Published in final edited form as:

Am J Surg Pathol. 2013 July ; 37(7): 1067–1075. doi:10.1097/PAS.0b013e31828309c4.

A Multicenter Study Directly Comparing the Diagnostic Accuracy of Gene Expression Profiling and Immunohistochemistry for Primary Site Identification in Metastatic Tumors

Charles R. Handorf, MD, PhD^{*}, Anand Kulkarni, MD^{*}, James P. Grenert, MD, PhD[†], Lawrence M. Weiss, MD[‡], William M. Rogers, MD[§], Oliver S. Kim, MD^{||}, Federico A. Monzon, MD[¶], Meredith Halks-Miller, MD[#], Glenda G. Anderson, MBA^{**}, Michael G. Walker, PhD^{††}, Raji Pillai, PhD[#], and W. David Henner, MD, PhD[#]

^{*}Department of Pathology and Laboratory Medicine, University of Tennessee College of Medicine, Memphis, TN

[†]Department of Pathology, San Francisco General Hospital and Molecular Diagnostics Laboratory, University of California, San Francisco

[‡]Clariant Inc., Aliso Viejo

[§]Department of Pathology, El Camino Hospital, Mountain View

[#]Pathwork Diagnostics, Redwood City

^{**}5Degrees Biosciences, San Jose

^{††}WalkerBioscience, Carlsbad, CA

^{||}Department of Pathology, Advocate Good Shepherd Hospital, Barrington, IL

[¶]Baylor College of Medicine, Houston, TX

Abstract

Metastatic tumors with an uncertain primary site can be a difficult clinical problem. In tens of thousands of patients every year, no confident diagnosis is ever issued, making standard-of-care treatment impossible. Gene expression profiling (GEP) tests currently available to analyze these difficult-to-diagnose tumors have never been directly compared with the diagnostic standard of care, immunohistochemistry (IHC). This prospectively conducted, blinded, multicenter study compares the diagnostic accuracy of GEP with IHC in identifying the primary site of 157 formalin-fixed paraffin-embedded specimens from metastatic tumors with known primaries, representing the 15 tissues on the GEP test panel. Four pathologists rendered diagnoses by selecting from 84 stains in 2 rounds. GEP was performed using the Pathwork Tissue of Origin Test. Overall, GEP accurately identified 89% of specimens, compared with 83% accuracy using IHC ($P = 0.013$). In the subset of 33 poorly differentiated and undifferentiated carcinomas, GEP accuracy exceeded that of IHC (91% to 71%, $P = 0.023$). In specimens for which pathologists rendered their final diagnosis with a single round of stains, both IHC and GEP exceeded 90% accuracy. However, when the diagnosis

required a second round, IHC significantly underperformed GEP (67% to 83%, $P < 0.001$). GEP has been validated as accurate in diagnosing the primary site in metastatic tumors. The Pathwork Tissue of Origin Test used in this study was significantly more accurate than IHC when used to identify the primary site, with the most pronounced superiority observed in specimens that required a second round of stains and in poorly differentiated and undifferentiated metastatic carcinomas.

Keywords

metastatic cancer; poorly differentiated cancer; tumors of uncertain origin; immunohistochemistry; primary site of tumor; gene expression profiling; diagnostic accuracy

Over the past decade, multiplexed genomic tests to improve the diagnosis of cancer and decisions about its treatment have been developed. Studies that demonstrate clinical validity and utility are critical to the appropriate integration of these tests into standard clinical practice. To date, this level of evidence has been met by very few gene expression–based tests.¹ The blinded, multicenter validation study presented here was prospectively designed to objectively compare the diagnostic accuracy of immunohistochemistry and gene expression–based tests for the identification of the tissue of origin in metastatic tumors.

Primary site determination in metastatic cancer can be problematic. An estimated 100,000 tumors per year in the United States are of uncertain origin.² These tumors may have features that are completely undifferentiated, or they may be relatively well-differentiated, but without an obvious primary source. Determination of the tissue of origin is usually based on clinical information, radiography, and histopathology,³ but in some cases tumor classification remains undefined even after blood tests, imaging, endoscopy, and advanced microscopic and molecular evaluation.⁴ These cases are known as cancers of unknown primary (CUP) and are estimated to number 31,000 per year in the United States.⁵ Accurate determination of a tumor's origin has clinical utility, because it directs evidence-based treatment.^{2,6,7} Most chemo-therapeutics and targeted biological agents are tested, approved, and reimbursed within the context of tissue type.⁸

Immunohistochemistry (IHC) is the current standard in pathologic evaluation of tumors of uncertain origin.² The availability of new antibodies has progressively expanded the pathologist's toolbox, and recent efforts to systematize IHC staining have reduced subjectivity.^{2,9,10} Still, IHC is acknowledged to lack uniform specificity or sensitivity in this setting,^{2,11} and a small fraction of metastatic cancers with occult primaries remain unidentified despite exhaustive IHC investigations.¹²

Gene expression profiling (GEP) tests have been developed to assist in the evaluation of difficult-to-diagnose tumors.¹³ Reported accuracy of these new tests, established mainly in known primary and metastatic tumors, has been in the range of 75% to 90%.^{14–17} Retrospective studies have suggested the ability of GEPs to predict the correct primary site in CUPs.^{18,19} In prospective studies of true CUPs, however, the accuracy of any tissue-of-origin test is by definition impossible to gauge. This diagnostic lacuna plus the lack of large direct comparisons between IHC and GEP in metastatic tumors have raised questions about

the relative diagnostic accuracies of these 2 methodologies and their potential areas of practical interface in clinical practice.

In this blinded, multicenter, prospective study, we compared the accuracy of IHC and GEP in identifying the tissue of origin in formalin-fixed, paraffin-embedded specimens from metastatic tumors with known primaries.

MATERIALS AND METHODS

Study Design

The study was designed to be a prospective assessment of diagnostic accuracy. The study plan, inclusion criteria, primary end point, and statistical analysis plan were established in a written protocol before the first specimen was selected or tested. The goal of the study was to compare the diagnostic accuracy of GEP to that of IHC-guided methods in a setting that resembles current clinical practice, with metastatic specimens that are representative of the tumors seen in cancers of uncertain primary. The study protocol was developed with inputs from a practicing pathologist and was tested in a pilot study, published separately.²⁰ The study design and workflow are illustrated in Figure 1.

Patient Specimens

A total of 160 formalin-fixed, paraffin-embedded metastatic tumor specimens from 153 patients, obtained under an Institutional Review Board–approved protocol, were sourced from 4 institutions: the University of Tennessee; Virginia Commonwealth University; Indiana University-Purdue University, Indianapolis; and Folio Biosciences. Seven patients provided >1 specimen, each from a different biopsy site. The principal investigator (C.H.), a board-certified pathologist, screened all specimens for inclusion in the study, on the basis of 5 criteria: (1) metastatic tumor with a known primary site determined by clinical and imaging information; (2) known primary site was 1 of 15 available on the GEP panel; (3) sufficient tissue for at least twenty-five 5- μ m-thick sections; (4) 60% non-necrotic tumor tissue; and (5) consistency between reported histology and that found on review by board-certified pathologists. Within these constraints, the specimens were selected to resemble uncertain primaries seen in clinical practice. Although all specimens included in the study were from metastatic tumors, true CUPs were excluded because the end point or diagnostic truth for any CUP is by definition unknowable. Table 1 summarizes patient and specimen characteristics.

Slide Preparation and Staining

All specimens were assigned masked identifiers at the University of Tennessee, Memphis, where all hematoxylin and eosin (H&E), IHC, and histochemical staining analyses were performed in a Clinical Laboratory Improvement Amendments (CLIA)-certified laboratory. Stains could be ordered from a panel of 84 stains (73 IHC and 11 histochemical stains), agreed upon in advance by all investigators (Table 2).

Digital Imaging and Review of Stained Slides

Each stained slide was digitized using a Whole Slide Imaging system (ScanScopeXT; Aperio Technologies Inc., San Diego, CA) at 0.25 $\mu\text{m}/\text{pixel}$ resolution and the quality verified by the principal investigator before being posted for the ordering evaluating pathologist (EP). Through a web-based interface, each EP was able to order stains, view requested slides as digital images, and record both diagnosis and confidence level after the initial H&E staining, after 1 round of staining, and after 2 rounds of staining. Although the digital image review process was not formally validated, all pathologists were comfortable with its use. When >1 EP ordered a given stain for a given specimen, each was provided the same digitized image.

Immunohistochemistry

The EPs were provided patient sex, biopsy site, and gross descriptions for each specimen, but were otherwise blinded to the clinical history and primary site. The EPs operated independently, with the freedom to use their own professional judgment and experience to select stains to diagnose the primary site. The 4 EPs were chosen to represent a mix of experience and practice settings. All were board-certified with 3 to 30+ years of experience in academic centers, community practice, and/or pathology reference laboratories. Each specimen was independently evaluated by 3 pathologists.

Gene Expression Profiling

GEP was performed at Pathwork Diagnostics Laboratory (PWDL) using the Pathwork Tissue of Origin Test (Pathwork Diagnostics Inc., Redwood City, CA). The personnel at PWDL were blinded to the clinical history, gross description, biopsy site, and primary site of the specimens. The processing steps described previously¹⁴ included microdissection, RNA extraction, cDNA preparation, and microarray processing. The test reports similarity to the 15 tissues on the test panel as a Similarity Score, ranging from 0 to 100 and summing to 100 across all 15 tissues. As previously established,¹⁴ specimens with a maximum similarity score <20 are considered “nonevaluable.” Two specimens failed by this criterion. A third specimen yielded inadequate cDNA. For the remaining 157 specimens, the tissue corresponding to the highest Similarity Score was reported as the GEP result.

Statistical Analysis

Statistical analyses were performed using SAS 9.3 and R 2.14. For all analyses, 2-sided P -values <0.05 were considered statistically significant.

The primary study end point was agreement of IHC with reference diagnosis compared with agreement of GEP with reference diagnosis. This end point was evaluated using conditional logistic regression and a one-to-many comparison (GEP:pathologists). The odds ratio (OR) and 95% confidence interval for the association between specimen classification accuracy of the 2 classifiers (GEP, EP/IHC; correct or incorrect) were calculated. The null hypothesis (OR= 1) represents equivalent accuracy.

Power calculations indicated that a sample size of 160 specimens would provide 90% power (2-tailed, P -value <0.05) in determining an accuracy difference of at least 15% between the

IHC and GEP, assuming (i) an accuracy of 70% for IHC¹¹ and 85% for GEP¹⁴ in the tissue of origin, (ii) a 1:1 matching of GEP to EP, and (iii) nonagreements between the call and the truth were not consistent between GEP and EP. The higher observed accuracy, the use of 1:3 matching (GEP compared with 3 pathologists), and the consistency between GEP and EP with respect to correct calls (eg, in well-differentiated specimens) resulted in greater power.

RESULTS

The study enrolled 160 formalin-fixed, paraffin-embedded metastatic tumor specimens from 153 patients. Cases were selected to represent those in which the diagnosis was not obvious upon morphology review; 33% were poorly differentiated or undifferentiated. Three of the specimens failed GEP, and thus 157 specimens evaluable by both IHC and GEP are reported.

Overall Accuracy

Overall, GEP was in agreement with the reference diagnosis in 89.2% of the specimens tested, compared with 83.3% agreement with IHC-based methods. This difference is statistically significant ($P = 0.013$) with an OR of 2.9 and a 95% confidence interval of 1.2 to 6.7.

Accuracy varied by known primary site (Table 3). Both methods identified non-Hodgkin lymphoma, melanoma, and thyroid specimens with 100% accuracy, whereas gastric metastases proved most challenging, achieving <30% accuracy. IHC achieved a positive predictive value >98% for 4 of 15 tissue sites, compared with 8 of 15 for GEP (Fig. 2).

IHC Accuracy

Overall accuracies of the individual EPs were similar (84.1%, 82.2%, 83.2%, and 85.2%). Although the distribution of incorrect calls varied somewhat from EP to EP by primary site, pairwise κ ranged from 0.76 to 0.83, indicating good to excellent concordance between participants.²¹

Slide and Stain Usage

Across all 157 specimens, EPs requested an average of 8.3 stains (range, 0 to 20), whereas GEP micro-dissection and processing used 3.8 on average (range, 1 to 8) ($P < 0.001$). More slides were used in the first round of stain assessment (median of 6; range, 1 to 10) than in the second round (median of 4; range, 1 to 10).

Subgroup Analysis

A systematic subgroup analysis was performed to identify characteristics of the “difficult-to-diagnose specimen,” which might be readily assessed a priori by a pathologist. Four candidate subgroups were evaluated: (1) poorly versus well-differentiated tumors; (2) adenocarcinoma versus nonadenocarcinoma; (3) specimens for which only a first round of stains was ordered; and (4) those for which a second round of stains was required. The results are presented in a forest plot of comparative accuracy (Fig. 3). Biopsy site subgrouping was not possible because of the small specimen counts within each subgroup.

In all 51 poorly differentiated and undifferentiated tumors, GEP accuracy was 94.1%, whereas IHC accuracy was 79.1% ($P=0.016$). In the subset of 33 poorly differentiated and undifferentiated carcinomas, GEP accuracy exceeded IHC (91% to 71%, $P=0.023$). By contrast, the IHC and GEP performed similarly in the well-differentiated and moderately differentiated tumors (85.3% and 86.8%, respectively; $P=0.52$). Within the adenocarcinoma subgroup, GEP accuracy was 86.4% compared with that of IHC of 79.8% ($P=0.022$).

A potentially useful metric for distinguishing the “difficult-to-diagnose specimen” is the number of stains an individual pathologist intends to order for the first round of assessment—a median of 6 in this study. The EPs evaluated 157 specimens, for a total of 472 specimen reviews. In 88 specimen reviews, 5 or fewer stains were ordered, and both IHC and GEP performed similarly in this group (93% and 91%, respectively; $P=0.40$). When 6 or more first-round stains were used ($n=142$), GEP was significantly more accurate than IHC (87.5% and 79.4%, respectively; $P=0.012$).

A final subgroup analysis examined those specimen reviews requiring a second round of stains. In 283 specimen reviews the EP made a final diagnosis within the first round of stains ordered with an accuracy of 94.3% compared with a GEP accuracy of 93.3% on the same specimens ($P=0.47$). Of these 283 specimen reviews, 9 were completed using only H&E, and 274 were completed after the first round of stains. For these specimen reviews completed with 1 round, the number of stains ordered ranged from 1 to 10, with a median of 6 and an average of 6.6. The remaining 189 specimen reviews required a second round of stains, with an IHC accuracy of 67.0% compared with a GEP accuracy of 83.1% ($P<0.001$). For these specimen reviews completed after a second round, the number of stains ordered ranged from 1 to 10, with a median of 4 and an average of 4.6.

DISCUSSION

Few assays have been rigorously validated for use in diagnostic pathology; fewer still have been rigorously validated for use in diagnosing the primary site of meta-static tumors. This is the first prospectively conducted study to directly compare the diagnostic accuracy of GEP with that of IHC methods in a controlled, multicenter, clinical trial setting.

Our results indicate that GEP is more accurate than the current standard of care, IHC, in a representative sample of metastatic tumors. The comparative accuracy advantage of GEP was greatest in the poorly differentiated and undifferentiated carcinomas. These tumors have lost some or all of the morphologic features that pathologists use to identify the likely primary tissue type, and therefore it is not surprising that these cases would be more challenging to diagnose using morphology-based methods.

These findings also confirm the ability of IHC to diagnose the majority of metastatic tumors. The results also suggest that GEP may be a reasonable alternative to exhaustive IHC in the workup of occult primaries when the initial round of staining proves inconclusive. Under the conditions of this study, there was no gain in diagnostic accuracy with a second round of staining.

The goal of the pathologic workup is to arrive at the correct diagnosis and allow more patients to receive tissue-specific treatment, which is presumed to improve survival in patients who present with cancers of uncertain origin.² Although this presumption has yet to be fully verified, recent studies do in fact indicate that successful profiling of highly ambiguous metastatic tumors can lead to more specific treatment and improved survival.^{22,23} Thus, site-directed therapy remains the standard of care for these patients.

As more targeted anticancer agents are introduced, accurate and efficient classification of the tissue type becomes essential. Identification of tumor lineage and primary site allows the pathologist to focus the search for biomarkers that dictate targeted therapy in a particular tumor type.^{4-6,22-25} In this emerging diagnostic algorithm, appropriate use of GEP on the most challenging cases may help conserve tumor tissue in the hunt for the primary site of the neoplasm. Although not an issue in this trial, conserving biopsy material for further biomarker evaluation is often important when limited diagnostic material is available.

In this study, we sought to minimize elements of variability. For IHC analysis, we centralized specimen handling and slide staining, utilized a set catalog of antibodies, removed fiscal constraints on stain selection, used a common digitized slide for duplicate stain orders, and blinded pathologists to all clinical information other than biopsy site, patient sex, and gross specimen description. To further minimize variability and to enable a direct comparison of accuracy between IHC and GEP, the study was restricted to the 15 tissue types represented on the GEP test panel. These 15 tissue types represent approximately 90% of the human solid tumor malignancies seen each year in the United States^{26,27}; thus the study is able to mimic the diversity of the most common specimens processed in the typical hospital pathology laboratory. Although it is possible that less common tumor types could be disproportionately represented in the CUP category, there are little data to support this notion. A recent study suggests that neuroendocrine carcinomas, which are not on the test panel, represent 3% to 5% of all CUPs.²⁸

Although this study was not designed to evaluate the GEP performance of tumor types that are off panel, the behavior of many of the more common off-panel tissue types on the test is known.²⁹ In general, off-panel tumor types result in lower scores and often reflect the expression patterns of embryologically related tissue types that are on the panel, as in an example reported for a suspected mesothelioma in a previously published study.³⁰ For this reason, clinical use of the test includes interpretation by a pathologist who is trained in the use of the GEP test. This training includes an understanding of cross-reactivity patterns.

The general concordance of IHC results among EPs indicates that we were successful in probing mainly the core cognitive domain of the IHC workup (hypothesis generation, stain selection, slide interpretation) rather than the more variable communicative and site-specific aspects of the diagnostic process.³¹ To maximize objectivity in GEP, we used automatic prediction based on the highest expression score and made no allowance for score interpretation by PWDL's trained pathology staff, which is the common clinical practice when reporting test results. The availability of complete clinical information might have influenced the results of pathologists' interpretations in both the IHC and the GEP approaches.

The study has several limitations. We used a broad range of metastatic specimens with known primaries instead of a collection of CUPs, because a gold-standard diagnosis of the primary site is required to compare the accuracy of the 2 methods. We limited the study to on-panel tumors. This limitation biased estimates of both the GEP test performance and the pathologists' performance, as the EPs also knew the tumors belonged to 1 of the 15 types.

Differentiation is another possible confounder, as the relative accuracy of the 2 methods may depend on the mix of well-differentiated and poorly differentiated specimens seen in clinical practice. Some specimens are not evaluable by GEP for reasons including RNA quality; 3 specimens were excluded for nonevaluability. A sensitivity analysis performed to assess this limitation found that GEP accuracy remains superior to that of IHC even if all 3 of these GEP results were imputed to be incorrect when included in the analysis ($P = 0.047$). The study results might be only applicable to the specific GEP test that was evaluated in this study; other tests that use alternative approaches may perform differently. The use of telepathology was not formally validated; however, all pathologists were trained and felt proficient in its use.²⁰

Important clinical implications can be drawn from this study. First, several of the tumor types for which GEP shows superior accuracy (breast, ovarian) are those for which treatment outcomes are most promising. A retrospective outcome study of 107 true CUP patients diagnosed with the GEP test between July and December 2009 found a median survival of 14 months, 33% of patients alive at 2 years, and 30% alive at 3 years, with the longest-surviving patients among those diagnosed with breast, lung, and ovarian cancers.²³ Although historical survival comparisons bring inherent bias, these findings suggest a provocative contrast to the 12 study compilations of 928 CUP patients treated between 1997 and 2008, which found a median survival of 8.9 months and 12.3% of patients alive at 2 years.³² Second, although several of the tumor types (pancreas, stomach, and bladder) have poor prognosis and gain little benefit from correct identification under today's treatment paradigm, such identification does enable the patient to participate in an appropriate clinical trial, contributing to the overall advance of potentially promising treatment for advanced-stage disease.

In summary, the controlled, prospective, multi-center study reported here finds that the GEP test used in this study was significantly more accurate than IHC when used to identify the primary site of metastatic tumors, with best performance observed in specimens that required a second round of stains, used 6 or more stains in the first round, or were poorly differentiated or undifferentiated carcinomas. These results can be used to guide the appropriate integration of GEP into clinical practice to increase diagnostic accuracy, conserve biopsy tissue, and allow the selection of site-specific therapy.

Acknowledgments

Conflicts of Interest and Source of Funding: Pathwork Diagnostics was the study sponsor. C.R.H. and A.K. received research grants from Pathwork Diagnostics for the performance of this study. J.P.G., L.M.W., W.M.R., O.S.K., G.G.A., and M.G.W. were compensated by Pathwork Diagnostics for study participation. R.P., M.H.-M., and W.D.H. are employees of Pathwork Diagnostics. F.A.M. has received honoraria from Pathwork Diagnostics for speaking engagements and consulting.

The authors thank Allen Gown for advice on study design; Paul Courter for editorial assistance; Ashley Ezekiel for technical support; Jing Shi, MD, PhD, and Rebecca Panos for statistical support; Catherine Dumur, PhD (Virginia Commonwealth University), Michael Idowu, MD (Virginia Commonwealth University), and George Sandusky, MD (Indiana University) for supplying some of the specimens; Eloisa Fuentes, MD, and Andrea Pingitore, MD for secondary review of tissue specimens; personnel from the University Pathology Group for their assistance in tissue processing and staining; Judith Loebel for assistance with conducting the study under Good Clinical Practice; staff at PWDL for processing specimens through the Tissue of Origin test; and Shawn Becker, MD, for helpful discussions.

References

1. Febbo PG, Ladanyi M, Aldape KD, et al. NCCN task force report: evaluating the clinical utility of tumor markers in oncology. *J Natl Compr Cancer Netw*. 2011; 9(suppl 5):S1–32.
2. Greco FA, Oien K, Erlander M, et al. Cancer of unknown primary: progress in the search for improved and rapid diagnosis leading toward superior patient outcomes. *Ann Oncol*. 2012; 23:298–304. [PubMed: 21709138]
3. National Comprehensive Cancer Network (NCCN). [Accessed August 22, 2012] Clinical Practice Guidelines in Oncology: Occult Primary (Cancer of Unknown Primary [CUP]) Version 1.2012. Available at: http://www.nccn.org/professionals/physician_gls/pdf/occult.pdf
4. Greco, FA., Hainsworth, JD. Cancer of unknown primary. In: DeVita, VT, Hellman, S., Rosenberg, SA., editors. *Cancer: Principles and Practice of Oncology*. 9. Philadelphia: Lippincott; 2011. p. 2022–2051.
5. American Cancer Society. Cancer—Unknown Primary, 1/19/1012. Atlanta, GA: American Cancer Society; 2012. Available at: <http://www.cancer.org/Cancer/CancerofUnknownPrimary/DetailedGuide/index> [Accessed June 19, 2012]
6. Pavlidis N, Fizazi K. Cancer of unknown primary (CUP). *Crit Rev Oncol Hematol*. 2005; 54:243–250. [PubMed: 15890271]
7. Varadhachary GR, Talantov D, Raber MN, et al. Molecular profiling of carcinoma of unknown primary and correlation with clinical evaluation. *J Clin Oncol*. 2008; 26:4442–4448. [PubMed: 18802157]
8. National Comprehensive Cancer Network (NCCN). [Accessed June 8, 2012] Clinical Practice Guidelines in Oncology: NCCN Guidelines for Treatment of Cancer By Site. 2012. Available at: http://www.nccn.org/professionals/physician_gls/f_guidelines.asp
9. Oien KA. Pathologic evaluation of unknown primary cancer. *Semin Oncol*. 2009; 36:8–37. [PubMed: 19179185]
10. Hashimoto K, Sasajima Y, Ando M, et al. Immunohistochemical profile for unknown primary adenocarcinoma. *PloS One*. 2012; 7:e31181. [PubMed: 22299055]
11. Anderson GG, Weiss LM. Determining tissue of origin for metastatic cancers: meta-analysis and literature review of immunohistochemistry performance. *Appl Immunohistochem Mol Morphol*. 2010; 18:3–8. [PubMed: 19550296]
12. Bugat R, Bataillard A, Lesimple T, et al. Summary of the standards, options and recommendations for the management of patients with carcinoma of unknown primary site (2002). *Br J Cancer*. 2003; 89(suppl 1):S59–S66. [PubMed: 12915904]
13. Monzon FA, Koen TJ. Diagnosis of metastatic neoplasms: molecular approaches for identification of tissue of origin. *Arch Pathol Lab Med*. 2010; 134:216–224. [PubMed: 20121609]
14. Pillai R, Deeter R, Rigl CT, et al. Validation and reproducibility of a microarray-based gene expression test for tumor identification in formalin-fixed, paraffin-embedded specimens. *J Mol Diagn*. 2011; 13:48–56. [PubMed: 21227394]
15. Ma XJ, Patel R, Wang X, et al. Molecular classification of human cancers using a 92-gene real-time quantitative polymerase chain reaction assay. *Arch Pathol Lab Med*. 2006; 130:465–473. [PubMed: 16594740]
16. Monzon FA, Lyons-Weiler M, Buturovic LJ, et al. Multicenter validation of a 1,550-gene expression profile for identification of tumor tissue of origin. *J Clin Oncol*. 2009; 27:2503–2508. [PubMed: 19332734]

17. Rosenwald S, Gilad S, Benjamin S, et al. Validation of a microRNA-based qRT-PCR test for accurate identification of tumor tissue origin. *Mod Pathol*. 2010; 23:814–823. [PubMed: 20348879]
18. Monzon FA, Medeiros F, Lyons-Weiler M, et al. Identification of tissue of origin in carcinoma of unknown primary with a micro-array-based gene expression test. *Diagn Pathol*. 2010; 5:3. [PubMed: 20205775]
19. Greco FA, Spigel DR, Yardley DA, et al. Molecular profiling in unknown primary cancer: accuracy of tissue of origin prediction. *Oncologist*. 2010; 15:500–506. [PubMed: 20427384]
20. Kulkarni A, Pillai R, Ezekiel AM, et al. Comparison of histopathology to gene expression profiling for the diagnosis of metastatic cancer. *Diagn Pathol*. 2012; 7:110. [PubMed: 22909314]
21. Jacob, Cohen. A coefficient of agreement for nominal scales. *Educ Psychol Meas*. 1960; 20:37–46.
22. Nystrom SJ, Hornberger JC, Varadhachary GR, et al. Clinical utility of gene-expression profiling for tumor-site origin in patients with metastatic or poorly differentiated cancer: impact on diagnosis, treatment, and survival. *Oncotarget*. 2012; 3:620–628. [PubMed: 22689213]
23. Hainsworth JD, Rubin MS, Spigel DR, et al. Molecular gene expression profiling to predict the tissue of origin and direct site-specific therapy in patients with carcinoma of unknown primary site: a prospective trial of the Sarah Cannon Research Institute. *J Clin Oncol*. 2012 [Epub ahead of print].
24. Handorf CR. Gene expression analysis and immunohistochemistry in evaluation of cancer of unknown primary: time for a patient-centered approach. *J Natl Compr Cancer Netw*. 2011; 9:1415–1420.
25. Varadhachary GR. Carcinoma of unknown primary: focused evaluation. *J Natl Compr Cancer Netw*. 2011; 9:1406–1412.
26. American Cancer Society. [Accessed on August 21, 2012] Cancer Facts & Figures. 2012. Available at: <http://www.cancer.org/acs/groups/content/@epidemiologysurveillance/documents/document/acspc-031941.pdf>
27. Hess KR, Varadhachary GR, Taylor SH, et al. Metastatic patterns in adenocarcinoma. *Cancer*. 2006; 106:1624–1633. [PubMed: 16518827]
28. Spigel DR, Hainsworth JD, Greco FA. Neuroendocrine carcinoma of unknown primary site. *Semin Oncol*. 2009; 36:52–59. [PubMed: 19179188]
29. U.S. FDA. [Last accessed January 2, 2013] 510(k) Decision Summary for the Pathwork Tissue of Origin Test Kit – FFPE. Jun 9. 2012 Available at: http://www.accessdata.fda.gov/cdrh_docs/reviews/K092967.pdf
30. Beck AH, Rodriguez-Paris J, Zehnder J, et al. Evaluation of a gene expression microarray-based assay to determine tissue type of origin on a diverse set of 49 malignancies. *Am J Surg Pathol*. 2011; 35:1030–1037. [PubMed: 21602661]
31. Pena GP, de Andrade-Filho JS. How does a pathologist make a diagnosis? *Arch Pathol Lab Med*. 2009; 133:124–132. [PubMed: 19123724]
32. Greco FA, Pavlidis N. Treatment for patients with unknown primary carcinoma and unfavorable prognosis factors. *Semin Oncol*. 2009; 36:65–74. [PubMed: 19179190]

FIGURE 1.

Study design: metastatic tumor specimens meeting the criteria were obtained as archived formalin-fixed, paraffin-embedded specimens from 4 institutions: the University of Tennessee (UT); Virginia Commonwealth University (VCU); Indiana University-Purdue University, Indianapolis (IUPUI); and Folio Biosciences (FB). The CLIA-certified laboratory at UT sectioned each block and sent digitized H&E images to 4 different EPs and to the pathology team at PWDL. The EPs and PWDL were blinded to the primary site. The EPs ordered stains from a panel of 84 IHC and histochemical stains. All staining was performed at the UT CLIA-certified laboratory, and digitized images of the requested stained slides were provided to only the ordering EP using the web interface. The EPs, who were trained on the Whole Slide Imaging system and the web interface before the trial and during a pilot study,²⁰ could issue a final diagnosis at any stage in the workup. Coded slides sent to PWDL were analyzed using the Pathwork Tissue of Origin Test.¹⁴ The accuracy of IHC and GEP were compared using statistical methods described.

Known Primary Site	BL	BR	CO	GA	GC	KI	LI	LU	LY	ME	OV	PA	PR	SC	TH	TOTAL
BL	6		1	1		1							1			10
BR		25														25
CO			25													25
GA			3	2		1						1				7
GC					2						1			1		4
KI				1		12								1		14
LI							6									6
LU								5			1					6
LY									3							3
ME										12						12
OV											7	1				8
PA				2								3				5
PR													3			3
SC														17		17
TH															12	12
Total	6	25	29	6	2	14	6	5	3	12	9	5	4	19	12	157
PPV%	100	100	86	33	100	86	100	100	100	100	78	60	75	89	100	
BL	13	1	2	2		3		4				5				30
BR		63		2				6			1	3				75
CO			69	5								1				75
GA			3	6								12				21
GC			1	2	4		1	2		1		1				12
KI						42										42
LI		2					14				1			1		18
LU								18								18
LY									9							9
ME										36						36
OV		1				1					18	3			1	24
PA			1	3								11				15
PR				1			1	2					5			9
SC						1		1					1	48		51
TH															37	37
Total	13	66	76	21	4	47	16	33	9	37	20	36	6	49	38	472
PPV%	100	95	91	29	100	89	88	55	100	97	90	31	83	98	97	

FIGURE 2.

Confusion matrix: GEP and IHC. Known primary site versus test prediction/final diagnosis. Test predictions and final diagnoses with positive predictive value (PPV) of 98% or higher are highlighted in green. Bladder (BL), breast (BR), colorectal (CO), gastric (GA), hepatocellular (LI), kidney (KI), melanoma (ME), non-Hodgkin's lymphoma (LY), non-small cell lung (LU), ovarian (OV), pancreas (PA), prostate (PR), sarcoma (SC), testicular germ cell (GC), and thyroid (TH).

FIGURE 3.

Forest plot of comparative accuracy, IHC compared with GEP, overall, and subgroup analysis. ORs <1 favor IHC, >1 favor GEP. Statistically significant subgroups are shown in blue; lines indicate the 95% confidence interval.

TABLE 1

Patient and Tumor Characteristics (n = 160) Enrolled in the Study

Characteristics	N	%
Sex (patients = 153)		
Male	82	54
Female	71	46
Age (y)		
< 49	36	24
50–59	37	24
60–69	57	37
>70	23	15
Ethnicity		
African American	41	27
White	108	71
Other	4	2
Biopsy site (specimens = 160)		
Bone (nondecalfied)	8	5
Brain	34	21
Liver	22	14
Lung	18	11
Lymph node	43	27
Peritoneum/omentum	11	7
Soft/skin tissue	24	15
Tumor & differentiation		
Metastatic specimens	160	100
Poor/undifferentiated	51	32
Well/moderate	109	68
Primary specimens	0	0
Specimen source		
University of Tennessee	81	51
Indiana University-Purdue University	43	27
Virginia Commonwealth University	32	20
Folio Biosciences	4	3

TABLE 2

Stains Available to EPs in Study

No.	IHC Stain	Clone
(A) List of IHC stains		
1	Actin, muscle specific	1A4
2	AE1/AE3	Mouse monoclonal
3	AFP	Rabbit polyclonal
4	hCG	Rabbit polyclonal
5	CA-125	Mouse monoclonal
6	Calcitonin	Rabbit polyclonal
7	Caldesmon	E89
8	Calretinin	Rabbit polyclonal
9	CAM 5.2	Cam 5.2
10	CD10	270
11	CD117	Rabbit monoclonal
12	CD138	VS38c
13	CD20	L-26
14	CD3	Rabbit polyclonal
15	CD30	BerH2
16	CD34	QBEnd-10
17	CD43	L-60
18	CD45RO	A6
19	CD56	1B6
20	CD99	Mouse monoclonal
21	CDX2	CDX2-88
22	CEA-polyclonal	Mouse monoclonal, II-7
23	Chromogranin	LK2H10(2)
24	CK20	Ks 20.8
25	CK7	K72
26	CK17	Mouse monoclonal
27	CK19	Mouse monoclonal
28	CK5/6	D5 & 16B4
29	Pancytokeratin	AE1, AE3, PCK26
30	Desmin	NCL-DE-R-11
31	E-cadherin	ECH-6
32	EMA	Mouse monoclonal
33	Estrogen Receptor	6F11
34	GCDFP-15	Mouse monoclonal
35	GFAP	Polyclonal
36	Glypican-3	1G12
37	Hep-Par1	OCHIES
38	HMB45	HMB45

No.	IHC Stain	Clone
39	HMW keratin	AE3
40	HNF-1	SAB2105022-5OUG
41	Inhibin	Alpha (R1)
42	CK, Oscar	Oscar
43	Leucocyte common antigen	RP2/18
44	Mammaglobin	Rabbit monoclonal
45	Melan-A	MC-7C10
46	MOC-31	MOC-31
47	MUC1	MRQ-17
48	MUC2	MRQ-18
49	MUC5AC	MRQ-19
50	Myeloperoxidase	Rabbit polyclonal
51	Myogenin	F5D
52	Napsin A	Rabbit polyclonal
53	NSE	E27
54	OCT-4	MRQ-10
55	P504S	13H4
56	p53	BP 53-11
57	p63	Tap63a
58	PAX2	2E4-11063
59	PAX5	Rabbit monoclonal SP34
60	PAX8	Rabbit polyclonal
61	PLAP	NB10
62	Progesterone receptor	Y85
63	Prostate specific antigen	Rabbit polyclonal
64	RCC	Mouse monoclonal PN-15
65	S-100	4C4.9
66	Synaptophysin	Rabbit, SP11
67	Thrombomodulin	1009
68	Thyroglobulin	1D4
69	TTF1	8G7G3/1
70	Uroplakin	AU-1
71	Villin	CWWB1 (mouse monoclonal)
72	Vimentin	V9
73	WT-1	6F-H2 (mouse monoclonal)

(B) List of histochemical stains

No.	Histochemical stain
1	Alcian blue-PAS
2	Argentaffin
3	Argyrophil
4	Colloidal iron stain
5	Elastic, Verhoeff

No.	IHC Stain	Clone
6	Mucicarmine	
7	PAS	
8	PASD	
9	PTAH	
10	Reticulum, Gomori	
11	Trichrome, Masson	

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

TABLE 3

Accuracy of IHC and GEP Stratified by Known Primary Site

Known Primary Site	No. Specimens	IHC Accuracy* (%)	GEP Accuracy*
Bladder	10	43	60
Breast	25	84	100
Colorectal	25	92	100
Gastric	7	29	29
Testicular germ cell	4	33	50
Kidney	14	100	86
Liver	6	78	100
Non-small cell lung	6	100	83
Non-Hodgkin lymphoma	3	100	100
Melanoma	12	100	100
Ovarian	8	75	88
Pancreatic	5	73	60
Prostate	3	56	100
Soft tissue sarcoma	17	94	100
Thyroid	12	100	100
Overall	157	83.3	89.2 [†]

* IHC accuracy calculated as weighted average agreement with known primary site, includes 3 IHC final diagnoses for each case. GEP accuracy is calculated as GEP agreement with known primary site, using the highest Similarity Score for 157 specimens.

[†] Difference in overall accuracy is statistically significant ($P=0.013$); OR (95% confidence interval), 2.9 (1.3–6.7).