

# Validation of a gene expression-based tissue of origin test applied to poorly differentiated and undifferentiated cancers

FA Monzon<sup>1</sup>, CI Dumur<sup>3</sup>, M Lyons-Weiler<sup>2</sup>, CM Sciulli<sup>2</sup>, CT Garrett<sup>3</sup>, JM Hagenkord<sup>2</sup>, L Buturovic<sup>4</sup>, R Deeter<sup>4</sup>, SH Becker<sup>4</sup>, CT Rigl<sup>4</sup>, GG Anderson<sup>4</sup>

<sup>1</sup>The Methodist Hospital, Houston, TX; <sup>2</sup>University of Pittsburgh, Pittsburgh, PA

<sup>3</sup>Virginia Commonwealth University, Richmond, VA; <sup>4</sup>Pathwork Diagnostics, Sunnyvale, CA

## ABSTRACT

**Background:** The Pathwork™ Tissue of Origin (TOO) Test uses the expression of 1550 distinct genes to quantify the molecular similarity between a tissue biopsy sample and fifteen tissues of origin of clinical relevance. Previous studies described the proprietary microarray (Pathchip™), prediction algorithm and performance across laboratories. Here we describe the results of the clinical validation studies evaluating the performance of the Pathwork TOO Test in a clinical setting.

**Methods:** Fresh frozen tissue biopsy samples for 477 poorly differentiated and undifferentiated primary and metastatic human tumors were obtained from the Health Sciences Tissue Bank at the University of Pittsburgh and commercial tissue repositories. All samples were of known origin (reference diagnosis) as established by a pathologist. Tissue samples were processed in three laboratories with a recommended protocol. Data from the gene expression assay was analyzed with the Pathwork TOO algorithm and a Tissue of Origin report was generated for each sample. The TOO results were compared against the reference diagnoses to establish the performance characteristics of the Pathwork TOO Test.

**Results:** The test revealed overall agreement of 89% with the reference diagnoses for all 15 tissue sites and >92% agreement in at least 8 tissues of origin. If the quantitative result (Similarity Score) exceeded the recommended threshold, the probability that the indicated tissue is present was 95% across all tissue types. If the Similarity Score was less than 5, the probability that the indicated tissue is absent was 98% across all tissue types.

**Conclusions:** The Pathwork Tissue of Origin (TOO) Test successfully identified the origin of primary tumors in 89% of samples tested and provided >92% agreement for at least 8 tissues of origin. This test has the potential to be an effective aid in the diagnosis of cancer patients presenting with poorly differentiated and undifferentiated tumors.

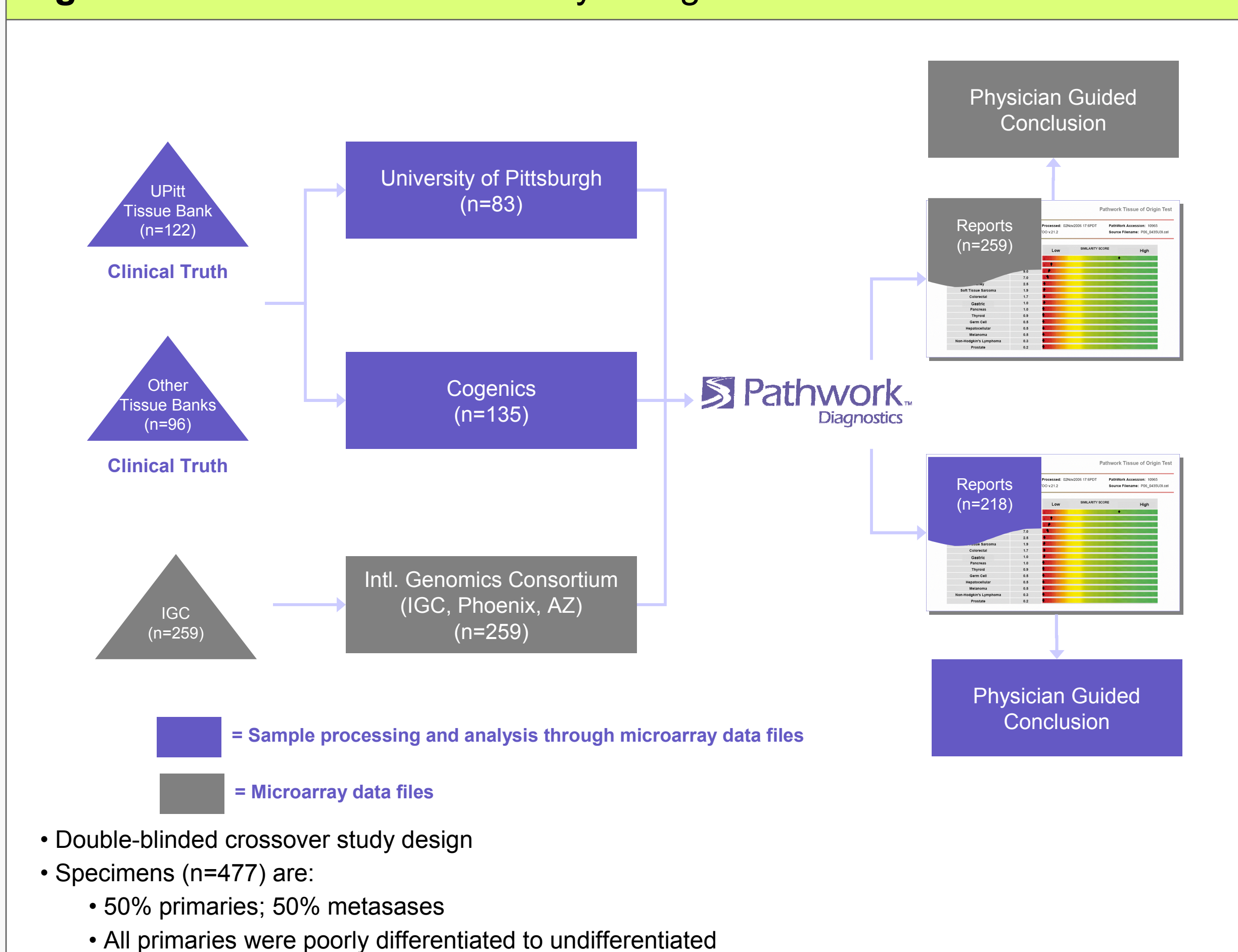
## BACKGROUND

The Pathwork™ Tissue of Origin Test (Pathwork Diagnostics, Sunnyvale, CA) is a microarray-based gene expression diagnostic test for determining the similarity of tumors of unknown origin to cancers from one of 15 known tissues of origin (TOO). The test uses proprietary normalization and classification algorithms and a high-density oligonucleotide microarray (Pathchip™, manufactured by Affymetrix, Inc., Santa Clara, CA) to measure the expression of 1668 gene probe sets or markers. The molecular similarity of each tumor specimen's expression pattern is compared to 15 distinctive patterns from different tissue types: bladder, breast, colorectal, gastric, germ cell, hepatocellular, kidney, non-small-cell lung, non-Hodgkin's lymphoma, melanoma, ovarian, pancreatic, prostate, soft tissue sarcoma, and thyroid. For each specimen, the algorithm reduces the highly complex expression data into 15 separate "Similarity Scores", one score for each potential tissue type. These scores (scale 0 to 100) are reported to the pathologist, who then establishes whether or not a particular tissue type is present in the specimen (termed a Physician Guided Conclusion (PGC), a clinical call). A Similarity Score higher than 30 suggests the presence of the markers that are characteristic of that tissue.

In this study, we evaluate the performance characteristics of the Pathwork TOO Test using a set of 477 clinical samples (metastatic or poorly differentiated or undifferentiated primary malignancies) in a multicenter clinical validation study.

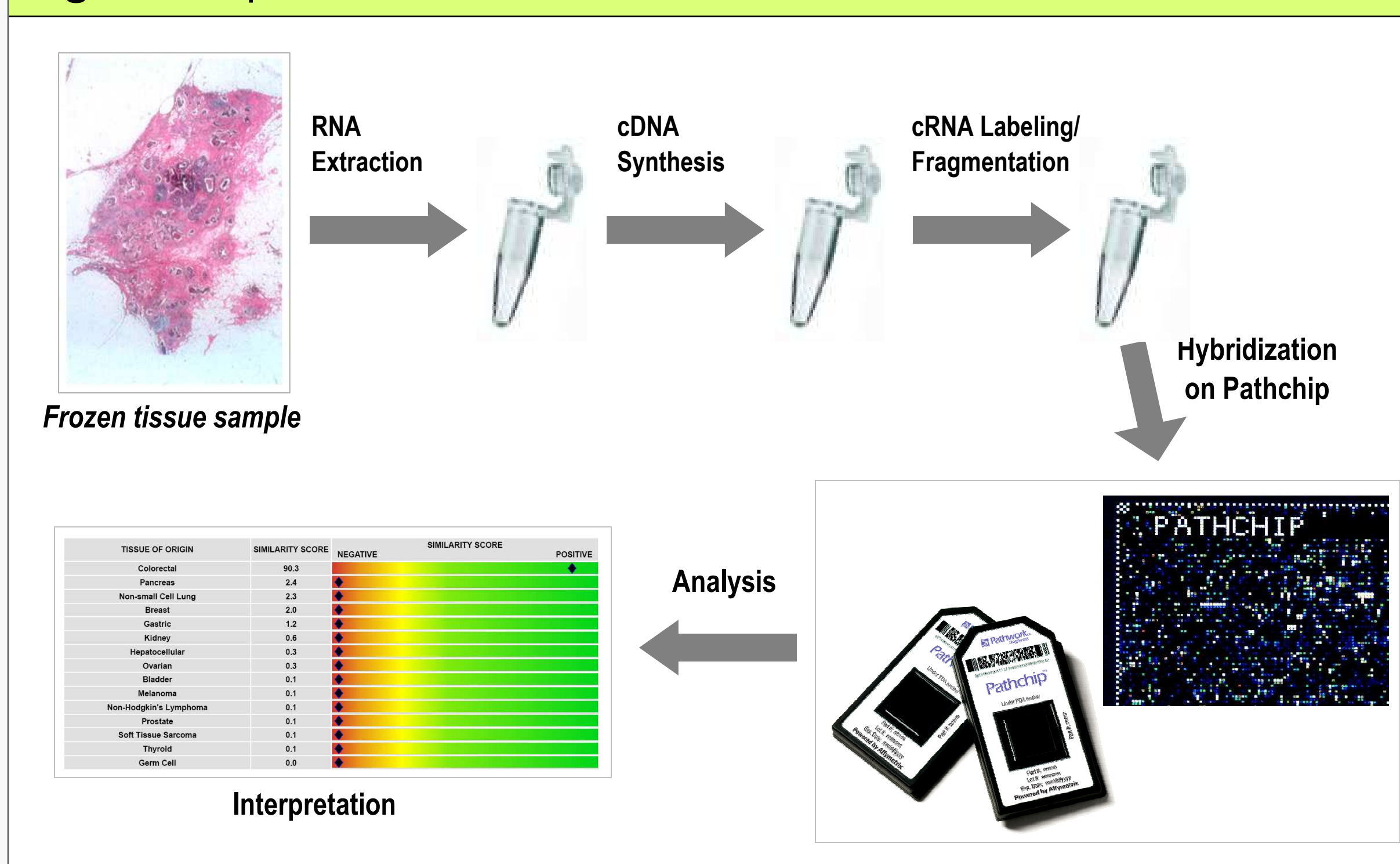
## MATERIALS AND METHODS

**Figure 1. Clinical Validation Study Design**



## RESULTS

**Figure 2. Specimen Workflow**



**Table 1. Tissue of Origin Performance**

Tissue Type	True Positive	Sensitivity (PPA) (95% CI)	Specificity (NPA) (95% CI)
Bladder	28	78.6 [59.0, 91.7]	100.0 [99.2, 100.0]
Breast	68	94.1 [85.6, 98.4]	98.3 [96.5, 99.3]
Colorectal	56	92.9 [82.7, 98.0]	99.5 [98.3, 99.9]
Gastric	17	82.4 [56.6, 96.2]	99.3 [98.1, 99.9]
Germ Cell	12	75.0 [42.8, 94.5]	100.0 [99.2, 100.0]
Hepatocellular	17	88.2 [63.6, 98.5]	99.8 [98.8, 100.0]
Kidney	40	95.0 [83.1, 99.4]	99.8 [98.7, 100.0]
Non-Hodg. Lymph.	24	83.3 [62.6, 95.3]	99.8 [98.8, 100.0]
Non-Sm. Cell Lung	33	93.8 [79.2, 99.2]	99.5 [98.4, 99.9]
Melanoma	31	87.1 [70.2, 96.4]	98.9 [97.4, 99.6]
Ovarian	69	92.8 [83.9, 97.6]	99.8 [98.6, 100.0]
Pancreas	13	76.9 [46.2, 95.0]	99.8 [98.8, 100.0]
Prostate	25	88.0 [68.8, 97.5]	100.0 [99.2, 100.0]
Soft Tiss. Sarcoma	31	83.9 [66.3, 94.5]	99.3 [98.0, 99.9]
Thyroid	13	100.0 [75.3, 100.0]	99.6 [98.4, 99.9]
<b>Overall</b>	<b>477</b>	<b>89.5 [86.4, 92.1]</b>	<b>99.6 [99.4, 99.7]</b>

Sensitivity (PPA) – Positive Percent Agreement,  $100 \times \text{TP} / \text{POS}$ , where TP is the number of True Positive calls for the given tissue of origin and POS is the total number of positive specimens for the given tissue of origin.

Specificity (NPA) – Negative Percent Agreement,  $100 \times (1 - \text{FP} / \text{NEG})$ , where FP is the number of False Positive calls for the given tissue of origin and NEG is the number of negative specimens for the given tissue of origin.

**Table 2. Conditional Probabilities**

Tissue Type	Probability that a Similarity Score $\geq 30$ represents a true positive call for that tissue type in the specimen.		Probability that a Similarity Score $< 5$ represents a true negative call for that tissue type in the specimen.	
	Probability	95% CI	Probability	95% CI
Bladder	100.0	[ 84.6, 100.0 ]	99.5	[ 98.4, 99.9 ]
Breast	90.1	[ 80.7, 95.9 ]	99.4	[ 98.0, 99.9 ]
Colorectal	94.6	[ 85.1, 98.9 ]	93.7	[ 90.7, 96.0 ]
Gastric	82.4	[ 56.6, 96.2 ]	98.6	[ 97.0, 99.5 ]
Germ Cell	100.0	[ 66.4, 100.0 ]	99.3	[ 98.1, 99.9 ]
Hepatocellular	100.0	[ 79.4, 100.0 ]	96.4	[ 94.2, 97.9 ]
Kidney	97.4	[ 86.5, 99.9 ]	99.7	[ 98.6, 100.0 ]
Non-Hodg. Lymph.	100.0	[ 89.4, 100.0 ]	91.1	[ 87.9, 93.6 ]
Non-Sm. Cell Lung	94.1	[ 80.3, 99.3 ]	97.0	[ 94.6, 98.5 ]
Melanoma	95.5	[ 77.2, 99.9 ]	99.8	[ 98.7, 100.0 ]
Ovarian	100.0	[ 94.6, 100.0 ]	99.7	[ 98.2, 100.0 ]
Pancreas	91.7	[ 61.5, 99.8 ]	100.0	[ 99.1, 100.0 ]
Prostate	100.0	[ 84.6, 100.0 ]	99.8	[ 98.8, 100.0 ]
Soft Tiss. Sarcoma	86.7	[ 69.3, 96.2 ]	98.1	[ 96.3, 99.2 ]
Thyroid	93.3	[ 68.1, 99.8 ]	99.6	[ 98.4, 99.9 ]
<b>Overall</b>	<b>95.9</b>	<b>[ 93.6, 97.5 ]</b>	<b>98.1</b>	<b>[ 97.8, 98.5 ]</b>

## RESULTS (Continued)

Rates of agreement were similar at Sites 1, 2, and 3 (88.1%, 90.3%, and 89.2%, respectively). Sensitivity of the test ranged from 100% for thyroid cancer specimens (n=13) to 75% for germ cell malignancy specimens (n=12)(Table 1). Performance of the test was somewhat better with poorly and undifferentiated primary tumors (92.6% Agreement) as compared to metastatic specimens (86.3% Agreement). Conditional probabilities illustrate the likelihood that a true positive tissue call was obtained when a Similarity Score of  $\geq 30$  was reported, or that a true negative tissue call was obtained when a Similarity Score of  $< 5$  was reported (Table 2).

In Table 3, eight out of ten carcinoma of unknown primary (CUP) specimens showed positive results for a tissue of origin. In 6 cases, the tumor origin identified by the TOO test was part of the differential diagnosis of the pathologists. In two cases the identified TOO was not within the differential but is possible. In all cases the TOO test reduces uncertainty and provides information that complements the pathologic workup.

**Table 3. Application to Carcinoma of Unknown Primary**

Case ID	Sex	Sample Site	Grade/ Morphology/Subtype	Diagnostic Possibilities	TOO Result (PGC)	Similarity Score (SS)	Reduces uncertainty?
A	F	right femur	Metastatic well to moderately differentiated adenocarcinoma with focal micropapillary pattern	Breast, ovary, uterine	Breast	96.4	Yes
B	F	left femur	Poorly differentiated metastatic adenocarcinoma	Pancreas, stomach, ovary	Breast	95.0	Yes
C	F	N/A	Metastatic Adenocarcinoma	Ovary, lung	Ovary	87.5	Yes
D	F	lung, left upper lobe	Moderately differentiated papillary adenocarcinoma with psammoma bodies (3.5cm) with focal angiolymphatic invasion and involvement of 1/7 N1 lymph nodes.	Lung, ovary	Lung	66.9	Yes
E	F	omentum	Metastatic poorly differentiated adenocarcinoma	Pancreas, stomach, ovary	Pancreas	39.5	Yes
F	M	abdominal wall	Metastatic poorly differentiated adenocarcinoma	Colon, prostate, stomach, lung	Pancreas	37.3	Yes
G	M	right groin	Metastatic Adenocarcinoma	Esophagus, stomach, pancreas, biliary	Breast	36.1	Yes
H	M	colon, omentum	Poorly differentiated adenocarcinoma, primary site unknown, with extensive seeding of the peritoneal cavity and deposits within the mesentery	Colon	Colorectal	31.5	Yes
I	M	right clavicle	Metastatic moderately differentiated mucinous adenocarcinoma with areas of necrosis, involving the bone and adjacent soft tissue	Esophagus, stomach, pancreas, biliary	Indeterminate	n/a	Yes, rule out nine tissues
J	F	liver, lymph node	Liver - Poorly differentiated carcinoma. Lymph node - Metastatic poorly differentiated carcinoma	Liver, cholangio-carcinoma	Indeterminate	n/a	Yes, rule out nine tissues

## CONCLUSIONS

Our results show that the Pathwork test could reliably identify TOO in approximately nine of every ten metastatic or poorly differentiated specimens. Only in approximately one of every 16 specimens did the TOO Test make an incorrect call of the tissue of origin. In these specimens, the test's overall accuracy across 15 tissue types using conditional probabilities was approximately 95% for positive calls and 98% for negative calls. The diagnostic performance of this test compares favorably with other expression-based TOO classifiers, especially because test performance does not decrease when evaluating poorly differentiated or undifferentiated tumors.

Although the test's performance reported in this study was studied with tumors of known origin (n = 477), when the test was applied to ten metastatic tumors with unknown primary origin after routine pathologic examination it successfully identified a plausible tissue of origin for 8 of these samples.

## REFERENCES

- Affymetrix GeneChip Expression Analysis Technical Manual. Santa Clara, CA. Affymetrix Inc. November 2004. Available at: [http://www.affymetrix.com/support/technical/manual/expression\\_manual.afx](http://www.affymetrix.com/support/technical/manual/expression_manual.afx). Accessed February 5, 2007.
- Dumur C et al. Interlaboratory Performance of a Microarray-Based Gene Expression Test to Determine Tissue of Origin in Poorly Differentiated and Undifferentiated Cancers. J. Molecular Diagnostics (in press).
- Monzon FA, et al. Clinical validation of a gene expression microarray-based tissue of origin test applied to primary and metastatic tumors. Annual Meeting of the Association for Molecular Pathology; November, 2006; Orlando, FL.
- Monzon FA, et al. Clinical Applications of Gene Expression Microarrays: Reproducibility of a Tissue of Origin Test for Metastatic Tumors of Unknown Origin. Annual Meeting of the United States and Canadian Academy of Pathology. March, 2007; San Diego, CA.
- National Comprehensive Cancer Network available at: [http://www.nccn.org/professionals/physician\\_gls/](http://www.nccn.org/professionals/physician_gls/) Accessed March 1, 2007.

