Challenges and unanswered questions for the next decade of circulating tumour cell research in lung cancer

Sumitra Mohan, Francesca Chemi, Ged Brady

Clinical and Experimental Pharmacology Group, CRUK Manchester Institute, University of Manchester, Manchester, UK

Contributions: (I) Conception and design: All authors; (II) Administrative support: All authors; (III) Provision of study materials or patients: None; (IV) Collection and assembly of data: None; (V) Data analysis and interpretation: All authors; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Ged Brady, Clinical and Experimental Pharmacology Group, CRUK Manchester Institute, University of Manchester, M20 4BX, Manchester, UK. Email: ged.brady@cruk.manchester.ac.uk.

Abstract: Since blood borne circulating tumour cells (CTCs) initially shed from the primary tumour can seed and initiate metastasis at distant sites a better understanding of the biology of CTCs and their dissemination could provide valuable information that could guide therapeutic intervention and real time monitoring of disease progression. Although CTC enumeration has provided a reliable prognostic readout for a number of cancers, including lung cancer, the precise clinical utility of CTCs remains to be established. The rarity of CTCs together with the vanishingly small amounts of nucleic acids present in a single cell as well as cell to cell heterogeneity has stimulated the development of a wide range of powerful cellular and molecular methodologies applied to CTCs. These technical developments are now enabling researchers to focus on understanding the biology of CTCs and their clinical utility as a predictive and pharmacodynamics markers. This review summarises recent advances in the field of CTC research with focus on technical and biological challenges as well the progress made towards clinical utility of characterisation of CTCs with emphasis on studies in lung cancer.

Keywords: Circulating tumour cells (CTCs); lung cancer; liquid biopsy; biomarkers; metastasis

Submitted Apr 08, 2017. Accepted for publication Jun 01, 2017.
doi: 10.21037/tlcr.2017.06.04
View this article at: http://dx.doi.org/10.21037/tlcr.2017.06.04

Introduction

Circulating tumour cells (CTCs) were identified and first proposed as part of the fundamental process of cancer dissemination in the mid-19th century by the Australian pathologist, Thomas Ashworth (1). Since more than 90% of cancer-related deaths are thought to be due to metastatic disease (2,3), it is not surprising that the presence of CTCs indicates metastatic disease is generally linked with poor prognosis (4,5) and dissemination of tumour cells starts at early stages of disease progression (5,6). However, the mechanism of dissemination and the accompanying biology of CTCs are still poorly understood.

In this review, we will present our view of the technical challenges of working with CTCs as well as some of the key biological and clinical findings including the potential use of CTCs in the clinical management of lung cancer including their use as prognostic, pharmacodynamic and predictive biomarkers. We will reference CTC studies across all cancer types but will highlight studies addressing lung cancer CTCs. Given the broad scope of CTC research and the large numbers of CTC publications we cannot provide an exhaustive coverage of the subject so for those requiring more detail we would recommend recent excellent CTC reviews (7-9).

Technical challenges

The biggest challenge for CTC research is the rarity of CTCs which is estimated at a level of one CTC per $10^6$–$10^7$. 
white blood cells (WBCs). Detecting and isolating rare CTCs from the bloodstream requires sensitive technologies that can efficiently pick out these cells amongst the millions of other blood cells. Several of the key technical challenges that face CTC researchers and potential solutions are addressed in the following section.

**Identification and isolation of CTCs**

There are three broad basic strategies for CTC enrichment, identification and analysis from whole blood: (I) negative selection through removal of blood cells thereby enriching the remaining CTCs; (II) positive selection of CTCs; and (III) analyse all cells accompanied by CTC identification through intensive image analysis. Each of these strategies has its own strengths and weaknesses so the choice of approach should be guided primarily by the clinical/biological aim of the study as well as pragmatic considerations such as cost and time available.

Many negative selection approaches are based on the availability of antibodies recognising blood specific cell surface epitopes and the fact that these epitopes are not widely expressed in cancer cells. The attraction of this approach is that by targeting the unwanted cells, in this case red blood cells (RBCs) and WBCs, the CTCs are left untouched and are less likely to be influenced or damaged by the enrichment approach. The commercial RosetteSep™ kit (StemCell Technologies) provides a simple method to remove the bulk of mature blood cells including both RBCs and WBCs, and has been shown to leave behind viable CTCs that can be cultured (10,11) or implanted into mice and grown as tumours (12-14). Recently, a new enhanced negative depletion strategy named Multi-marker Immunomagnetic Negative Depletion Enrichment of CTCs (MINDEC) has been developed where blood cell fraction is isolated by density gradient centrifugation followed by the use of superparamagnetic beads coupled with a cocktail of antibodies against different blood classes (15). Although negative selection approaches have proven useful, due to the overwhelming numbers of WBCs compared to CTCs, the output from negative selection is often heavily contaminated with WBCs which may interfere with downstream CTC based assays.

For positive selection of CTCs there are two major approaches based either on physical properties such as size, density and deformability or biological properties such as cell surface epitopes. The most successful and widely used CTC technology is the FDA recognised CellSearch® system (Jannsen Diagnostics, Raritan, NJ, USA) which is a semi-automated platform that employs ferromagnetic beads coated with the epithelial cell surface marker, EpCAM to “pull out” or enrich epithelial CTCs (4). Following the initial CTC enrichment CellSearch® utilises automated fluorescence staining to identify and enumerate CTCs based on their cell/nuclear morphology along with detectable expression of epithelial cytokeratins and low expression of the WBC marker CD45 (4). Since CellSearch® has been shown to be both reliable and reproducible it has been widely used in multi-center trials and CellSearch® CTC counts have been shown to be prognostic for metastatic breast (4), lung (16), colorectal (CRC) (17) and prostate cancer (18). A key feature of the CellSearch® platform which led to its widespread uptake has been the use of the CellSave® preservative blood collection tube (Jannsen Diagnostics, Raritan, NJ, USA) which allows whole blood to be stored or shipped at room temperature for up to 96 h prior to analysis (4,17). Other systems have been described which also employ an EpCAM-based enrichment technologies including the CTC-Chip (19), MagSweeper (Stanford University, USA) (20), iCHIP (Massachusetts General Hospital Center, USA) (21) and GILUPI CellCollector® (GILUPI Nanomedizin, Germany) (22) (Table 1) Since EpCAM-based enrichment technologies are limited to epithelial CTC detection alternative strategies are required for detection of CTCs which are not epithelial. For example, CellSearch® has developed a melanoma tailored enrichment and staining kit, where melanoma CTCs are captured via the CD146 cell surface marker (also known as melanoma cell adhesion molecule or MCAM) and positively stained for the high molecular weight melanoma associated antigen (HMW-MAA) while CD45 and CD34 are used as WBC markers (23).

Although positive selection utilising cell surface markers has been very successful it is also clear that there are CTCs which change or lack these markers (see further discussion regarding EMT below) and this has prompted the development of marker-independent CTC enrichment platforms. Marker-independent CTC enrichment platforms are generally devices which retain CTCs on the basis of physical properties such as size (ISET®, Rarecells Diagnostics, France) (24), cell size and deformability (Parsortix™ Cell Separation, ANGLE plc, UK) (25,26) (Table 1) and flow properties (ClearCell®, Clearbridge Biomedics, Singapore) (27) (Table 1). Captured CTCs can be then recovered for further molecular analysis, however CTCs may not be always larger and stiffer than WBCs, thus

---

**Table 1**

<table>
<thead>
<tr>
<th>Platform Name</th>
<th>Company/Location</th>
<th>Year of Introduction</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CellSearch®</td>
<td>Jannsen Diagnostics, Raritan, NJ, USA</td>
<td>2004</td>
<td>Semi-automated system that employs ferromagnetic beads coated with EpCAM.</td>
</tr>
<tr>
<td>CellSearch®</td>
<td>Jannsen Diagnostics, Raritan, NJ, USA</td>
<td>2004</td>
<td>Positive selection of CTCs based on expression of EpCAM.</td>
</tr>
<tr>
<td>CellSave®</td>
<td>Jannsen Diagnostics, Raritan, NJ, USA</td>
<td>2004</td>
<td>Preservative blood collection tube.</td>
</tr>
<tr>
<td>iCHIP</td>
<td>Massachusetts General Hospital Center, USA</td>
<td>2011</td>
<td>Enrichment and staining kit for melanoma cells.</td>
</tr>
<tr>
<td>CellCollector®</td>
<td>GILUPI Nanomedizin, Germany</td>
<td>2016</td>
<td>Negative selection strategy for both epithelial and non-epithelial CTCs.</td>
</tr>
<tr>
<td>CTC-Chip</td>
<td>Stanford University, USA</td>
<td>2018</td>
<td>Magnetic beads, ISET technology.</td>
</tr>
<tr>
<td>MagSweeper</td>
<td>Stanford University, USA</td>
<td>2019</td>
<td>Magnetic beads, ISET technology.</td>
</tr>
<tr>
<td>ClearCell®</td>
<td>Clearbridge Biomedics, Singapore</td>
<td>2019</td>
<td>Flow separation technology for CTCs.</td>
</tr>
</tbody>
</table>

© Translational lung cancer research. All rights reserved. tlcr.amergroups.com Transl Lung Cancer Res 2017;6(4):454-472
Table 1  Technologies for enrichment and isolation of CTCs from a blood sample

<table>
<thead>
<tr>
<th>CTC isolation technologies</th>
<th>Approach</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Antibody-dependent enrichment</td>
<td>CellSearch®</td>
<td>EpCAM-coated ferromagnetic beads</td>
<td>FDA clearance and clinical relevance in many cancer types; semi-automated; visual confirmation of CTCs; highly reproducible and automated</td>
</tr>
<tr>
<td></td>
<td>MagSweeper</td>
<td>Flow through EpCAM based immunomagnetic capture</td>
<td>Automated; high purity; can process whole blood, possibility of live cell capture</td>
</tr>
<tr>
<td></td>
<td>GILPUI CellCollector®</td>
<td>EpCAM-coated Seldinger guide wire into the cubital vein</td>
<td>Allows for analysis of all peripheral blood in an individual</td>
</tr>
<tr>
<td></td>
<td>IsoFlux</td>
<td>EpCAM-coated magnetic beads combined with microfluidic processing</td>
<td>Automated; Enriched cells can be recovered in low volume (3 µL) and hence suitable for downstream molecular analysis</td>
</tr>
<tr>
<td></td>
<td>Herringbone chip</td>
<td>Microfluidic device consisting of EpCAM-coated microposts</td>
<td>Visual confirmation of CTCs; potential for further staining using immunofluorescence-conjugated antibodies; high detection rate; potential for detection of CTMs</td>
</tr>
<tr>
<td></td>
<td>HD-SCA</td>
<td>Peripheral blood mononuclear cells (PBMCs) are applied onto a slide and subsequently stained using markers of choice</td>
<td>Potential for tailoring a selection of antibodies; visual confirmation of CTCs with automated digital microscopy; potential for downstream molecular analysis</td>
</tr>
<tr>
<td></td>
<td>RareCyte</td>
<td>PBMCs are applied onto a slide and subsequently stained using markers of choice</td>
<td>Potential for tailoring a selection of antibodies; visual confirmation of CTCs; potential for downstream molecular analysis; plasma fraction of the blood sample is removed and replaced with buffer, hence potential for ctDNA analysis</td>
</tr>
<tr>
<td>Physical properties based enrichment</td>
<td>Parsortix™</td>
<td>Cell size and deformability based enrichment</td>
<td>Label free approach; avoids staining protocols; broad population of CTCs enriched; plasma fraction of the blood sample is removed and replaced with buffer, hence potential for ctDNA analysis</td>
</tr>
<tr>
<td></td>
<td>ISET®</td>
<td>Filtration based on cell size based</td>
<td>Label free approach; avoids staining protocols; broad population of CTCs enriched</td>
</tr>
<tr>
<td></td>
<td>ClearCell®</td>
<td>Microfluidic approach based on cell size, deformability and inertia</td>
<td>Label free approach; avoids staining protocols; broad population of CTCs enriched</td>
</tr>
<tr>
<td></td>
<td>CTC iChip</td>
<td>Size-based separation followed by negative and positive selection using magnetic beads</td>
<td>Enrichment of broad population of CTCs by positive selection by EpCAM or negative selection by CD45;</td>
</tr>
</tbody>
</table>

In this table we summarise technologies used for isolation of CTCs using antibody-dependent and physical property-dependent enrichment technologies. CTCs, circulating tumour cells; ctDNA, circulating tumour DNA.
can be washed away through the filters (28).

More recently, imaging and scanning technologies have been developed to a point where CTCs can be visualised, identified and isolated without any prior enrichment (8). A number of integrated systems have been described for the interrogation of whole blood including high definition single cell analysis (HD-SCA) (29) and the RareCyte system (RareCyte, Seattle, WA, USA) (30) (Table 1). Both are based on the principle of “no cell left behind” where all nucleated blood cells are deposited on microscope slides and stained with markers allowing discrimination of blood and non-blood cells thereby increasing the possibility to detect rare CTCs subpopulations (8). An additional feature for both the HD-SCA and RareCyte® platforms is that, once candidate CTCs have been identified, they have an integrated capacity to isolate target cells directly from the slide in a format that is compatible with single cell molecular analysis. In contrast, most CTC enrichment technologies require an additional step for CTC visualisation and isolation. This has been achieved in a variety of ways including fluorescence-activated cell sorting (FACS) (31), CTC visualisation followed by isolation via micromanipulation (21,32-34), and the DEPArray™ automated system which incorporates cell manipulation via electrostatic charge (35,36) (Table 1).

CTC molecular analysis

The striking increase in the sensitivity of DNA and RNA technologies have been pivotal for CTC research since they have enabled researchers to confirm the tumour identity of cells identified by morphology or tissue specific surface markers by establishing that they harbour tumour specific molecular signatures. Early observations generally using targeted sequencing or PCR demonstrated the presence of tumour specific mRNAs or DNA changes in either whole blood or samples enriched for CTCs. Statopoulos et al. developed an assay for detection of cytokeratin-19 (CK-19) mRNA in peripheral blood of patients with breast cancer (37-39). Similarly, in CRC a panel of genes including carcinoembryonic antigen (CEA), CK-19, CK-20 and CD133 have been used to verify the presence of colorectal tumour cells in enriched blood samples (40). However, any study carried out on enriched CTC populations will inevitably suffer from reduced sensitivity through the effective dilution caused by varying levels of contaminating WBCs.

Over the last decade, advances in both CTC enrichment protocols and molecular methodologies has culminated in a range of successful strategies for genome wide amplification and analysis of single CTCs which have extended many of the observations made with enriched CTCs. There are several advantages of single CTC analysis including: (I) single CTC analysis is 100% pure (assuming a CTC was isolated) and avoids any complications or dilution of signal due to contaminating cells; (II) since a cell has on average only two DNA copies of each gene and around 100,000 mRNA molecules (41) the degree of next generation sequencing (NGS) required for generating a representative single cell snapshot is relatively low; (III) single cell genomic analysis can unambiguously identify co-existing genetic changes; (IV) the integration of RNA and DNA analysis from the same cell can establish linkage between the genetic alterations and alterations in RNA pathways (42); (V) single cell data provides an ideal means of establishing the degree of heterogeneity and estimating tumour evolution (43,44). For single CTC molecular analysis the main problems encountered are largely technical, financial and logistical. Since the amplification process necessary for single CTC analysis is extremely sensitive it is critical that steps are taken to avoid or identify potential contamination such as defining separate clean areas and use of negative controls.

Given the vanishingly small amounts of DNA (6.6 pg) present in a single cell, whole genome amplification (WGA) is a pre-requisite for extensive genomic analysis of CTCs. The widespread use of CTC WGA was pioneered by Christoph Klein and co-workers through the establishment of a simple single tube amplification system (available as the AMPLI1™ WGA kit, Silicon Biosystems) based on restriction enzyme digestion, ligation of primer and subsequent PCR amplification (45). There are now many commercially available single cell WGA systems to choose from generally based on PCR (46,47) or linear amplification (48-50). There are obvious pros and cons for each WGA system (51) and the choice of which WGA system to use needs to be based on the clinical/scientific aim of the study, whether the WGA system is compatible with the CTC enrichment process used and the resources available.

Following WGA of CTCs one of the most widely used approach for CTC molecular analysis has been establishing genome copy number alterations (CNA) which provides a simple snapshot of each CTC which will allow comparisons between CTCs and, when available, the matching tumour [Figure 1 provides examples of non-small cell lung cancer (NSCLC) CTC CNA profiles]. Initial single cell studies compared targeted mutation profiles and copy number profiles of CTCs with the primary tumours and metastases of the corresponding patients. One such study of circulating
Figure 1 Determining CTC copy number alterations (CNA). (A) Shows a schematic outline of the CTC NGS analysis workflow used. (B) Presents genome wide CNA profiles and accompanying images of single CTCs and WBCs isolated by epitope dependent and independent technologies from two patients with NSCLC. For all four cells the CNA pattern is presented in a linear format with marked copy number increases shown in red and losses in blue alongside the fluorescent stained images of the starting cells. The top panel presents the CNA patterns and images of a NSCLC cytokeratin (CK) positive, CD45 negative CTC and a CK negative, CD45 positive WBC obtained following CellSearch® EpCAM enrichment and subsequent CK and CD45 staining with CK staining depicted as yellow, CD45 as green and nuclear staining (DAPI) as blue. The bottom panel presents the CNA patterns and images of a NSCLC vimentin (VIM) positive, CD45/CD31 negative CTC and a VIM positive, CD45/CD31 positive WBC obtained following Parsortix™ size enrichment and subsequent VIM plus combined CD45/CD31 staining with VIM staining depicted as red, CD45/CD31 as green and nuclear staining (DAPI) as blue. CTC, circulating tumour cell; NGS, next generation sequencing; NSCLC, non-small cell lung cancer.
melanoma cells using single cell comparative genomic hybridisation (single-cell CGH) revealed multiple CNA consistent with melanoma which was in sharp contrast to the normal cells that revealed balanced CGH profiles (52). In addition, the use of hierarchical clustering provided a means of establishing linkages between each cell and suggested a clonal origin of the melanoma cells examined (52).

For de novo detection of variants in single cells there is a profound challenge of how to distinguish the artefactual “variants” that accrue during the million fold or more amplification during WGA from the actual mutations present in the starting cell. Although it is difficult to be completely certain of true biological origin of variants called in a single cell or CTC, the use of suitable controls, technical replicates, biological replicates and bioinformatics can be used to establish the extent of amplification errors (53). It may be possible to reduce the impact of amplification errors by incorporating molecular bar codes (54) or using low error amplification processes (53,55) but it is not yet clear how much benefit these strategies will bring. For the moment the most effective means of identifying true variants in CTCs is the use of biological replicates, i.e., identify variants seen in two or more independent CTCs (56).

In addition to the advances in genomic analysis of CTCs there have been similar exciting advances in CTC RNA profiling and particularly in the plethora of single cell RNA-Seq methods that have emerged in the last 5 years (57). The attraction of genome-wide RNA profiling of CTCs is that it can give a broad picture of the biological activity of the cell including pathway activation and the presence of expressed drug targets. Transcriptomic profiles of CTCs can be viewed as an emergent property of all of the genetic and epigenetic changes within the cell and as such may be easier to characterise than the often complex genomic changes seen in CTCs. The ability to carry out effective single CTC RNA profiling is largely driven by CTC enrichment protocols and the impact they have on the RNA within the final enriched cell. Since RNA is chemically more labile than DNA and subject to enzymatic degradation it is clear that CTC enrichment and identification strategies that work well for DNA isolation and genomic analysis may not be suitable for RNA analysis. For example, the CellSearch® platform that utilises the CellSave® preservative blood tube works well for DNA but it is clear that both the preservative and the downstream cell processing steps have a negative impact on RNA quality. For this reason most genome wide CTC RNA profiling methods have avoided the use of preservative and have isolated viable cells suitable for direct lysis and processing. A number of groups have now described a broad range of effective genome wide and targeted RNA approaches applied to either enriched pools or single CTCs (36,58,59). Transcriptomic profiling of CTCs is now providing valuable insights into CTC biology (see discussion below) and further advances in methodologies such as combined RNA/DNA SCA (42) and droplet based analysis of many 1000s of cells (60) will provide an even greater depth of understanding.

**Blood sampling considerations**

Since CTCs are generally very rare it is also important to consider the volume of blood sampled and the implications this has for effective CTC identification and isolation (61). Fischer et al. in 2008 demonstrated the possibility of combining cell detection approach such as CellSearch® along with diagnostic leukapheresis (DLA) to increase the possibility of detecting CTCs (62). In this study, they increased the CTC detection rate in patient samples from a range of cancer types by 44% by using DLA followed by CellSearch® as compared to using 7.5 mL of blood on CellSearch®. DLA is a clinically safe approach that enables detection of rare CTCs from larger volume of blood. The other approach that can increase effective blood sampling volumes is the GILUPI CellCollector® which introduces a EpCAM-coated Seldinger guide wire into the cubital vein allowing CTC enumeration and isolation from the entire circulating blood in an individual patient (22).

**Understanding CTC biology**

Having established an impressive tool kit for isolating and analysing CTCs recent research is now beginning to unravel the complex biology of CTCs. In addition to examining the relationship between CTCs, primary tumour and metastasis it is important to establish that CTCs are themselves tumourigenic and if properties of CTCs can shed light on the process of metastatic spread. Two key hallmarks of the malignant tumour cells responsible for the initiation of the metastatic process are motility and invasiveness (63) and it is likely that there will be features in CTCs that reflect these characteristics. The dissemination of cells from the primary tumour and resultant metastatic seeding may well start early even at stages where primary tumour is thought to be localised with no overt metastases (5,64–66). Tumour cells that enter the circulatory system to become CTCs do so through a process known as intravasation and the rate

© Translational lung cancer research. All rights reserved. tlcramgroups.com Transl Lung Cancer Res 2017;6(4):454-472
of intravasation from solid tumours in patients is largely unknown although several clinical and animal model studies suggest that millions of tumour cells are shed even from small tumours (67). Although most cells that intravasate die (68,69), a small proportion of cells exit the vasculature to enter organs, a process known as extravasation and colonise distant organs. CTCs as well as their pre-cursors and descendants must overcome many obstacles such as immune surveillance and physical/cellular barriers in order to survive in the circulatory system and to colonise the host organ (70,71). The biological processes associated with propagation, survival and subsequent colonisation of organs distant to the primary tumour are discussed in the following section.

Tumour and CTC heterogeneity

An early model of clonal evolution linked to the generation of genetic diversity in tumours was proposed in 1976 (72) and the emergence of intra-tumour heterogeneity was confirmed experimentally in murine models (73). More recently, NGS technologies have been used to understand the breadth and depth of this overwhelming genetic diversity in primary and the metastatic tumours (74-79).

NGS and bioinformatic strategies have been developed and applied to tumour tissue, which have allowed both a measure of heterogeneity and the generation of phylogenetic models which provide a picture of the inferred development and branched evolution of the tumour (78,80-82). Advances in NGS technologies now allow routine analysis of single cells or nuclei and this has been fruitfully applied to dissociated tumour cells to provide unambiguous detection of tumour sub-clones (81,83-86). The picture that has emerged from both whole tissue and single cell tumour analysis is one of considerable intra and inter patient tumour heterogeneity that is most likely contributing to metastatic spread and the emergence of therapy resistance. Studies in animal models have provided evidence that this cellular and molecular heterogeneity is matched by functional heterogeneity whereby only a subset of cells, often known as tumour initiating cells or tumour stem cells, are capable of initiating tumours in solid tumours such as colon (87,88), pancreas (89), melanoma (90), liver (91) and brain (92).

Disseminated tumour cells (DTCs) provide another fascinating view of cellular heterogeneity since in some malignancies, notably in breast cancer, DTCs are often detectable in apparently healthy organs such as bone marrow and can reside there for many years after surgical removal of the primary tumour until signs of overt malignant growth (93,94). Bone marrow can function as a reservoir for DTCs (71,95) not only for breast cancer but also for other malignancies such as lung cancer and CRC which do not usually metastasise in the bone (65,96). It has been proposed that in the vast majority of cancer patients, DTCs can enter state of proliferative quiescence, process known as cellular dormancy which can represent the main reason why DTCs evade systemic therapy persisting as minimal residual disease (71,97,98). Genetic and transcriptomic characterisation of DTCs could help understand the nature, origin and biology of this cancer cell type in the individual patients and could subsequently provide extremely valuable information of the future course of the patients’ disease. In breast and prostate cancer, DTCs were found with fewer copy number changes compared to the primary tumour, suggesting that the dissemination may happen at an early tumour stage (99). In contrast, Demeulemeester et al. demonstrated by single cell sequencing of DTCs isolated from the bone marrow of breast cancer patients that they originate from either the main primary tumour clone, primary tumour subclones or subclones in an axillary lymph node metastasis, indicating that breast cancer dissemination may arise relatively late in tumour evolution (64).

Although several groups have also reported the presence of DTCs in the bone marrow of patients with localised NSCLC, and shown a correlation with worse clinical outcome (65,100,101) these studies included relatively small numbers of patients (particularly in comparison to breast cancer) making it unclear to what extent NSCLC is driven by DTCs. DTCs were also detected in the bone marrow of 18% of patients with SCLC using a RT-PCR approach to detect levels of gastrin-releasing peptide, which is preferentially expressed in SCLC (102). However, given the rapid aggressive nature of SCLC and the high numbers of CTCs observed in this disease it is not clear if there is a role for a long term DTCs located in the bone marrow.

Although it is likely that the first CTCs arise from the primary tumour and play a role in metastasis, once metastasis are established they can also be a source of CTCs. In one SCLC study CNA and mutational profiles were established from CTCs, primary tumor as well metastasis and the comparison indicated that there was more overlap between CTCs and metastasis than between CTCs and the primary tumor (103). Given the aggressive nature of SCLC it is not surprising that metastasis can also contribute to CTC based spread and further studies are required to establish if CTC contribution from developing
metastasis is a general feature of all cancers. A CRC study revealed similar results with CNA and mutational profiles of CTCs presenting with more overlap with liver metastases than the primary tumour (43).

**CTC-derived tumour xenografts (CDX)**

One of the major challenges for CTC research is to establish if there is any functional capacity of CTCs and trace this back to molecular changes that are shared by CTCs and tumour. Recently it has been shown that CTCs isolated from patient blood can form tumours when introduced into immune compromised mice firmly establishing that CTCs are viable and capable of forming tumours (12-14). In a study investigating breast cancer, CTCs isolated from 110 patients were transplanted into the femoral medullar cavity of immunocompromised mice and three patient-derived CTCs were able to form multiple metastases (104). In a SCLC CDX study it was also established that CTCs from both patients who responded to chemotherapy (chemosensitive) and CTCs from patients who did not (chemorefractory) were capable of forming tumours in immunocompromised mice and that when mice harbouring CDXs were treated with platinum and etoposide the response seen mirrored the response observed in the corresponding patient (12). In the same SCLC study, genomic analyses revealed that CellSearch® CTCs and the matching CTC-derived CDX are highly related in terms of both copy number and TP53 and RBL mutations (12). CTCs were also shown to be tumorigenic in melanoma, where CDXs were established with a success rate of 13% and they had similar metastatic tropism as the corresponding patient (14).

**Epithelial-to-mesenchymal transition (EMT)**

The key feature of EMT relevant to CTC biology is the switch from relatively restricted or anchored epithelial cell to a more motile flexible mesenchymal cell type (105). EMT was first coined by embryologists and describes cell migration from the ectoderm during the gastrulation process to produce the mesoderm and is now also referred to as type I EMT (106,107). A similar process was also observed in wound healing and fibrosis and this is referred to as type II EMT (108). For CTCs it has been proposed that EMT facilitates intravasation from primary epithelial tumour through the switch to a mesenchymal phenotype associated with increased cell motility and migration (105,109). However, histopathological examination of metastatic tumours reveals a predominantly epithelial tumour and relatively fewer mesenchymal characteristics (110) indicating that, for CTCs, there is a reciprocal mesenchymal-to-epithelial transition (MET) associated with extravasation and reestablishment of an epithelial metastasis (5,109,111,112).

During EMT there is a loss or redistribution of the epithelial adhesion molecule, E-cadherin, a trans-membrane glycoprotein that is responsible for the tight junction between epithelial cells and integral to apical-basal polarity (111). Along with the modulation of cell-cell contact markers, EMT is also associated with remodelling of the cytoskeleton with an increase in vimentin-rich intermediate filaments providing a more flexible mesenchymal cellular structure as opposed to the more rigid cytokeratin-rich starting point (113). The process of EMT can be triggered by paracrine signalling of several factors such as TGF-beta, WNT or interleukin-6 (105,114,115) and these triggers in turn activate transcription factors such as Snail, Twist and Zeb that maintain the mesenchymal phenotype by autocrine signalling (115). These EMT associated alterations have also been observed in CTC sub-populations in breast cancer by single cell RNA-seq analysis (116). Determining the biological relevance of EMT for CTC mediated metastatic spread is confounded by the fact that EMT and MET are processes that occur over time (it is not clear how long either process will be) and that metastatic spread may occur either continually or sporadically. In this context, several studies have reported identification of CTCs with mesenchymal markers in patients with advanced stage of breast cancer as compared to patients with localised breast cancer (117). To get a clearer picture of the relevance of CTC biology including EMT it is possible to take advantage of the simplicity of blood sampling and establish multiple longitudinal blood samples to build a picture of CTC biology during the course of disease progression. Yu et al. applied this strategy to monitor breast cancer CTCs and identified emergence of a mesenchymal CTC phenotype in samples from patients undergoing chemotherapy when compared to samples taken at baseline (116). This study utilised RNA in situ hybridization to detect epithelial and mesenchymal genes and highlighted the role of TGF-beta activation along with the transcription factor FOXC1 in EMT in breast cancer CTCs. EMT and MET switching in CTCs can also pose a challenge for detecting the resultant mesenchymal CTCs since they are in a hematopoietic cell environment consisting of many mesenchymal cells this
will be expanded on in the section summarising technical challenges.

One of the earliest reports of EMT in CTCs from patients with NSCLC compared CTC identification and enumeration using epithelial epitopes to enrich and stain CTCs (CellSearch®) to an epitope independent CTC identification system based on size selection (ISET®) and showed both increased CTC numbers as well as heterogeneous expression of EMT markers detected using the ISET® approach (112). Subsequently, other studies in SCLC (112), NSCLC (118), head and neck (119), breast (117,120) and prostate (120) cancer supported the observation of heterogeneous expression of mesenchymal and epithelial markers in CTCs. Although, in patients with NSCLC the CellSearch®-ISET® comparison revealed more NSCLC CTCs detected by ISET® than CellSearch® both systems provided CTC numerical readouts that were prognostic suggesting both EpCAM positive and negative CTCs may play a role in tumour spread (121).

A recent examination of blood samples from a patient with NSCLC revealed the presence of both cytokeratin positive CTCs expressing or not expressing the epithelial marker EpCAM suggesting that epithelial markers alone may not be sufficient for complete CTC enumeration (122). Interestingly, the authors went on to show only the EpCAM positive CTCs were associated with clinical outcome suggesting that, for this patient cohort, the EpCAM negative CTCs may not play a role in propagating the disease (122). Also, a study in a mouse model of SCLC identified a population of long term tumour propagating cells and this population presented with marked high expression of EpCAM and CD24 which is also prevalent in human primary SCLC tumours (123). Moreover, in the 47 attempts to generate CDX models (12), the 15 human blood samples that generated a mouse model also presented with more than 50 EpCAM positive CTCs in 7.5 mL suggesting that EpCAM positive CTC subpopulation that has tumour-initiating capacity (123).

There is an unmet requirement of comprehensive evaluation of markers associated with mesenchymal and epithelial characteristics that will allow assess the degree of EMT in different solid tumour types. This type of study has the potential to explain the low yield of EpCAM positive NSCLC CTCs on CellSearch® (5) and further highlights the need to combine both CTC enrichment technologies and analytical methods to more clearly understand the underlying CTC biology.

**Vascular mimicry (VM)**

In addition to CTCs adopting a mesenchymal phenotype which may facilitate intravasation it is also been recently reported that CTCs may also take on a vascular phenotype through a process known as VM (124). The term VM was first coined by Mary Hendrix and described the formation of tumour-derived blood vessels and it was hypothesised that this may provide an easy access for tumour cells to the bloodstream (125). VM was first described in human melanoma where it was shown that tumour cells shown to co-express endothelial and tumour markers and formed blood vessels (125) and further studies have given insights in VM induction across a variety of cancer types including breast, ovary, lung, prostate, bladder cancer (126). VM has been associated with tumour dissemination and metastasis (127) and in SCLC it has been shown that a subpopulation of CTCs co-express VE-cadherin (a VM marker) as well as epithelial cytokeratins indicating VM is involved in CTCs dissemination (124).

**Circulating tumour microemboli (CTM)**

Once CTCs arrive in the blood stream the next problem they face is to survive in the circulation. Reports of tumour cell clusters or CTM in the blood stream of patients with lung, colon, kidney and prostate cancer (16,128-130) have indicated that inclusion within a CTM may provide protection for CTCs particularly since the presence of CTMs is linked with poor prognosis (16,131). Circulating tumour emboli have been defined as groups of tumour cells either alone or associated with fibroblasts (132), leukocytes (133,134), endothelial cells (135-137) or platelets (138-140). There is growing evidence suggesting the co-operative behaviour exhibited by cells within CTMs can promote mutual survival and increased tumorigenic potential (16,33,141-144). It has also been suggested that interactions between blood cell and platelets can benefit both CTMs and single CTCs by protecting them from the host immune system (134,145-147).

Cellular cooperation has been elegantly demonstrated using a model derived from hamster oral keratinocytes of both mesenchymal and epithelial cells (148). At the subcutaneous site, injection of either pure population of cells formed tumours. However, only the mesenchymal cells entered the blood stream but neither cell type alone was capable of forming lung metastases. If pure populations of cells were injected intravenously, only the epithelial
cell population formed lung metastases. When mixed population of cells were subcutaneously injected, although both cell types entered the blood stream the epithelial cells formed the lung metastases. This study shed light on the cooperation between cell types in transit which could be crucial in the metastatic cascade. In SCLC, a mouse model was used to examine the relevance and potential co-operation between neuroendocrine and mesenchymal tumour cells (144). The introduction of either cell type individually into immunocompromised mice resulted in the formation of local tumours without any liver metastasis. However, both cell types administered together resulted in formation of liver metastases along with the local tumour (144) and preformed CTMs injected into mice have a higher propensity to initiate metastases when compared to single cells suggesting a survival advantage of the CTMs (16,133,149).

Given the growing body of evidence it seems likely that CTMs play a pivotal role in tumour dissemination. However a number of questions remain unanswered regarding CTMs such as what is the role of cell heterogeneity (including blood cells) within CTMs and how do CTMs traverse narrow capillaries? Intriguingly, recent modelling using microscale devices, computational simulations and animal studies provides evidence that clusters can transit through capillaries by unfolding into single-file conga lines providing a potential explanation for how CTMs can travel throughout a patient’s body and seed distant sites (150).

CTCs and clinical utility
Along with helping us to understand the biology of primary and metastatic cancer, an examination of CTCs may also provide clinical benefit in the areas of: early detection; prognostic/predictive biomarkers; drug target identification; monitoring disease response/progression and identifying mechanisms of drug resistance. In this section, we address the potential clinical utility of CTCs.

Early detection
Early detection of disease is crucial for effective treatment of cancer. For example, SCLC is generally diagnosed when the disease is well advanced and treatment does not provide long term benefit. In general, the numbers of CTCs present in a 7.5 mL blood sample correlates with clinical staging; with the highest numbers seen in patients with metastases and lower numbers in patients without overt metastases (151,152). However, a recent study has demonstrated the predictive value of CTCs in early NSCLC where patients with chronic obstructive pulmonary disease (COPD) harbour detectable CTCs and then developed lung nodules 1–4 years later with four patients diagnosed with invasive adenocarcinoma and a fifth diagnosed with squamous cell carcinoma (153). However, in the same study there was a clear false positive rate since three patients who harboured detectable CTCs did not develop overt cancer (153). Clearly the potential for using CTCs in early detection needs to be further explored particularly using broader CTC detection systems in large screening programs.

Prognostic biomarkers
Prognostic biomarkers are indicators of the patient’s overall clinical outcome including progression free survival (PFS) and overall survival (OS) (154). With a number of emerging therapeutic options for the treatment of oligometastatic disease including local radiotherapy and immunotherapy that could potentially have an impact on the PFS in patients, the potential of CTCs as prognostic markers has been explored by a multitude of groups. The FDA have recognised the prognostic utility of the CellSearch® platform in many solid cancers such as breast (4), prostate (18) and CRC (17). The presence of more than 5 CTCs in breast (4) and prostate (18) and 3 CTCs in CRC (17) per 7.5 mL of blood correlates positively with shorter OS. This prognostic utility of CellSearch® CTCs has also been demonstrated for NSCLC where the CTC count cut-off is 5 CTCs per 7.5 mL of peripheral blood (5). However, SCLC is one of the solid tumours where the CTCs are more abundant than any other solid tumour thus far reported with a range of 0–44,896 cells per 7.5 mL blood and hence the corresponding CellSearch® CTC count cut-off for SCLC is set at a higher value of 50 CTCs per 7.5 mL blood (16). A NSCLC study using the ISET® CTC technology (CTCs enriched based on cell size) examined bloods from 208 patients with stages I–IV identified CTCs in 50% of patients (155). Although, this study found no correlation between the numbers of CTCs and the staging of the disease, it concluded that a threshold of >50 CTCs corresponded to shorter PFS and OS (155). However, a similar ISET® CTC study of 101 stage III/IV NSCLC patients showed a threshold of 5 CTCs in 7.5 mL blood corresponded with poor prognosis (5) indicating potential differences in the patients recruited to each study and also
highlighting the need for standardisation of CTC calling.

The prognostic relevance of CTC clusters or CTMs has also been discussed in many studies. For example, CTMs were detected by ISET® technology in 43% of patients with stage IIIB/IV NSCLC (16) and in an independent study using HD-CTC assays CTMs were seen in 50% of patients with stage I–IV NSCLC (156) Both studies highlighted the prognostic value of CTMs. Furthermore, analysis of pulmonary vein blood in patients with NSCLC also revealed the existence of CTCs and CTMs in 33% of patients which predicted tumour recurrence and poorer disease-free survival (33,66).

**Predictive biomarkers**

Predictive biomarkers provide information on the likelihood of response to therapy and, in the clinic, can be used to support the selection of suitable personalised therapies (154). For example, over 64% of lung adenocarcinomas have an underlying driver mutation (157) and around 30% of these mutations are linked to an approved targeted therapy with the most common genetic alterations being EGFR mutations and translocations of ALK or ROS1 (158).

Another potential for CTCs in personalised therapy is the longitudinal monitoring of therapy linked mutations as exemplified in a NSCLC study where an examination of 12 patients with NSCLC with EGFR mutated tumour revealed the same detectable EGFR mutations in CTCs in 11 patients (159). Similarly, a study by Marchetti et al. detected EGFR mutation in CTCs of 84% (31 out of 37 patients) EGFR mutant primary tumour (160). Several similar studies were able to reproduce similar results in detecting mutations in CTCs that matched the primary tumour and in cases where the mutations were not detected was likely due to the tumour heterogeneity between primary tumour and the metastatic tumours.

A number of studies have examined the feasibility of identifying rearrangements or translocations of ALK and ROS1 genes in CTCs as a means of guiding therapy (161,162). In one study fluorescence in-situ hybridisation (FISH) has been used to generate the filter-adapted FISH (FA-FISH) technique which was able to identify unique ALK rearrangement patterns in mesenchymal CTCs (105). It was suggested that the mesenchymal CTCs detected by FA-FISH could represent a highly invasive subpopulation with the potential to generate metastatic lesions (105). A similar FA-FISH approach was also used to detect both ROS1 translocations and CNA in NSCLC CTCs which correlated with poor response to crizotinib treatment (163).

In contrast to NSCLC, molecular analysis of SCLC has shown that the overall genomic landscape is chaotic with many mutations and widespread CNA but, despite the near ubiquitous inactivation of RB1 and TP53 frequent “actionable druggable” mutations have not been found (164,165). However, in a recent SCLC study a biomarker was developed based on CNA patterns detectable in single and pooled CTCs from 31 pre-treatment patients and was able to classify patients as either chemosensitive or chemorefractory to standard cisplatin etoposide with an accuracy of 83.3% (44). Interestingly, in the same study five patients who responded to therapy and relapsed with chemorefractory disease showed unaltered CTC CNA patterns at relapse indicating that mechanism of acquired chemoresistance differs from de novo chemoresistance (44).

**Treatment response and mechanisms of resistance**

Besides use of CTCs as prognostic or predictive markers, they have also been used as a pharmacodynamic marker where a change in the numbers of CTCs or a particular subset of CTCs following therapy indicates response to therapy. For example, in SCLC the total numbers of CTCs were significantly decreased following a single dose of chemotherapy (166,167). Similar results were observed in patients with NSCLC (5,168-170). In contrast, an increase in CTC numbers is indicative of failure to respond to therapy. Changes in CTC numbers were also examined in patients with localised NSCLC undergoing radiation treatment the change in CTC numbers post treatment was evaluated as a surrogate for disease response (171).

Additional studies have investigated subsets of CTCs linked to the therapy used. For example, CTCs expressing insulin-like growth factor-1 receptor (IGF-1R) were enumerated in patients with prostate cancer treated with a monoclonal human antibody, CP-751,871, targeting IGF-IR (172) and HER2 expressing CTCs were examined in patients with breast cancer treated in the neoadjuvant GeparQuattro trial (173).

Two clinical studies evaluated targeting the Notch ligand, delta-like ligand 3 (DLL3), and the T-cell checkpoint receptors programmed cell-death protein 1 (PD-1) and cytotoxic T-lymphocyte-associated protein 4 (CTLA-4) for SCLC. Rovaalpituzumab tesirine, DLL3 targeting agent has shown encouraging results in phase I clinical trials and is now moving on to phase II studies. Several other agents are in clinical trials for treatment of lung cancer such as new...
In the recent past, immune checkpoint inhibitors have emerged as promising therapeutic targets in both NSCLC and SCLC yielding unprecedented improvements in OS and quality of life. More recently, expression of PDL-1 in CTCs has been examined in trials with immune checkpoint inhibitors and has linked with patients who experienced progressive disease (177).

**New drug target identification via CTC analysis**

Since CTCs have been shown to be tumourigenic and can be obtained from a simple blood sample taken at pre-treatment, on-treatment or post-treatment CTCs also represent an attractive source for drug target identification. Although as yet there has been no systematic use of CTCs for de novo drug target identification the recent advances in single cell profiling methodologies which now enable genetic and epigenetic analysis of CTCs (178) now provide this opportunity. For example, putative drug targets could be identified by examining genes encompassed in the CNA patterns seen in single and pooled CTCs which are linked to treatment response (44).

**Conclusions**

The study of CTCs represents an important field of research in cancer biology and also holds great promise as diagnostic, prognostic and predictive biomarkers. Currently the use of a blood samples as a “liquid biopsy” in oncology has been dominated by the striking advances in examining the cell-free component of blood and particularly circulating tumour DNA (ctDNA) (179,180). However, there are several features of CTCs that are not shared by ctDNA that justify continued CTC research including the generation of CDX and CTC culture models that can provide a direct means of testing new drugs/drug combinations (181). In addition, RNA analysis of CTCs can provide insights into their biological status and serial CTC RNA profiling can be used to identify underlying mechanisms of resistance. Furthermore, DNA analysis of individual CTCs provides an unambiguous means of understanding the mutational burden of the tumour as well as providing additional means for establishing the degree of tumour heterogeneity. Since a single blood sample encompasses both cell-free and cellular components, CTC and ctDNA can be readily combined to provide an enhanced understanding of the patient’s disease status (26,180,182).

In the field of oncology CTCs remain a key focus of biological and clinical research. With the recent improvements in CTC enrichment researchers are now able to explore key ongoing biological questions in CTC research such as: how many CTCs are required for a representative snapshot of the donor; can CTCs be detected reliably in early disease and can CTCs be used to routinely guide cancer patient care.

**Acknowledgements**

We would like to thank Sarah Hilton, Dawar Pasha and Jackie Pierce for developing the staining protocol for identification of CTCs post enrichment from an epitope independent system (Parsortix™), Debbie Burt for isolation of single cells via the DEPArray and Sakshi Gulati for bioinformatic analysis and generation of CNA plots.

**Funding**

We would like to acknowledge the help of the CR-UK Clinical and Experimental Pharmacology Group (C5759/A20971) for their support of this study, as well as the financial support from Medical Research Council Manchester Single Cell Research Centre (MR/M008908/1). F Chemi is funded by Cancer-ID (115749-Cancer-ID) and S Mohan is funded by AstraZeneca (D1330N00013). Additional research funding was also provided by the companies Menarini Silicon Biosysytems (provider of DEPArray) and ANGLE plc UK (provider of Parsortix).


