MODERN PATHOLOGY

USCAP UNITED STATES AND CANADIAN ACADEMY OF PATHOLOGY
Creating a Better Pathologist

Journal homepage: https://modernpathology.org/

Research Article

Canadian Multicentric Pan-TRK (CANTRK) Immunohistochemistry Harmonization Study

Martin D. Hyrcza[a], Sebastiao N. Martins-Filho[b,c], Alan Spatz[d], Han-Jun Wang[d], Bibianna M. Purgina[e], Patrice Desmeules[f], Paul C. Park[g], Gilbert Bigras[b], Sungmi Jung[h], Jean-Claude Cutz[i], Zhaolin Xu[j], David M. Berman[k], Brandon S. Sheffield[l], Carol C. Cheung[c], Charles Leduc[m], David M. Hwang[n], Diana Ionescu[o], Paul Klonowski[p], Myriam Chevarie-Davis[q], Rose Chami[r], Bryan Lo[e], Tracy L. Stockley[c], Ming-Sound Tsao[c], Emina Torlakovic[s,*]

[a] *Department of Pathology and Laboratory Medicine, University of Calgary, Arnie Charbonneau Cancer Institute, Calgary, Alberta, Canada;* [b] *Department of Laboratory Medicine and Pathology, University of Alberta, Edmonton, Alberta, Canada;* [c] *University Health Network, Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, Ontario, Canada;* [d] *McGill University Health Center, Lady Davis Institute, McGill University, Montreal, Quebec, Canada;* [e] *Department of Pathology and Laboratory Medicine, Ottawa Hospital, University of Ottawa, Ottawa, Ontario, Canada;* [f] *Institut Universitaire de Cardiologie et de Pneumologie de Québec, Université Laval, Quebec City, Quebec, Canada;* [g] *Shared Health, Department of Pathology, Rady Faculty of Health Sciences, University of Manitoba, Winnipeg, Manitoba, Canada;* [h] *Department of Pathology, McGill University Health Centre, McGill University, Montreal, Quebec, Canada;* [i] *Department of Pathology and Molecular Medicine, McMaster University, Hamilton, Ontario, Canada;* [j] *Department of Pathology, Dalhousie University, Halifax, Nova Scotia, Canada;* [k] *Department of Pathology and Molecular Medicine, Queen's University, Kingston, Ontario, Canada;* [l] *Department of Pathology, William Osler Health System, Brampton, Ontario, Canada;* [m] *Centre Hospitalier de l'Université de Montréal, Montreal, Quebec, Canada;* [n] *Department of Laboratory Medicine and Pathobiology, Sunnybrook Health Sciences Centre, University of Toronto, Toronto, Ontario, Canada;* [o] *Department of Pathology and Laboratory Medicine, British Columbia Cancer Agency, University of British Columbia, Vancouver, British Columbia, Canada;* [p] *Department of Pathology and Laboratory Medicine, University of Calgary Cumming School of Medicine Diagnostic and Scientific Centre, Calgary, Alberta, Canada;* [q] *Département de Pathologie et Biologie Cellulaire, Hôpital Maisonneuve-Rosemont, Université de Montréal, Montreal, Quebec, Canada;* [r] *Department of Laboratory Medicine and Pathobiology, The Hospital for Sick Children, University of Toronto, Toronto, Ontario, Canada;* [s] *Department of Pathology and Laboratory Medicine, Royal University Hospital, Saskatchewan Health Authority, and College of Medicine, University of Saskatchewan, Saskatoon, Saskatchewan, Canada*

ARTICLE INFO

Article history:
Received 16 May 2023
Revised 19 October 2023
Accepted 9 November 2023
Available online 14 November 2023

Keywords:
analytical sensitivity
Canadian ring trial
IHC calibrators
immunohistochemistry (IHC)
NTRK
reference standards

ABSTRACT

Tumor-agnostic testing for *NTRK1-3* gene rearrangements is required to identify patients who may benefit from TRK inhibitor therapies. The overarching objective of this study was to establish a high-quality pan-TRK immunohistochemistry (IHC) screening assay among 18 large regional pathology laboratories across Canada using pan-TRK monoclonal antibody clone EPR17341 in a ring study design. TRK-fusion positive and negative tumor samples were collected from participating sites, with fusion status confirmed by panel next-generation sequencing assays. Each laboratory received: (1) unstained sections from 30 cases of TRK-fusion-positive or -negative tumors, (2) 2 types of reference standards: TRK calibrator slides and IHC critical assay performance controls (iCAPCs), (3) EPR17341 antibody, and (4) suggestions for developing IHC protocols. Participants were asked to optimize the IHC protocol for their instruments and detection systems by using iCAPCs, to stain the 30 study cases, and to report the percentage scores for membranous, cytoplasmic, and nuclear staining. TRK calibrators were used to assess the analytical sensitivity of IHC protocols developed by using the 2 reference standards. Fifteen of 18 laboratories achieved diagnostic sensitivity of 100% against

These authors contributed as co-senior authors: Ming-Sound Tsao and Emina Torlakovic.
* Corresponding author.
E-mail address: emina.torlakovic@usask.ca (E. Torlakovic).

0893-3952/© 2023 THE AUTHORS. Published by Elsevier Inc. on behalf of the United States & Canadian Academy of Pathology. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).
https://doi.org/10.1016/j.modpat.2023.100384

next-generation sequencing. The diagnostic specificity ranged from 40% to 90%. The results did not differ significantly between positive scores based on the presence of any type of staining vs the presence of overall staining in ≥1% of cells. The median limit of detection measured by TRK calibrators was 76,000 molecules/cell (range 38,000 to >200,000 molecules/cell). Three different patterns of staining were observed in 19 TRK-positive cases, cytoplasmic-only in 7 samples, nuclear and cytoplasmic in 9 samples, and cytoplasmic and membranous in 3 samples. The Canadian multicentric pan-TRK study illustrates a successful strategy to accelerate the multicenter harmonization and implementation of pan-TRK immunohistochemical screening that achieves high diagnostic sensitivity by using laboratory-developed tests where laboratories used centrally developed reference materials. The measurement of analytical sensitivity by using TRK calibrators provided additional insights into IHC protocol performance.

© 2023 THE AUTHORS. Published by Elsevier Inc. on behalf of the United States & Canadian Academy of Pathology. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Introduction

The TRK proteins (TrkA, TrkB, and TrkC), encoded by *NTRK* genes (*NTRK1*, *NTRK2*, and *NTRK3*), are transmembrane receptor tyrosine kinases that bind neurotrophins and regulate proliferation, differentiation, and maturation of a variety of neural tissues.[1,2] Oncogenic fusions involving TRK proteins have been documented across a wide spectrum of malignancies including sarcomas, gliomas, and carcinomas.[1,3-5] Their prevalence varies from less than 1% in common cancer types such as lung and colorectal adenocarcinomas,[6,7] to more than 80% in secretory carcinomas of the breast and salivary glands.[8-10] Oncogenic TRK fusion proteins result from chromosomal rearrangements involving *NTRK* genes containing the C-terminal tyrosine kinase domain of *NTRK* and the N-terminus of one of nearly 100 fusion partners described to date.[1,2,11,12]

Selective TRK tyrosine kinase inhibitors have been studied in clinical trials involving advanced and/or metastatic tumors with *NTRK* rearrangements, with promising results.[13-15] Consequently, there is a need to effectively screen for TRK-fusion-containing tumors. Current recommendations include direct molecular testing in tumors with high prevalence of TRK fusions,[16-18] yet this approach may be impractical and cost-prohibitive in malignancies with low prevalence. Pan-TRK immunohistochemistry (IHC) with monoclonal antibody EPR17341 was suggested as a screening tool to identify cases in which abnormal TRK protein expression would result in fusion testing being pursued.[19] This strategy has been successful in multiple cancer types at various institutions.[20-24]

Despite multiple studies, there is currently no standard recommendation for scoring and interpreting pan-TRK IHC. Some studies accepted the presence of any staining above background, in any compartment (cytoplasmic, nuclear, or membranous) as positive,[22] whereas others required at least 1% of cells showing staining in any of the 3 compartments.[20,23,24] Furthermore, there is no consensus on the acceptable sensitivity and specificity of pan-TRK IHC in detecting *NTRK* rearrangements or what reference standards could be helpful in the process of laboratory-developed test (LDT) development.

The objective of the Canadian multicentric pan-TRK (CANTRK) ring study was to optimize testing protocols to detect cancers with TRK fusions across 18 pathology and molecular diagnostics laboratories across Canada, using pan-TRK IHC for screening and panel NGS for detecting *NTRK* gene fusions. The companion CANTRK-NGS study has recently been reported.[25] The current manuscript details the design, execution, and results of the IHC part component of the CANTRK project.

Materials and Methods

Ring Study Design

The study was approved by the ethics review boards of the University Health Network and participating centers that contributed samples. Ten of the participating institutions used the BenchMark ULTRA platform, 5 used Dako OMNIS, 2 used Leica Bond III, and 1 laboratory used Dako Autostainer.

At the start of the study, selected participating institutions contributed tumor blocks from 40 cases that satisfied one of the following criteria: (1) have a high prevalence for *NTRK* fusion (eg, secretory carcinoma of salivary gland or breast, ganglioblastoma, infantile fibrosarcoma), (2) are known to harbor an *NTRK* fusion from molecular workup, or (3) are from existing tissue microarrays of various tumor types that have been screened by pan-TRK immunohistochemistry and were suspicious for TRK positivity. Ten cases were excluded to avoid diagnostic/molecular redundancy or due to insufficient material. The remaining 30 cases were analyzed using NGS panel assays (Oncomine Comprehensive Assay v3, ThermoFischer or FusionPlex Lung, ArcherDX) at 2 molecular diagnostic laboratories as described in the CANTRK-NGS study.[25] Altogether 19 cases were initially confirmed as positive and 11 were negative for TRK fusions. One positive sample was further excluded for tissue quality and technical reasons (CANTRK-22, see Results). The final study set is summarized in Table 1.

Each participating laboratory received: (1) duplicate sets of 30 unstained slides (4 μM) from the 19 fusion-positive and 11 fusion-negative cases; (2) 2 unstained slides with pan-TRK iCAPCs, produced for this study and consisting of sections of appendix, kidney, and TRK-expressing cell lines (KM12 obtained from NCI-DTCD and CUTO-3 obtained courtesy of Dr R.C. Doebele, University of Colorado Cancer Centre; (3) TRK IHC calibrator slides designed specifically for this study and containing 10 different levels of TRK peptide allowing for measurement of limits of detection (LOD) (Boston Cell Standards); and (4) EPR17341 antibody (either Roche/Ventana Assay or Abcam), together with a rabbit monoclonal negative control Ig (Cat. No. 790-4795). The laboratories running the BenchMark instruments were provided the VENTANA IHC pan-TRK Assay (see Supplementary Table S1). The laboratories using other staining instruments were provided with the Abcam EPR17341 pan-TRK antibody concentrate (ab181560). Laboratories that used Ventana IHC pan-TRK Assay could choose to either follow the recommended protocol as per the manufacturer specification sheet of the assay (recommended by study protocol) or could modify the IHC protocol conditions and develop their LDT

Table 1
The composition of the 30-sample test set listing the diagnoses, gene rearrangements, and immunohistochemistry staining patterns

Fusion (N = 19)	Diagnosis	C+M	C	C+N
Pattern of staining by fusion				
ETV6::NTRK3	Secretory breast carcinoma			1
	Infantile fibrosarcoma			1
	Papillary thyroid carcinoma			1
	Salivary gland secretory carcinoma			8
LMNA::NTRK1	Dysplastic-type nevus		1	
WNK2::NTRK2	Ganglioblastoma		1	
STRN3::NTRK2	Ganglioblastoma		1	
TPR::NTRK1	Lung adenocarcinoma		1	
TPM3::NTRK1	Lung adenocarcinoma	1		
	Papillary thyroid carcinoma	2		
SPECC1L::NTRK3	Sarcoma, NOS		1	
Total		3	5	11
Pattern of staining by *NTRK* gene				
Involving *NTRK1*		3	2	
Involving *NTRK2*			2	
Involving *NTRK3*			1	11

C, cytoplasmic only; C+M, cytoplasmic and membranous; N+C, nuclear and cytoplasmic.

with the Ventana IHC pan-TRK prediluted Ab by using iCAPCs as reference standards. All laboratories that used antibody concentrate developed new IHC protocols (LDT) by using reference standards (tissue samples and cell lines) with their own or centrally suggested LDT IHC protocols. The stained slides were assessed, and the presence of staining and staining percentage were scored for membranous, cytoplasmic, and nuclear compartments by individual site pathologists before they were returned to the central laboratory. The readouts were also centrally reviewed by a single pathologist (E.T.).

Reference Standards/Materials

1. IHC Critical Assay Performance Controls

The iCAPCs design included appendix, kidney, liver, and cell lines with and without TRK fusions (KM12, CUTO-3, and A549). For this study, the iCAPCs included a descriptive lower LOD (d-LOD) defined as ganglion cells in the myenteric plexus of the appendix and/or juxtaglomerular apparatus of the kidney, or perivascular (autonomic nerve) plexus in small arteries of the kidney. The acceptability criteria for iCAPCs' results included achieving strong staining of the positive cell line and variable but at least focally strong staining of the d-LOD (as per above). Other types of cells/tissues tested as well as negative cell lines were expected to be negative. If the developed protocols successfully demonstrated the above acceptability criteria, they would then be used to stain the 30 study cases.

2. TRK Calibrators

The slides with calibrators were manufactured by the Boston Cell Standards (BCS) specifically for this study. The calibrators included glass beads with 10 different levels of peptide that were used to generate the EPR17341 antibody. Negative control beads were also included. The calibrators were used to lower LOD or the lowest amount of TRK molecules/cell, which resulted in positive staining. The results were assessed centrally as per a previously described protocol.[26]

Immunohistochemistry Protocols

Two of the participating Canadian laboratories have previously validated their pan-TRK IHC protocols against their local NGS results. One used an Omnis instrument and the relevant detection system from Dako, and one used Bond III (Leica Biosystems). The pan-TRK IHC protocol for Ventana instruments was also available on the manufacturer's IHC pan-TRK Assay instruction sheet. These protocols were tested and made available for consideration as a starting point for all other participating laboratories. Additional alternative LDT IHC protocols were developed by using the "signal-to-noise" ratio approach by the CANTRK central laboratory and were available for consideration and guidance to participating laboratories. The initial staining protocol was then optimized using locally developed control tissues, as well as centrally provided iCAPCs that included various tissues and cell lines. Once the acceptability criteria were met for the iCAPC results, the laboratories were asked to stain the study sample cohort of 30 tumor slides and calibrators. Criteria for acceptability for study cases were 100% sensitivity and >80% specificity against central NGS results. Laboratories that did not reach acceptability criteria for diagnostic sensitivity and specificity, were asked to consider further optimization of their pan-TRK IHC protocols; however, this was considered optional as long as the diagnostic sensitivity of the protocols was 100% against NGS.

Results

TRK Calibrators Results

The TRK calibrator results are presented in Table 2. All but 3 laboratories developed protocols with 100% sensitivity, using a wide range of IHC protocols that differed greatly in the quality and intensity of staining (Fig. 1). This was evident even in cases that would be considered strongly positive with optimal protocols. The results of TRK calibrators indicated that when the measured LOD was ≤60,000 TRK molecules/cell, no NGS-positive cases would be missed by IHC. This cutoff was especially important for cases with overall weak expression of TRK where differing protocols meant falling just above or just below the cutoff (Fig. 2). There was an excellent correlation between the percent positive cells and the ability to detect weak nuclear or membranous staining when the above-defined LOD was achieved. The assay with the worst sensitivity (highest LOD) was the Roche Pan-TRK Assay when performed strictly adhering to the manufacturer's protocol (Table 2). When the Roche Pan-TRK Assay IHC protocol was modified by extending the incubation time for the RTU antibody, the analytical sensitivity of the IHC protocols was improved.

IHC Staining Patterns

Three different patterns of staining were observed, including cytoplasmic, nuclear, and membranous with various combinations in different cases. A study case that showed all 3 patterns of staining is illustrated in Figure 3. There was a tendency for the least-sensitive assays to miss weak staining in one or more of the cellular compartments, especially weak nuclear and focal weak membranous staining (Figs. 3 and 4). By self-assessment, participating pathologists reported cytoplasmic-only staining in 7 samples, nuclear and cytoplasmic in 9 samples, and cytoplasmic and membranous in 3 samples. No pathologist reported

Table 2
The measured LOD and overall sensitivity and specificity of pan-TRK IHC protocols against NGS results

Laboratory	Phase 2 LOD ($\times 10^3$ molecules/cell)	Any positivity readout sensitivity (%)	Any positivity readout specificity (%)	1% Readout sensitivity (%)	1% Readout specificity (%)
1	>200	100	56	100	67
2	70	100	70	100	70
3	76	100	67	100	67
4	151	100	56	100	56
5	40	100	50	100	50
6	64	100	70	100	70
7	69	100	60	100	70
8	49	95 (100)[a]	40	95 (100)[a]	40
9	146	79	90	79	90
10	173	84	80	84	80
11	48	89 (100)[a]	80	89 (100)[a]	80
12	78	100	60	100	60
13	38	100	50	100	50
14	>200	84	80	84	80
15	169	100	80	100	80
16	52	100	60	100	60
17	152	100	60	100	60
18	67	100	70	100	70

LOD, limits of detection.
[a] The results in the parentheses indicate corrected sensitivity results following expert readout/assessment.

membranous-only, nuclear-only, or combined membranous and nuclear pattern without cytoplasmic staining. The distribution of patterns of staining in relation to NGS results is shown in Table 1.

Instrument-Based Results

Optimal protocols could be reached on all instruments used by participating laboratories. Similar or same LDT IHC protocols developed on the Roche or Leica Biosystems instruments achieved comparable results when tested in different laboratories using the corresponding instruments. This indicated that methodology transfer for this assay appears to be broadly feasible even for LDT protocols with these platforms. However, LDT IHC protocols were not transferable when developed/performed on a Dako Omnis instrument. Although it is entirely possible to develop optimal protocols on the Dako Omnis, our experience showed that laboratories that developed significantly different protocols for this instrument sometimes achieved the same or similar results (Fig. 5), whereas when they used the same protocols they sometimes achieved significantly different results. It is not clear whether this may be due to differences in instrument calibration or some other reason. The final protocols used by laboratories with Ventana/Roche platform are depicted in Supplementary Table S2.

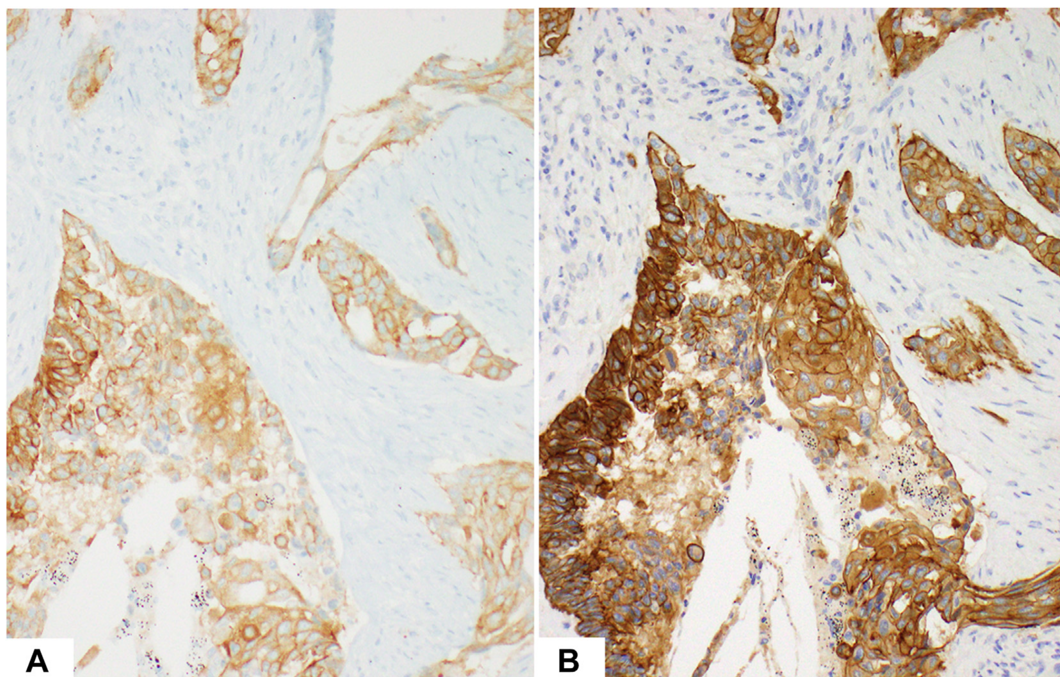

Figure 1.
Tumors with the highest expression of TRK were positive in IHC protocols with low analytical sensitivity (A) and high analytical sensitivity (B). IHC, immunohistochemistry.

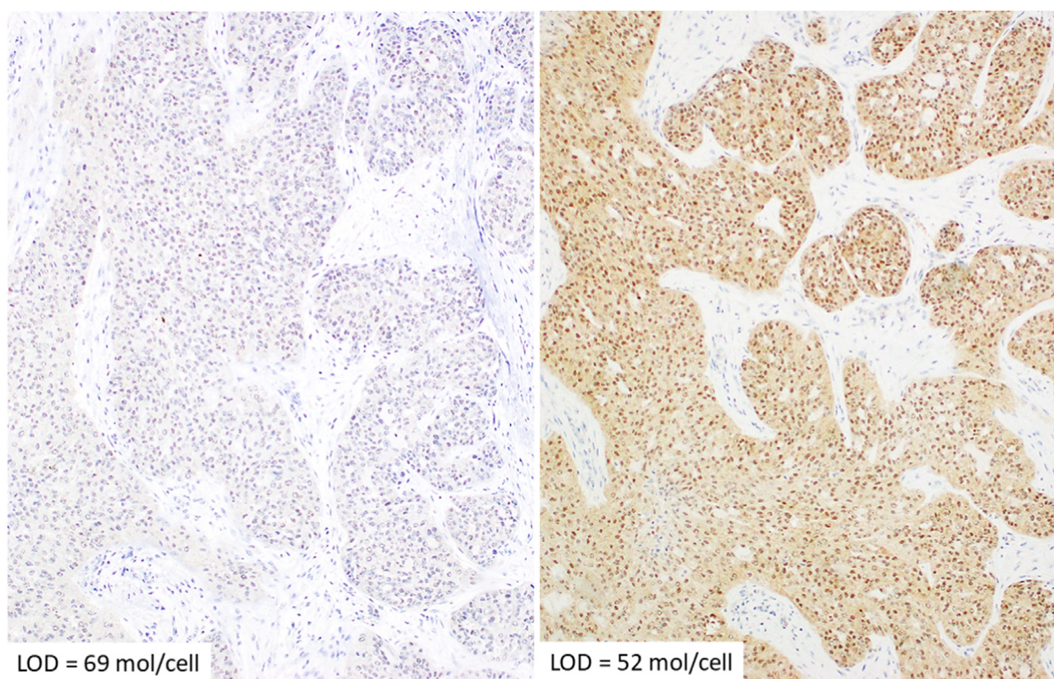

Figure 2.
Even smaller differences in measured analytical sensitivity seemed to be critical for low-positive tumors. In our study, 60,000 molecules/cell was a threshold that needed to be achieved to demonstrate positivity in weakly positive tumors unequivocally. The same tumor when stained with the IHC protocol with an LOD of 69,000 molecules/cell was negative and when stained with the IHC protocol with an LOD of 52,000 molecules/cell was positive. IHC, immunohistochemistry; LOD, limit of detection.

Optimal protocols for the other 2 platforms were also included in Supplementary Table S2.

Preanalytical Issues

Some heterogenous staining patterns observed in large surgical resections suggested poor tissue preservation secondary to delayed fixation, especially in the center of samples. This is illustrated in Figure 6.

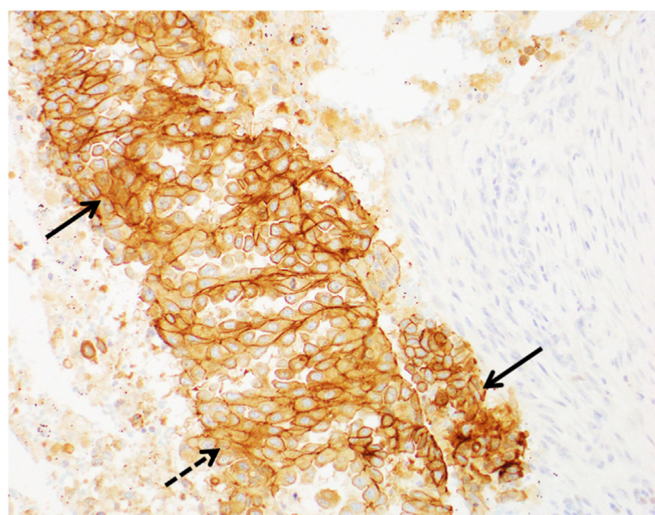

Figure 3.
This tumor shows all 3 patterns of staining including cytoplasmic, membranous, and nuclear (arrows). Weak nuclear staining is easily missed if not specifically searched for (dashed arrow).

Diagnostic Sensitivity and Specificity Against NGS Results

During the analysis it became apparent that 1 sample (CANTRK-22) gave a “false-negative” result by all participants/laboratories. This was a case of infantile fibrosarcoma of the lung, a tumor well known for its association with *ETV6::NTRK3* and which had previously undergone molecular diagnosis for this fusion at the original laboratory. In the parallel CANTRK-NGS study,[25] this fusion could not be detected by all laboratories, whereas all other fusions were confirmed. The review showed that the tissue provided by the contributing laboratory showed prominent autolysis and did not appear to be representative of the tumor. Therefore, this sample was eliminated from the analysis and the sensitivity and specificity values presented in Table 2 were calculated after the exclusion of sample CANTRK-22.

The overall results are shown in Tables 1 and 2. Defining a positive result either as any positivity or as the presence of staining in any of the 3 compartments in >1% of cells gave similar results with minimal differences. Specifically, the diagnostic sensitivity values remained the same, whereas the diagnostic specificity values increased slightly for 2 of the laboratories. The results of the sensitivities and specificities are also summarized in Figure 7. The overall agreement between the study center pathologists and the central pathologist (E.T.) was very high; it was 100% in 8 laboratories, 97% in 3 laboratories, 96% in 1 laboratory, 93% in 5 laboratories, and 90% in 1 laboratory (Supplementary Fig. S1.). The histogram of all self-assessment results for all cases is summarized in Supplementary Figure S2. There was no association in the level of overall agreement and the number of positive vs negative cases. For example, there was 100% agreement in the readout for the laboratory with the highest number of negative results.

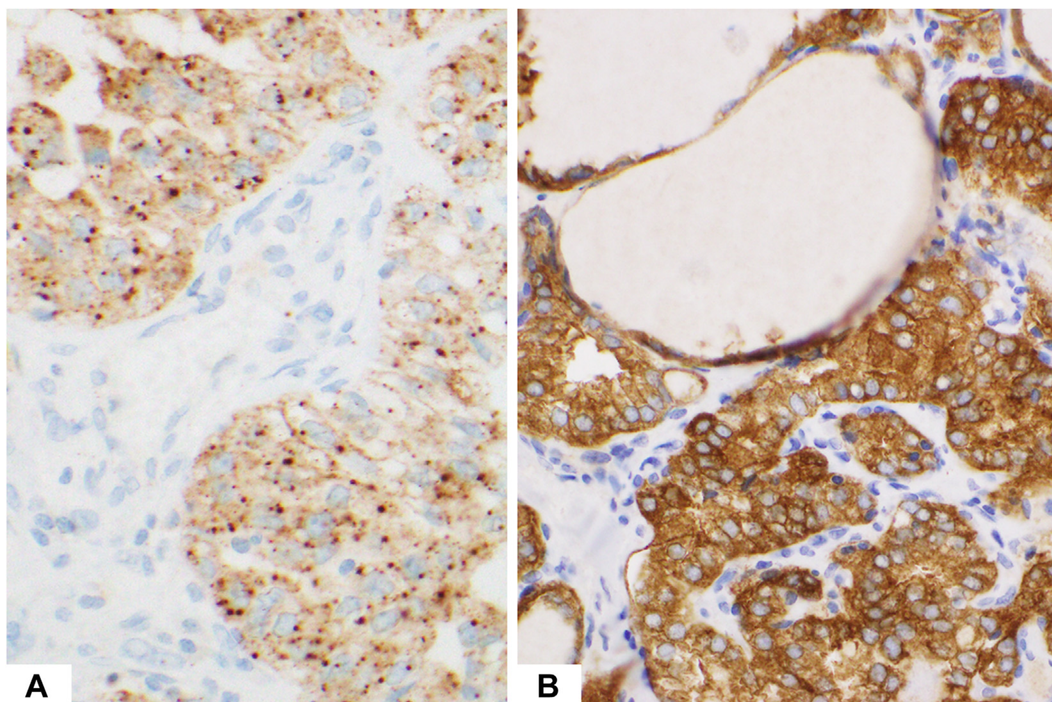

Figure 4.
Granular cytoplasmic positivity may be an artifact of IHC protocol (A) but also real (B). The artifactual granular and often dot-like positivity not only in the cytoplasm but also in the nuclei and the plasma membrane are typically seen on the Roche/Ventana platform with OptiView detection system when OptiView Amplification is added at the end. This is illustrated in (A) with a tumor that showed diffuse cytoplasmic positivity with any of the other detection systems. Some tumors appear to have genuine cytoplasmic granularity, which is illustrated in (B) and was demonstrated by most IHC protocols irrespective of the platform. IHC, immunohistochemistry.

Discussion

The objective of this study was to harmonize the pan-TRK IHC screening testing among 18 regional Canadian pathology laboratories using pan-TRK monoclonal antibody clone EPR17341. Once optimized, the pan-TRK IHC test would then be used to screen groups of tumors known for their low prevalence of TRK fusions and perform confirmatory molecular testing on IHC-positive tumors. The ring study design was selected as it allows for determination of performance characteristics of a diagnostic test using identical samples in multiple laboratories under the control of a supervising institution(s). This study design has previously been used successfully to implement *ALK*, *ROS1*, and PDL1 testing across Canadian laboratories.[27-29] A similar approach has also been recently used for pan-TRK IHC testing in Belgium.[30]

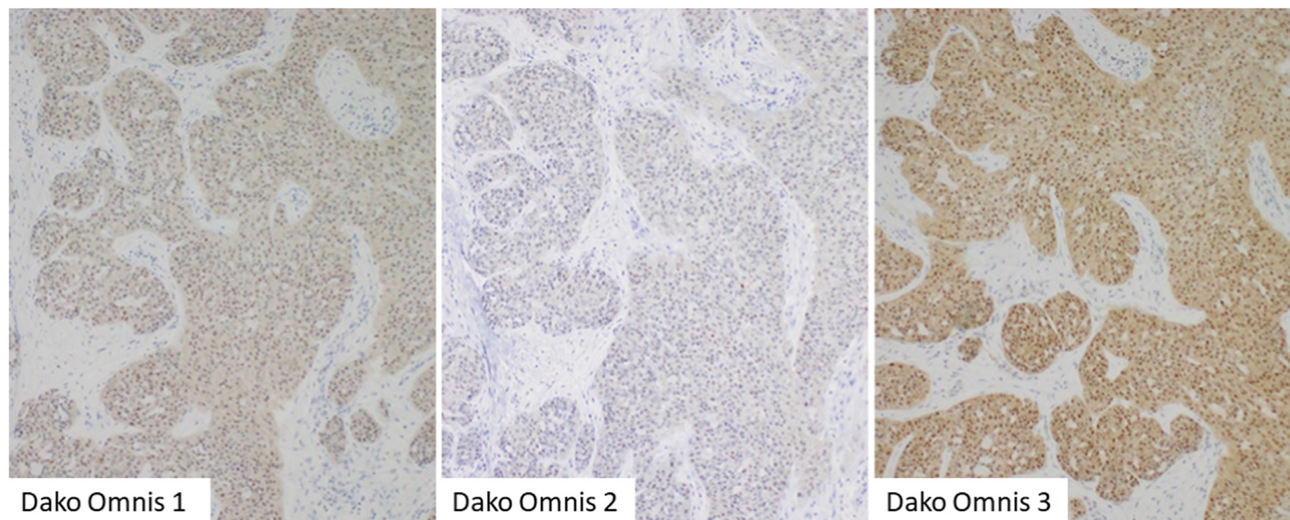

Figure 5.
Same tumor stained on a DAKO Omnis instrument in 3 different laboratories. Omnis 1 uses a protocol with 1:1000 primary Ab dilution and only 1 linker, whereas Omnis 2 and Omnis 3 use the same protocol with 1:200 primary Ab dilution and 2 linkers.

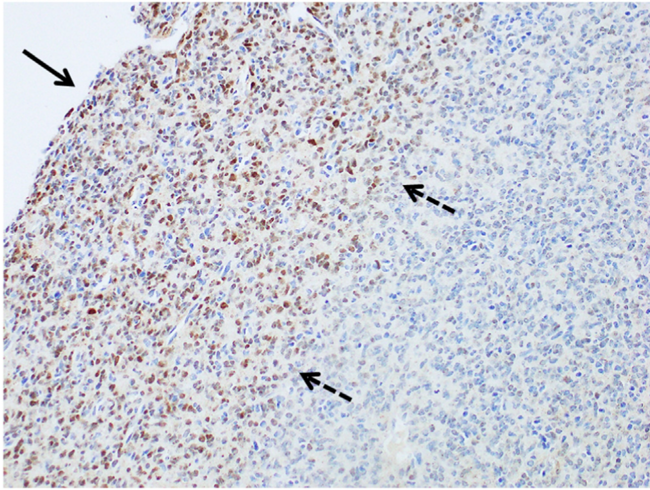

Figure 6.
This tumor shows a sharp line (dashed arrows) between the surface (arrow) that was in contact with formalin sooner and the deeper areas of the tumor. This result shows that the epitope detected by EPR17341 may be sensitive to prolonged cold ischemia, which needs to be taken into consideration for all resection specimens.

Although different studies, including ours, clearly indicate that it is entirely possible to develop 100% sensitivity for a pan-TRK IHC assay, even with cases of *ETV6::NTRK3* fusion known to produce less-intense expression/staining with IHC testing, the current approach to the readout of IHC slides for this biomarker calls for improvement because of a strong trend for low diagnostic specificity with highly sensitive protocols. A small number of negative samples were the major limitation in our study and therefore, the results of the specificity cannot be extrapolated to IHC protocol performance in clinical practice. Another limitation of the study is that because of the need to include as broad a representation of different tumor and fusion types as possible, the positivity rate of the case set (18/29 or 62%) does not reflect the real-world positivity rate in most tumors, which is below 2%. This may limit the applicability of our observations. However, because our positive case set included several low-positive cases, it clearly shows that the reportable range of protein expression of TRK that is of clinical importance is very wide and that highly sensitive protocols are required to detect all true positive cases.

The purpose of a pan-TRK IHC screening assay is not to directly identify patients for targeted therapy but to identify tumors that should be further tested by NGS for the confirmation of the presence of TRK fusions. The main objective of the assay is to decrease the cost of testing and save tissue that would be wasted in a large number of cases from unnecessary testing with NGS. The expected level of diagnostic sensitivity and specificity are therefore different than in other screening assays (eg, testing asymptomatic population for the presence of disease). There is an important distinction between these different types of screening assays, and the term "*technical screening assay*" could be applied to pan-TRK IHC to emphasize this distinction. Therefore, a "*technical screening assay*" should have 100% sensitivity, so as not to miss any positive cases, and as good a specificity as possible. In this study, 15 of 18 laboratories achieved a diagnostic sensitivity of 100%, whereas the diagnostic specificity ranged from 40% to 90%. It should also be noted that achievement of 100% diagnostic specificity is not realistic as some tumors express TRK proteins in the absence of gene rearrangements, most likely due to upregulation of *NTRK1-3* gene expression[31] or due to normal expression of the TRK protein. Overexpression of wild-type TRK protein is a possible explanation for the 2 control samples in our study showing significant cytoplasmic staining (CANTRK-29 and -30). Previous studies have reported pan-TRK IHC specificity levels from 30% to 70%.[19-24,30] Our results using the same cutoffs were nearly identical with specificity levels ranging from 40% to 80%. However, our study also shows that based on the large variation in analytical sensitivity, it should be possible to modify the IHC protocols to optimize diagnostic sensitivity and specificity, as shown by several laboratories that achieved superior results. Supplementary Table S2, in addition to Ventana/Roche protocols, also details the best protocols for the other 2 IHC platforms used in this study. Dako Omnis protocols are not specifically included, because there was a great variation in IHC protocol conditions producing identical results and the other way around, leading to uncertainty about which protocols could be recommended.

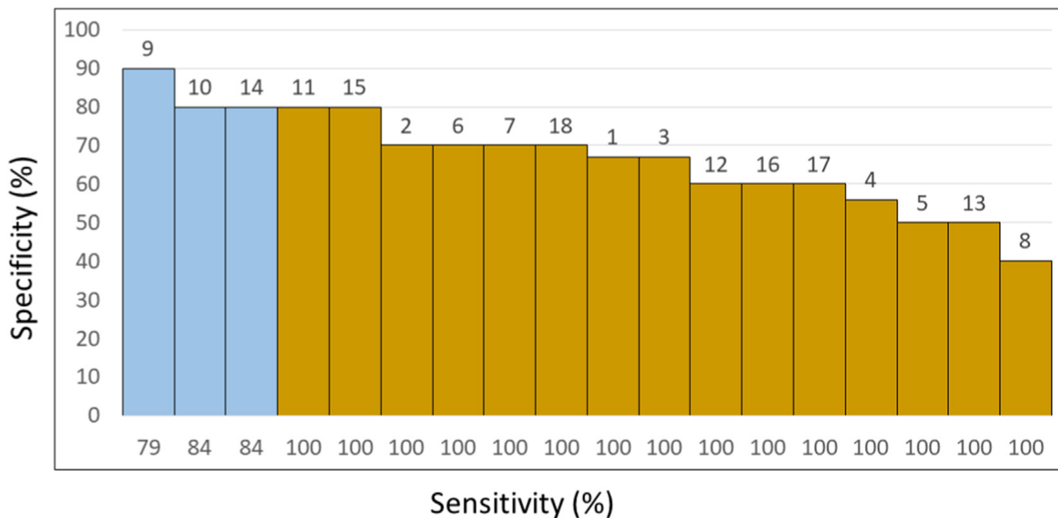

Figure 7.
Results of sensitivity vs specificity are shown for all laboratories. A specificity of 80% was achieved by only 2 laboratories that had 100% sensitivity. However, the number of negative samples was very small in this study, and the overall results of specificity cannot be extrapolated to protocol performance in clinical practice. The numbers above the bars represent the laboratory number, which is also listed in Table 2.

The determination of diagnostic specificity and sensitivity values of the pan-TRK IHC screening test is complicated by several parameters: (1) the presence of cases without TRK fusions that show expression of TRK, (2) IHC protocols that cause nonspecific staining (in tumors or cells that do not express any TRK protein), (3) the impact of preanalytical conditions (specifically delayed fixation) that can radically change the IHC results with EPR17341, (4) potential interobserver variation in the pathologist's readout, and (5) the absence of a reproducible IHC scoring method or scoring method with a cutoff that would help exclude most cases that show TRK protein expression but do not have TRK fusions.

There is no established "gold standard" for immunohistochemical scoring of TRK protein expression, and therefore, we used NGS as a gold standard in our study, with all samples previously sequenced as described in a companion CANTRK publication.[25] Previous studies have used either the presence of any staining above the background as positive[22] or the presence of any level of staining in 1% of cells or more.[20,23,24] Both methods were compared in our study and the differences in the resulting sensitivity and specificity values were minimal. However, this approach also led to acceptance in clinical practice of protocols that have a satisfactory sensitivity of 100% but quite low specificity (eg <80%). Ideally, a higher threshold for positivity could be used with optimal IHC protocols with high analytical sensitivity and specificity. For example, 5 of 6 laboratories that employed VENTANA pan-TRK (EPR17341) Assay following the IHC protocols proposed in the specification sheet demonstrated very low sensitivity with a measured LOD range of 146,000 to >200,000 molecules/cell. However, 3 of 4 laboratories that developed their LDT with the pan-TRK RTU antibody achieved a better sensitivity with LOD of less than 80,000 molecules/cell; one of the laboratories that achieved an LOD of 40,000 showed 100% sensitivity (see this protocol in Supplementary Table S2).

Given that the participating pathologists have not received any training for establishing "any positivity" or "1% cutoff" and that all pathologists provided readouts only on their own stained slides, it was expected that there would be a degree of personal bias of what is considered positive and what is negative based on personal previous experience with this biomarker. There was also a clear tendency to subtract background staining, which varied greatly depending on the level of nonspecific staining between different protocols. However, the central review was conducted by a single experienced pathologist (E.T.) and performed by using the same criteria with the same threshold for all slides irrespective of protocol characteristics. Despite the above, the concordance with central readout review for positive vs negative results was very high (between 90% and 100%) (Supplementary Fig. S1). The distribution of the scores by local pathologists on their slides is shown in Supplementary Fig. S2. One result of the central readout review was the demonstration that a low diagnostic sensitivity of the pan-TRK assay may occur due to the pathologist ignoring weak membranous and/or nuclear staining resulting in a false-negative call. This issue should be preventable by conducting training of the pathologists who will interpret TRK IHC at their institutions, similar to training received for other biomarkers.

The current study also shows that somewhat surprisingly, all laboratories that used the Ventana pan-TRK Assay (Roche) according to the original protocol conditions in the specification sheet of the assay, achieved the lowest analytical sensitivity and overall had the lowest diagnostic sensitivity against NGS results as detailed above. Based on these results, we suggest that all laboratories that use the Ventana pan-TRK assay modify the IHC protocol to develop an IHC assay of higher sensitivity. Based on the results with TRK BCS calibrators, we suggest developing LDT protocols with an LOD of 60,000 molecules/cell or less. The protocols that achieved such results while maintaining acceptable specificity are illustrated in Supplementary Table S2 and could be used as a starting point in protocol development. We therefore recommend that all laboratories that use Ventana pan-TRK Assay should undertake to optimize it first to increase its sensitivity. Although LDTs are generally not readily transferable or reproducible from one laboratory to another, for this particular IHC assay on this particular instrument platform, the set of IHC protocol conditions required for optimal results were either identical or nearly identical. Alternatively, laboratories could consider using the Abcam antibody on Ventana platform.[30,32] We have tried this in the central laboratory on reference materials (cell lines, tissue controls, and calibrators) with satisfactory results; however, as this was not part of the study design, this alternative was not tested by the participating laboratories.

The use of reference standards in the form of TRK calibrators enabled us to determine the LOD (60,000 molecules/cell) required for optimal results with any EPR17341 Ab IHC assay used for the purpose of detecting expression of TRK protein secondary to TRK fusions. We recommend as the first step in pan-TRK IHC protocol development with EPR17341 Ab to develop IHC protocol conditions that will reach analytical sensitivity at this level first, before the protocol is further validated for its technical or diagnostic accuracy with tumor cases. The results of our study also show that LTD protocols could differ significantly from each other but produce the same LOD and subsequently the same accuracy against the NGS results. Although this conclusion is generally applicable, it was particularly the case for Dako Omnis instruments.

TRK staining can be observed in any of the 3 cellular compartments: nuclear, cytoplasmic, and membranous, and any of these 3 patterns of staining or their combination is acceptable as positive. Weak nuclear and weak membranous staining were not specifically searched for by the site pathologists and it was missed especially when present only focally. This was most frequently the case when there was strong cytoplasmic positivity in the tumors, which obscured the weak staining in the nuclei and cell membranes. We also emphasize that when weak cytoplasmic staining is noted (just above background), a careful search for either weak nuclear or membranous positivity may be helpful in confirming the specificity of the weak cytoplasmic staining, which could otherwise be dismissed as nonspecific staining by pathologists. It was noted that in one laboratory with an analytically highly sensitive assay, self-assessment showed less than 100% diagnostic sensitivity due to an arbitrarily high threshold for what was interpreted as "positive" vs "negative" results. We also noted that laboratories that used dark hematoxylin staining produced slides that were more difficult to assess for weak staining, especially weak nuclear and membranous staining, which ultimately decreased their diagnostic sensitivity despite optimal measured LOD. Indeed, the laboratory that developed a protocol with the highest analytical sensitivity (measured LOD of 38,000 molecules/cell) produced diagnostic sensitivity of less than 100% by self-assessment because 2 positive cases were missed due to hematoxylin concealing the staining in the nuclei and the cytoplasmic membranes. Therefore, we highly recommend using less-intense hematoxylin staining for pan-TRK IHC LDT protocols.

When no guidelines for the readout/interpretation/scoring exist, it is typical that pathologists adjust the thresholds based on the overall performance of the assay as well as their experience with other scoring systems. This is a suboptimal approach as each predictive biomarker assay may require different considerations for the readout based on many different potentially important parameters. These considerations are beyond the scope of this

work and will be reported in a follow-up publication. However, for the reasons discussed above, false-positive readouts are unavoidable, and therefore, we recommend that any IHC result that appears positive be reported instead as “Pending confirmatory assay results” to avoid distress and confusion when the confirmatory tests are negative. For the same reasons, we recommend against using a result of technical screening assays such as pan-Trk IHC to guide treatment decisions directly in cases in which the amount or quality of RNA is insufficient for confirmatory NGS assays. Instead, a repeat biopsy should be considered. It is possible to develop and analytically harmonize multi-institutional pan-TRK IHC testing with optimal diagnostic sensitivity for the technical screening assay by using reference standards that enable either their measured (TRK IHC calibrators) or descriptive (tissues and cell lines) analytical sensitivity. Irrespective of the IHC instrument platform or detection system, the recommended LOD for the IHC assays using EPR17341 Ab is 60×10^3 molecules/cell.

Delayed fixation may have a significant impact on IHC results with EPR17341. If the impact of delayed fixation can be determined by pathologists, the results of the staining should not be generalized, but assessed on the area with best fixation.

Acknowledgments

We thank Bayer Canada, and especialy Haya Sarras and Bahiyyih Schmalenbe, for providing critical logistic support during the conduct of this study.

Authors Contributions

M.S.T., E.T., and S.N.M.F. designed the study. M.D.H., A.S., H.J.W., B.M.P., P.D., P.C.P., G.B., S.J., J.C.C., and Z.X. contributed samples for the study; all contributed to data generation and review of results. E.T., M.D.H., and M.S.T. performed data analysis. M.D.H., S.N.M.F., M.S.T., and E.T. contributed to primary writing of the manuscript. All authors reviewed and contributed to the editing of the manuscript.

Data Availability

The original data from this study is available for review by request.

Funding

This study was funded by Bayer Inc. in the capacity of institution-sponsored collaborative research funding. Bayer has no role in the design, conduct, analysis, or interpretation of the results of this study, and in the drafting of this manuscript.

Declaration of Competing Interests

The authors declare the following financial interests/personal relationships that may be considered as potential competing interests: Dr Hyrcza received consultation and speaker’s fees from Bayer Canada and consultation fees from Roche. Dr Tsao received a research grant, consultation fees, and speaker’s fees from Bayer Canada. Dr Torlakovic received consultation fees from Bayer Canada and Roche. Dr Cheung received grants from Bayer Canada. Dr Ionescu received advisory board consultation fees from Bayer Canada and Roche. Dr Spatz received advisory board consultation fees from Bayer Canada and Roche; he also participated in supported biomarker programs organized by Bayer Canada and Roche. Dr Stockley received advisory board consultation fees from Bayer Canada. Dr Leduc received advisory board consultation fees from Bayer Canada. Dr Desmeules received advisory board consultation fees and research support from Bayer Canada and Roche. Dr Hwang received advisory board consultation fees from Bayer Canada and Roche and received a quality improvement grant from Bayer Canada. Dr Sheffield received consulting and advisory board honoraria as well as research support from Bayer Canada and Roche. Dr Cheung has received advisory board fees from Bayer Canada. The remaining authors declare no conflict of interest.

Ethics Approval

Ethics approvals were obtained from the coordinating site (University Health Network) for the entirety of the study and locally from all participating institutions.

Supplementary Material

The online version contains supplementary material available at https://doi.org/10.1016/j.modpat.2023.100384

References

1. Amatu A, Sartore-Bianchi A, Bencardino K, Pizzutilo EG, Tosi F, Siena S. Tropomyosin receptor kinase (TRK) biology and the role of NTRK gene fusions in cancer. *Ann Oncol*. 2019;30(Suppl_8):viii5−viii15. https://doi.org/10.1093/annonc/mdz383
2. Chao MV. Neurotrophins and their receptors: a convergence point for many signaling pathways. *Nat Rev Neurosci*. 2019;4:299−309. https://doi.org/10.1038/nrn1078
3. Frattini V, Trifonov V, Chan JM, et al. The integrated landscape of driver genomic alterations in glioblastoma. *Nat Genet*. 2013;45:1141−1149. https://doi.org/10.1038/ng.2734
4. Knezevich SR, McFadden DE, Tao W, Lim JF, Sorensen PH. A novel ETV6-NTRK3 gene fusion in congenital fibrosarcoma. *Nat Genet*. 1998;18:184−187. https://doi.org/10.1038/ng0298-184
5. Skálová A, Vanecek T, Sima R, et al. Mammary analogue secretory carcinoma of salivary glands, containing the ETV6-NTRK3 fusion gene: a hitherto undescribed salivary gland tumor entity. *Am J Surg Pathol*. 2010;34:599−608. https://doi.org/10.1097/PAS.0b013e3181d9efcc
6. Farago AF, Taylor MS, Doebele RC, et al. Clinicopathologic features of non−small-cell lung cancer harboring an NTRK gene fusion. *JCO Precis Oncol*. 2018;2018:PO.18.00037. https://doi.org/10.1200/PO.18.00037
7. Bocciarelli C, Caumont C, Samaison L, et al. MSI-High RAS-BRAF wild-type colorectal adenocarcinomas with MLH1 loss have a high frequency of targetable oncogenic gene fusions whose diagnoses are feasible using methods easy-to-implement in pathology laboratories. *Hum Pathol*. 2021;114:99−109. https://doi.org/10.1016/j.humpath.2021.05.006
8. Skálová A, Hyrcza MD, Vaneček T, Baněčková M, Leivo I. Fusion-positive salivary gland carcinomas. *Genes Chromosomes Cancer*. 2022;61:228−243. https://doi.org/10.1002/gcc.23020
9. Cocco E, Scaltriti M, Drilon A. NTRK fusion-positive cancers and TRK inhibitor therapy. *Nat Rev Clin Oncol*. 2018;15:731−747. https://doi.org/10.1038/s41571-018-0113-0
10. Forsythe A, Zhang W, Phillip Strauss U, Fellous M, Korei M, Keating K. A systematic review and meta-analysis of neurotrophic tyrosine receptor kinase gene fusion frequencies in solid tumors. *Ther Adv Med Oncol*. 2020;12:1758835920975613. https://doi.org/10.1177/1758835920975613
11. Hsiao SJ, Zehir A, Sireci AN, Aisner DL. Detection of tumor NTRK gene fusions to identify patients who may benefit from tyrosine kinase (TRK) inhibitor therapy. *J Mol Diagn*. 2019;21:553−571. https://doi.org/10.1016/j.jmoldx.2019.03.008
12. Koehne de González A, Mansukhani MM, Fernandes H, Hsiao SJ. Pan-tumor screening for NTRK gene fusions using pan-TRK immunohistochemistry and RNA NGS fusion panel testing. *Cancer Genet*. 2022:262−263:47−52. https://doi.org/10.1016/j.cancergen.2021.12.010

13. Drilon A, Laetsch TW, Kummar S, et al. Efficacy of larotrectinib in TRK fusion-positive cancers in adults and children. *N Engl J Med*. 2018;378:731–739. https://doi.org/10.1056/NEJMoa1714448
14. Hong DS, DuBois SG, Kummar S, et al. Larotrectinib in patients with TRK fusion-positive solid tumors: a pooled analysis of three phase 1/2 clinical trials. *Lancet Oncol*. 2020;21:531–540. https://doi.org/10.1016/S1470-2045(19)30856-3
15. Drilon A, Siena S, Ou SI, et al. Safety and antitumor activity of the multitargeted pan-TRK, ROS1, and ALK inhibitor entrectinib: combined results from two phase I trials (ALKA-372-001 and STARTRK-1). *Cancer Discov*. 2017;7:400–409. https://doi.org/10.1158/2159-8290.CD-16-1237
16. Bebb DG, Banerji S, Blais N, et al. Canadian consensus for biomarker testing and treatment of TRK fusion cancer in adults. *Curr Oncol*. 2021;28:523–548. https://doi.org/10.3390/curroncol28010053
17. Yoshino T, Pentheroudakis G, Mishima S, et al. JSCO-ESMO-ASCO-JSMO-TOS: international expert consensus recommendations for tumor-agnostic treatments in patients with solid tumors with microsatellite instability or NTRK fusions. *Ann Oncol*. 2020;31:861–872. https://doi.org/10.1016/j.annonc.2020.03.299
18. Marchiò C, Scaltriti M, Ladanyi M, et al. ESMO recommendations on the standard methods to detect NTRK fusions in daily practice and clinical research. *Ann Oncol*. 2019;30:1417–1427. https://doi.org/10.1093/annonc/mdz204
19. Hechtman JF, Benayed R, Hyman DM, et al. Pan-Trk immunohistochemistry is an efficient and reliable screen for the detection of NTRK fusions. *Am J Surg Pathol*. 2017;41:1547–1551. https://doi.org/10.1097/PAS.0000000000000911
20. Van Bockstal MR, Beniuga G, Craciun L, et al. The use of pan-tropomyosin receptor kinase immunohistochemistry as a screening tool for the detection of neurotrophic tropomyosin-related kinase fusions: real-world data from a national multicentric retrospective study. *Pathobiology*. 2022;29:1–14. https://doi.org/10.1159/000522426
21. Koopman B, Kuijpers CCHJ, Groen HJM, et al. Detection of NTRK fusions and TRK expression and performance of pan-TRK immunohistochemistry in routine diagnostics: results from a nationwide community-based cohort. *Diagnostics*. 2022;12:668. https://doi.org/10.3390/diagnostics12030668
22. Mohamed F, Kurdi M, Baeesa S, et al. The diagnostic value of pan-Trk expression to detect neurotrophic tyrosine receptor kinase (NTRK) gene fusion in CNS tumors: A study using next-generation sequencing platform. *Pathol Oncol Res*. 2022;28:1610233. https://doi.org/10.3389/pore.2022.1610233
23. Brčić I, Godschachner TM, Bergovec M, et al. Broadening the spectrum of NTRK rearranged mesenchymal tumors and usefulness of pan-TRK immunohistochemistry for identification of NTRK fusions. *Mod Pathol*. 2021;34:396–407. https://doi.org/10.1038/s41379-020-00657-x
24. Solomon JP, Linkov I, Rosado A, et al. NTRK fusion detection across multiple assays and 33,997 cases: diagnostic implications and pitfalls. *Mod Pathol*. 2020;33:38–46. https://doi.org/10.1038/s41379-019-0324-7
25. Stockley TL, Lo B, Box A, et al. CANTRK: a Canadian ring study to optimize detection of NTRK gene fusions by next-generation RNA sequencing. *J Mol Diagn*. 2023;25:168–174. https://doi.org/10.1016/j.jmoldx.2022.12.004
26. Sompuram SR, Vani K, Tracey B, Kamstock DA, Bogen SA. Standardizing immunohistochemistry: a new reference control for detecting staining problems. *J Histochem Cytochem*. 2015;63:681–690. https://doi.org/10.1369/0022155415588109
27. Cutz JC, Craddock KJ, Torlakovic E, et al. Canadian anaplastic lymphoma kinase study: a model for multicenter standardization and optimization of ALK testing in lung cancer. *J Thorac Oncol*. 2014;9:1255–1263. https://doi.org/10.1097/JTO.0000000000000239
28. Cheung CC, Smith AC, Albadine R, et al. Canadian ROS proto-oncogene 1 study (CROS) for multi-institutional implementation of ROS1 testing in non-small cell lung cancer. *Lung Cancer*. 2021;160:127–135. https://doi.org/10.1016/j.lungcan.2021.08.003
29. Torlakovic E, Albadine R, Bigras G, et al. Canadian multicenter project on standardization of programmed death-ligand 1 immunohistochemistry 22C3 laboratory-developed tests for pembrolizumab therapy in NSCLC. *J Thorac Oncol*. 2020;15:1328–1337. https://doi.org/10.1016/j.jtho.2020.03.029
30. De Winne K, Sorber L, Lambin S, et al. Immunohistochemistry as a screening tool for NTRK gene fusions: results of a first Belgian ring trial. *Virchows Arch*. 2021;478:283–291. https://doi.org/10.1007/s00428-020-02921-6
31. Guibourg B, Cloarec E, Conan-Charlet V, et al. EPR17341 and A7H6R pan-TRK immunohistochemistry result in highly different staining patterns in a series of salivary gland tumors. *Appl Immunohistochem Mol Morphol*. 2020;28:719–724. https://doi.org/10.1097/PAI.0000000000000825
32. Adam J, Stang NL, Uguen A, et al. Multicenter harmonization study of pan-Trk immunohistochemistry for the detection of NTRK3 fusions. *Mod Pathol*. 2023;36:100192. https://doi.org/10.1016/j.modpat.2023.100192