlighted by our 2019 study where we developed a miRNA ratio trained model from miRNAs enriched in microdissected melanoma cells compared to nevus cells. While

TABLE 3 | Proposed melanoma diagnosis biomarkers and prospective uses.

Category	Method class	Method specifics	Total study number	Study breakdown	Estimated tissue requirement	Study objective	Number of pathologists confirming cohort	AUC	S/Sp/PPV/NPV (%), LR+/LR-‡	Qualitative vs. quantitative	Citation
DNA Methylation	40-CpG classifier	Next generation sequencing	162	89 melanoma, 73 nevi	250–300 ng of DNA from microdissected slides	Classify uncertain samples as melanoma or nevi	4	0.996	96.6/100/100/96.2, >90/0.03	Quantitative	(52)
	Promoter methylation	1,505 CpG site microarray, analyzed 12 CpG loci highly predictive of melanoma	49	22 melanoma, 27 nevi	250 ng of DNA from FFPE tissue	Differentiate nevi and malignant melanoma	NR	0.89–1.0	≤100/ ≤100/NR/NR	Quantitative	(53)
		CpG island hypermethylation in promoter of CLDN11	405	199 melanoma, 208 nevi	Not reported, extraction from tissue	Differentiate dysplastic nevi and malignant melanoma	NR	0.806	52/94/ /NR/NR, 8.67/0.51	Quantitative	(54)
		RASSF1A promoter methylation in serum cell-Free DNA	152	84 melanoma, 68 nevi	Serum draw	Differentiate healthy vs. in-situ, invasive, and metastatic melanoma	NR	0.905	NR	Quantitative	(55)
Gene expression	microRNA ratio	Sequencing or qRT-PCR to obtain microRNA Ratio Trained Model score	179	106 melanoma, 73 nevi	Single tissue section	Differentiate nevi and malignant melanoma	5+	0.9	81/88/ NR/NR, 6.75/0.22	Quantitative	(56)
	RT-PCR	qRT-PCR for SILV	193	47 melanoma, 48 nevi	3–12 tissue sections	Differentiate dysplastic nevi and malignant melanoma	NR	0.94	NR	Quantitative	(57)
	Microarray	DNA microarray assay using 14 genes	120	62 melanoma, 58 nevi	8 µM tissue section, microdissected with laser capture	Differentiate nevi and malignant melanoma	2–4	NR	90/86/NR/ NR*, 6.43/0.10	Quantitative	(58)
Serum nucleic acid	Circulating miRNA panel	qRT-PCR for six microRNAs in serum	65	Serum from 50 uveal melanoma, 5 metastatic uveal melanoma, 10 uveal nevi	Serum draw	Differentiate uveal melanoma from nevi	NR	NR for panel	93/100*/100/76, 20/0.07	Quantitative	(59)
Protein Expression	Ciliation index	Immunofluorescence for acetylated alpha-Tubulin	124	26 melanoma, 42 nevi	Single tissue section	Classification of Spitzoid tumors as malignant or benign	3+	0.84	81/65/NR/NR, 2.31/0.29	Semi-quantitative	(60)
	IHC	IHC for 5-hydroxymethyl-cystosine	190	126 melanoma, 45 nevi	Single tissue section	Differentiate nevi from melanoma	2	0.78	93/98/NR/NR, 46.5/0.07	Semi-quantitative	(61)
	Multi-IHC semi-quantitative scoring system	IHC scoring of Ki-67, p16, and HMB45	308	234 melanoma, 74 nevi	Three tissue sections	Differentiate nevi and malignant melanoma	NR	0.987	97.4/97.3/NR/NR, 36.1/0.03	Semi-quantitative	(62)

AUC, area under (receiver operating characteristic) curve; FFPE, formalin-fixed paraffin embedded; LR+, positive likelihood radio; LR−, negative likelihood radio; NPV, negative predictive value; NR, not reported; PPV, positive predictive value; S, sensitivity; Sp, specificity; qRT-PCR, quantitative real time polymerase chain reaction.
*Leave 1-out validation.
‡All LR values are calculated based on reported sensitivity and specificity.

the initial training model had an AUC of 0.98, the clinical accuracy of the qRT-PCR validation cohort containing whole FFPE sections dropped to a sensitivity of 81%, specificity of 88%, and AUC of 0.90 (56). Likewise, a microarray study on microdissected FFPE tissue found that the expression profiles of 36 genes could be used to identify melanoma with sensitivity (90%) and specificity (88%) in the validation cohort (58). While the gene expression panels described above have test characteristics that are theoretically compatible with a useful clinical test, their reliability is uncertain until larger studies confirm clinical utility.

Of the different classes of proposed biomarkers, DNA methylation testing is not widely available outside of reference laboratory testing for X chromosome disorders and specific genes, such as *MLH1* and *MGMT*. In melanoma, promoter methylation status of two genes have been examined as potential biomarkers. The Ras associated domain family 1 isoform A (*RASSF1A*) gene is a tumor suppressor found to be methylated in melanomas compared to healthy subjects (AUC of 0.905) by cell-free DNA (cfDNA) analysis (55). While cfDNA allows for a non-invasive blood draw and detects *in situ* melanomas with an AUC of 0.945, no studies have been published validating the *RASSF1A* marker. In comparison, the promoter for *CLDN11*, a component of the tight junction, undergoes CpG island methylation in approximately 50% of primary melanomas compared to 3–6% of dysplastic nevi. *CLDN11* promoter methylation status was determined by extraction from FFPE tissue and methylation-specific PCR in a large cohort of melanomas and nevi. This yielded a 52% sensitivity and 94% specificity (AUC of 0.806), suggesting utility for ruling in melanoma in the context of a suspicious dysplastic nevus. Another group developed a CpG site microarray panel and, ultimately, a 40-CpG next generation sequencing (NGS) panel to assess the methylation of status of multiple genes from microdissected FFPE slides. While both of these studies had characteristics compatible with a useful clinical test (sensitivity up to 96.6%, specificity up to 100%, AUC up to 0.996), they are limited by small cohort sizes and lack of long-term patient outcome data (52, 53).

Overall, there are several potential biomarkers that could be used to distinguish early melanoma from dysplastic nevi. All of the above tests have LRs (**Table 3**) that could be beneficial in distinguishing atypical melanocytic proliferations from melanoma based on our model (**Figure 4**). Of these biomarkers, we speculate that the most accessible, least complex, minimally invasive, and relatively affordable test that could be readily adopted into the laboratory is 5-hmC IHC (61). Prior to its adoption, however, large cohort studies are required to confirm the promising initial results. Of all of the tests included in **Tables 2**, **3**, the 40-CpG classifier next generation sequencing test for DNA methylation has the best overall combined characteristics (52). If future multi-institutional studies show good concordance and validation of the test characteristics, the 40-CpG classifier test could be useful for assessing the most challenging cases of atypical melanocytic proliferations. However, the major limitation of this test is the NGS approach: ample tumor tissue will be required, the method is expensive, and turn-around-time will likely be at least 8–14 days. Thus, other tests that require less tissue, cost, time, and expert interpretation, such as the multi-IHC panel or ciliation index, also warrant further consideration (62, 68).

Potential Biomarkers for Staging, Prognosis, and Treatment Effect Monitoring

In addition to indications for biopsy and diagnostic accuracy, potential biomarkers were assessed for other important clinical decision points related to the staging and treatment of invasive melanomas (**Figure 1C**, **Table 4**).

Prognostic Biomarkers

Prognostic biomarkers compose the broadest category of biomarkers. In terms of clinical decision impact, they can range from nebulous indicators of survival odds to potential heralds of disease recurrence. The principal issue with survival indicators is that they may require adoption by the AJCC or other analogous staging protocols prior to widespread utilization. Otherwise, the ordering clinician may be unsure how an unfavorable result will affect their patient's stage, choice of intervention, and ultimately outcome. A biomarker with excellent test characteristics may therefore be effectively useless if it is not incorporated into a staging framework from which clinical decisions can be made.

One commercially available prognostic test is the Castle DecisionDx assay. While the DecisionDx assay is not included in AJCC staging nor NCCN treatment guidelines, the test aims to identify patients with stage I and II disease who may be at increased risk of metastasis and death. Analytically, this gene expression profile (GEP) has demonstrated reproducibility across instruments in separating "class 1" (low-risk) from "class 2" (high-risk) probability scores and in technical concordance between runs (82). Results from small retrospective and prospective clinical cohorts have suggested improved ability to stratify risk of relapse and distant metastasis independent of SLNB status (83–87). However, each of these studies are limited by the lack of stage or substage breakdown of test performance, rendering it difficult to determine how much additional information is imparted by the biomarker beyond existing staging criteria. Podlipnik et al. studied the additive utility of DecisionDx probability scores in conjunction with combined AJCC stages ("low-risk": IB-IIA, "high-risk": IIB-IIC) and found disease free survival was lower in melanoma patients with DecisionDx high-risk "class 2" scores, regardless of grouped AJCC staging (87). Furthermore, in a 205 patient cohort, comprised of combined Stage I and II melanoma cases, where the DecisionDx GEP was combined with the AJCC Individualized Melanoma Patient Outcome Prediction Tool, the authors reported improved sensitivity but worse specificity for predicting disease recurrence, distant metastasis, and death compared to either metric alone (78). However, because neither positive predictive value/negative predictive values (PPV/NPV), nor AJCC substage breakdown of DecisionDx test performance were reported, it is challenging to determine the utility of this test beyond AJCC staging criteria, and it remains to be determined if and how this test can be integrated into current standard of care.

TABLE 4 | Proposed prognostic, staging, and treatment-monitoring biomarkers.

Category	Method class	Method specifics	Total study number	Study breakdown	Estimated tissue requirement	Study objective	Number of pathologists confirming cohort	AUC	S/Sp/PPV/NPV (%), LR+/LR-‡	Qualitative vs. quantitative	Citation
Prediction of sentinel lymph node biopsy status											
Castle DecisionDx	31 GEP	TaqMan qRT-PCR from FFPE tissue	1421	By T: 613 T1, 452 T2, 240 T3, 116 T4	macrodissected FFPE sections	Identify T1–T2 melanomas at low risk for positive SLNB	NR	NR	NR	Quantitative	(70)
Skyline	GEP and clinico-pathologic features	TaqMan qRT-PCR from FFPE tissue	506	160 discovery, 360 development, 146 validation	FFPE blocks	Develop and validate a GEP panel to identify positive SLNB	2+ dermato-pathologists	0.78	89/76/NR/NR, 3.71/0.14	Quantitative GEP, qualitative clinicopathologic features	(71)
			754	128 positive SLNB, 626 negative SLNB	FFPE blocks	Identify patients who can forgo SLNB	Initial diagnosis made with 2+ dermato-pathologists. No other review	0.82	TIb: 41/82/7/98, 2.28/0.72; TIIa: 80/53/21/95, 1.7/0.24; TIIb: 94/27/21/96, 1.29/0.22; TIIIa: 99/12/33/98, 1.13/0.08; TIIIb: 100/7/45/100, 1.08/<0.01		(72)
Metastasis identification											
Circulating Tumor Cells	Flow and fluorescent microscopy	CD146 cell sorting. DAPI+, HMW-MAA+, CD45/CD34– Immuno-fluorescence	15	Tumor: 15, 0. Metastasis 6, 9	Whole blood	Identification of early mets	NR	NR	CTC: 33/100/100/69, >30/0.67; CTC+5-S-cysteinyldopa: 67/100/100/82, >60/0.33	Semi-quantitative	(73)
Serum nucleic acid	Circulating non-coding RNA by qRT-PCR	hsa-miR-1246 miScript assay for RNU2-1f	122	33 training, 16 distant metastasis validation, 16 nevi, 57 healthy control	Serum draw	Use RNU2-1f non-coding RNA serum level to diagnose melanoma mets	NR	0.9375 (validation)	Regional mets: 69.6/87.2/NR/NR Distant mets: 70.0/87.2/NR/NR	Quantitative	(74)
Protein Expression	ELISA	ELISA for S100 and OPN	106	6 squamous cell carcinomas, 76 melanomas, 24 metastatic mela-nomas, 3 normal	Plasma draw	Distinguish metastatic melanoma from local disease with OPN +/– S100 and LDH	NR	OPN and S100: 0.97	95.5/85.9/NR/NR, 6.77/0.05	Quantitative	(75)
Prognosis											
Castle DecisionDx	31 GEP	TaqMan qRT-PCR from FFPE tissue	479	107 discovery, 268 development, 104 validation	macrodissected FFPE sections	Identification of stage I/II tumors with worse prognosis	NR	0.93	100/78/NR/NR, 4.54/<0.01	Quantitative	(76)
			217	All post-SLNB: 58 positive SLN, 159 negative SLN	macrodissected FFPE sections	Prognostic accuracy of GEP and SLNB in predicting RFS, distant mets, and OS	NR	NR	N.R./N.R./50/82		(77)

(Continued)

TABLE 4 | Continued

Category	Method class	Method specifics	Total study number	Study breakdown	Estimated tissue requirement	Study objective	Number of pathologists confirming cohort	AUC	S/Sp/PPV/NPV (%), LR+/LR-‡	Qualitative vs. quantitative	Citation
			205	By stage: 2 I, 68 IA, 39 IB, 1 II, 40 IIA, 41 IIB, 14 IIC	Biopsy	Compare GEP to AJCC in predicting 5 yr RFS, distant mets, and OS	NR	NR	GEP+AJCC for Recurrence: 90/71/NR/NR, 3.10/0.14; Distant Mets: 88/63/NR/NR., 2.38/0.19; Death: 82/62/NR/NR, 2.16/0.29		(78)
Gene Expression	RT-PCR	qRT-PCR for PAX3d to monitor recurrent in stage II-IV disease	198	111 melanoma, 87 healthy controls	Plasma draw	Identify recurrence in stage II-IV disease with plasma PAX3d mRNA	NR	0.823	Stage II–III relapse: 67/75/67/75, 2.68/0.44; Stage IV relapse: 75/93/43/98, 10.7/0.27	Quantitative	(79)
Serum nucleic acid	Circulating tumor DNA	PCR of BRAF and NRAS mutants	125	20 with progression, 9 with pseudo-progression	Plasma draw	Differentiate radiologic pseudo-progression and true progression	NR	NR	90/100/100/82, >90/0.1	Quantitative	(80)
Monitoring of treatment effect											
Serum nucleic acid	Circulating tumor DNA	BRAF V600mut ctDNA by qRT-PCR	36	16 before therapy, 20 after therapy	Plasma draw	Monitor patient response to BRAF/MEK inhibition therapy with BRAF V600mut ctDNA	NR	NR	70/100/NR/NR., >70/0.30	Quantitative	(81)

AUC, area under (receiver operating characteristic) curve; CTC, circulating tumor cells; ctDNA, circulating tumor DNA; ELISA, enzyme-linked immunosorbent assay. FFPE, formalin-fixed paraffin embedded; GEP, gene expression panel; LR+, positive likelihood radio; LR–: negative likelihood radio; Mets, metastasis; NPV, negative predictive value; NR, not reported; OS, overall survival; PPV, positive predictive value; RFS, recurrence-free survival; S, sensitivity; qRT-PCR, quantitative real time polymerase chain reaction; SLNB, sentinel lymph node biopsy; Sp, specificity.
‡All LR values are calculated based on reported sensitivity and specificity.

The AJCC melanoma staging and substaging criteria are built upon robust collection of patient clinicopathologic features and have been refined over decades using long-term outcomes data for more than 45,000 patients with melanoma (1). As such, this is the standard to which prognostic biomarkers must be measured, and full transparency of clinical characteristics and AJCC substages of patients in future studies will be necessary to meaningfully compare prognostic outcomes to AJCC staging (88, 89). Because of these limitations, a recent consensus group statement from the melanoma prevention working group members recommended against routine use of prognostic GEPs, including the DecisionDx, outside of clinical trials, without thorough discussion of testing limitations with the patient, or outside of multidisciplinary conference recommendations (5).

Stratifying Risk of Lymph Node Metastasis

A major clinical decision point is to perform a sentinel lymph node biopsy (SLNB) for patients with T1b tumors or in AJCC stages IB and II, which will upstage the patient to stage III if positive (**Table 1**). At present, the benefit of SLNB is in determining who is eligible for adjuvant systemic therapy (90). Examples of adjuvant therapy include the combination of dabrafenib plus trametinib in resected Stage III BRAFV600E/K mutant melanoma, which improves 5 year overall survival (91), and pembrolizumab which improves relapse free survival in high risk, resected, Stage III melanoma patients (92). Since current NCCN guidelines recommend SLNB if the odds of a positive node are >5% (90), the positive rate of SLNB is low (93). In theory, there may be significant clinical utility for risk-stratifying biomarkers to determine which patients may benefit from this procedure (**Figure 5**). Because of the large degree of uncertainty as to which patients will benefit from SLNB, minimal changes in LR of LN metastasis may be sufficient to alter clinical decision making in T1b melanomas (**Figures 5A,D**). Further, the risk to the patient of an undetected LN metastasis (a missed opportunity to treat Stage III disease before progression) must be incorporated into the decision. Indeed, improved efficacy of adjuvant systemic therapy alters the risk-benefit analysis of SLNB, potentially shifting the decision to favor surgery in some patients.

In theory, a prognostic assay with high LR- or NPV could rule out SNLB eligible patients who would not benefit from the procedure. Conversely, a test with high LR+ or PPV could identify patients who may benefit from SLNB but otherwise would not have been offered (i.e., melanoma *in situ*; **Figure 5B**, or T1a melanoma; **Figure 5C**). In reality, NPV and PPV are often inversely related, and as such, the goals of the test must be clearly defined. For a molecular assay to replace the gold standard SLN procedure in patients without clinically positive nodes (94), which at present is the only determinant of patient eligibility for systemic therapy (90), the assay must have an NPV higher than the SLN procedure [NPV of SLNB reported up to 94% (95)]. While the value of the assay remains uncertain, SkylineDx has conducted validation studies for this indication. In their initial study, SkylineDx combined established prognostic factors, such as patient age and Breslow depth, with their 11 GEP to predict SLN metastasis

(71), thereby convoluting the contribution of the GEP to the classification. Concerningly, the initial gene set was defined using a cohort of only seven patients—three with and four without regional metastasis (72)—calling into question the likelihood the GEP will perform well across a greater diversity of patients.

A recent consensus statement expressed concern about the lack of test performance metrics in SkylineDx's studies (5). These values have since been released (72), demonstrating a specificity of 53% and LR- of 0.24 for TIIa disease with otherwise unremarkable test characteristics (**Table 4**). Based on these results and our model (**Figure 5**), there is a narrow window of utility for determining which TIIa tumors would benefit from a SLNB and a larger range of utility in determining which patients could safely forego SLNB using this test. However, the combined low number of positive SLNB (128 of 754 biopsies) and small discovery cohort (*n* = 7) are concerning given the heterogeneity of melanoma. Extensive validation on diverse cohorts that report a NPV higher than SLNB are required to determine whether the SkylineDx GEP is capable of answering the clinical questions for which these tests are ordered.

Similarly, a 2019 Castle DecisionDx study claimed the 31 GEP assay distinguished individuals >55 years of age with T1-T2 disease who have a <5% chance of SLN involvement. However, no measures of accuracy or predictive value were presented, making it challenging to assess the clinical utility of the test for this application (70). Of greater concern, the Castle DecisionDx test was developed to assess for metastatic risk associated with a primary melanoma, not to predict SLN involvement. Although the repurposing of a clinical test developed as an indicator of one condition to another is possible, successful performance would be fortuitous, especially since previous validation studies assessing metastasis risk groups are irrelevant when considering this test's utility for assessing the value of SLNB. Thus, while the Castle DecisionDx test may eventually improve the NPV of SLNB in prognosticating relapse or distant metastasis (83–85), without further study this test should not be considered to supplant SLNB, which remains the most important prognostic staging element guiding further treatment (90).

To determine the impact of these GEPs, and other tests, on decision to biopsy SLNs and their role in staging (see below), a large retrospective study and/or longitudinal studies will be necessary to determine how these biomarkers predict clinically meaningful outcomes like distant metastasis free-survival and all-cause mortality. Meta-analysis of existing studies has demonstrated that currently available GEP tests have limited ability to prognosticate recurrence in early stage melanomas (96). Due to the cost of conducting such trials, in terms of economic considerations and irreplaceable patient samples, the melanoma community will need to identify and prioritize the coordinated validation of only the most promising biomarker assays. Considerations for test selection need to include not only test characteristics, but also the details by which those characteristics were generated (size and diversity of cohorts, independence of studies, selection bias) and the likelihood to change clinical decision making when considering current guidelines.

FIGURE 5 | Utility of biomarkers to augment Breslow depth in determining which patients are offered sentinel lymph node biopsy (SLNB). **(A)** As depth of melanoma increases, from *in situ* to greater than one millimeter, the probability of sentinel lymph node positivity, and thereby stage and treatment selection, increases. Sentinel lymph node biopsy (SLNB) is positive in ∼5% of patients with T1b melanoma, resulting in a decision threshold posterior odds of ∼1:20 for many surgical oncologists to offer SLNB and reflecting NCCN guideline recommendations (90). As SLNB is not routinely offered for melanoma *in situ* or T1a melanomas, the incidence of SLNB positivity is unknown. We therefore estimate that these are substantially and progressively less positive as tumor thickness decreases. SLNB is standard of care for T2+ melanomas (gray dot) and thus no biomarker is likely to alter this recommendation without exceptionally low LR- test characteristics. **(B)** Currently patients with melanoma *in situ* are not generally offered SLNB since the likelihood of positivity (prior odds) is very low. A very high LR+, estimated at 25, would be necessary to recommend SLNB. **(C)** As the melanoma Breslow depth increases, the prior odds of sentinel lymph node metastasis increases (becomes closer to 1:1). Therefore, a lower LR+, such as LR+ ≥10, would be necessary to prompt SLNB. **(D)** Biomarkers may have the greatest utility for T1b malignant melanomas. Current NCCN guidelines recommend thoughtful consideration of SLNB based on patient-specific factors, and any test that could offer relatively small changes in certainty could be of great clinical utility. Thus, a modest LR+ of 2.0 or LR- of 0.5 may be sufficient for a clinically-actionable biomarker.

Monitoring for Disease Recurrence

Monitoring for disease recurrence via serial imaging is costly and, depending on modality, can lead to significant cumulative radiation exposure in patients. Use of blood- or serum-based biomarker assays (termed liquid biopsies) may be able to identify residual or recurrent disease prior to demonstrable evidence of such disease with imaging studies (97). Prediction or early identification of distant metastasis would enable clinicians to begin adjuvant systemic therapy prior to confirmation of macroscopic metastatic disease by radiology.

Current AJCC staging criteria and NCCN guidelines incorporate one serum-based prognostic biomarker in the treatment of melanoma: lactate dehydrogenase (LDH), a proxy for tumor lysis and disease burden (1, 90). Several studies have demonstrated that a baseline LDH level >2 times above the laboratory-established upper reference range is associated with poor survival outcomes in patients with stage IV melanoma, despite the advent of targeted immunological agents (98–101). However, due to lack of sensitivity and biological specificity for melanoma, LDH has been relegated to advanced metastatic disease staging and treatment effect monitoring (1, 90). Other

studies have proposed biomarkers with test characteristics superior to LDH. These include U2 small nuclear RNA, which is reported to correlate with increased risk of regional (AUC 0.80) and distant (AUC 0.94) metastasis (74), and serum *PAX3d* mRNA which has potential to monitor recurrence in stage IV disease, with specificity up to 93% and NPV up to 98% (79). However, further prospective testing is necessary to determine if and how these tests could be incorporated into current treatment recommendations for the prediction of early metastasis.

In a proof-of-concept study, circulating tumor cells (CTC) quantification from peripheral blood was a specific indicator of peripheral metastasis with a PPV of 100% (73). Besides the small cohort size ($n = 15$), one major limitation of this study was the unusual combination of immunomagnetic cell sorting and four-color immunofluorescence microscopy. A second challenge for CTC approaches in general is the propensity for melanoma cells to de-differentiate and thereby down-regulate lineages-specific cell markers used for selecting CTCs (S100, MelanA, SOX10). If CTC quantification could be adapted to a more widely available and objective platform, such as flow cytometry, and invariant melanoma markers identified, then it might find utility as a routine liquid biopsy. Alternatively, plasma concentrations of proteins other than LDH, such as osteopontin and S100 (75), may have functionality in predicting early metastasis, but have yet to be validated.

Identifying and validating biomarkers for minimal residual disease burden (MRD) have the potential to reduce imaging frequency, assess post-treatment disease burden, and identify microscopic disease relapse that is undetectable with standard imaging. One promising assay type is the quantification of common melanoma mutations as circulating tumor DNA (ctDNA) from peripheral blood. While there are ctDNA assays approved to identify actionable mutations in a variety of solid tumor types, these assays usually have a limit of detection (LOD) between 0.1 and 1.0% of ctDNA, which corresponds to 1 tumor sequence per 100–1,000 normal sequences (102). Current molecular techniques including digital droplet PCR, allele-specific oligonucleotide primer PCR, and NGS have the ability to further increase sensitivity by detecting known mutant alleles, resulting in an LOD between 0.001 and 0.1% of sequences for hematologic malignancies such as multiple myeloma (103). Detecting solid tumor ctDNA in peripheral blood has been more challenging due to the lower estimated amount of ctDNA present in non-metastatic cancer patients (102), but recent advances using techniques such as targeted digital sequencing (TARDIS) have improved the LOD of tumor-specific sequences to 0.01% of circulating DNA (104), likely representing the absolute minimum sensitivity required to detect a 5 mm primary tumor in breast cancer (102). Since the amount of ctDNA in patients with melanoma is on average 2 orders of magnitude less compared to breast cancer (105, 106), it is expected that a more sensitive assay with a lower LOD will be required to assess MRD in melanoma.

Several pilot studies have assessed the use of ctDNA as a MRD marker to monitor disease progression and relapse, with a few achieving a competitive LOD. One example demonstrated that monitoring *BRAF* V600 mutant sequences could identify disease progression prior to radiographic evidence in 44% of cases and was co-identified with imaging in an additional 26% (sensitivity 70%, specificity 100%, LOD 0.01%) (81). Similarly, a second study examining the most common *NRAS* and *BRAF* mutants found that detectable levels of ctDNA did not occur in cases with radiologic pseudoprogression (sensitivity 90%, specificity 100%, PPV 100%, NPV 82%, LOD <10 copies/ml or ∼1%) (80). The only study with a prospective validation cohort ($n = 36$) found that increasing mutant *TERT/NRAS/BRAF* ctDNA preceded radiologic evidence of progression by an average of 3.5 months (AUC of 0.85–0.87, LOD: $BRAF^{V600E}$ and $NRAS^{Q61}$ 0.01%, $TERT^{C250T}$ 1 copy/reaction and TER^{C228T} 10 copies/reaction) and were superior to LDH and S100 serology for this indication (107). Additional studies are needed to confirm these findings in a larger cohort, to determine the impact of observed mutational heterogeneity on test performance (107), and to find surrogates for melanomas that do not bear a *TERT*, *NRAS*, or *BRAF* mutation.

The Wild West of Predictive Biomarkers for Systemic Therapy

For Stage III melanomas that have metastasized to sentinel lymph nodes and for those in which nodal or in transit metastatic disease has been surgically resected, adjuvant therapy using anti PD-1 immunotherapy or BRAF/MEK inhibition is currently recommended for consideration based on recent clinical trial data showing improved relapse-free survival (RFS) (91, 108). There is significant interest in using adjuvant therapy in Stage II high-risk melanomas, as evidenced by a large, double-blind, prospective, randomized clinical trial currently evaluating the safety and efficacy of immune checkpoint blockade in resected Stage II melanoma (KEYNOTE-716: https://clinicaltrials.gov/ct2/show/NCT03553836). Recently an 11-gene GEP was published that demonstrated the ability to stratify patients with stage II melanoma into high GEP score groups and low GEP score groups, with the latter having significantly higher melanoma specific survival and relapse free survival rates (109). The authors of this study speculate that patients with higher GEP scores may be better candidates for adjuvant therapy due to the increased risk for relapse and death, potentially reflecting ambitions to test this hypothesis in a future study (109).

If an actionable *BRAF* V600 mutation is identified by sensitive and specific molecular testing (110), then BRAF plus MEK inhibitor therapy is an option for first-line therapy in resected Stage III and in Stage IV melanoma patients (90). Regardless of BRAF mutation status, immune checkpoint inhibitors (ICI) against PD-1 and CTLA-4 are frequently used due to improved progression-free and overall survival compared to both chemotherapy and combined BRAF/MEK inhibition (111–113). However, ∼65% of Stage IV metastatic melanomas do not significantly respond to ICI therapy (114–116), indicating an unfilled niche for future biomarkers to discriminate responders from non-responders prior to therapy initiation. Furthermore, accurately predicting which Stage III, Stage II, and even Stage I patients, who comprise the majority of cases of melanoma specific mortality (117), have the potential to benefit from

adjuvant therapy is a major goal of predictive biomarkers in melanoma.

With the development of new ICIs, kinase inhibitors, combination therapy regimens, and other novel treatment modalities, there will be an increasing need for biomarkers to match patients to optimal adjuvant therapies. Successful predictive biomarkers will have the potential to decrease patient morbidity by avoiding side-effects of less helpful treatments, boost response rates by identifying the most appropriate therapy upfront, and increase the cost-effectiveness of treatment by avoiding expensive ICIs when not indicated. When there are multiple treatment options considered first- or second-line treatments, there is even more value in the role of biomarkers guiding therapy selection, however the dichotomous framework that we have presented for other key decision-making points in the treatment of melanoma is less applicable. The complex decisions surrounding adjuvant therapy selection would necessitate multi-dimensional LR+/LR- analyses not attempted in this review due to exponentially increasing complexity. We await the results of ongoing studies and clinical trials in this field and refer the reader to several recent reviews that cover the impressive breadth of proposed biomarkers for use with ICI (118–121).

DISCUSSION

Many studies have addressed the potential utility of biomarkers in the diagnosis and treatment of melanoma, highlighting the drive to improve melanoma patient care through novel technologies. In this review, we have discussed several key clinical decision points in the diagnosis and treatment of melanoma including determining when to biopsy, aiding in the histopathologic diagnosis of melanoma, and stratification of prognosis of malignant melanoma. We have applied fundamental theory of clinical test utility to model these decision points and identify test characteristics required for a biomarker to influence decision making. This discussion is not intended to be an ultimate determination of the utility of any specific biomarker but rather to highlight how clinical-use-assumptions affect the characteristics required of a test to be potentially incorporated into practice. We hope this framework and analysis will assist basic science researchers in designing investigations with a greater likelihood of true translation and aid clinical researchers in identifying which candidate biomarker tests are worth allocating resources for validation.

Similar to other malignancies, melanoma is best treated when diagnosed early. When a melanoma is identified prior to invasion, treatment has high rates of cure with wide local excision alone (1, 122) but as invasion depth increases toward eventual metastasis, treatment becomes increasingly morbid and survival rate decreases. An important observation that arises from the modeling applied here is that across this spectrum, the characteristics of an ideal biomarker for melanoma change to reflect the changing clinical considerations surrounding treatment morbidity and efficacy. For early melanomas (T0-T1a), specificity (and thus LR−) is emphasized for diagnosis to avoid missing melanomas that are treatable with excision alone. At this stage of diagnosis, an intermediate lesion will default to treatment given the relatively low morbidity of excision, and as a result, sensitivity is generally less important. However, for primary malignant melanomas with a more significant invasive component (T1b-T4), the question of staging and assessing SLNB utility becomes important. Here the utility of a biomarker with high sensitivity (LR+), specificity (LR−) or both become quite useful. With Stage III+ melanomas demonstrating metastatic behavior, biomarkers with optimized specificity characteristics (LR+) suggesting response to specific treatment regimens will potentially be more useful given the higher morbidity and cost associated with adjuvant therapies.

Ideal biomarkers would prioritize neoplasia with true malignant potential for appropriate treatment, while providing confidence that intermediate neoplasia with reassuring biomarker findings can safely avoid further treatment. However, due to a lack of an ideal gold standard for verification, rigorous clinical validation of candidate tests by conventional standards is difficult for ambiguous lesions. Practical considerations of cost, ease of use, and tissue requirement will also factor into which biomarkers can be incorporated into clinical practice. Thus, the field should take great care when considering which biomarkers to prioritize for validation. While immunohistochemistry, copy number variation, and FISH represented the mainstay of biomarker use in this space for many years, more recently reported techniques including gene expression profiling, DNA methylation, and ciliation index have the potential to provide superior test characteristics and objective interpretation at lower cost and tissue requirement and thus represent exciting new potential avenues for further improvement in the diagnosis of melanoma (44, 52, 60).

The challenges of developing novel biomarkers for use in melanoma diagnosis are significant. One of the most substantial obstacles is the inherent subjectivity of histopathologic diagnosis of melanoma. When validation studies rely on histopathologic diagnosis, ambiguous lesions are often excluded, though these are the exact cases most in need of adjudication with an objective biomarker. Because of the imperfect nature of this gold standard, several studies have used consensus diagnosis from three or more experienced dermatopathologists (45, 46), which we would recommend of future studies. At a minimum, reporting the number of pathologists diagnosing each case, and whether or not they are experts in pigmented lesions, is helpful to determine the reliability of the study. Given the subjectivity of histopathologic diagnosis, we should strive to review the performance of biomarkers compared to long-term patient follow up data, despite the large sample sizes, cost, and time necessary to conduct such studies. Recurrence-free survival and melanoma-specific survival are perhaps the true gold standards, and gathering these data is essential to incorporate biomarkers into the standard of care and meaningfully improve patient outcomes.

Because of the high level of uncertainty surrounding the determination of which patients will benefit from SLNB, relatively small changes in certainty imparted by a biomarker may result in large changes in clinical management. The models applied here suggest that biomarkers with high LR+/sensitivity characteristics have the potential to confirm the utility of SLNB, while biomarkers with high LR−/specificity characteristics may prove useful in determining which patients can safely forego SLNB. Due to the high cost and morbidity associated with SLNB compared to wide local excision alone, and the potential for biomarkers to prove clinically useful with either high LR+, LR− or both, significant attention has rightfully been paid to biomarkers that address the question of which patients should undergo SLNB. Compared to diagnostic biomarkers, those targeting the necessity of SLNB require a relatively smaller change in certainty (smaller LR+ or LR−) to be clinically actionable. This, coupled with the costs of SLNB, financial and otherwise, suggest that biomarker tests in this space are worthy of the significant resources needed for further development.

Regardless of the purpose of a candidate biomarker, large prospective validation studies conducted on diverse and relevant patient populations that demonstrate clinically meaningful outcomes over time are needed to propel biomarkers into the mainstay of treatment. The time, large sample sizes, potential coordination across multiple trial sites, and significant expense of such studies present major barriers to generalizable validation studies. Because of these challenges, the vast majority of published biomarkers do not have multiple validation studies. Even those that do (such as Myriad myPath Melanoma and Castle Biosciences DecisionDx-Melanoma) (44, 76) have yet to be incorporated into NCCN guidelines due to lack of evidence in ambiguous melanocytic neoplasms and lack of evidence of actionable prognostic information across the melanoma spectrum (90). Due to these lengthy development and validation phases, combined with the pressing need for clinical tests, novel biomarkers are sometimes ordered by clinicians prior to robust establishment of clinical utility (35, 123). It is generally recommended that clinical tests should not be ordered that do not have the potential to change management or prognosis. Unnecessary testing can result in unanticipated results that complicate an otherwise straightforward diagnosis, cause undue physical and emotional hardship on the patient, lead to additional unwarranted testing, and increases costs borne by the patient and medical infrastructure (124–127).

CONCLUSIONS

The challenges of clinically validating a biomarker are significant. Given the substantial resources necessary to comprehensively validate a biomarker, clarity regarding the clinical question being addressed and the practical considerations that would confer clinical utility should be fundamental considerations prior to initiating a study. This review applies the fundamental theory of biomarker utility to estimate the necessary test characteristics required to influence clinical decision making in the diagnosis and treatment of melanoma, and summarizes the test characteristics and rigor of existing and candidate tests. It is clear that there is significant room for improvement in the validation of biomarkers in this space, and we hope that the models and considerations summarized here will help guide the community in study design and investment that will yield the maximum clinical utility.

AUTHOR CONTRIBUTIONS

DD and RJ-T conceptualized this work. DD, ES, and RJ-T performed review of the literature and wrote the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

ACKNOWLEDGMENTS

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. doi: 10.3322/caac.21409

2. Chapman PB, Hauschild A, Robert C, Haanen JB, Ascierto P, Larkin J, et al. Improved survival with vemurafenib in melanoma with BRAF V600E mutation. *N Engl J Med.* (2011) 364:2507–16. doi: 10.1056/NEJMoa1103782

3. Hauschild A, Grob JJ, Demidov LV, Jouary T, Gutzmer R, Millward M, et al. Dabrafenib in BRAF-mutated metastatic melanoma: a multicentre, open-label, phase 3 randomised controlled trial. *Lancet.* (2012) 380:358–65. doi: 10.1016/S0140-6736(12)60868-X

4. Menzer C, Menzies AM, Carlino MS, Reijers I, Groen EJ, Eigentler T, et al. Targeted therapy in advanced melanoma with rare BRAF mutations. *J Clin Oncol.* (2019) 3142–51. doi: 10.1200/JCO.19.00489

5. Grossman D, Okwundu N, Bartlett E, Marchetti M, Othus M, Coit D, et al. Prognostic gene expression profiling in cutaneous melanoma: identifying the knowledge gaps and assessing the clinical benefit. *JAMA Dermatol.* (2020) 156:1004–011. doi: 10.1001/jamadermatol.2020.1729

6. Kwak M, Farrow NE, Salama AKS, Mosca PJ, Hanks BA, Slingluff CL, et al. Updates in adjuvant systemic therapy for melanoma. *J Surg Oncol.* (2019) 119:222–31. doi: 10.1002/jso.25298

7. Karimkhani C, Green AC, Nijsten T, Weinstock MA, Dellavalle RP, Naghavi M, et al. The global burden of melanoma: results from the

Global Burden of Disease Study 2015. *Br J Dermatol.* (2017) 177:134–40. doi: 10.1111/bjd.15510

8. Ferlay J, Colombet M, Soerjomataram I, Mathers C, Parkin DM, Piñeros M, et al. Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods. *Int J Cancer.* (2019) 144:1941–53. doi: 10.1002/ijc.31937

9. USCSWG (United States Cancer Statistics Working Group). *U.S. Cancer Statistics Data Visualizations Tool, Based on 2019 Submission Data* (1999-2017). U.S. Department of Health and Human Services, Centers for Disease Control and Prevention and National Cancer Institute (2020). Available online at: www.cdc.gov/cancer/dataviz

10. Krensel M, Schäfer I, Augustin M. Cost-of-illness of melanoma in Europe – a systematic review of the published literature. *J Eur Acad Dermatology Venereol.* (2019) 33:504–10. doi: 10.1111/jdv.15315

11. Piepkorn MW, Barnhill RL, Elder DE, Knezevich SR, Carney PA, Reisch LM, et al. The MPATH-Dx reporting schema for melanocytic proliferations and melanoma. *J Am Acad Dermatol.* (2014) 70:131–41. doi: 10.1016/j.jaad.2013.07.027

12. Elmore JG, Elder DE, Barnhill RL, Knezevich SR, Longton GM, Titus LJ, et al. Concordance and reproducibility of melanoma staging according to the 7th vs 8th edition of the AJCC cancer staging manual. *JAMA Netw Open.* (2018) 1:e180083. doi: 10.1001/jamanetworkopen.2018.0083

13. Dinehart MS, Dinehart SM, Sukpraprut-Braaten S, High WA. Immunohistochemistry utilization in the diagnosis of melanoma. *J Cutan Pathol.* (2020) 47:446–50. doi: 10.1111/cup.13648

14. Gershenwald JE, Scolyer RA. Melanoma staging: American Joint Committee on Cancer (AJCC) 8th edition and beyond. *Ann Surg Oncol.* (2018) 25:2105–10. doi: 10.1245/s10434-018-6513-7

15. Elmore JG, Barnhill RL, Elder DE, Longton GM, Pepe MS, Reisch LM, et al. Pathologists' diagnosis of invasive melanoma and melanocytic proliferations: observer accuracy and reproducibility study. *BMJ.* (2017) 357:2813. doi: 10.1136/bmj.j2813

16. Carney PA, Reisch LM, Piepkorn MW, Barnhill RL, Elder DE, Knezevich S, et al. Achieving consensus for the histopathologic diagnosis of melanocytic lesions: use of the modified Delphi method. *J Cutan Pathol.* (2016) 43:830–7. doi: 10.1111/cup.12751

17. Lott JP, Elmore JG, Zhao GA, Knezevich SR, Frederick PD, Reisch LM, et al. Evaluation of the Melanocytic Pathology Assessment Tool and Hierarchy for Diagnosis (MPATH-Dx) classification scheme for diagnosis of cutaneous melanocytic neoplasms: results from the International Melanoma Pathology Study Group. *J Am Acad Dermatol.* (2016) 75:356–63. doi: 10.1016/j.jaad.2016.04.052

18. Elder DE, Piepkorn MW, Barnhill RL, Longton GM, Nelson HD, Knezevich SR, et al. Pathologist characteristics associated with accuracy and reproducibility of melanocytic skin lesion interpretation. *J Am Acad Dermatol.* (2018) 79:52–9.e5. doi: 10.1016/j.jaad.2018.02.070

19. Piepkorn MW, Longton GM, Reisch LM, Elder DE, Pepe MS, Kerr KF, et al. Assessment of second-opinion strategies for diagnoses of cutaneous melanocytic lesions. *JAMA Netw Open.* (2019) 2:e1912597. doi: 10.1001/jamanetworkopen.2019.12597

20. Geller BM, Frederick PD, Knezevich SR, Lott JP, Nelson HD, Titus LJ, et al. Pathologists' use of second opinions in interpretation of melanocytic cutaneous lesions: policies, practices, and perceptions. *Dermatol Surg.* (2018) 44:177–85. doi: 10.1097/DSS.0000000000001256

21. Titus LJ, Reisch LM, Tosteson ANA, Nelson HD, Frederick PD, Carney PA, et al. Malpractice concerns, defensive medicine, and the histopathology diagnosis of melanocytic skin lesions. *Am J Clin Pathol Oct.* (2018) 150:338–45. doi: 10.1093/ajcp/aqy057

22. Gardner LJ, Strunck JL, Wu YP, Grossman D. Current controversies in early-stage melanoma: questions on incidence, screening, and histologic regression. *J Am Acad Dermatol.* (2019) 80:1–12. doi: 10.1016/j.jaad.2018.03.053

23. Schmidt R. *Testing a Test: Beyond Sensitivity and Specificity.* (2015). Available online at: https://arup.utah.edu/education/testing_test.php/.

24. Bewick V, Cheek L, Ball J. Statistics review 13: receiver operating characteristics curves. *Crit Care.* (2004) 8:508–12. doi: 10.1186/cc3000

25. Phillips M, Greenhalgh J, Marsden H, Palamaras I. Detection of malignant melanoma using artificial intelligence: an observational study of diagnostic accuracy. *Dermatol Pract Concept.* (2019) 10:e2020011. doi: 10.5826/dpc.1001a11

26. Glazer AM, Rigel DS. Analysis of trends in geographic distribution of us dermatology workforce density. *JAMA Dermatol.* (2017) 153:472–3. doi: 10.1001/jamadermatol.2016.6032

27. Guitera P, Menzies SW, Longo C, Cesinaro AM, Scolyer RA, Pellacani G. *In vivo* confocal microscopy for diagnosis of melanoma and basal cell carcinoma using a two-step method: analysis of 710 consecutive clinically equivocal cases. *J Invest Dermatol.* (2012) 132:2386–94. doi: 10.1038/jid.2012.172

28. Pampena R, Borsari S, Lai M, Benati E, Longhitano S, Mirra M, et al. External validation and comparison of four confocal microscopic scores for melanoma diagnosis on a retrospective series of highly suspicious melanocytic lesions. *J Eur Acad Dermatol Venereol.* (2019) 33:1541–6. doi: 10.1111/jdv.15617

29. Kittler H, Pehamberger H, Wolff K, Binder M. Diagnostic accuracy of dermoscopy. *Lancet Oncol.* (2002) 3:159–65. doi: 10.1016/S1470-2045(02)00679-4

30. Binder M. Epiluminescence microscopy. A useful tool for the diagnosis of pigmented skin lesions for formally trained dermatologists. *Arch Dermatol.* (1995) 131:286–91. doi: 10.1001/archderm.1995.01690150050011

31. Nelson KC, Swetter SM, Saboda K, Chen SC, Curiel-Lewandrowski C. Evaluation of the number-needed-to-biopsy metric for the diagnosis of cutaneous melanoma: a systematic review and meta-analysis. *JAMA Dermatol.* (2019) 155:1167–74. doi: 10.1001/jamadermatol.2019.1514

32. Shah A, Hyngstrom J, Florell SR, Grossman D. Use of the pigmented lesion assay to rapidly screen a patient with numerous clinically atypical pigmented lesions. *JAAD Case Rep.* (2019) 5:1048–50. doi: 10.1016/j.jdcr.2019.10.004

33. Yao Z, Allen T, Oakley M, Samons C, Garrison D, Jansen B. Analytical characteristics of a noninvasive gene expression assay for pigmented skin lesions. *Assay Drug Dev Technol.* (2016) 14:355–63. doi: 10.1089/adt.2016.724

34. Gerami P, Yao Z, Polsky D, Jansen B, Busam K, Ho J, et al. Development and validation of a noninvasive 2-gene molecular assay for cutaneous melanoma. *J Am Acad Dermatol.* (2017) 76:114–120.e2. doi: 10.1016/j.jaad.2016.07.038

35. Varedi A, Gardner LJ, Kim CC, Chu EY, Ming ME, Leachman SA, et al. Use of new molecular tests for melanoma by pigmented-lesion experts. *J Am Acad Dermatol.* (2020) 82:245–7. doi: 10.1016/j.jaad.2019.08.022

36. Lee KC, Peacock S, Weinstock MA, Zhao GA, Knezevich SR, Elder DE, et al. Variation among pathologists' treatment suggestions for melanocytic lesions: a survey of pathologists. *J Am Acad Dermatol.* (2017) 76:121–8. doi: 10.1016/j.jaad.2016.07.029

37. Hilliard NJ, Krahl D, Sellheyer K. P16 Expression differentiates between desmoplastic Spitz nevus and desmoplastic melanoma. *J Cutan Pathol.* (2009) 36:753–9. doi: 10.1111/j.1600-0560.2008.01154.x

38. Mason A, Wititsuwannakul J, Klump VR, Lott J, Lazova R. Expression of p16 alone does not differentiate between Spitz nevi and Spitzoid melanoma. *J Cutan Pathol.* (2012) 39:1062–74. doi: 10.1111/cup.12014

39. Lezcano C, Jungbluth AA, Nehal KS, Hollmann TJ, Busam KJ. PRAME expression in melanocytic tumors. *Am J Surg Pathol.* (2018) 42:1456–65. doi: 10.1097/PAS.0000000000001134

40. Lezcano C, Jungbluth AA, Busam KJ. *Comparison of Immunohistochemistry for PRAME With Cytogenetic Test Results in the Evaluation of Challenging Melanocytic Tumors.* (2020). Available online at: www.ajsp.com

41. Bastian BC, Olshen AB, LeBoit PE, Pinkel D. Classifying melanocytic tumors based on DNA copy number changes. *Am J Pathol.* (2003) 163:1765–70. doi: 10.1016/S0002-9440(10)63536-5

42. Gaiser T, Kutzner H, Palmedo G, Siegelin MD, Wiesner T, Bruckner T, et al. Classifying ambiguous melanocytic lesions with FISH and correlation with clinical long-term follow up. *Mod Pathol.* (2010) 23:413–9. doi: 10.1038/modpathol.2009.177

43. Gerami P, Li G, Pouryazdanparast P, Blondin B, Beilfuss B, Slenk C, et al. A highly specific and discriminatory FISH assay for distinguishing between benign and malignant melanocytic neoplasms. *Am J Surg Pathol.* (2012) 36:808–17. doi: 10.1097/PAS.0b013e31824b1efd

44. Clarke LE, Warf BM, Flake DD, Hartman AR, Tahan S, Shea CR, et al. Clinical validation of a gene expression signature that differentiates benign nevi from malignant melanoma. *J Cutan Pathol.* (2015) 42:244–52. doi: 10.1111/cup.12475

45. Clarke LE, Flake DD, Busam K, Cockerell C, Helm K, McNiff J, et al. An independent validation of a gene expression signature to differentiate malignant melanoma from benign melanocytic nevi. *Cancer.* (2017) 123:617–28. doi: 10.1002/cncr.30385

46. Ko JS, Matharoo-Ball B, Billings SD, Thomson BJ, Tang JY, Sarin KY, et al. Diagnostic distinction of malignant melanoma and benign nevi by a gene expression signature and correlation to clinical outcomes. *Cancer Epidemiol Biomarkers Prev.* (2017) 26:1107–13. doi: 10.1158/1055-9965.EPI-16-0958

47. Clarke LE, Pimentel JD, Zalaznick H, Wang L, Busam KJ. Gene expression signature as an ancillary method in the diagnosis of desmoplastic melanoma. *Hum Pathol.* (2017) 70:113–20. doi: 10.1016/j.humpath.2017.10.005

48. Clarke LE, Mabey B, Flake DD, Meek S, Cassarino DS, Duncan LM, et al. Clinical validity of a gene expression signature in diagnostically uncertain neoplasms. *Per Med.* (2020) 17:361–71. doi: 10.2217/pme-2020-0048

49. Wang L, Rao M, Fang Y, Hameed M, Viale A, Busam K, et al. A genome-wide high-resolution array-cgh analysis of cutaneous melanoma and comparison of array-cgh to fish in diagnostic evaluation. *J Mol Diagn.* (2013) 15:581–91. doi: 10.1016/j.jmoldx.2013.04.001

50. Minca EC, Al-Rohil RN, Wang M, Harms PW, Ko JS, Collie AM, et al. Comparison between melanoma gene expression score and fluorescence *in situ* hybridization for the classification of melanocytic lesions. *Mod Pathol.* (2016) 29:832–43. doi: 10.1038/modpathol.2016.84

51. Reimann JDR, Salim S, Velazquez EF, Wang L, Williams KM, Flejter WL, et al. Comparison of melanoma gene expression score with histopathology, fluorescence *in situ* hybridization, and SNP array for the classification of melanocytic neoplasms. *Mod Pathol.* (2018) 31:1733–43. doi: 10.1038/s41379-018-0087-6

52. Conway K, Edmiston SN, Parker JS, Kuan PF, Tsai YH, Groben PA, et al. Identification of a robust methylation classifier for cutaneous melanoma diagnosis. *J Invest Dermatol.* (2019) 139:1349–61. doi: 10.1016/j.jid.2018.11.024

53. Conway K, Edmiston SN, Khondker ZS, Groben PA, Zhou X, Chu H, et al. DNA-methylation profiling distinguishes malignant melanomas from benign nevi. *Pigment Cell Melanoma Res.* (2011) 24:352–60. doi: 10.1111/j.1755-148X.2011.00828.x

54. Gao L, Van Den Hurk K, Moerkerk PTM, Goeman JJ, Beck S, Grui NA, et al. Promoter CpG Island hypermethylation in dysplastic nevus and melanoma: CLDN11 as an epigenetic biomarker for malignancy. *J Invest Dermatol.* (2014) 134:2957–66. doi: 10.1038/jid.2014.270

55. Salvianti F, Orlando C, Massi D, De Giorgi V, Grazzini M, Pazzagli M, et al. Tumor-related methylated cell-free DNA and circulating tumor cells in melanoma. *Front Mol Biosci.* (2016) 2:76. doi: 10.3389/fmolb.2015.00076

56. Torres R, Lang UE, Hejna M, Shelton SJ, Joseph NM, Shain AH, et al. MicroRNA ratios distinguish melanomas from nevi. *J Invest Dermatol.* (2019) 140:164–73.e7. doi: 10.1016/j.jid.2019.06.126

57. Alexandrescu DT, Kauffman CL, Jatkoe TA, Hartmann DP, Vener T, Wang H, et al. Melanoma-specific marker expression in skin biopsy tissues as a tool to facilitate melanoma diagnosis. *J Invest Dermatol.* (2010) 130:1887–92. doi: 10.1038/jid.2010.61

58. Koh SS, Opel ML, Wei JPJ, Yau K, Shah R, Gorre ME, et al. Molecular classification of melanomas and nevi using gene expression microarray signatures and formalin-fixed and paraffin-embedded tissue. *Mod Pathol.* (2009) 22:538–46. doi: 10.1038/modpathol.2009.8

59. Stark MS, Gray ES, Isaacs T, Chen FK, Millward M, McEvoy A, et al. A panel of circulating microRNAs detects uveal melanoma with high precision. *Transl Vis Sci Technol.* (2019) 8:12. doi: 10.1167/tvst.8.6.12

60. Lang UE, Torres R, Cheung C, Vladar EK, Mccalmont TH, Kim J, et al. Ciliation index is a useful diagnostic tool in challenging spitzoid melanocytic neoplasms. *J Invest Dermatol.* (2020) 140:1401–9.e2. doi: 10.1016/j.jid.2019.11.028

61. Rodic N, Zampella J, Sharma R, Burns KH, Taube JM. Diagnostic utility of 5-hydroxymethylcytosine immunohistochemistry in melanocytic proliferations. *J Cutan Pathol.* (2015) 42:807–14. doi: 10.1111/cup.12564

62. Uguen A, Talagas M, Costa S, Duigou S, Bouvier S, De Braekeleer M, et al. A p16-Ki-67-HMB45 immunohistochemistry scoring system as an ancillary diagnostic tool in the diagnosis of melanoma. *Diagn Pathol.* (2015) 10:195. doi: 10.1186/s13000-015-0431-9

63. Lezcano C, Pulitzer M, Moy AP, Hollmann TJ, Jungbluth AA, Busam KJ. *Immunohistochemistry for PRAME in the Distinction of Nodal Nevi From Metastatic Melanoma.* (2019). Available online at: www.ajsp.com

64. Kim W, Han TH, Kim HJ, Park MY, Kim KS, Park RW. An automated measurement of ciliary beating frequency using a combined optical flow and peak detection. *Healthc Inform Res.* (2011) 17:111–9. doi: 10.4258/hir.2011.17.2.111

65. Snedecor ER, Sung CC, Moncayo A, Rothstein BE, Mockler DC, Tonnesen MG, et al. Loss of primary cilia in melanoma cells is likely independent of proliferation and cell cycle progression. *J Invest Dermatol.* (2015) 135:1456–8. doi: 10.1038/jid.2015.22

66. Love NR, Lang UE, Cheung C, Kim J. Depletion of primary cilium in acral melanoma. *J Cutan Pathol.* (2019) 46:665–71. doi: 10.1111/cup.13484

67. Kim J, Dabiri S, Seeley ES. Primary cilium depletion typifies cutaneous melanoma *in situ* and malignant melanoma. *PLoS ONE.* (2011) 6:e27410. doi: 10.1371/journal.pone.0027410

68. Lang UE, Cheung C, Vladar EK, Swetter SM, Kim J. Loss of primary cilia correlates with cytologic severity in dysplastic melanocytic nevi. *J Cutan Pathol.* (2016) 43:113–9. doi: 10.1111/cup.12612

69. Lang UE, Love NR, Cheung C, McCalmont TH, Kim J. Use of the ciliation index to distinguish invasive melanoma from associated conventional melanocytic nevi. *Am J Dermatopathol.* (2020) 42:11–5. doi: 10.1097/DAD.0000000000001459

70. Vetto JT, Hsueh EC, Gastman BR, Dillon LD, Monzon FA, Cook RW, et al. Guidance of sentinel lymph node biopsy decisions in patients with T1-T2 melanoma using gene expression profiling. *Futur Oncol.* (2019) 15:1207–17. doi: 10.2217/fon-2018-0912

71. Meves A, Nikolova E, Heim JB, Squirewell EJ, Cappel MA, Pittelkow MR, et al. Tumor cell adhesion as a risk factor for sentinel lymph node metastasis in primary cutaneous melanoma. *J Clin Oncol.* (2015) 33:2509–15. doi: 10.1200/JCO.2014.60.7002

72. Bellomo D, Arias-Mejias SM, Ramana C, Heim JB, Quattrocchi E, Sominidi-Damodaran S, et al. Model combining tumor molecular and clinicopathologic risk factors predicts sentinel lymph node metastasis in primary cutaneous melanoma. *JCO Precis Oncol.* (2020) 4:319–34. doi: 10.1200/PO.19.00206

73. Hida T, Yoneta A, Wakamatsu K, Yanagisawa K, Ishii-Osai Y, Kan Y, et al. Circulating melanoma cells as a potential biomarker to detect metastasis and evaluate prognosis. *Australas J Dermatol.* (2016) 57:145–9. doi: 10.1111/ajd.12455

74. Kuhlmann JD, Wimberger P, Wilsch K, Fluck M, Suter L, Brunner G. Increased level of circulating U2 small nuclear RNA fragments indicates metastasis in melanoma patients. *Clin Chem Lab Med.* (2015) 53:605–11. doi: 10.1515/cclm-2014-1064

75. Maier T, Laubender RP, Sturm RA, Klingenstein A, Korting HC, Ruzicka T, et al. Osteopontin expression in plasma of melanoma patients and in melanocytic tumours. *J Eur Acad Dermatol Venereol.* (2012) 26:1084–91. doi: 10.1111/j.1468-3083.2011.04210.x

76. Gerami P, Cook RW, Wilkinson J, Russell MC, Dhillon N, Amaria RN, et al. Development of a prognostic genetic signature to predict the metastatic risk associated with cutaneous melanoma. *Clin Cancer Res.* (2015) 21:175–83. doi: 10.1158/1078-0432.CCR-13-3316

77. Gerami P, Cook RW, Russell MC, Wilkinson J, Amaria RN, Gonzalez R, et al. Gene expression profiling for molecular staging of cutaneous melanoma in patients undergoing sentinel lymph node biopsy. *J Am Acad Dermatol.* (2015) 72:780–5.e3. doi: 10.1016/j.jaad.2015.01.009

78. Ferris LK, Farberg AS, Middlebrook B, Johnson CE, Lassen N, Oelschlager KM, 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) 76:818–25.e3. doi: 10.1016/j.jaad.2016.11.051

79. Autilio C, Paolillo C, Lavieri MM, Pocino K, De Paolis E, Di Stasio E, et al. PAX3d mRNA over 2.76 copies/μL in the bloodstream predicts cutaneous malignant melanoma relapse. *Oncotarget.* (2017) 8:85479–91. doi: 10.18632/oncotarget.20177

80. Lee JH, Long GV, Menzies AM, Lo S, Guminski A, Whitbourne K, et al. Association between circulating tumor DNA and pseudoprogression in patients with metastatic melanoma treated with anti–programmed cell death 1 antibodies. *JAMA Oncol.* (2018) 4:717–21. doi: 10.1001/jamaoncol.2017.5332

81. Schreuer M, Meersseman G, Van Den Herrewegen S, Jansen Y, Chevolet I, Bott A, et al. Quantitative assessment of BRAF V600 mutant circulating cell-free tumor DNA as a tool for therapeutic monitoring in metastatic melanoma patients treated with BRAF/MEK inhibitors. *J Transl Med.* (2016) 14:1–11. doi: 10.1186/s12967-016-0852-6

82. Cook RW, Middlebrook B, Wilkinson J, Covington KR, Oelschlager K, Monzon FA, et al. Analytic validity of DecisionDx-Melanoma, a gene expression profile test for determining metastatic risk in melanoma patients. *Diagn Pathol.* (2018) 13:13. doi: 10.1186/s13000-018-0690-3

83. Zager JS, Gastman BR, Leachman S, Gonzalez RC, Fleming MD, Ferris LK, et al. Performance of a prognostic 31-gene expression profile in an independent cohort of 523 cutaneous melanoma patients. *BMC Cancer.* (2018) 18:130. doi: 10.1186/s12885-018-4016-3

84. Gastman BR, Zager JS, Messina JL, Cook RW, Covington KR, Middlebrook B, et al. Performance of a 31-gene expression profile test in cutaneous melanomas of the head and neck. *Head Neck.* (2019) 41:871–9. doi: 10.1002/hed.25473

85. 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) 80:149–57.e4. doi: 10.1016/j.jaad.2018.07.028

86. Keller J, Schwartz TL, Lizalek JM, Chang E, Patel AD, Hurley MY, et al. Prospective validation of the prognostic 31-gene expression profiling test in primary cutaneous melanoma. *Cancer Med.* (2019) 8:2205–12. doi: 10.1002/cam4.2128

87. Podlipnik S, Carrera C, Boada A, Richarz NA, López-Estebaranz JL, Pinedo-Moraleda F, 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) 33:857–62. doi: 10.1111/jdv.15454

88. Wong SL, Kattan MW, McMasters KM, Coit DG. A nomogram that predicts the presence of sentinel node metastasis in melanoma with better discrimination than the American joint committee on cancer staging system. *Ann Surg Oncol.* (2005) 12:282–8. doi: 10.1245/ASO.2005.05.016

89. Lo SN, Ma J, Scolyer RA, Haydu LE, Stretch JR, Saw RPM, et al. Improved risk prediction calculator for sentinel node positivity in patients with melanoma: the melanoma institute australia nomogram. *J Clin Oncol.* (2020) 38:2719–27. doi: 10.1200/JCO.19.02362

90. Swetter SM, Thompson JA, Albertini MR, Barker CA, Boland G, Carson WE, et al. NCCN Guidelines Version 1.2021 Melanoma: Cutaneous. National Comprehensive Cancer Network (2020).

91. Dummer R, Hauschild A, Santinami M, Atkinson V, Mandalà M, Kirkwood JM, et al. Five-year analysis of adjuvant dabrafenib plus trametinib in stage III melanoma. *N Engl J Med.* (2020) 383:1139–48. doi: 10.1056/NEJMoa2005493

92. 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. doi: 10.1056/NEJMoa1802357

93. Morton DL, Thompson JF, Cochran AJ, Mozzillo N, Nieweg OE, Roses DF, et al. Final trial report of sentinel-node biopsy versus nodal observation in melanoma. *N Engl J Med.* (2014) 370:599–609. doi: 10.1056/NEJMoa13 10460

94. Karakousis GC, Gimotty PA, Czerniecki BJ, Elder DE, Elenitsas R, Ming ME, et al. Regional nodal metastatic disease is the strongest predictor of survival in patients with thin vertical growth phase melanomas: a case for SLN staging biopsy in these patients. *Ann Surg Oncol.* (2007) 14:1596–603. doi: 10.1245/s10434-006-9319-y

95. Saltman BE, Ganly I, Patel SG, Coit DG, Brady MS, Wong RJ, et al. Prognostic implication of sentinel lymph node biopsy in cutaneous head and neck melanoma. *Head Neck.* (2010) 32:1686–92. doi: 10.1002/hed.21390

96. Marchetti MA, Coit DG, Dusza SW, Yu A, McLean LT, Hu Y, et al. Performance of gene expression profile tests for prognosis in patients with localized cutaneous melanoma: a systematic review and meta-analysis. *JAMA Dermatol.* (2020) 156:953–62. doi: 10.1001/jamadermatol.2020.1731

97. Pantel K, Alix-Panabières C. Liquid biopsy and minimal residual disease — latest advances and implications for cure. *Nat Rev Clin Oncol.* (2019) 16:409–24. doi: 10.1038/s41571-019-0187-3

98. Finck SJ, Giuliano AE, Morton DL. LDH and melanoma. *Cancer.* (1983) 51:840–3.

99. Weide B, Elsässer M, Büttner P, Pflugfelder A, Leiter U, Eigentler TK, et al. Serum markers lactate dehydrogenase and S100B predict independently disease outcome in melanoma patients with distant metastasis. *Br J Cancer.* (2012) 107:422–8. doi: 10.1038/bjc.2012.306

100. Long GV, Grob JJ, Nathan P, Ribas A, Robert C, Schadendorf D, et al. Factors predictive of response, disease progression, and overall survival after dabrafenib and trametinib combination treatment: a pooled analysis of individual patient data from randomised trials. *Lancet Oncol.* (2016) 17:1743–54. doi: 10.1016/S1470-2045(16)30578-2

101. Nosrati A, Tsai KK, Goldinger SM, Tumeh P, Grimes B, Loo K, et al. Evaluation of clinicopathological factors in PD-1 response: derivation and validation of a prediction scale for response to PD-1 monotherapy. *Br J Cancer.* (2017) 116:1141–7. doi: 10.1038/bjc.2017.70

102. Diamandis EP, Fiala C. Can circulating tumor DNA be used for direct and early stage cancer detection? *F1000Research.* (2017) 6:2129. doi: 10.12688/f1000research.13440.1

103. Romano A, Palumbo GA, Parrinello NL, Conticello C, Martello M, Terragna C. Minimal residual disease assessment within the bone marrow of multiple myeloma: a review of caveats, clinical significance and future perspectives. *Front Oncol.* (2019) 9:699. doi: 10.3389/fonc.2019.00699

104. McDonald BR, Contente-Cuomo T, Sammut SJ, Odenheimer-Bergman A, Ernst B, Perdigones N, et al. Personalized circulating tumor DNA analysis to detect residual disease after neoadjuvant therapy in breast cancer. *Sci Transl Med.* 11:eaax7392. doi: 10.1126/scitranslmed.aax7392

105. Chin RI, Chen K, Usmani A, Chua C, Harris PK, Binkley MS, et al. Detection of solid tumor molecular Residual Disease (MRD) using circulating tumor DNA (ctDNA). *Mol Diagn Ther.* (2019) 23:311. doi: 10.1007/s40291-019-00390-5

106. Bratman SV, Yang SYC, Iafolla MAJ, Liu Z, Hansen AR, Bedard PL, et al. Personalized circulating tumor DNA analysis as a predictive biomarker in solid tumor patients treated with pembrolizumab. *Nat Cancer.* (2020) 1:873–81. doi: 10.1038/s43018-020-0096-5

107. Váraljai R, Wistuba-Hamprecht K, Seremet T, Diaz JMS, Nsengimana J, Sucker A, et al. Application of circulating cell-free tumor DNA profiles for therapeutic monitoring and outcome prediction in genetically heterogeneous metastatic melanoma. *JCO Precis Oncol.* (2019) 3:PO.18.00229. doi: 10.1200/PO.18.00229

108. Eggermont AMM, Blank CU, Mandala M, Long GV, Atkinson VG, Dalle S, et al. Longer follow-up confirms recurrence-free survival benefit of adjuvant pembrolizumab in high-risk stage III melanoma: updated results from the EORTC 1325-MG/KEYNOTE-054 trial. *J Clin Oncol.* (2020) 38:3925–36. doi: 10.1200/JCO.20.02110

109. Amaral TMS, Hoffmann MC, Sinnberg T, Niessner H, Sülberg H, Eigentler TK, et al. Clinical validation of a prognostic 11-gene expression profiling score in prospectively collected FFPE tissue of patients with AJCC v8 stage II cutaneous melanoma. *Eur J Cancer.* (2020) 125:38–45. doi: 10.1016/j.ejca.2019.10.027

110. Ritterhouse LL, Barletta JA. BRAF V600E mutation-specific antibody: a review. *Semin Diagn Pathol.* (2015) 32:400–8. doi: 10.1053/j.semdp.2015.02.010

111. Larkin J, Chiarion-Sileni V, Gonzalez R, Grob J-J, Rutkowski P, Lao CD, et al. Five-year survival with combined nivolumab and ipilimumab in advanced melanoma. *N Engl J Med.* (2019) 381:1535–46. doi: 10.1056/NEJMoa1910836

112. Luke JJ. Comprehensive clinical trial data summation for BRAF-MEK inhibition and checkpoint immunotherapy in metastatic

melanoma. *Oncologist.* (2019) 24:e1197. doi: 10.1634/theoncologist.2018-0876

113. Robert C, Grob JJ, Stroyakovskiy D, Karaszewska B, Hauschild A, Levchenko E, et al. Five-Year Outcomes with dabrafenib plus trametinib in metastatic melanoma. *N Engl J Med.* (2019) 381:626–36. doi: 10.1056/NEJMoa1904059

114. Topalian SL, Sznol M, McDermott DF, Kluger HM, Carvajal RD, Sharfman WH, et al. Survival, durable tumor remission, and long-term safety in patients with advanced melanoma receiving nivolumab. *J Clin Oncol.* (2014) 32:1020–30. doi: 10.1200/JCO.2013.53.0105

115. Ribas A, Hamid O, Daud A, Hodi FS, Wolchok JD, Kefford R, et al. Association of pembrolizumab with tumor response and survival among patients with advanced melanoma. *JAMA.* (2016) 315:1600–9. doi: 10.1001/jama.2016.4059

116. Betof Warner A, Palmer JS, Shoushtari AN, Goldman DA, Panageas KS, Hayes SA, et al. Long-term outcomes and responses to retreatment in patients with melanoma treated with PD-1 blockade. *J Clin Oncol.* (2020) 38:1655–63. doi: 10.1200/JCO.19.01464

117. Lo SN, Scolyer RA, Thompson JF. Long-term survival of patients with thin (T1) cutaneous melanomas: a breslow thickness cut point of 0.8 mm separates higher-risk and lower-risk tumors. *Ann Surg Oncol.* (2018) 25:894–902. doi: 10.1245/s10434-017-6325-1

118. Buder-Bakhaya K, Hassel JC. Biomarkers for clinical benefit of immune checkpoint inhibitor treatment-a review from the melanoma perspective and beyond. *Front Immunol.* (2018) 9:1474. doi: 10.3389/fimmu.2018.01474

119. Morrison C, Pabla S, Conroy JM, Nesline MK, Glenn ST, Dressman D, et al. Predicting response to checkpoint inhibitors in melanoma beyond PD-L1 and mutational burden. *J Immunother Cancer.* (2018) 6:32. doi: 10.1186/s40425-018-0344-8

120. Yi M, Jiao D, Xu H, Liu Q, Zhao W, Han X, et al. Biomarkers for predicting efficacy of PD-1/PD-L1 inhibitors. *Mol Cancer.* (2018) 17:129. doi: 10.1186/s12943-018-0864-3

121. Arora S, Velichinskii R, Lesh RW, Ali U, Kubiak M, Bansal P, et al. Existing and emerging biomarkers for immune checkpoint immunotherapy in solid tumors. *Adv Ther.* (2019) 36:2638–78. doi: 10.1007/s12325-019-01051-z

122. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. *CA Cancer J Clin.* (2020) 70:7–30. doi: 10.3322/caac.21590

123. Grossman D, Kim CC, Hartman RI, Berry E, Nelson KC, Okwundu N, et al. Prognostic gene expression profiling in melanoma: necessary steps to incorporate into clinical practice. *Melanoma Manag.* (2019) 6:MMT32. doi: 10.2217/mmt-2019-0016

124. Plebani M. Exploring the iceberg of errors in laboratory medicine. *Clin Chim Acta.* (2009) 404:16–23. doi: 10.1016/j.cca.2009.03.022

125. Houben PHH, Winkens RAG, Van Der Weijden T, Vossen RCRM, Naus AJM, Grol RPTM. Reasons for ordering laboratory tests and relationship with frequency of abnormal results. *Scand J Prim Health Care.* (2010) 28:18–23. doi: 10.3109/02813430903281758

126. Fang C, Otero HJ, Greenberg D, Neumann PJ. Cost-utility analyses of diagnostic laboratory tests: a systematic review. *Value Heal.* (2011) 14:1010–8. doi: 10.1016/j.jval.2011.05.044

127. Kelly SG, Yarrington M, Zembower TR, Sutton SH, Silkaitis C, Postelnick M, et al. Inappropriate *Clostridium difficile* testing and consequent overtreatment and inaccurate publicly reported metrics. *Infect Control Hosp Epidemiol.* (2016) 37:1395–400. doi: 10.1017/ice.2016.210

Conflict of Interest: RJ-T is an inventor on international patent application no. PCT/US2019/023834 concerning the use of microRNA ratios as molecular diagnostics for melanoma.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.