The detection of circulating breast cancer cells in blood

A M Gilbey, D Burnett, R E Coleman, I Holen

At present, sampling of the lymph nodes or bone marrow for the detection of regions of metastatic disease in patients with breast cancer can only be undertaken at the time of initial diagnosis and surgery. However, the sampling of these tissues and the methods used are inaccurate, time consuming, and cannot be used for easy routine screening to determine disease recurrence and response to treatment. Because of the problems encountered with current methods and tissues sampled at the time of breast cancer diagnosis, this review discusses the urgent requirement for and potential development of a quick, simple, and accurate diagnostic test utilising the haematogenous system, a source of circulating tumour cells in patients with breast cancer, and highly sensitive molecular biological techniques, such as reverse transcription polymerase chain reaction. In addition, this review also highlights potential problems that may be encountered and should be avoided when devising such a test.

Breast cancer remains an important public health problem. The incidence of this disease and the resulting deaths continue to increase and are higher among women who reside in Western, developed countries. In 1995, 182 000 new cases of breast cancer were diagnosed in the USA, and in 1997 approximately 36 000 new cases were diagnosed in the UK. The overall incidence of breast cancer in these two countries totalled approximately 30% of all cancers identified in women, and was more than twice as high as the second most common cancers, colorectal cancer and lung/bronchus cancer in the UK and USA, respectively.

"The incidence of breast cancer and the resulting deaths continue to increase and are higher among women who reside in Western, developed countries."

The steady increase in the incidence of breast cancer has been attributed to many factors. These include increased screening and thus detection of the disease, an increasingly aging population, and changes in lifestyle risk factors, in particular the use of oral contraceptives and alcohol consumption. The improvements in the detection and treatment of breast cancer have resulted in a steady increase in the longterm survival of patients with breast cancer. Consequently, the increase in breast cancer incidence ultimately results in a growing demand for supportive, palliative, and medical services.

Breast cancer metastases

Metastasis is a cascade of linked sequential steps involving multiple host–tumour interactions. Cancer cells possessing multiple genetic abnormalities grow unregulated and eventually lose the ability to adhere to one another. Thus, tumour cells detach from the primary tumour, migrate through the basement membrane and extracellular matrix, intravasate, and travel in the lymphatic and/or blood systems to a new site, before attachment, extravasation, and the development of a new focus and neovascularisation. When cancer cells enter the lymphatic system, they travel to the lymph nodes (LN) in the case of breast cancer, to the sentinel nodes in the axilla and intercostal spaces—before entering the bloodstream and subsequent progression to other organs.

Axillary LN metastases

Although most patients with newly diagnosed breast cancer have operable disease and are therefore considered potentially curable, the probability of all patients who possess histologically confirmed negative LN remaining disease free after five years was calculated to be 85%. Thus, 15% of "node negative" patients will probably develop metastatic disease. There are two possible reasons for the spread of the disease in these node negative patients, namely: (1) metastatic regions within the LN are not detected by currently used techniques, and (2) metastatic spread of this disease occurs by an alternative route to that of the LN.

Detection of axillary LN metastases

Patients with breast cancer less than 2 cm in diameter have also been found to possess LN metastases by routine histochemical examination, a process that involves staining sectioned tissues, previously embedded in paraffin wax, with two dyes, haematoxylin and eosin. These patients are classed as "node positive" at initial diagnosis. However, Nasser et al and Cote et al showed that immunohistochemistry (IHC), a technique using antibodies to detect epithelial specific antigens, was able to detect regions of...
metastases in LN that were previously undetected by haematoxylin and eosin staining.\textsuperscript{11, 12} Considering that Gusterson and Ott calculated that a pathologist has a 1% chance of identifying a metastatic focus in a patient with breast cancer which is three cells in diameter,\textsuperscript{13} and that Cote \textit{et al} determined that the identification of regions of metastasis by haematoxylin and eosin staining in node negative LNs required the analysis of up to 144 slides/patient,\textsuperscript{1} it is not surprising that regions of metastasis within the LN are not detected.

**Axillary LN removal and metastatic spread of breast cancer**

In contrast to those studies detailed above, Fisher \textit{et al} compared the incidence of metastatic disease in two sets of patients with breast cancer, those who had their LN removed at the time of surgery and those whose LN were not removed.\textsuperscript{15} This study revealed that the removal of LN did not alter significantly the subsequent incidence of metastatic disease in patients with breast cancer.\textsuperscript{13} This suggests that in these patients the haematogenous route and not the lymphatic system may have aided tumour cell dissemination and the metastatic spread of breast cancer.

Although IHC is a more accurate technique than histochemical examination for the diagnosis of metastases,\textsuperscript{11, 12} It is too cumbersome and costly and is therefore not used routinely in clinical practice.\textsuperscript{10} Hence, there is a need for a quick and accurate test for the diagnosis of metastatic breast cancer cells in such tissues, which could assist in determining the possibility of the disease recurring or metastasising.

**Bone marrow metastases**

As reviewed previously, bone is the single most common site of metastasis in patients with breast cancer,\textsuperscript{16, 17} and patients with recurring tumours will develop bone marrow (BM) metastases at some point during the evolution of their disease. Although LN involvement is thought to be the most important prognostic factor for tumour recurrence,\textsuperscript{2} the detection of tumour cells by IHC within the BM has emerged as an independent marker of disease recurrence or survival in patients with breast cancer.\textsuperscript{18–20}

Tumour cells isolated from BM have been shown to be capable of clonogenic growth in vitro,\textsuperscript{21} thus indicating that tumour cells present in the BM could participate in the recrudescence of the tumour and the development of metastatic disease in vivo. Indeed, several studies have shown that the presence of tumour cells in the BM correlates strongly with an early relapse of disease and decreased patient survival.\textsuperscript{19, 20} Consequently, analysis of the BM may help in the diagnosis of the spread of breast cancer and may also define those patients at higher risk of recurrence and death.

However, it should be noted that 15% of patients who relapsed previously exhibited no immunohistochemical evidence of BM metastases,\textsuperscript{20} and sampling of BM for subsequent analysis is both a painful and costly procedure. Thus, the analysis of BM aspirates may not represent the best method for the detection of tumour cells as an indicator of the spread of breast cancer. Furthermore, this technique is not amenable to multiple sampling, particularly for monitoring patients’ responses to treatment.

**The fate of circulating tumour cells**

The haematogenous route offers a potential source of circulating tumour cells and blood sampling can be done at frequent intervals and is relatively painless. Thus, the development of a painless, quick, repeatable, and accurate test to detect metastatic breast cancer cells is a worthwhile objective.

Several methods used routinely in the treatment of breast cancer have been shown to increase the numbers of breast cancer cells circulating in the blood, which persist for varying lengths of times in different patients.\textsuperscript{22–24} However, the fate of these cells shed into the bloodstream and the impact this has upon the treated patient and their recovery from the disease remains unknown.

\textbf{The haematogenous route offers a potential source of circulating tumour cells and blood sampling can be done at frequent intervals and is relatively painless.} \textsuperscript{22–24}

Although most circulating tumour cells have been shown to be apoptotic,\textsuperscript{25} a proportion of these circulating cells may be capable of clonogenic growth in vitro, as seen previously with breast cancer cells isolated from the BM.\textsuperscript{22, 23} Most recently, the breast cancer cells capable of establishing tumours were identified as those that expressed CD44 but not CD24.\textsuperscript{22} Thus, the cells capable of clonogenic growth in vitro\textsuperscript{24} may be CD44$^+$CD24$^-$ and thus represent the small proportion of circulating breast cancer cells capable of establishing metastases in vivo. Therefore, the development of a diagnostic test to detect these CD44$^+$CD24$^-$ circulating tumour cells may indicate a patient’s potential for tumour recurrence and development of metastatic disease, particularly during treatment.

**Detection of circulating tumour cells and nucleic acids**

Various molecular and cellular approaches, differing in their sensitivity and specificity, have been used in the detection of circulating tumour cells or circulating tumour cell nucleic acids within the blood. IHC and flow cytometry (FC) have been used to detect circulating cells based upon the expression of specific cell surface markers. Circulating nucleic acids and tumour cells have been detected using the polymerase chain reaction (PCR), which can detect specific regions of DNA, or reverse transcription PCR (RT-PCR) and variations of RT-PCR, which can detect mRNA, thus indicating the expression of specific genes. These techniques are discussed in the following sections.

**Expression of cell surface markers**

\textit{Immunohistochemistry}

IHC has been used for the detection of metastases in BM and LN and, more recently, the presence of circulating tumour cells in blood smears.\textsuperscript{19, 20, 27} This technique has proved to be more accurate in the detection of metastatic regions in the LN than haematoxylin and eosin staining.\textsuperscript{19, 20} However, there are several drawbacks in using this method for the detection of regions of metastasis.

Metastatic breast cancer cells show considerable heterogeneity in the expression of many carcinoma associated cell surface molecules,\textsuperscript{28} and this may affect the efficacy of this method in the detection of circulating metastatic cells. In addition, IHC is a time consuming method,\textsuperscript{29} particularly because at least 100 000 cells need to be analysed for a reliable assessment of the presence of tumour cells by IHC.\textsuperscript{19} Furthermore, IHC requires a trained cytologist to confirm the identity of the stained cells.\textsuperscript{30} However, most importantly, IHC is unable to make an accurate measurement of the circulating tumour cell load within the blood. Therefore, IHC could not be used for the routine screening of patients with breast cancer to determine disease recurrence and response to treatment.

**Flow cytometry**

A flow cytometric assay has been developed that can detect rare breast cancer cells in spiked blood and bone marrow...
samples, and has subsequently been used to detect circulating epithelial cells within the blood of patients with breast cancer. FC, in combination with immunomagnetic separation, has been shown to be capable of detecting less than one epithelial cell/ml of blood. However, although this technique is very sensitive, there are several disadvantages in the use of such an approach to detect circulating breast cancer cells in blood.

At present, FC alone cannot morphologically characterise the epithelial cells detected within a given blood sample. Consequently, further IHC analysis of the identified epithelial cells using specific breast cancer marker antibodies is required to confirm that these epithelial cells are breast cancer cells. This will invariably result in an increase in the time taken and costs incurred in performing such an assay. In addition, the use of immunomagnetic separation to enrich the epithelial cell content before FC can also result in the detection of epithelial cells within control samples. This therefore dictates the subsequent analysis of isolated epithelial cells to confirm that they are authentic tumour cells.

Although FC, in combination with immunomagnetic separation and IHC analysis, has shown a good correlation between changes in tumour cell load within the blood and treatment and clinical status of the patient, the development of a test to detect circulating tumour cells within the blood using such methods will be time consuming and costly.

Detection of circulating nucleic acids
Increased quantities of nucleic acids derived from various tumours have been detected in the plasma of patients with cancer by various molecular techniques including PCR and RT-PCR.

Circulating DNA
Chromosomal abnormalities, including microsatellite instability, loss of heterozygosity, and gene mutations present in tumour DNA have been detected in the circulating DNA in the blood of many patients with breast cancer, lung cancer, colorectal cancer, and head and neck cancer. Although the routes by which tumour DNA in such patients enters the circulation currently remain unknown, cell lysis, tumour necrosis, apoptosis, and active cell shedding have been hypothesised to be involved in this process.

A significant correlation has been reported between the presence of circulating tumour DNA and a poor prognosis in patients suffering from small cell lung cancer, breast cancer, and cancer of the oesophagus. This suggests that the detection of tumour cell DNA may be a useful indicator of the development of metastatic disease. Nevertheless, the detection of circulating tumour DNA has several limitations. As discussed previously, tumour cell DNA can be amplified by PCR, thus indicating the presence of small numbers of tumour cells shedding nucleic acid within a heterogeneous population of cells. DNA, however, is relatively stable in human tissues, and furthermore, the fraction of circulating tumour DNA to which tumours contribute varies between patients. Therefore, the detection and measurement of circulating tumour DNA may not be indicative of viable tumour cells actively shedding nucleic acids and thus the potential of a particular patient to develop metastatic disease.

An additional point that requires consideration in the detection of circulating tumour DNA is the particular gene target used in the PCR. To use the PCR technique efficiently, the tumour cell DNA must possess specific changes that would distinguish it from the DNA shed from surrounding haemopoietic cells. Tumour specific DNA sequence abnormalities are uncommon in solid tumours. Thus, the detection of nucleic acids from circulating tumour cells by PCR would require the molecular analysis of the primary tumour from each patient to determine any individual genomic alterations.

Circulating mRNA
Telomerase, cytokeratin 19 (CK19), and mammaglobin mRNA have been identified in plasma from patients with breast cancer, and tyrosinase mRNA in plasma from patients with metastatic melanoma. Although two of these studies did not relate the presence of tumour specific mRNA in the plasma of the patient to the clinicopathological characteristics of the tumours, an association has been shown between the presence of circulating mammaglobin and CK19 mRNA in patients with breast cancer and poor prognostic features, including increased tumour size, proliferative index, and histological stage.

"An association has been shown between the presence of circulating mammaglobin and cytokeratin 19 mRNA in patients with breast cancer and poor prognostic features"

The detection of circulating mRNA by RT-PCR allows the amplification of genes that are specifically expressed or the expression of which is significantly upregulated in tumour cells. The use of this technique would eliminate the requirement for screening primary tumours for specific chromosomal aberrations and genomic alterations. In addition, mRNA is more fragile than DNA and is presumed to be highly susceptible to degradation by blood RNases. Thus, the detection of tumour specific mRNA could be indicative of viable tumour cells actively shedding nucleic acids.

Detection of circulating cells
It is generally thought that molecular techniques are more sensitive than the cellular technique IHC, and that the implementation of such techniques has greatly improved the ability to detect low numbers of breast cancer cells.

Two studies showed that the detection of tumour cells expressing CK19 in both BM and blood mononuclear cells was 10 times more sensitive using nested RT-PCR rather than IHC. However, in contrast, a similar level of detection was seen for carcinoembryonic antigen and maspin using both nested RT-PCR and IHC. Thus, it is clear that the difference in the sensitivity of these two techniques depends upon the particular tumour marker being assayed. Although the sensitivities of these two techniques in these studies were comparable, the detection rate of the nested RT-PCR was greater than that of IHC and the test detected circulating tumour cells in samples that were shown to be negative by IHC.

In addition to the implementation of more sensitive molecular techniques, it has been suggested that enriching the tumour cell population from the blood before analysis by such molecular methods may increase the sensitivity for the detection of circulating tumour cells. Immunomagnetic separation, used before RT-PCR for tumour cell enrichment from BM, resulted in the detection of one CK19 expressing tumour cell in 10^4–10^5 mononuclear cells and was therefore at least 10 times more sensitive than using RT-PCR alone in this study. The detection of tumour cells by this method was also approximately 100 times more sensitive than the detection of CK19 by IHC. However, the use of such a technique to enrich the circulating tumour cell population from blood samples would ultimately increase both the cost and time taken to prepare and analyse blood samples from patients with breast cancer.

RT-PCR and variations of this technique have been used in the detection of specific tumour cell markers, thus indicating the presence of circulating breast cancer cells in blood, with varying degrees of sensitivity and specificity (table 1). For
example, circulating tumour cells expressing CK19 have been detected by nested RT-PCR,27 58–60 62 64 65 competitive nested RT-PCR,41 and quantitative RT-PCR41 46 with varying levels of sensitivity (table 1). These assays, detecting circulating tumour cells expressing CK19, have been shown to be both specific27 59–62 and non-specific,60 61 63 64 66. In contrast, the detection of the β chain of human chorionic gonadotrophin,64 parathyroid hormone related protein,65 mammaglobin,57–70 and Her2/neu62 expressing tumour cells appears to be 100% specific, regardless of the molecular technique used. Consequently, these studies indicate that the development of a quick, sensitive, and specific diagnostic test to detect circulating tumour cells is feasible. Nevertheless, to avoid non-specific amplification and thus false positive results, the particular tumour marker selected requires careful consideration.

**MULTIPLE MARKER ASSAYS**

Breast cancers are composed of a heterogeneous collection of cells with differing degrees of tumour marker expression. Thus, circulating tumour cells within a particular patient with breast cancer may not express the particular tumour marker being assayed. Consequently a multiple marker assay, taking into account tumour heterogeneity and mRNA expression variability, would improve the detection of circulating cells compared with the detection of a single marker in patients with breast cancer.71

“Multiple marker assays may significantly improve the sensitivity of detecting heterogeneous tumour cells compared with single marker assays”

Several studies, using two or more breast cancer markers to detect circulating cells in blood, were able to detect varying numbers of circulating cells (table 2). Of these studies, eight commented upon the combined positivity when using more than one tumour marker.28 49 71 77 79 82–84 The heterogeneity of breast cancer cells meant that these multiple marker assays detected circulating tumour cells in the blood more frequently than the individual markers alone. Consequently, multiple marker assays may significantly improve the sensitivity of detecting heterogeneous tumour cells compared with single marker assays. However, this is dependant upon the selection of markers.

**Selection of markers**

The ability to identify circulating breast cancer cells requires a sensitive and specific test. The ideal marker for such a test would be universally but uniquely expressed on all breast cancer cells. However, no single specific marker linked to breast cancer that meets these criteria has been identified. Consequently, a multimarker assay has to be developed to increase the sensitivity of disease detection.85 The development of a multimarker diagnostic test requires careful consideration and is dependant upon the questions to be addressed. Such questions are as follows:

1. Are these circulating cells epithelial? CK19 is a cytoplasmic intermediate filament that exists primarily in epithelial cells and tumour cells derived from epithelial cells, and is not expressed in blood. Subsequently, the detection of CK19 within the blood of a patient with

<table>
<thead>
<tr>
<th>Marker</th>
<th>Amplification method</th>
<th>PCR cycling conditions</th>
<th>Sensitivity</th>
<th>Specificity in control blood</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>CK20</td>
<td>Nested RT-PCR</td>
<td>One 35 cycle reactions</td>
<td>10⁵ ZR75.1 or 10⁵ MCF-7/T47-D cells detected in 10⁶ CK20 negative PBMCs or 1–100 ng of total tumour RNA, depending upon tumour or cell line</td>
<td>Not specific</td>
<td>30</td>
</tr>
<tr>
<td>c-erbB-2</td>
<td>Semi-nested RT-PCR</td>
<td>One 30 cycle reaction and one 20 cycle reaction</td>
<td>1 MCF-7 cell in 10⁶ CCRF-CEM cells</td>
<td>Not specific</td>
<td>52</td>
</tr>
<tr>
<td>Maspin</td>
<td>Semi-nested RT-PCR, Southern blot</td>
<td>One 35 cycle reaction</td>
<td>1 BC cell in 10⁶ PBL</td>
<td>NA</td>
<td>53</td>
</tr>
<tr>
<td>EGFR</td>
<td>RT-PCR</td>
<td>One 30 cycle reaction</td>
<td>1 BC cell in 10⁶ PBL</td>
<td>Specific</td>
<td>54</td>
</tr>
<tr>
<td>PTHR</td>
<td>Nested RT-PCR</td>
<td>One 20 cycle reaction</td>
<td>Detect EGFR transcripts in 30 fg</td>
<td>Not specific</td>
<td>55</td>
</tr>
<tr>
<td>MUC1</td>
<td>One 45 cycle reactions</td>
<td>1 MCF-7 cell in 10⁶ PBL</td>
<td>Not specific</td>
<td>57</td>
<td></td>
</tr>
<tr>
<td>CK19</td>
<td>Nested RT-PCR</td>
<td>One 35 cycle reaction</td>
<td>Detect 5 pg of MCF7 RNA or 1 MUC1 expressing cell in 5 ml blood</td>
<td>Specific</td>
<td>56</td>
</tr>
<tr>
<td>CK19</td>
<td>Nested RT-PCR</td>
<td>Two 35 cycle reactions</td>
<td>1 MCF-7 cell detected in 10⁶ BM cells</td>
<td>Not specific</td>
<td>60</td>
</tr>
<tr>
<td>CK19</td>
<td>Nested RT-PCR</td>
<td>Two 35 cycle reactions</td>
<td>Detect EGFR transcripts in 30 fg</td>
<td>Specific</td>
<td>62</td>
</tr>
<tr>
<td>CK19</td>
<td>Nested RT-PCR</td>
<td>One 20 cycle reaction and one 22 cycle reaction</td>
<td>&lt;1 MCF-7 cell in 10⁶ PBL</td>
<td>Specific</td>
<td>63</td>
</tr>
<tr>
<td>CK19</td>
<td>Q RT-PCR</td>
<td>One 30 cycle reaction</td>
<td>1 MCF-7 cell detected in 10⁶ BM cells</td>
<td>Not specific</td>
<td>64</td>
</tr>
<tr>
<td>CK19</td>
<td>Nested RT-PCR</td>
<td>Two 35 cycle reactions</td>
<td>1 MCF-7 cell detected in 10⁶ BM cells</td>
<td>Specific</td>
<td>65</td>
</tr>
<tr>
<td>CK19</td>
<td>Nested RT-PCR</td>
<td>Two 35 cycle reactions</td>
<td>Detect EGFR transcripts in 30 fg</td>
<td>Specific</td>
<td>66</td>
</tr>
<tr>
<td>CK19</td>
<td>Q RT-PCR</td>
<td>One 30 cycle reaction</td>
<td>1 MCF-7 cell equivalent</td>
<td>Not specific</td>
<td>67</td>
</tr>
<tr>
<td>Mammaglobin</td>
<td>Semi-nested RT-PCR</td>
<td>Two 35 cycle reactions</td>
<td>1 MCF-7 cell detected in 10⁶ BM cells</td>
<td>Specific</td>
<td>68</td>
</tr>
<tr>
<td>Mammaglobin</td>
<td>Nested RT-PCR</td>
<td>Two 35 cycle reactions</td>
<td>Detect EGFR transcripts in 30 fg</td>
<td>Specific</td>
<td>69</td>
</tr>
<tr>
<td>Mammaglobin</td>
<td>Q RT-PCR</td>
<td>One 45 cycle reaction</td>
<td>1 MCF-7 cell detected in 10⁶ BM cells</td>
<td>Specific</td>
<td>70</td>
</tr>
<tr>
<td>Mammaglobin</td>
<td>Nested RT-PCR</td>
<td>Two 35 cycle reactions</td>
<td>Detect EGFR transcripts in 30 fg</td>
<td>Specific</td>
<td>71</td>
</tr>
<tr>
<td>Mammaglobin</td>
<td>RT-PCR</td>
<td>One 35 cycle reaction</td>
<td>1 MCF-7 cell detected in 10⁶ BM cells</td>
<td>Specific</td>
<td>72</td>
</tr>
<tr>
<td>Markers</td>
<td>Amplification method</td>
<td>PCR cycling conditions</td>
<td>Sensitivity</td>
<td>Specificity in control blood</td>
<td>Combined positivity</td>
</tr>
<tr>
<td>---------</td>
<td>----------------------</td>
<td>------------------------</td>
<td>-------------</td>
<td>-----------------------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>CEA, CK19, maspin</td>
<td>Semi-nested or nested RT-PCR</td>
<td>CEA: one 30 and one 20 cycle reaction CK19: two 35 cycle reactions Maspin: two 35 cycle reactions</td>
<td>CEA/maspin: 10 T47-D cells/10^6 MNC CK19: 1 T47-D cell/10^6 MNC</td>
<td>CEA: specific CK19 and maspin: not specific</td>
<td>50% of samples were positive for CK19 and CEA 83% were positive for at least one of the two markers CEA and CK19 90% and 70% of maspin positive samples were positive for CEA and CK19 respectively 1 sample positive for both markers 49% patients positive for 1 marker 15% positive for 2 markers 3% positive for 3 markers 2% positive for all four markers</td>
</tr>
<tr>
<td>M'globin, CK19, p-hCG, c-Met, MAGE-A3, j-NaxGAT</td>
<td>Nested RT-PCR</td>
<td>One 43 and one 40 cycle reaction One 35 cycle reaction for each marker</td>
<td>1 tumour cell/ml of blood 1–5 tumour cells in 10^7 PBL for each marker Each marker could be detected at total RNA concentrations &gt;1.0 ng/ml</td>
<td>Specific Specific</td>
<td></td>
</tr>
<tr>
<td>CEA, CK19, CK20, GA733.2, MUC1</td>
<td>RT-PCR and southern blot</td>
<td>One 35 cycle reaction for each marker</td>
<td>10 ng RNA of each of the markers could be detected by southern blotting</td>
<td>CEA: specific for CK19 and CEA 83% were positive for at least one of the two markers CEA and CK19 90% and 70% of maspin positive samples were positive for CEA and CK19 respectively 1 sample positive for both markers 49% patients positive for 1 marker 15% positive for 2 markers 3% positive for 3 markers 2% positive for all four markers</td>
<td></td>
</tr>
<tr>
<td>MUC1, CK19, CD44</td>
<td>RT-PCR and nested RT-PCR</td>
<td>MUC1: one 30 cycle reaction CK19: one 40 cycle reaction and one 35 cycle reaction CD44: one 20 cycle reaction and one 30 cycle reaction Two 35 cycle reactions</td>
<td>MUC1: 1 MCF-7 cell/ml of blood OK19: 10^5 MCF-7 cells/10 ml of blood</td>
<td>MUC1: specific CK19: not specific</td>
<td></td>
</tr>
<tr>
<td>Maspin, CK19</td>
<td>Nested RT-PCR</td>
<td>One 35 cycle reaction and one 30 cycle reaction</td>
<td>10 MCF-7 cell in 10^5 PBWN</td>
<td>Maspin: specific CK19: not specific</td>
<td></td>
</tr>
<tr>
<td>CK19, MUC1, CEA</td>
<td>RT-PCR, semi-nested, and nested RT-PCR</td>
<td>CK19: One 35 and one 30 cycle reaction MUC1: one 35 cycle reaction</td>
<td>CEA: one 20 cycle reaction and one 30 cycle reaction</td>
<td>CEA (A): 1 MCF-7 cell in 10^5 PBWN; CEA (B): 1 MCF-7 cell in 10^5 PBWN</td>
<td>All CEA positive patients were positive for CK19 Not all CK19 positive patients were CEA positive</td>
</tr>
<tr>
<td>CK19, MUC1, CEA</td>
<td>RT-PCR, semi-nested, or nested RT-PCR</td>
<td>CK19: two 35 cycle reactions MUC1: one 35 cycle reaction CEA (A): one 20 cycle reaction and one 30 cycle reaction CEA (B): 1 T47-D cell/10^6 MNC</td>
<td>CEA (A): 1 MCF-7 cell in 10^5 PBWN; CEA (B): 1 MCF-7 cell in 10^5 PBWN</td>
<td>CEA: specific unless treated with G-CSF</td>
<td></td>
</tr>
<tr>
<td>M'globin, CK19, EGFR</td>
<td>Nested RT-PCR</td>
<td>M'globin and EGFR: two 40 cycle reactions CK19: two 30 cycle reactions MUC1/EGFR/CEA: two 30 cycle reactions</td>
<td>1 MD-MBA-361 cell in every 10^6 PBWN</td>
<td>M'globin: specific</td>
<td>5 patients M'globin/CK19 positive and 2 patients positive for all 3 markers</td>
</tr>
<tr>
<td>Maspin, m'globin, CK19, CK20, MUC1, EGFR</td>
<td>Nested RT-PCR</td>
<td>MUC1/EGFR/CEA: two 30 cycle reactions CK19: two 35 cycle reactions CK20: two 32 cycle reactions Mammaglobin: one 32 cycle and one 35 cycle reaction Maspin: one 30 cycle and one 35 cycle reaction</td>
<td>1–10 MCF7 or SKBR-3 cells/10^6 PBWN cells for all assays</td>
<td>Maspin, CK19, and m'globin: not specific CK19, CEA, MUC1, EGFR: not specific</td>
<td></td>
</tr>
<tr>
<td>Maspin, m'globin, c-ErbB2</td>
<td>Qi RT-PCR</td>
<td>One 40 cycle reaction for each marker</td>
<td>Maspin and m'globin: 1 MCF-7 cell in 10^5 PBWN Maspin and m'globin: 10 ng MCF-7 total cDNA M'globin: 100 pg MCF-7 total cDNA</td>
<td>Maspin and m'globin: specific c-ErbB2-2: not specific</td>
<td></td>
</tr>
</tbody>
</table>
Table 2 Continued

| Markers Amplification method PCR cycling conditions Sensitivity Specificity in control blood Combined positivity Ref |
|--------------------------------------------------|--|-------------------|---------------------|-----------------|-----------------|
| CK19, CK20, EGFR RT-PCR CK19: one 40 cycle reaction ND CK19 and CK20: not specific EGFR: specific CK19/CK20: 5% positive for both |  |  |  | 82 |
| CK19, CK20, β-hCG RT-PCR CK20: one 42 cycle reaction EGFR: one 33 cycle reaction Not detailed CK19/CK20: 9.72% |  |  |  | 83 |
| MUC1, CK19, PIP, m’globin, m’globin B, PSE Q RT-PCR One 40 cycle reaction for each marker One 40 cycle reaction for each marker 1 MDA-361 cell in 5 x 10⁸ PBL CK19, CK20, β-hCG: specific |  |  |  | 84 |

β-hCG, β chain of human chorionic gonadotrophin; β-NacGAT, β-N-acetylgalactosaminyl transferase; CEA, carcinoembryonic antigen; CK, cytokeratin; c-met, hepatocyte growth factor receptor; EGFR, epidermal growth factor receptor; MAGE A3, melanoma associated antigen; m’globin, mammaglobin; MNC, mononuclear cells; MUC1, membrane associated mucin 1; ND, not determined; PBL, peripheral blood leucocytes; PB MN, peripheral blood mononuclear cells; PIP, prostate inducible protein; PSE, prostate specific ets factor; Q, quantitative; RT-PCR, reverse transcription polymerase chain reaction.

THE PROBLEM OF FALSE POSITIVE RESULTS

The apparent false detection of target mRNA by RT-PCR in blood could also be the direct result of illegitimate expression of the target gene. Recently, two CK19 pseudogenes, CK19a and CK19b, have been identified, which have significant homology to the real CK19 gene. They are expressed in a very limited number of tissue types and only to a very low level, but they are easily distinguished from the real CK19 gene by RT-PCR. However, the CK19 pseudogenes are not expressed in any of the tissues that express the real CK19 gene. Therefore, the detection of CK19 mRNA in blood cells is more likely to be a result of the expression of the CK19 pseudogenes, which are not expressed in any of the tissues that express the real CK19 gene.

Non-specific detection of tumour specific targets

Although the sensitivity of molecular techniques has greatly improved, there is also the potential for false positives. The detection of false positives is often difficult to determine because several factors may contribute to this problem, some of which are discussed in further detail below.

The apparent false detection of target mRNA by RT-PCR in blood could also be the direct result of illegitimate expression of the target gene. Recently, two CK19 pseudogenes, CK19a and CK19b, have been identified, which have significant homology to the real CK19 gene. They are expressed in a very limited number of tissue types and only to a very low level, but they are easily distinguished from the real CK19 gene by RT-PCR. However, the CK19 pseudogenes are not expressed in any of the tissues that express the real CK19 gene. Therefore, the detection of CK19 mRNA in blood cells is more likely to be a result of the expression of the CK19 pseudogenes, which are not expressed in any of the tissues that express the real CK19 gene.

Non-specific detection of tumour specific targets

Although the sensitivity of molecular techniques has greatly improved, there is also the potential for false positives. The detection of false positives is often difficult to determine because several factors may contribute to this problem, some of which are discussed in further detail below.
sequence homology to CK19 mRNA, which has been used in the detection of circulating tumour cells in blood (tables 1 and 2). Subsequently, attempts to detect the expression of the authentic CK19 may result in the detection of either or both of these pseudogenes.

**Solutions to the amplification of pseudogenes and illegitimate gene expression**

**DNase treatment and primer design**

To eliminate the problem of amplification of pseudogenes from contaminating genomic DNA, DNase could be used before RT-PCR. However, RNA extractions have varying degrees of DNA contamination and DNase may not eliminate it all. This may result in the detection of false positives if the primers have not been designed carefully to eliminate all possible amplification of pseudogenes.

Furthermore, the sequence of any amplified product should be determined to confirm the specificity of the assay. Although this is not practical for every individual test, sequencing of the amplified product when devising the assay could initially determine the specificity of the primers to amplify the correct product. Aerts et al used DNase to remove contaminating DNA, and stated that the primers used were designed to prevent amplification of the CK19 pseudogenes. In spite of these precautions, false positives were still detected and it was hypothesised that this was the result of low level expression in non-malignant cells in the blood. However, the false positives that were detected in this study were not subjected to sequencing. Thus, it cannot be definitely concluded that these false positives were the result of expression by non-malignant cells in the blood, particularly because many reports using specific CK19 primers report an absence of ectopic transcription in haematological tissues (tables 1 and 2).

**Measurement of target gene expression**

If the illegitimate transcription of targets did occur in non-tumour cells, the measurement of target gene expression in patients with cancer compared with normal volunteers would help minimise this problem. For example, the measurement of CK19 mRNA in the blood of patients with colorectal and breast cancer showed that values were significantly higher in patients with colorectal cancer above that of the normal volunteers, previously attributed to illegitimate expression, could genuinely indicate the presence of circulating tumour cells within the blood. Thus, increased CK19 mRNA values in patients with colorectal cancer, above that of the normal volunteers, could genuinely indicate the presence of circulating tumour cells within the blood.

In addition, target gene measurement could determine a patient’s circulating tumour cell load during the clinical course and could therefore aid the clinician in the patient’s subsequent treatment. The amount of CK19 mRNA in the blood of patients with breast and colorectal cancer increased as their disease developed. Thus, regular measurement of CK19 mRNA in patients with breast cancer (as an assessment of circulating tumour cell load) may indicate their disease progression or potential for disease relapse.

**Contamination of epithelial cells or nucleic acids**

**Epithelial cell contamination**

It has been suggested that blood sampling for subsequent analysis introduces contaminating epithelial cells into the blood sample. Although the problem of the release of cells into the blood sample collected as a result of invasive diagnostic procedures cannot be completely eliminated, potential contamination could be minimised by discarding the first sample of blood collected. Several studies considered the problem of epithelial cell contamination upon sample taking and took necessary precautions by discarding the first sample of blood taken (tables 1 and 2). Thus, it could be assumed that in all other studies detailed in these tables, the first blood sample was not discarded, was used in subsequent sample testing, and therefore was a source of potential epithelial cell contamination. Nevertheless, in one study, the analysis of the first and second blood samples taken from normal volunteers showed that there was no significant difference in the detection of CK19 positive cells between the two samples. It was concluded that discarding the first sample did not significantly alter the extent of epithelial cell contamination.

**Sample preparation**

Techniques in blood collection and preparation may cause changes in gene expression levels ex vivo. Several investigations have shown that the addition of anticoagulants to blood samples affect leucocyte viability and cause ex vivo changes in cytokine production. These results suggested that the addition of anticoagulants significantly alters the ability to detect accurately the presence of a particular target gene because of gene upregulation or downregulation during the time taken between sample collection and processing. In contrast, the addition of RNA stabilisers to blood samples has been shown to preserve cytokine RNA values for seven days. Thus, it may be better to preserve RNA for subsequent testing by the addition of RNA stabilisers rather than anticoagulants.

**“Techniques in blood collection and preparation may cause changes in gene expression levels ex vivo”**

Several studies (tables 1 and 2) specified the exact length of time taken between sample collection and processing, and a third study stated that samples were processed quickly. However, none of the studies detailed analysed the expression of the mRNA targets over time. Hence, it may be reasonable to assume that variation exists, within and between studies, in the time between sample collection and processing, which may affect the expression of the target gene and the ability to detect it. Thus, the presence of anticoagulants does affect target gene expression and may contribute to the reported differences in the levels of detection of expression of each marker.

**CONCLUSIONS**

In conclusion, the histochemical or IHC techniques routinely used at present to diagnose metastatic regions within the LN and/or BM of patients with breast cancer are relatively unreliable, time consuming, costly, and require specialist staff. In addition, a proportion of patients who develop metastatic disease either do not have positive LN at the time of analysis or they are not detected by the pathological techniques used. Thus, there is a growing need for a more reliable test that is comparatively simple, accurate, quick, and inexpensive, in addition to causing little discomfort to the patient.

The haematogenous system offers a source of circulating tumour cells in patients with breast cancer and sampling of the blood is relatively painless and can be done at frequent intervals to allow an assessment of the patient’s recovery or potential to develop metastatic disease. In addition, the increased sensitivity of molecular biological techniques, notably RT-PCR, when compared with IHC, indicates that the development of a reliable molecular diagnostic test to aid clinicians in the management and treatment of their patients is a promising prospect. However, this review has highlighted several issues that require careful consideration when developing such a test to ensure that it is as accurate and reliable as possible.
Take home messages

- The histochemical and immunohistochemical techniques used routinely to diagnose lymph node and bone marrow metastases in patients with breast cancer are relatively unreliable, time consuming, costly, uncomfortable for the patient, and do not detect all those who go on to develop metastatic disease.

- Molecular biological tests that are comparatively simple, accurate, quick, inexpensive, and less invasive are needed.

- Reverse transcription polymerase chain reaction, using tumour cells circulating in the blood, offers increased sensitivity and would be useful for patient treatment and management.

- A multimarker test that detects cytokeratin 19, magmoglobin (or another suitable breast cancer marker), and CD44 would be useful to determine whether a patient with breast cancer had breast cancer cells with metastatic potential circulating within the blood.

- However, several issues, such as problems of pseudo-gene expression, illegitimate gene expression, epithelial cell contamination, and RNA breakdown ex vivo require careful consideration to ensure that such a test is accurate and reliable as possible.

ACKNOWLEDGEMENTS
This work was supported by a TCS funded programme with additional financial support provided by Micropathology Ltd, UK.

Authors’ affiliations
A M Gilbey, D Burnett, Micropathology Ltd, University of Warwick Science Park, Barclays Venture Centre, Sir William Lyons Road, Coventry CV4 7EZ, UK
R E Coleman, Cancer Research Centre, Weston Park Hospital, Sheffield S10 2SJ, UK
J Holen, University of Sheffield, Sheffield S10 2JF, UK

REFERENCES