EUROPEAN JOURNAL OF CANCER 43 (2007) 2026–2036

available at www.sciencedirect.com

ScienceDirect

journal homepage: www.ejconline.com

Review

Switching benchmarks in cancer of unknown primary: From autopsy to microarray

*George Pentheroudakis, Vassilios Golfinopoulos, Nicholas Pavlidis**

Department of Medical Oncology, Ioannina University Hospital, Niarxou Avenue, 45500 Ioannina, Greece

ARTICLE INFO

Article history:
Received 23 June 2007
Accepted 27 June 2007
Available online 14 August 2007

Keywords:
Cancer of unknown primary
Autopsy
Microarray
Gene expression

ABSTRACT

Introduction: Cancer of unknown primary (CUP) is associated with unknown biology and dismal prognosis. Information on the primary site of origin is scant and has never been analysed. We systematically reviewed all published evidence on the CUP primary site identified by two different approaches, either autopsy or microarray gene expression profiling.
Methods: Published reports on identification of CUP primary site by autopsy or microarray-based multigene expression platforms were retrieved and analysed for year of publication, primary site, patient age, gender, histology, rate of primary identification, manifestations and metastatic deposits, microarray chip technology, training and validation sets, mathematical modelling, classification accuracy and number of classifying genes.
Results: From 1944 to 2000, a total of 884 CUP patients (66% males) underwent autopsy in 12 studies after presenting with metastatic or systemic symptoms and succumbing to their disease. A primary was identified in 644 (73%) of them, mostly in the lung (27%), pancreas (24%), hepatobiliary tree (8%), kidneys (8%), bowel, genital system and stomach, as a small focus of adenocarcinoma or poorly differentiated carcinoma. An unpredictable systemic dissemination was evident with high frequency of lung (46%), nodal (35%), bone (17%), brain (16%) and uncommon (18%) deposits. Between the 1944–1980 and the 1980–2000 series, female representation increased, 'undetermined neoplasm' diagnosis became rarer, pancreatic primaries were found less often while colonic ones were identified more frequently. Four studies using microarray technology profiled more than 500 CUP cases using classifier set of genes (ranging from 10 to 495) and reported strikingly dissimilar frequencies of assigned primary sites (lung 11.5%, pancreas 12.5%, bowel 12%, breast 15%, hepatobiliary tree 8%, kidneys 6%, genital system 9%, bladder 5%) in 75–90% of the cases.
Conclusions: Evolution in medical imaging technology, diet and lifestyle habits probably account for changing epidemiology of CUP primaries in autopsies. Discrepant assignment of primary sites by microarrays may be due to the presence of 'sanctuary sites' in autopsies, molecular misclassification and the postulated presence of a pro-metastatic genetic signature. In view of the absence of patient therapeutic or prognostic benefit with primary identification, gene expression profiling should be re-orientated towards unraveling the complex pathophysiology of metastases.

© 2007 Elsevier Ltd. All rights reserved.

* *Corresponding author:* Tel./fax: +30 2651099394.
E-mail address: npavlid@cc.uoi.gr (N. Pavlidis).
0959-8049/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ejca.2007.06.023

1. Introduction

Cancer of unknown primary site (CUP) ranks among the 10 most common malignancies of developed societies. Resistance to antineoplastic therapy and poor patient outcome represent hallmarks of this clinical entity.[1] Failure to identify the primary tumour, though probably not relevant for prognosis in the presence of metastatic dissemination, is a psychological burden for the patient and physician alike. Moreover, patient management with tailored chemotherapeutic regimens and targeted agents has increasingly become primary site-dependent over the last decade. Accordingly, failure to identify the primary tumour theoretically incurs the risk of failure to administer appropriate systemic combination chemotherapy as well as targeted molecular therapy.[2] Finally, it is generally believed that patient prognosis can be better predicted when the primary site of metastatic tumours is known. The latter concept has been recently challenged by the definition of distinct clinicopathologic subsets of CUP patients with more favourable prognosis with appropriate therapy (20% of CUP cases).[3] Disappointingly, the vast majority (80%) of CUP patients harbour widespread visceral or bony deposits and belong to poor prognosis subsets.

Autopsy rates have been declining steadily over the last three decades in hospitals across Europe and America.[4] Better definition of cancer prognosis, progress in medical imaging technologies and interventional bioptic procedures probably contributed towards this trend. Although autopsy had a significant contribution to our understanding of diseases, lack of appreciation of its value coupled to concern regarding infectious disease transmission may also have played a role. In published autopsy series in patients with CUP, the rate of primary tumour identification ranged from 50% to 80%.[5] Valuable information regarding the location and incidence of the primary in addition to clinicopathologic data have only been reported in a segmented manner in relatively small cohorts and have never been critically analysed.

Recently, profiling of multiple gene expression in tumours has been made possible thanks to development of high-throughput DNA microarray platforms.[6] Some molecular biology data suggest that each tumour maintains at least the basic genetic signature of the tissue of origin throughout clonal evolution and metastatic dissemination.[7,8] Consequently, multigene expression profiling via microarray platforms is anticipated to be useful either for class prediction (biologic assignment of a primary in metastatic tumours) or for class discovery (identification of distinct disease subgroups with different biological behaviour). Data on molecular identification of the primary tumour in metastatic malignancies of known and unknown primary are neither mature nor abundant, but are slowly accumulating. In this systematic review, we present all published data on the primary tumour of CUP, identified either via autopsy or molecular array profiling and we screen for similarities and differences in incidence patterns over time and between methods.

2. Methods

Online databases (PubMed, EmBase) were searched using the key words: cancer of unknown primary; autopsy; DNA microarray; gene expression profiling. Full-text papers were obtained as well as conference abstracts. When deemed necessary, authors were contacted for supplemental information. We restricted our search to papers describing autopsy findings in patients with CUP in an adequate fashion (number of autopsies, gender, histology, primary tumours found and metastatic sites) as well as papers on DNA microarray methodologies that assign sites of origin in metastatic tumours of known and unknown primary.

Comparison of variable distribution between groups was performed with the χ^2 Exact test, multiple comparisons done via the Bonferroni adjustment method. Reported p-values are two-sided and considered significant when less than 0.05.

3. Results

3.1. Autopsy series

Twelve cohort studies containing necessary data on autopsy number, gender, histological diagnosis, identified primary tumours and metastatic sites were retrieved.[5,9–19] In those studies, a total of 884 patients diagnosed with cancer of unknown primary site underwent autopsy post-mortem from years 1944 to 2000. The median age at the time of CUP diagnosis ranged from 58 to 66 years. Biopsy of metastases disclosed

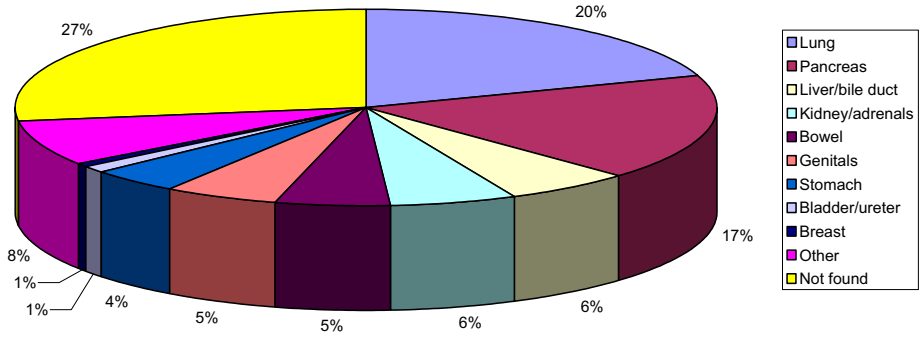

Fig. 1 – Relative proportion of autopsy-found primaries in published series.

well to moderately differentiated adenocarcinomas in more than half of the cases, while poorly differentiated or undifferentiated carcinoma, squamous carcinoma and extremely anaplastic 'undifferentiated neoplasias' were diagnosed in a quarter to a tenth of the cases. Melanomas, small cell carcinomas, neuroendocrine carcinomas, and sarcomas collectively accounted for less than 5% of CUP diagnoses ('other' category). Autopsy resulted in identification of the primary in 73% of procedures, most commonly in the lung (27% of identified primaries), pancreas (24%), liver or bile duct system (8%), kidney or adrenals (8%), colon (7%), genital organs (prostate 2%, testis 0.3%, ovary/uterus 4%, uterine cervix 0.7%) (Fig. 1). A primary tumour was identified in the breast in only five cases, while other tissues of origin accounted for a total of 10% of identified primaries (thyroid, mesothelial/peritoneal membranes, mediastinum, skin and bone). Patients had presented with symptoms or signs due to metastases in more than three quarters of cases with co-existence of non specific systemic symptoms like malaise, weight loss, anorexia in a third of them. Paraneoplastic syndromes or infectious manifestations were present at the time of diagnosis less often. High-volume widespread metastases were the norm: the most common sites involved were the lungs, lymph nodes, liver, bone, brain and abdominal sites (peritoneum, adrenals). Quite interestingly, metastatic seedlings in uncommon or peculiar sites (spleen, stomach, bowel, ovary, skin, soft tissues, parotid, thyroid, scalp, heart and breast) were observed in 115 patients (18%). Autopsy data are summarised in Table 1. Detailed autopsy data for older cohort studies spanning from 1944 to 1980 as well as for more recent series of autopsies performed between years 1980 and 2000 are shown in Tables 2 and 3, respectively.

Temporal trends of study variables were screened for by comparing summary characteristics between the 1944–1980 series and the 1980–2000 series (Table 4). The rate of occurrence of autopsy increased from less than half of CUP cases to approximately 75% in the last quarter of the 20th century, though inherent bias in the published series may interpret this finding. With passing time, the female CUP population represented an increasing proportion of the patient pool while the incidence of histological diagnosis of an undifferentiated neoplasm has declined. The rate of identification of the primary tumour at autopsy and the relative frequency of tumour sites seem to have remained constant over time, with the exception of a marginal decline of pancreatic and rise of colonic tumours.

Overall, even when the primary tumour is one with high incidence in the general population, such as lung cancer, published autopsy series establish its peculiar natural history and biological behaviour: the tumour is a small nodule that fails to grow or cause typical symptoms but metastasises early and widely. Almost all patients presented with non-specific or metastatic symptoms/signs of relatively short duration. Moreover, the epidemiology of autopsy-found primaries is strikingly different from that of known primary tumours affecting the general population: incidences of breast, bowel, bladder and prostate cancers are very low in comparison to the ones quoted in patients with known primary tumours. On the other hand, primary tumours of the lung, pancreas, liver/bile duct, stomach and kidneys that are affecting general

Table 1 – Summary data from all published autopsy CUP cases

Total number of autopsied CUP patients	Median age at diagnosis	Male/female	Histology	Primary site found	Primary sites		Manifestations	Metastatic sites (relative incidence)	
884	58–66	586 (66%)/298 (34%)	Adenocarcinoma 480 (54%)	644/884 (73%)	Lung	175 (27%)	Metastatic 85%	Lung	295(46%)
			Undifferentiated Ca 204 (23%)		Pancreas	153 (24%)	Non-specific	LN	228 (35%)
					Liver/bile duct	53 (8%)	(systemic) 30%	Liver	149 (23%)
			Squamous Ca 91 (10%)		Kidney/adrenals	52 (8%)	Paraneoplastic	Bone	112 (17%)
			Undetermined neoplasm 77 (9%)		Bowel	46 (7%)	12%	Brain	101 (16%)
					Genital system	46 (7%)	Infectious 7%	Abdomen	64 (10%)
			Other 32 (4%)		Stomach	38 (6%)		Uncommon (spleen, ovary, skin, soft tissues, parotid, thyroid, scalp, heart and breast)	115 (18%)
					Bladder/ureter	8 (0.01%)			
					Breast	5 (0.007%)			
					Other	68 (10%)			

Ca: carcinoma, LN: lymph nodes.

Table 2 – Older autopsy series 1944–1980

Author	Study period	Autopsies/ total no. of patients	Median age at diagnosis	M/F	Histology	Primary site found	Site (*N*)	Metastatic sites
Snyder et al.[9]	1944–1975	25/49	56	28/21	Adenocarcinoma by definition	0/25	None found	Lung (52), Liver (36), Brain (24), Ovary (40), Bone (40), Adrenal (4), Spleen (16)
Jordan and Shildt[10]	1971–1981	18/56	58	33/23	Adenocarcinoma (30) Undifferentiated Ca (16) Undetermined neoplasm (10)	13/18	Lung (5), Pancreas (3), Stomach (2), Bile duct (2), Mesothelioma (1)	Non available
Steckel and Kagan[11]	1968–1974	34/255	NA	NA	*In autopsied cases:* Adenocarcinoma (7) Squamous Ca (3) Undifferentiated Ca (1) Other (3)	14/34	Lung (7), Pancreas (2), Kidney(1) Bladder (1), Bile duct (1), Mediastinal neuroendocrine (1), Kaposi sarcoma (1)	Abdomen (10), Bone (3), LNs (8), Lung (7), Liver (2), Brain(2), Skin/soft tissue (2)
Didolkar et al.[12]	1950–1973	97/254	58	135/119	Adenocarcinoma (102) Undifferentiated Ca (84) Squamous Ca (37) Undetermined neoplasm (28)	71/97	Lung (31), Pancreas (5), Stomach (5), Kidney/adrenals (6), Bowel (4), Ovary (3), Lymphoma (3), Liver (3), Uterus (2)	Non available
Nystrom et al.[13]	1967–1974	130/266	NA	NA	Adenocarcinoma (178) Undifferentiated Ca (88)	107/130	Pancreas (30), Lung (28), Liver (16), Bowel (15), Stomach (12), Kidney/ adrenal (10), Ovary (4), Prostate (4), Breast (3), Other (6)	Lung (85), Liver (38), LNs (38), Brain (24), Bone (9), Skin (15), Peritoneum (6)
Le Chevalier et al.[14]	1959–1980	302/302	53	225/77	Adenocarcinoma (137) Undifferentiated Ca (83) Squamous Ca (46) Other (36)	255/302	Pancreas (80), Lung (52), Kidney/adrenal (24), Head/neck(19), Stomach (15), Bowel (11), Prostate (10), Ovary (10), Liver (8), Bladder (5),Other (21)	LNs (112), Lung (56), Bone (38), Brain (29), Skin (27), Liver (14), Pleura (7), Stomach (4), Other (15)
Stewart et al.[15]	1977–1978	16/87	60	45/42	Adenocarcinoma by definition	14/16	Lung (5), Bowel (3), Lymphoma (2), Pancreas (1), Liver (1), Stomach (1), Melanoma (1)	Non available

Table 3 – More recent autopsy series 1980–2000

Author	Study period	Autopsies/total number of patients	Median age at diagnosis	M/F	Histology	Primary site found	Site (*N*)	Metastatic sites
Hamilton and Langlands[16]	1979–1983	38/87	59	47/40	Adenocarcinoma (46) Undifferentiated Ca (29) Squamous Ca (4) Other (6)	30/38	Lung (14), Pancreas (4), Stomach (4), Bowel (2), Liver (2), Kidney (1), Ovary (1), Other (2)	LNs (20), Bone (18), Liver (18), Lung (17), Abdomen (14), CNS (8), Other (5)
Al Brahim et al.[5]	1980-2000	53/53	66	31/22	Adenocarcinoma (37) Undifferentiated Ca (5) Other (6)	27/53	Lung (7), Pancreas (4), Stomach (3), Bile duct (1), Appendix (1)	LNs (34), Liver (32), Lung (20), Adrenals (17), Bone (10), Brain (9), Peritoneum (9)
Blaszyk et al.[17]	1984–1999	64/64	64	34/30	Adenocarcinoma (51) Squamous Ca (3) Other (2)	35/64	Pancreas (13), Bowel (11), Lung (8), Ovarian (1), Prostate (1), Other (1)	Non available
Maiche[18]	1981–1988	64/109	62	62/47	Adenocarcinoma (37) Squamous Ca (33) Undifferentiated Ca (31) Other (8)	43/64	Lung (13), Kidney (6), Pancreas (4), Bowel (4), Liver (3), Breast (2), Ovarian (2), Bladder (2), Other (7)	Including liver (39), Bone only (25), LNs (28), Brain only (8), Liver only (4), Lung (24)
Mayordano et al.[19]	1974–1990	43/43	62	24/19	Adenocarcinoma (23) Undifferentiated Ca (4) Squamous Ca (3) Other (12)	35/43	Bile duct (7), Pancreas (6), Lung (4), Prostate (3), Stomach (2), Kidney (2)	Non available

Ca: carcinoma, LN: lymph nodes.

Table 4 – Comparison of autopsy data between older and recent series

Parameter	Autopsy series 1944–1980	Autopsy series 1980–2000	P-value
Relative frequency of autopsy in CUP patients	622/1269 (49%)	262/356 (73%)	0.032
Female population	282/748 (38%)	158/356 (44%)	0.045
Relative frequency of diagnosis of 'undetermined neoplasm'	38/892 (4%)	10/1077 (1%)	0.038
Identification of primary at autopsy	474/622 (76%)	170/262 (65%)	NS
Relative frequency of identified primary	Total 474	Total 170	
Tumours			
Lung	25%	27%	NS
Pancreas	25%	19%	0.05
Kidney/adrenals	9%	6%	NS
Stomach	7%	5%	NS
Bowel	7%	11%	0.05
Liver/bile duct	6%	6%	NS
Ovary/uterus	4%	3%	NS
Prostate	3%	2%	NS
Other	14%	21%	

NS: non significant at the two-sided $p < 0.05$ level.

population patients less often, make up half of the CUP cases.[13,14,20] In addition to the distinct epidemiology of CUP primaries, the pattern of metastatic dissemination carries intriguing characteristics. Apart from the high frequency of 'peculiar' metastatic deposits (18% of autopsied patients with available information), the incidence of pulmonary (46%), brain (16%) and bony (17%) deposits lies in the high end of the incidence range reported for metastatic primary tumours.[13,14,20,21] The relatively low incidence (23%) of liver metastases in CUP tumours that lie hidden in the gastrointestinal tract in half of the cases, further emphasises the unpredictable systemic dissemination of these tumours, which are adenocarcinomas in the majority of cases. Though several biases may partly explain these differences, the hypothesis of existence of a distinct genetic programme of CUP tumours preferentially arising in certain organ sites and seeding other secondary sites is generated by these data and remains to be tested.

3.2. *Molecular platform series*

Eight studies that reported use of multiple gene expression profiling platforms for identification of tissue of origin of malignant neoplasms were retrieved.[22–30] All were published after year 2000 and used either complementary DNA (cDNA) chips (six studies) or oligonucleotide chips (two studies). The basic concept was identical: messenger RNA (mRNA) was extracted from fresh frozen (six studies) or formalin-fixed paraffin-embedded (FFPE, four studies) tumour tissue, was purified, amplified and tagged with fluorescent stains. Subsequently, either after reverse transcription to cDNA or as is, it was hybridised to the microarray chips that harboured hundreds or thousands of probes complementary to multiple gene sequences. Analysis and quantitation of gene expression was obtained by study of fluorescent signal intensity generated by each chip probe.

In all studies, the investigators relied on differential multi-gene expression profiles that were generated by application of microarray technology in solid tumours of known primaries (training set). Thanks to these experiments, each tumour type was assigned a 'unique', specific multi-gene signature. The specificity of each genetic signature was confirmed by its application in a second, independent set of solid tumours of known primary (validation set), so as to test successful prediction of the tissue of origin of each tumour. The second assumption that investigators relied on, was that every tumour throughout its evolution and metastatic dissemination maintains its distinctive molecular signature. The final third step involved comparison of the multi-gene expression pattern of a CUP case to each of the tumour type-specific patterns and assignment of the CUP specimen to the primary tumour type to which the CUP molecular signature most closely resembles.

Crucial to this approach is the gene signature that will be selected for biological classification of tumours. In all studies, several thousands of genes were screened for differential expression in various solid tumour types. Complex mathematical modelling approaches were used so as to identify an 'optimal-minimal' set of genes that can discriminate between tumours of different tissues of origin. In all cases, unsupervised analysis (not taking into account sample classification) was followed by some method of supervised analysis (knowledge of sample classification) for the grouping of multi-gene expression patterns to tumour types. The mathematical approaches used were not standard across all studies published and are, therefore, a significant source of heterogeneity, in addition to the technical platforms and tumours used.

Table 5 summarises four published reports of gene signatures ranging from 110 to more than 16,000 genes that predict the site of origin of solid tumours with a 78–85% accuracy. Though these genetic signatures were able to correctly pinpoint the origin of metastatic deposits in metastatic tumours of known primary in blinded experiments with 84% accuracy, they were not tested in CUP. Table 6 shows data on three microarray genetic signatures that were found to identify the primary site in 76–96% of known primary tumours and were subsequently used for the biological classification of

Table 5 – DNA microarray platforms for biological assignment of primary without CUP validation

Author	Tissue	Microarray platform	Training set	Validation	Test sample accuracy	No. of genes	Validation in metastases from known primaries	Validation in metastases from CUP
Bloom[22]	Frozen tissue (no LCD)	Oligonucleotide and cDNA	400 primary tumours of 21 tumour types	140 tumours	85%	400	50 metastatic samples from known primaries (84% accuracy)	No
Ramaswamy[23]	Frozen tissue (no LCD)	Oligonucleotide	144 primary tumours of 14 types	46 primary tumours and 8 metastases	78% (30% in poorly differentiated carcinomas)	16,063	Correct classification of 6 out of 8 metastatic samples	No
Su[24]	Frozen tissue (no LCD)	cDNA	100 tumours of 11 types	63 primary tumours and 12 metastases	85%	110	Correct classification of 9 out of 12 metastatic samples	No
Ma[25]	Frozen tissue plus FFPE (no LCD)	cDNA	466 FFPE tumours of 39 types (75% primary tumours, 25% metastases)	112 FFPE tumours of 31 types (70% primary, 30% metastases)	82% (71% for poorly differentiated tumours)	979	Accuracy in predicting the known primary site of metastases 84%	No

LCD, laser capture microdissection, FFPE, formalin-fixed paraffin-embedded, cDNA: complementary deoxyribonucleic acid.

more than 500 CUP cases to neoplastic tissue of origin groups. The CUPPrint chip (Agendia, Amsterdam) used 495 genes, whereas data on the other two multi-gene expression chips were used for the generation of more easy to use low-density RT-PCR arrays with 79 and 10 genes, respectively. Breast adenocarcinoma was the most common primary tumour of origin of CUP cases, followed by pancreatic, colorectal, lung, hepatobiliary, renal, bladder, ovarian and gastric cancer (Fig. 2). Biological prediction of the primary was not possible with high likelihood in 10–15% of CUP specimens.

Comparison of the relative incidences of CUP primary tumours identified at autopsy or biologically assigned by microarray gene expression profiling in published series is summarised in Table 7. Although formal statistical comparison cannot be performed due to non-availability of data on the exact sample size of CUP cases studied by means of CUPPrint, some interesting points should be highlighted. Molecularly-assigned primaries tend to be lung or pancreatic cancer less often, while a breast, urothelial or colorectal primary is incriminated more often. Other tumour types are undentified as the neoplasm of origin with equal frequency in autopsy and molecular studies. The differences in the distribution of autopsy-found and molecularly-assigned primaries may be partly due to known or unknown biases. Still, they can also serve as hypothesis-generating clues suggestive of distinct biology of neoplastic clones arising in several tissues.

4. Discussion

The clinical practice-changing technological breakthrough of computerised tomography (CT) and the widespread adoption of diet and lifestyle characterising developed societies occurred throughout North America and Europe during the seventies and eighties. Therefore, comparison of autopsy CUP data between the periods 1944–1980 and 1980–2000 provides interesting epidemiological hints. The increase of the frequency of autopsies from 49% to 73% after 1980 contradicts reported trends.[3–5] A time and publication bias may be apparent here: clinical interest in CUP emerged after 1980 when several specialised centres developed strict autopsy strategies and started publishing their results, whereas before the eighties autopsies were only occasionally performed for that purpose in community oncology hospitals. The increase in the female proportion of CUP patients in the 1980–2000 period may be due to women taking up smoking after the second World War (lung, pancreatic, cervical, renal and bladder cancer) as well as due to better access to health services and improved socioeconomic status. Improvement in experience, immunohistochemical techniques, caryotyping, electron microscopy studies and reporting from specialised centres resulted in the decline of the rate of pathological diagnosis of 'undetermined neoplasm' from 4% to 1% after 1980.[31,32] Still, success at picking up the primary during autopsy (65–76% of autopsies) remained stable over time. The inability to identify a primary tumour post-mortem in a third of CUP cases in spite of technological advances in necropsy procedures implies that these primary tumours may either be microscopic or regressed. Study of their biological behaviour offers better promise for therapeutic applications than study of their anatomic localisation.

Table 6 – DNA microarray platforms for biological assignment of primary with CUP validation

Author	Tissue	Microarray platform	Training set	Validation	Accuracy	No. of genes	Validation in metastases from known primaries	Validation in metastases from CUP	Primary sites
Erlander, Van Laar[26,27]	Frozen tissue plus FFPE (LCD)	cDNA	700 tumours of 53 types	101 tumours (FFPE)	88%	Initially 2387; optimised gene subset 495 (CupPrint, Agendia)	Correct classification of 53 out of 61 metastases of known primary	$N > 500$ CUP	Breast 16%, Pancreas 12%, Bowel 11%, Liver/bile duct 9%, Genital system 8.5%, Lung 8%, Bladder/ ureter 6%, Kidney/ adrenals 5%, Stomach 4%, Other 20.5%
Tothill[28]	Frozen tissue (no LCD)	cDNA	229 tumours of 14 types	59 tumours (for low-density array)	79-gene qPCR low-density array with 96% accuracy	Initially 600; 79-gene low density array	Correct classification of 8 metastatic deposits	For 11 out of 13 CUP cases strong decision margin predictions	Squamous (2), Lung (4), Breast (3), Kidney (2), Colorectal (1), Ovary (1)
Talantov 2006[29]	FFPE (no LCD)	cDNA	386 primary tumours of 6 types	48 metastases (33 CUP; 15 from known primary)	10-gene qRT-PCR platform (Veridex) with 76% accuracy	10	Accuracy in predicting 11 out of 15 metastases of known primary	Primary assignment in 37/48 CUP	Lung (9), Kidney (8), Breast (5), Colorectal (4), Pancreas (2), Ovary (2), Prostate (2), Other (5), CUP (11)
Hainsworth[30]		Veridex 10-gene qRT-PCR platform						Primary assignment in 42/69	Lung (15), Pancreas (11), Colon (12), Ovary (4)

LCD, laser capture microdissection; FFPE, Formalin-fixed paraffin-embedded, cDNA, complementary deoxyribonucleic acid, qRT-PCR, quantitative reverse-transcription polymerase chain reaction, *N*, sample size.

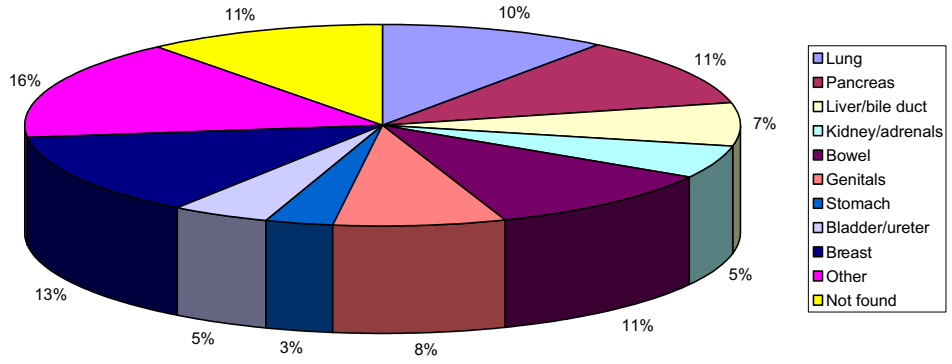

Fig. 2 – Relative proportion of molecularly-assigned primaries in published series.

Table 7 – Comparison of incidences of primary tumours found at autopsy and assigned by molecular platforms

Autopsy-found primaries, $N = 644$ (%)		DNA-assigned primaries, $N > 500$ (%)	
Lung	27	Lung	11.5
Pancreas	24	Pancreas	12.5
Liver/bile duct	8	Liver/bile duct	8
Kidney/adrenals	8	Kidney/adrenals	6
Bowel	7	Bowel	12
Genital system	7	Genital system	9
Stomach	6	Stomach	3
Bladder/ureter	0.01	Bladder/ureter	5
Breast	0.007	Breast	15
Other	10	Other	18

N: Sample size.

The epidemiology of autopsy-found CUP primary tumours seems to be shifting over time, partly in keeping with the general population cancer registry data.[20] Pancreatic cancer is identified less often. Asymptomatic pancreatic tumours are based in an inaccessible and poorly imaged location such as the retroperitoneum. Accordingly, the widespread use of abdominal CT after 1980 probably contributed to more frequent antemortem diagnosis of such tumours, eliminating them from the CUP group altogether. Smoking cessation, a turn to healthier diet, exercise and cervical smear screening may have also contributed to a decline in invasive pancreatic cancer incidence.[33] The increase in identification of colonic primaries in CUP patients is in keeping with the diet and lifestyle habits in developed societies. A diet rich in red meat, saturated lipids and poor in vegetable fibres and fruit, obesity and lack of exercise increase the relative risk of development of bowel malignancies.[34,35]

Overall, even modern autopsy series performed when antemortem investigation with CT was available show disproportionately high incidences of lung, pancreatic, hepatobiliary and renal primaries in CUP patients.[20] Since poor antemortem imaging no longer applies, one may reasonably hypothesise that neoplastic clones that regress or remain dormant in the primary site but metastasise early and widely, develop more often in these tissues. On the other hand, identification of breast, prostate, uterine/cervical and bladder primaries in CUP seldom occurs at least in part because of easier diagnosis of such small primaries in patients during their lifetime. Mammography, PSA testing, gynaecological examination and Pap test, cystoscopy coupled to early appearance of heamaturia actually pick up these tumours and 'clean up' the CUP cohort that includes only patients with microscopic tumours at these sites, possibly harbouring distinct genetic programmes.

Microarray platforms are not limited by problems such as inappropriate clinical investigation or accessibility/imaging biases since they rely on expression of multiple tissue-specific genes in order to classify a tumour. The classifying power of these platforms seems to be 75–90% with a very satisfactory 75–88% accuracy. Ability to identify the primary is maintained in metastatic deposits as well, though it drops with dedifferentiation.[7,23] The genes that make up the tissue-specific signatures used for primary assignment belong to the transcription factor family, protein translation apparatus, cellular membrane proteins, cell cycle checkpoints, cellular motility apparatus, chromosomal division regulators as well as stromal cell genes.

Intriguing discrepancies emerge when the relative incidence of primary tumours identified at autopsies or predicted by microarrays are compared in patients with CUP. The proportion of lung and pancreatic primaries decreases to half with the application of microarrays, a finding that is difficult to explain. Over-representation of these sites in the autopsy-studied CUP population due to the inability to properly diagnose the tumour antemortem does not suffice to explain the phenomenon, especially in series after 1980 when CT, flexible bronchoscopy and ERCP/MRCP became available. On the contrary, the relative incidence of breast and bladder cancer increases more than 100-fold when defined by means of microarrays instead of autopsy. The presence of 'sanctuary sites' during autopsy, sites that may harbour small or microscopic tumour foci but are not exhaustively evaluated pathologically (mammary glands, bladder/ureters) could explain this difference and should only be ruled out after meticulous audit of autopsy registries.[36] A bold, alternative hypothesis is that a constituent of a metastasis-prone genetic signature co-exists in all tumour primaries along with the tissue-specific genetic signature. This 'metastatic' genetic programme that transcends tissue-of-origin distinctions has been recently described and may lead to molecular

misclassification of various tumours to primary sites such as breast or bladder.[37]

Powerful as it is, microarray technology has several drawbacks. It is not known which genes are relevant for the biologic behaviour of a given tumour, since many may be surrogate markers for other genes performing pivotal functions. The selection of the classifying genes relies heavily on the number and type of primary solid tumours studied.[20,38] In the published series, the training set included less than 1000 tumour samples (typically 100–400) and less than 50 of them were metastatic deposits. These factors explain why many of the genes that define tissue-of-origin are not the same in the several microarray platforms devised. Moreover, in the way used today, data furnished by microarray platforms do not expand knowledge on the complex pathophysiology of CUP and metastatic dissemination in general. There is still significant heterogeneity in terms of materials, techniques, mathematical data manipulation and the need to rely on centralised analysis for reporting of chip results.[39,40] These facts add up to make microarray tests relatively labour-intensive and extremely costly. Moreover, it is not yet known if microarray analysis should be limited to neoplastic cells (laser capture microdissection) or should include non-neoplastic stroma.[37] Recent data showed that stromal cells carry differential gene expression that discriminates tumour types and contributes to survival of malignancy.[41]

Most worryingly, the information on primary furnished by microarrays does not provide any useful therapeutic or prognostic information. We do not possess any primary site-tailored antineoplastic therapy that will provide clinical benefit for the patient with metastatic disease. With the exception of a few favourable prognosis CUP subgroups, several investigators showed that tailored chemotherapy according to anticipated primary tumour did not lead to tumour control or survival analogous to that seen in patients with metastatic tumours of known primary site.[3] Microarrays were hoped to facilitate definition of prognosis by classifying CUP to a primary tumour type, but that may also prove false. Recently, Bishop et al. studied the outcome of more than 65,000 patients with metastatic tumours of known and unknown primaries and concluded that prognosis of CUP patients is different from that of patients with advanced tumours of equivalent primaries.[42]

In conclusion, their cost, lack of therapeutic or prognostic patient benefit and presence of methodological and analytical heterogeneity do not qualify microarray gene expression profiling platforms as the new benchmark in the hunt of the CUP primary. Autopsies should still be performed for registry, academic and patient families' benefit. There is an imperative need for insight into the pathophysiology of CUP and preliminary hints suggest that a pro-metastatic genetic signature exists. The amazing potential of microarray technology should be exploited in the search for genes implicated in the complex metastatic cascade. Encoded proteins could be targeted in order to develop new therapeutic approaches that may benefit patients with metastatic solid tumours. CUP constitutes an ideal metastatic model to be studied and targeted by means of microarrays and the opportunity should not be missed.

Conflict of interest statement

All authors of the paper disclose that there are no financial or personal relationships with other people or organisations that could inappropriately influence (bias) our work.

REFERENCES

1. Pentheroudakis G, Briasoulis E, Pavlidis N. Cancer of unknown primary site: missing primary or missing biology? *The Oncologist* 2007;**12**:418–25.
2. Pentheroudakis G, Pavlidis N. Perspectives for targeted therapies in cancer of unknown primary site. *Cancer Treat Rev* 2006;**32**:637–44.
3. Pavlidis N, Briasoulis E, Hainsworth J, Greco A. Diagnostic and therapeutic management of cancer of an unknown primary. *Eur J Cancer* 2003;**39**:1990–2005.
4. McPhee S. The autopsy: an antidote to misdiagnosis. *Medicine* 1996;**76**:41–3.
5. Al-Brahim N, Ross C, Carter B, Chorneyko K. The value of postmortem examination in cases of metastasis of unknown origin-20 year retrospective data from a tertiary care center. *Ann Diagnos Pathol* 2005;**9**:77–80.
6. Quackenbush J. Microarray analysis and tumour classification. *N Engl J Med* 2006;**354**:2463–72.
7. Dennis JL, Qien KA. Hunting the primary: novel strategies for defining the origin of tumours. *J Pathol* 2005;**205**:236–47.
8. Dennis JL, Vass K, Wit EC, Keith N, Oien KA. Identification from public data of molecular markers of adenocarcinoma characteristic of the site of origin. *Cancer Res* 2002;**62**:5999–6005.
9. Snyder RD, Mavligit GM, Valdivieso M. Adenocarcinoma of unknown primary site: a clinico-pathological study. *Med Pediatr Oncol* 1979;**6**:289–94.
10. Jordan WE, Shildt RA. Adenocarcinoma of unknown primary site: The Brooke Army Medical Center Experience. *Cancer* 1985;**55**:857–60.
11. Steckel RJ, Kagan R. Diagnostic persistence in working up metastatic cancer with an unknown primary site. *Radiology* 1980;**134**:367–9.
12. Didolkar MS, Fanous N, Elias EG, Moore RH. Metastatic carcinomas from occult primary tumours. A study of 254 patients. *Ann Surg* 1977;**186**:625–30.
13. Nystrom JS, Weiner JM, Heffelfinger-Juttner J, Irwin LE, Bateman JR, Meshnik-Wolf R. Metastatic and histologic presentations in unknown primary cancer. *Semin Oncol* 1977;**4**:53–8.
14. Le Chevalier T, Cvitkovic E, Caille P, et al. Early metastatic cancer of unknown primary origin at presentation. A clinical study of 302 consecutive autopsied patients. *Arch Intern Med* 1988;**148**:2035–9.
15. Stewart JF, Tattersall MHN, Woods RL, Fox RM. Unknown primary adenocarcinoma: incidence of overinvestigation and natural history. *Br Med J* 1979;**1**:1530–3.
16. Hamilton CS, Langlands AO. Adenocarcinoma of unknown primary site: a clinical and cost benefit analysis. *Int J Radiat Oncol Biol Phys* 1987;**13**:1497–503.
17. Blaszyk H, Hartmann A, Bjornsson J. Cancer of unknown primary: clinicopathologic correlations. *APMIS* 2003;**111**:1089–94.
18. Maiche AG. Cancer of unknown primary: a retrospective study based on 109 patients. *Am J Clin Oncol* 1993;**16**:26–9.
19. Mayordomo JL, Guerra JM, Guijarro C, et al. Neoplasms of unknown primary site: a clinicopathological study of autopsied patients. *Tumori* 1993;**79**:321–4.

20. Jemal A, Siegel R, Ward E, et al. Cancer statistics. *CA Cancer J Clin* 2006;**56**:106–30.
21. Leong SP, Cady B, Jablons DM, et al. Clinical patterns of metastasis. *Cancer Metastas Rev* 2006;**25**:221–32.
22. Bloom G, Yang IV, Boulware D, et al. Multi-platform, multi-site, microarray-based human tumour classification. *Am J Pathol* 2004;**164**:9–16.
23. Ramaswamy S, Tamayo P, Rifkin R, et al. Multiclass cancer diagnosis using tumour gene expression signatures. *PNAS* 2001;**98**:15149–54.
24. Su AI, Welsh JB, Sapinoso LM, et al. Molecular classification of human carcinomas by use of gene expression signatures. *Cancer Res* 2001;**61**:7388–93.
25. Ma XJ, Patel R, Wang X, et al. Molecular classification of human cancers using a 92-gene RT-PCR assay. *Arch Pathol Lab Med* 2006;**130**:465–73.
26. Erlander MG, Moore MW, Cotter P, et al. Molecular classification of carcinoma of unknown primary by gene tissues. *J Clin Oncol* 2004;**22**(14S): abstr 9545.
27. Van Laar RK, Horlings H, Warmoes M, et al. Implementation and validation of a clinically available cancer of unknown primary microarray-based diagnostic test. Proceedings of the 5th American Association of Cancer Research Oncogenomics Conference; 2007 Jan 31–Feb 3, Phoenix, Arizona; AACR 2007: abstr B13.
28. Tothill RW, Kowalczyk A, Rischin D, et al. An expression-based site of origin diagnostic method designed for clinical application to cancer of unknown origin. *Cancer Res* 2005;**65**:4031–40.
29. Talantov D, Baden J, Jatkoe T, et al. A quantitative RT-PCR assay to identify metastatic carcinoma tissue of origin. *J Mol Diagnost* 2006;**8**:320–9.
30. Hainsworth J, Talantov D, Jatkoe T, et al. Gene profiling of tumour tissue in the diagnosis of patients with carcinoma of unknown primary site: evaluation of the Veridex 10-gene molecular assay. *J Clin Oncol* 2007;**25**(18S): abstr 21109.
31. Battifora H. Recent progress in the immunohistochemistry of solid tumours. *Semin Diagn Pathol* 1984;**1**:251–71.
32. Dennis JL, Hvidsten TR, Wit EC, et al. Markers of adenocarcinoma characteristic of the site of origin: development of a diagnostic algorithm. *Clin Cancer Res* 2005;**11**:3766–72.
33. Lowenfels AB, Maisonneuve P. Epidemiology and risk factors for pancreatic cancer. *Best Pract Res Clin Gastroenterol* 2006;**20**:197–209.
34. Ahmed FE. Colon cancer: prevalence, screening, gene expression and mutation, and risk factors and assessment. *J Environ Sci Health C Environ Carcinog Ecotoxicol Rev* 2003;**21**:65–131.
35. Martinez ME. Primary prevention of colorectal cancer: lifestyle, nutrition, exercise. *Recent Results Cancer Res* 2005;**166**:177–211.
36. Tsujimura T, Yamada Y, Kubo M, Fushimi H, Kameyama M. Why couldn't an accurate diagnosis be made? An analysis of 1044 consecutive autopsy cases. *Pathol Int* 1999;**49**: 408–10.
37. Ramaswamy S, Ross KN, Lander ES, Golub TR. A molecular signature of metastasis in primary solid tumours. *Nat Genet* 2003;**33**:49–54.
38. Simon R, Radmacher MD, Dobbin K, McShane LM. Pitfalls in the use of DNA microarray data for diagnostic and prognostic classification. *J Natl Cancer Inst* 2003;**95**:14–8.
39. Brazma A, Hingamp P, Quackenbush J, et al. Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. *Nat Genet* 2001;**29**: 365–71.
40. Dupuy A, Simon R. Critical review of published microarray studies for cancer outcome and guidelines on statistical analysis and reporting. *J Natl Cancer Inst* 2007;**99**:147–57.
41. Zalatnai A. Molecular aspects of stromal-parenchymal interactions in malignant neoplasmas. *Curr Mol Med* 2006;**6**:685–93.
42. Bishop JF, Tracey E, Glass P, Jelfs P, Roder D. Prognosis of sub-types of cancer of unknown primary (CUP) compared to metastatic cancer. *J Clin Oncol* 2007;**25**(18S). abstr 21010.