Next-generation sequencing improves agreement and accuracy in the diagnosis of Spitz and spitzoid melanocytic lesions

Andrew Roth BS¹ | Nathaniel Lampley 3rd BS¹ | Anastasiya Boutko MS¹ | Jeffrey Zhao BA¹ | Sarah Benton BA¹ | Shantel Olivares BA¹ | Artur Zembowicz MD, PhD² | Pedram Gerami MD¹,³

1Department of Dermatology, Feinberg School of Medicine, Northwestern University, Chicago, Illinois, USA
2Department of Pathology and Laboratory Medicine, Tufts Medical School, Boston, Massachusetts, USA
3Robert H. Lurie Cancer Center, Feinberg School of Medicine, Northwestern University, Chicago, Illinois, USA

Abstract

Background: Spitzoid melanocytic neoplasms can be challenging to diagnose on histopathology alone. Next-generation sequencing (NGS) offers promise as a valuable aid in the diagnosis. Recently, one study reported increased inter-rater agreement in the diagnosis of spitzoid melanocytic neoplasms among 20 expert melanoma pathologists after incorporating NGS data. We hypothesized that NGS would carry a similar utility in a broader group of dermatopathologists and general pathologists.

Methods: Sixty-three participants of a live online (www.Dermpedia.org) CME course rendered a diagnosis on 70 cases composed of melanocytic neoplasms with spitzoid features. In Survey 1, cases included H&E slides and demographic information only, while Survey 2 included NGS data.

Results: With NGS information, inter-rater agreement significantly improved from “fair” to “almost perfect” and from “fair” to “substantial” for categorizing lesions as Spitz versus non-Spitz and conventional melanoma versus not, respectively. There was also an increase in diagnostic accuracy, evidenced by improved recognition of three metastatic tumors as being conventional melanomas.

Conclusion: The study supports the adoption of NGS as a valuable diagnostic adjunct for both expert and broader dermatopathologists in their assessments of spitzoid neoplasms.

KEYWORDS
dermatology, dermatopathology, melanoma, next-generation sequencing, Spitz

1 | INTRODUCTION

The current World Health Organization (WHO) Classification of Skin Tumors defines Spitz neoplasms as melanocytic neoplasms with spitzoid morphology accompanied by either a Spitz-associated genomic fusion or an HRAS-activating mutation.¹ Further, BRAF- and NRAS-mutated tumors should be excluded from the category of Spitz neoplasms.² The clinical relevance of this classification is illustrated by at least one previous study suggesting that most spitzoid melanocytic neoplasms diagnosed on histopathology resulting in distant metastasis are probably BRAF- or NRAS-mutated neoplasms misclassified as Spitz tumors.² Hence, it is logical that comprehensive genomic data

Andrew Roth and Nathaniel Lampley 3rd contributed equally to this study.

Received: 28 January 2022 | Revised: 29 June 2022 | Accepted: 4 July 2022
DOI: 10.1111/cup.14291

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2022 The Authors. Journal of Cutaneous Pathology published by John Wiley & Sons Ltd.
generated by next-generation sequencing (NGS) identifying the drivers and other relevant genomic aberrations in melanocytic tumors could improve the accuracy of diagnosis of Spitz neoplasms.

A recent study obtained diagnoses for a group of 70 spitzoid melanocytic neoplasms from 20 pathologists considered experts in melanoma pathology before and then again after having access to NGS studies with genomic data on the tumors. In that study, there was improved inter-rater agreement and accuracy following addition of the NGS data. The purpose of this current study was to test this same hypothesis in a broader group of dermatopathologists and general pathologists using participants in the Dermpedia (www.dermpedia.org) dermatology online CME course. Our study shows that as in expert melanoma pathologists, access to NGS data results in a statistically significant improvement in consensus and accuracy of diagnosis of Spitz neoplasms in broader dermatopathologists and general pathologists.

2 MATERIALS AND METHODS

This study was conducted with the approval of our University Institutional Review Board (IRB) (STU00001127). A set of 70 previously sequenced cases composed of melanocytic neoplasms with spitzoid morphologic features was used. All cases were sequenced for research purposes and were the same set of cases as those evaluated by the melanoma experts in the study referenced previously. Of note, 52 of the 70 cases contained a canonical structural variant including a Spitz-associated genomic fusion (n = 50) or HRAS mutation (n = 2) as defined by the fourth edition of the WHO Classification of Skin Tumors. Among the remaining 18 cases, seven had a BRAF mutation, seven had an NRAS mutation, two had an NF1 mutation, and one had both a BRAF and an NF1 mutation. The cases were sequenced using the Tempus Xt platform assay that included DNA sequencing with a 1711-gene panel and whole-transcriptome mRNA sequencing. Paired normal controls were also sent for sequencing. Variant calling was performed using Tempus bioinformatic pipelines reporting mutations with an allele frequency of 5% or more. There was a clinical follow-up documented for all cases with an average follow-up of 34 months. Diagnoses rendered were classified as Spitz nevus, Spitz tumor, Spitz melanoma, or melanoma with spitzoid features. Cases were presented two separate times to a group of 128 board-certified dermatopathologists and pathologists as part of an online course on challenging spitzoid melanocytic lesions through the online pathology education website (www.Dermpedia.org) streamed live in May 2021. The recording of this course including access to the cases are available at https://dermpedia.org/programs/collection-spitzoid-melanocytic-tumors-genetic-pathologic-correlations-wajapk2c00. Participants in the course were given the opportunity to view all scanned cases, which included H&E-stained slides and patient demographics and were able to render a preliminary diagnosis of each lesion into one of the following categories: (a) atypical nevus (non-Spitz); (b) severely atypical melanocytic tumor (of uncertain malignant potential/melanocytoma, non-Spitz); (c) conventional melanoma (with spitzoid features, non-Spitz); (d) Spitz
nevus/low-risk (atypical) Spitz tumor; (e) severely atypical Spitz tumor (of uncertain malignant potential); and (f) Spitz melanoma. While the term Spitz nevus and low-risk Spitz tumor have different meanings for many experts, for simplicity and to limit diagnostic categories, we used them synonymously in this study. The above was defined as Survey 1. After watching two expert dermatopathologist lectures (1-h long each) on incorporating genomics into the diagnosis of spitzoid lesions, participants in the course were given the opportunity to perform a second round of diagnoses, defined as Survey 2. Of the 128 participants, 63 completed all 70 cases on both surveys and were included in the statistical analysis. The specifics of the cases were not discussed until after both surveys were completed.

2.1 Statistical analysis

Data were collected and analyzed using IBM SPSS Statistics Version 27.0 (IBM Corp.). Inter-rater agreement was calculated using Cohen’s kappa (κ) for each pair of participants. An average Cohen’s κ was then calculated for each survey. Results were categorized as follows: ≤0 as poor; 0.01–0.2 as slight; 0.21–0.4 as fair; 0.41–0.6 as moderate; 0.61–0.8 as substantial; and 0.81–1 as almost perfect agreement. Two-sample Z tests were used to determine if the change in the average Cohen’s κ was statistically significant. McNemar and McNemar-Bowker chi-square tests were performed to determine if the difference in answer choices and answer categories (e.g., Spitz vs. non-Spitz) reached statistical significance. The significance level was initially set at $p = 0.05$ and subsequently adjusted with the Bonferroni correction to $p = 0.004$.

3 RESULTS

A total of 128 individuals enrolled in the online course were invited to participate in the study. Of these 128 individuals, 63 completed the pre-genomics and post-genomics portion of the diagnostic questions on all 70 cases in the study. Dermatopathologists comprised the largest portion of study participants (74%), followed by general pathologists (21.3%), dermatopathology fellows (0.8%), and all others (3.9%). The largest portion of participants came from non-academic settings (73.8%). All other participants came from academic institutions (26.2%). Answer selections from the survey results are displayed in Table 1. The diagnostic impression of the 70 cases was significantly affected by the genomic information provided in Survey 2. On average, the participants changed 40.3/70 diagnoses (57.6%) (SE 0.87 answers) after receiving genomic data. The number of diagnoses changed ranged from 21 to 54, with a median of 40. The difference between the two corresponding surveys was statistically significant ($n = 4410$, df = 15, $p < 0.004$).

We studied the diagnostic agreement between the participants for several parameters: Spitz versus non-Spitz, conventional melanoma versus not, and among all answer choices (Figure 1). After including genetic information, the average Cohen’s κ increased from 0.267 (SE 0.00271) to 0.835 (SE 0.00149) for the diagnosis of Spitz versus non-Spitz cases ($p < 0.004$). This corresponds to a significant improvement from fair to almost perfect agreement. For the diagnosis of conventional melanoma versus not, the average Cohen’s κ increased from 0.323 (SE 0.00271) to 0.689 (SE 0.00271), or from fair to substantial agreement ($p < 0.004$). There was also a significant increase in overall agreement among the six diagnostic choices from an average Cohen’s κ of 0.208 (SE 0.00149) to 0.359 (SE 0.00115); however, these results remained within the fair agreement category ($p < 0.004$).

Of the total 70 cases, there were three (Cases 6, 11, and 47) in which the clinical follow-up was that of either distant metastasis (Cases 11 and 47) or clinically bulky local metastasis (Case 6), indicating an unequivocally malignant tumor. The details of these cases are described in another study. The genomic information provided in Survey 2 helped in the identification of these cases as being malignant melanomas (Figure 2). For Case 6, which was an NRAS-mutated tumor with the addition of a TERT promoter (pTERT) mutation, 57% of participants correctly identified the case as conventional melanoma in Survey 1. Others chose Spitz melanoma (35%), severely atypical Spitz

![Figure 1](https://example.com/figure1.png)

**Figure 1** Diagnostic agreement for pre- and post-genomics for all 70 cases as measured by the average Cohen’s κ for Spitz versus non-Spitz, conventional melanoma versus not, and among all answer choices.
tumor (of uncertain malignant potential) (5%), and severely atypical melanocytic tumor (of uncertain malignant potential/melanocytoma, non-Spitz) (3%). After participants were provided with genomic information, the number of conventional melanoma diagnoses rose to 95% ($p < 0.004$). For Case 11, which was a $BRAF$-mutated tumor with a $pTERT$ mutation and loss of $CDKN2A$, 60% of participants chose

**FIGURE 3**  Case 58, a Spitz nevus/low-risk (atypical) Spitz tumor from the right upper arm of an 11-year-old male patient. NGS studies revealed a $KIF5B$-$RET$ fusion. The image displays H&E-stained sections at ×100 magnification with a ×400 inlay. Sections show a symmetric compound melanocytic neoplasm surmounted by epidermal hyperplasia. There are some small Kamino bodies in the epidermis. The junctional and dermal nests consist of monotonous appearing intermediate-sized melanocytes with vesicular nuclei and scant cytoplasm. Hence, the scant cytoplasm makes this case difficult to immediately recognize as part of the Spitz family. Mitotic activity is readily identified (arrowhead) and maturation is somewhat impaired. Prior to genomic data, 43% of participants selected a Spitz diagnosis (answers D, E, F). After incorporating genomic data, 92% selected a Spitz diagnosis.

**FIGURE 2**  Distribution of diagnostic selection for pre- and post-genomics for the three recurrent cases (Cases 6, 11, and 47) of conventional melanoma that resulted in metastatic disease.
conventional melanoma, 33% chose Spitz melanoma, and 6% chose severely atypical Spitz tumor (of uncertain malignant potential) in Survey 1. With genetic information, selection of conventional melanoma increased to 98% \((p < 0.004)\). For Case 47, which was a \(\text{BRAF}\)-mutated tumor with a \(p\text{TERT}\) mutation, 13% of participants diagnosed the lesion as conventional melanoma in Survey 1. Most participants selected Spitz nevus/low-risk (atypical) Spitz tumor (44%), followed by severely atypical Spitz tumor (of uncertain malignant potential) (21%), with the remaining selections distributed among the other answer choices (range: 3%–11%). Among these three metastatic cases (6, 11, and 47), there were three instances (1.5%) where a diagnosis was downgraded from a melanoma to a non-melanoma diagnosis after genomics. In Survey 2, the number of conventional melanoma diagnoses increased to 84% \((p < 0.004)\). For these three cases, the overall increased diagnosis of conventional melanoma from the first to the second survey was statistically significant \((p < 0.004)\).

Regard the 52 cases with genomic fusions or \(\text{HRAS}\) mutations, allocation of the tumor into the Spitz family (answer choices D, E, F) increased from 80.0% \((\text{SE} 1.3\%)\) to 95.8% \((\text{SE} 0.56\%)\) with genomic data \((n = 3276, p < 0.004)\) (Table 1, Figure 3). There was no significant difference between surveys categorizing these 52 cases as atypical (answers A and D), severely atypical (answers B and E), or malignant (answers C and F) lesions regardless of Spitz or non-Spitz classification \((p = 0.053)\). Among the 18 cases without fusions or \(\text{HRAS}\) mutations, classification as a non-Spitz neoplasm (answer choices A, B, C) increased from 47.7% \((\text{SE} 2.0\%)\) to 96.7% \((\text{SE} 0.81\%)\) with genomic data \((n = 1134, p < 0.004)\) (Table 1, Figure 4). Unlike the aforementioned 52 cases, the ability to categorize these 18 cases into atypical, severely atypical, or malignant lesions reached statistical significance after incorporating genomic information \((p < 0.004)\).

4 | DISCUSSION

The benefits of NGS in guiding targeted therapy for cancers are well established. In this study, we assessed the utility of NGS to assist in diagnosis of Spitz neoplasms, a historically difficult area of pathology because of overlapping morphological features of benign Spitz nevi and some malignant melanomas. Incorporating genetic information into the diagnostic process has become more important since the 2018 fourth edition of WHO Classification of Skin Tumours divided melanocytic neoplasms into nine categories based on genetic lineage.\(^1\) The WHO defines Spitz tumors as those with spitzoid morphology and containing a tyrosine kinase fusion or \(\text{HRAS}\) mutation.\(^1\) In contrast, melanocytic neoplasms with \(\text{BRAF}\), \(\text{NRAS}\), and \(\text{NF1}\) mutations are excluded from the category of Spitz neoplasms.

The diagnosis of most spitzoid neoplasms can be solved with morphology or morphology with some ancillary studies including immunohistochemistry (IHC). For example, classification as a Spitz neoplasm could be excluded by a positive \(\text{BRAF}\) \(V600E\) IHC. However,
There are undoubtedly a select proportion of melanocytic tumors that are unable to be solved with morphology and IHC alone. We hypothesize that some of these select cases may be solved by NGS studies. After one study reported increased diagnostic agreement with genomic data among a group of highly experienced dermatopathologists considered experts in melanocytic tumors, we aimed to determine if the same would be true among a broader group of dermatopathologists and general pathologists. To study this, we used participants of the www.Dermpedia.org dermatopathology online CME course. NGS data influenced the diagnosis in the majority (57.6%) of cases, improved consensus for categorizing lesions as Spitz versus non-Spitz, and, most importantly, improved the recognition of conventional melanoma.

Among the parameters measured, classification of the cases in a Spitz versus non-Spitz category had the greatest increase in the average Cohen’s κ from 0.267 to 0.835, which corresponds to an increase from fair to almost perfect agreement. Specifically, among the 52 cases with genomic fusions or HRAS mutations, the total number of Spitz answers increased from 80% to 95.8%. These results clearly indicate that NGS data result in greater consensus and more precise definitions of what constitutes a Spitz neoplasm. It is reasonable to expect that genomic-information-based classification systems will result in more clear communication between physicians, and ultimately, better patient care.

NGS data had the most dramatic impact on the ability of participants to identify conventional melanomas, especially the three cases (6, 11, and 47) associated with metastatic behavior. The three metastatic cases were associated with NRAS (one) or BRAF (two) mutations and all had hot spot pTERT mutations. With NGS data, the recognition of these three respective cases as a conventional melanoma increased from 57% to 95%, 60% to 98%, and 13% to 84%. Significantly, 79% of the participants favored a non-melanoma diagnosis for Case 47 prior to having the NGS information. Hence, it is our opinion and experience that the most valuable aspect of utilizing NGS in the assessment of melanocytic neoplasms involves distinguishing nevoid melanomas or other non-Spitz melanomas with MAPK-activating mutations (i.e., BRAF or NRAS, NF1) from Spitz neoplasms which carry tyrosine kinase fusions or HRAS mutations.

NGS also helped recognize some melanocytic tumors with Spitz-associated fusion drivers and relatively banal cytomorphic features but without the most classic spitzoid cytology as belonging to the Spitz family. This is valuable because without NGS data, many of such lesions could be classified as dysplastic nevi or melanocytic tumors of uncertain malignant potential concerning nevoid melanoma requiring different management, patient counseling, and/or follow-up.

The average Cohen’s κ for agreement among all answer choices also significantly increased from 0.208 (slight) to 0.359 (fair). The lesser agreement for all diagnostic categories is not surprising as criteria for subclassifying melanocytic lesions into non-melanoma subsets (Spitz nevus vs. atypical Spitz tumor; atypical nevus [non-Spitz] vs. severely atypical melanocytic tumor [of uncertain malignant potential/melanocytoma, non-Spitz]) are not well defined.

As expected because of greater level of expertise, the melanoma experts reviewing the same set of cases had greater pre-genomic agreement for Spitz versus non-Spitz and conventional melanoma versus not (κ = 0.36 and κ =0.25, respectively) than participants of the online course. NGS data had a larger impact on agreement among participants of this study, who showed greater improvement in each diagnostic category. Of interest, in the post-genomic survey, a higher percentage of participants accurately identified two of the three metastatic conventional melanomas compared to the expert group: 95% versus 60% for Case 6, 98% versus 90% for Case 11, and 84% versus 90% for Case 47. This could potentially be related to the lectures provided to the current participants on the subject matter in the www.Dermpedia.org CME course, or perhaps the expert cohort may place a greater weight on histopathologic features over genomic data because of their greater exposure to the spectrum of diagnostically challenging spitzoid cases. In selecting cases with canonical BRAF V600E mutations, using IHC for BRAF V600E could also solve this dilemma.

As for limitations, the data presented in this study were obtained during an educational exercise and the diagnoses rendered by the participants may not exactly reflect the diagnoses they would have made in a real-life situation. The participants were only provided restricted clinical and pathological information and had only limited time (about 5 min per case) to examine each lesion. No gross or dermoscopic images were presented and no other diagnostic tests or labs were given. The participants were only able to review a single H&E-stained slide from each case and were not provided any additional molecular findings, such as results of IHC studies or fluorescence in situ hybridization, which would have been available in the real diagnostic environment. Hence, while the study clearly shows that NGS data significantly improve diagnosis of spitzoid melanocytic tumors, the exact magnitude of its impact reported here may differ from that in clinical practice.

Our results make a case for adoption of NGS as a valuable diagnostic adjunct in the evaluation of select Spitz and spitzoid melanocytic lesions that remain unresolved after assessment with standard histomorphologic criteria and IHC. Additionally, NGS studies may allow for the development of more targeted assays that could also be used for similar purposes. In our study, NGS results in significantly increased consensus among experts and broader dermatopathologists. Most importantly, it facilitates distinguishing malignant conventional melanomas and dysplastic nevi from indolent Spitz tumors beyond what is possible with routine microscopy alone. It is reasonable to expect that wider adoption of NGS data may decrease the number of histopathologically ambiguous melanocytic lesions that are either reported as tumors of uncertain malignant potential or are misdiagnosed.

ACKNOWLEDGMENTS
This work was supported by the IDP Foundation and Dermatopathology Consultations LLC.

CONFLICTS OF INTEREST
Dr. Pedram Gerami has served as a consultant for Myriad Genomics, DermTech Int., Merck, and Castle Biosciences and has received honoraria for this work. The remaining authors declare no conflicts of interest.
DATA AVAILABILITY STATEMENT
The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES
1. Elder DE, Bastian BC, Cree IA, Massi D, Scolyer RA. The 2018 World Health Organization classification of cutaneous, mucosal, and uveal melanoma: detailed analysis of 9 distinct subtypes defined by their evolutionary pathway. *Arch Pathol Lab Med*. 2020;144(4):500-522. doi: 10.5858/arpa.2019-0561-RA

2. Quan VL, Zhang B, Zhang Y, et al. Integrating next-generation sequencing with morphology improves prognostic and biologic classification of Spitz neoplasms. *J Invest Dermatol*. 2020;140(8):1599-1608. doi: 10.1016/j.jid.2019.12.031

3. Benton S, Zhao J, Zhang B, et al. Impact of next-generation sequencing on interobserver agreement and diagnosis of spitzoid neoplasms. *Am J Surg Pathol*. 2021;45(12):1597-1605. doi: 10.1097/PAS.0000000000001753

4. Kumar N, Buzney C. Dermpedia.org: an online project created to advance the field of dermatology by organizing the knowledge and experience acquired by clinicians worldwide. *J Am Acad Dermatol*. 2013;68(4):6563. doi: 10.1016/j.jaad.2012.12.379

5. Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas*. 1960;20(1):37-46. 10.1177/00131644600200001753

6. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33(1):159-174. doi: 10.2307/2529310

7. Bowker AH. A test for symmetry in contingency tables. *J Am Stat Assoc*. 1948;43(244):572-574. doi: 10.1080/01621459.1948.10483284

How to cite this article: Roth A, Lampley N 3rd, Boutko A, et al. Next-generation sequencing improves agreement and accuracy in the diagnosis of Spitz and spitzoid melanocytic lesions. *J Cutan Pathol*. 2022;49(10):868-874. doi: 10.1111/cup.14291