Challenges in molecular testing in non-small cell lung cancer patients with advanced disease

Crispin T Hiley 1,2, John Le Quesne 3, George Santis 4, Rowena Sharpe 5, David Gonzalez de Castro 6,7, Gary Middleton 8,9 and Charles Swanton 1,10 *

1. Translational Cancer Therapeutics Laboratory, The Francis Crick Institute, London, UK.
2. Division of Cancer Studies, King’s College London, London, UK.
3. Department of Cancer Studies, University of Leicester, Leicester, UK.
4. Department of Respiratory Medicine and Allergy, Kings College London, UK.
6. Centre for Molecular Pathology, Royal Marsden Hospital, Sutton, UK.
7. School of Medicine, Dentistry and Biomedical Sciences, Queens University Belfast, UK.
8. Institute of Immunology and Immunotherapy, University of Birmingham, Birmingham, UK.
9. University Hospital Birmingham NHS Foundation Trust, Queen Elizabeth Hospital, Birmingham, UK.
10. CRUK Lung Cancer Centre of Excellence, UCL Cancer Institute, London, UK.

*Correspondence to: Professor Charles Swanton. The Francis Crick Institute, 44 Lincolns Inn Field. WC2A 3LY. charles.swanton@crick.ac.uk

Summary:

Lung cancer diagnostics have evolved significantly in the previous decade. The challenge of molecular testing to identify an increasing number of potentially clinically actionable genetic variants, using smaller samples obtained via minimally invasive techniques, is significant. Tumour heterogeneity and cancer evolution in response to therapy means that repeat biopsies or circulating biomarkers are likely to be increasingly useful to adapt treatment as resistance develops. We highlight some of the current challenges faced in clinical practice for molecular testing of EGFR, ALK and new biomarkers such as PDL1. Implementation of next generation sequencing (NGS) platforms for molecular diagnostics in non-small cell lung cancer is increasingly common allowing testing of multiple genetic variants from a single sample. The use of NGS to recruit to molecularly stratified clinical trials is discussed in the context of the UK Stratified Medicine Programme and The UK National Lung Matrix Trial.
Historical overview of lung cancer diagnostics

Lung cancer is the most common cause of mortality in the UK, accounting for 1 in 5 of all cancer deaths. With the estimated global incidence in 2012 of 1.83 million cases it is important to reflect that a century ago, lung cancer diagnosis was a rare event. In comparison to the current epidemic, in 1912 Isaac Adler’s collection of 374 case reports in his publication *Primary Malignant Growths of the Lungs and Bronchi* represented the entire known global incidence at the time. A century later the WHO histological classification of malignant epithelial tumours of the lung recognizes different histologies with many variants for each subtype and analyses from next generation sequencing (NGS) studies have divided this disease into molecular subtypes defined by distinct somatic alterations. This review will focus on key challenges faced in current clinical practice for molecular testing in non-small cell lung cancer (NSCLC). In broad terms the challenges are technical, logistical and related to tumour biology and some of the pertinent issues are highlighted (Figure 1).

Identification of tumour histology

Historically the treatment focus for those with advanced NSCLC was selection of an appropriate cytotoxic chemotherapy regimen irrespective of histological subtype. Several large studies were published that showed the efficacy of various platinum doublet combinations were comparable but with differing drug specific toxicities. However accurate classification of NSCLC subtype has become fundamental in the management of advanced NSCLC following the results of phase III clinical trials showing improved progression free survival in EGFR mutation positive adenocarcinoma treated with EGFR tyrosine kinase inhibitors (TKI), and improved overall survival with pemetrexed in the first line and maintenance setting for those with non-squamous histology. The number of tumours that cannot be given an accurate histological diagnosis (i.e. adenocarcinoma versus squamous cell carcinoma) has reduced significantly with the use of immunohistochemical markers. The use of markers for p63, p40 and cytokeratin CK 5/6 help to identify squamous cell carcinomas, while thyroid transcription factor 1 (TTF1), Napsin A and CK7, as well as mucin stains, are indicative of adenocarcinomas. However interpretation of immunohistochemistry panels still requires the expertise of an experienced histopathologist, as markers are not reliable in isolation. TTF1 for example, a marker
synonymous with adenocarcinoma, is expressed in only 80-90% of cases and is commonly expressed in neuroendocrine tumours.\(^{14,16}\) Immunohistochemistry can only be meaningfully interpreted in a detailed morphological context.

**Sampling challenges in advanced NSCLC**

The analysis of lung cancer tissue is particularly challenging as primary lung tumours often show much lower tumour cellularity than other tumour types. Even with macroscopic selection of areas of frank carcinoma the tumour purity (the fraction of a given region containing tumour cells) can often be <20% because of the high proportion of stromal cells, lymphocytic infiltration and necrosis (unpublished observations from the UK Lung TRACERx longitudinal cohort study).\(^{17}\).

This challenge is compounded by the nature of the specimen types routinely received by histopathology and molecular diagnostics laboratories. Presentation with metastatic disease is common and only a small proportion of patients with NSCLC undergo curative surgical resection.\(^{18}\) The large tissue samples obtained via open thoracotomy (wedge resection, lobectomy, pneumonectomy) are usually of sufficient quantity and quality for a number of histological and molecular assays if handled appropriately. However patients with advanced disease are predominately diagnosed with CT guided percutaneous or US guided endoscopic biopsy with 18 gauge needles or with fine needle aspiration. These patients are the cohort where molecular diagnostics are most important for determining the standard of care and enabling participation in clinical trials yet the sample quality and quantity from such needle biopsies is the most limiting for histological and molecular testing.

Obtaining adequate tissue for diagnosis, tissue sub-typing, molecular profiling and treatment planning are therefore key to patient management. The target tumour is not always easily accessible in patients presenting with a probable lung cancer. The development of endobronchial ultrasound transbronchial needle aspiration (EBUS-TBNA) is proving increasingly important in the investigation and management of thoracic malignancies as it offers a minimally invasive approach to sampling of mediastinal lymph nodes and masses. EBUS-TBNA is now increasingly embedded in routine clinical practice with wider use beyond high volume tertiary centres in the UK and USA. It is now generally accepted that EBUS-TBNA alone or in combination with endoscopic ultrasound fine needle aspiration (EUS-FNA) can potentially spare surgical mediastinoscopy or thoracotomy in the staging of NSCLC.\(^{19,20}\)
Importantly, EBUS-TBNA also offers the possibility of combining diagnosis and staging as a single procedure in patients with suspected lung cancer. In contrast to tissue biopsies or surgical samples that allow sub-typing (adenocarcinoma versus squamous cell carcinoma) on morphological criteria alone in the majority of cases, evaluation of cytological specimens obtained by EBUS-TBNA poses additional challenges that can be partly overcome with wider use of immunohistochemistry.\(^{21}\)

Identifying driver mutations, such as EGFR and ALK, in these small samples is central to management of patients with advanced disease. Whether molecular analysis is successfully performed depends on the absolute number of tumour cells, the proportion of tumour cells compared to total nucleated cells present and the method used for molecular analysis. In case of EBUS-derived samples, there is evidence to conclude that simple mutation analysis (EGFR, KRAS, ALK) can be successfully performed in most cases.\(^{22-24}\) The use of multi-gene targeted NGS panels, using only nanograms of DNA, to sequence fine needle aspiration samples is achievable and is becoming more commonly used in clinical practice.\(^{25,26}\) Whole exome sequencing (WES) and whole genome sequencing (WGS) analysis which require greater amounts of DNA, micrograms in the case of WGS, will be more challenging from EBUS-TBNA samples.

**Current challenges in molecular diagnostics for EGFR mutation analysis in clinical practice**

The initial randomised phase II studies of gefitinib demonstrated clinical activity,\(^{27,28}\) and phase III studies although negative for the primary outcome measure, suggested a benefit in patients with adenocarcinoma, those of Asian origin and never smokers.\(^{29,30}\) During this period a number of seminal case series identified EGFR mutations as a marker of sensitivity to EGFR TKIs,\(^{31-33}\) and analysis of samples from these early trials supported this.\(^{34}\) Subsequent phase III trials have incorporated EGFR mutation status and showed higher response rates and progression free survival (PFS) in patients with EGFR activating somatic mutations treated with EGFR TKIs compared to when treated with chemotherapy.\(^{8,9,35}\)

**Development of diverse mechanisms of resistance and selection of resistant clones in response to treatment**

The common EGFR mutations are located in the tyrosine kinase domain (exon 18-21) of the
EGFR gene, with detection of L858R and deletions in exon 19 the clinical priority as these determine sensitivity to first and second generation TKIs.\(^{(36)}\) The T790M mutation in exon 20 results in resistance to these therapies.\(^{(37)}\) Sensitive assays suggest that tumour clones harbouring the T790M mutation are often detectable prior to initiation of a first generation TKI but can also occur by genetic evolution in T790M mutation negative drug tolerant cells in response to treatment.\(^{(38-40)}\) Identification of this resistance mutation is more critical following the development of the third generation EGFR TKIs active against T790M mutation positive NSCLCs.\(^{(41,42)}\) But whether these T790M resistant clones pre-exist or evolve in response to treatment may have clinical implications with differing sensitivities to third generation TKIs.\(^{(40)}\) The capacity for tumours to evolve in response to first generation TKIs results in an additional diverse array of mechanisms of resistance such as amplification of MET, selection for PIK3CA or BRAF mutations and transformation to a small cell phenotype.\(^{(43)}\) Clearly, cancer evolution and selection of resistant subclones is not restricted to first generation TKIs. This is highlighted by recent reports of the emergence of T790M mutation negative disease and the development of novel secondary EGFR resistance mutations (C797S) after treatment of T790M mutation positive patients with third generation TKIs.\(^{(44,45)}\)

**EGFR mutation testing**

The nature of EGFR sensitizing mutations, being single nucleotide variants (SNV) or short deletions, lends themselves to molecular analysis of formalin fixed small samples which contain fragmented DNA.\(^{(46)}\) There are a variety of methods to detect EGFR mutations including conventional Sanger sequencing, amplification refractory mutations systems (ARMS), restriction fragment length polymorphisms and more recently as part of targeted NGS panels.\(^{(47,48)}\) Reporting the limitations of an assay along with the result is critical for clinical interpretation. Bi-directional Sanger sequencing without a mutation enrichment step has a lower limit of detection of 10-25% of total DNA meaning that the use of samples with low tumour cellularity can result in false negative mutation calls. Consequently the use of methods that can detect mutations in low tumour cellularity samples (<10%) is recommended. Polymerase chain reaction (PCR) based ultra sensitive and NGS methods can generate artefact mutations leading to false positive results. However techniques, such as duplex sequencing, are being developed to overcome the inherent error rate in sequencing technologies.\(^{(49)}\) Formalin fixed samples are particularly prone to DNA damage and display disproportionate levels of C>T/G>A changes in the 1-10% allele frequency range which can
result in false positive mutation calls.\textsuperscript{[50]} Publication of clinical trials results on response to EGFR TKI in patients with real but less common EGFR mutations can help guide clinical decision-making.\textsuperscript{[51]} Detection of EGFR mutations as part of a WES or WGS analysis allows multiple driver mutations to be queried simultaneously but the performance of bioinformatics tools to call mutations from NGS data varies. Such complexities need to be considered as these technologies are increasingly adopted into mainstream clinical practice.\textsuperscript{[52-54]}

**EGFR mutations, resistance and tumour heterogeneity**

There are very few reports of discordance of EGFR mutation status between primary disease and metastatic sites and these may be due to technical limitations of the assays used.\textsuperscript{[55]} Loss of the EGFR mutation was not a mechanism seen in seminal studies.\textsuperscript{[43]} Studies looking at the extent of intra tumour heterogeneity (ITH) in early lung cancer have shown EGFR to be exclusively a clonal event prevalent throughout the tumour.\textsuperscript{[56,57]} As resistance to EGFR TKI is usually due to acquisition of secondary mutations in EGFR or other driver genes the key challenge at the time of disease progression is to obtain a contemporaneous sample to inform selection of second line therapy. In general the most easily accessible lesion is used but in patients with a poor performance status this may not be a trivial task. Due to tumour heterogeneity it is possible that a single sample may be insufficient to accurately represent all the resistance mechanisms present or the breadth of clonal driver events present across multiple disease sites following progression on therapy.

**Current challenges in molecular diagnostics for ALK testing in clinical practice**

The discovery of an oncogenic anaplastic lymphoma kinase fusion gene (EML4-ALK) in 2007 identified another important molecular cohort in NSCLC.\textsuperscript{[58]} Present in 2-7% of NSCLC ALK fusion genes are restricted to adenocarcinoma subtypes and are more common in younger patients and never-smokers.\textsuperscript{[59-61]} Identification of this cohort is critical given the high response rates (57-74%) to ALK inhibition both as a first line therapy and after platinum-based chemotherapy.\textsuperscript{[59,62,63]} Subsequently other rare fusion genes have been identified involving ROS1, with similar exquisite sensitivity to kinase inhibition,\textsuperscript{[64]} but also RET and NTRK where objective response rates are lower.\textsuperscript{[65,66]}

**ALK fusion gene detection**
Testing for ALK fusion genes brings its own particular set of challenges. ALK is activated by genomic rearrangement, leading to the expression of a chimeric protein containing the effector part of the ALK tyrosine kinase fused to the proximal portion of another protein. In NSCLC cancer this is typically a balanced translocation with the ubiquitously highly expressed EML4 gene,\(^{[58]}\) although rarely other partner genes may be involved.\(^{[67,68]}\) Expression of the chimeric protein leads to upregulation of mitogenic signalling through the RAS/RAF pathway and interruption of this pathway by ALK inhibitors causes cancer cell death and tumour regression.\(^{[62]}\) ALK-mutated tumours often show unusual features on conventional microscopy, such as cribriform growth patterns and ‘signet ring’ cells with large vacuoles,\(^{[69]}\) but this is not sufficiently sensitive or specific to guide treatment.

The first widely adopted test for ALK-driven tumours was FISH (fluorescence in situ hybridisation), approved by the FDA (US Food and Drug Administration) in 2011.\(^{[70]}\) FISH is a technically demanding method, requiring specialised equipment and experienced practitioners. Tissue sections or cytology specimens are subjected to a protocol that labels either side of the ALK breakpoint locus with red and green fluorescent DNA probes. In non-transformed cell nuclei the coloured dots overlap and look yellow, while in translocated cells isolated red and/or green signals are seen. For a reliable FISH assay the tissue must be adequate in quantity and quality. This can be more challenging with small biopsy samples which may contain few cells or which show crushing artefacts that can impair interpretation.

In 2015, an immunohistochemistry method was approved by the FDA. This approach is simpler in principle, using an antibody stain to detect abnormal ALK antigen expression. However, currently available antibodies do not give a strong signal and so an additional signal amplification steps needs to be employed. This places the test beyond the capacity of many small labs. Nonetheless, the modified test is cheaper than FISH, easier to interpret, and has the theoretical advantage of additionally detecting ALK expression following rare atypical rearrangements. After much investigation, recent studies suggest immunohistochemistry can be an adequate stand-alone diagnostic, showing extremely high concordance with FISH.\(^{[71]}\) UK guidelines do not dictate which test should be applied, and practices vary regionally, though FISH is still often regarded as the ‘gold standard’ and is considered the definitive test in the US.\(^{[48]}\) As our understanding of tumour taxonomy and genotypes advances, it seems inevitable that some form of NGS platform will become the clinical standard for gene fusion detection.\(^{[72]}\) These methods have the potential to detect
ALK (and other) rearrangements in either a targeted panel or a WES or WGS approach.\textsuperscript{(73,74)} \textsuperscript{(72)}

**ALK fusion genes, resistance and tumour heterogeneity**

ALK fusion genes are considered to be clonal events with minimal discordance between primary and metastatic lesions.\textsuperscript{(75)} They were considered to be mutually exclusive with EGFR mutations however recent reports suggests a small minority of tumours can contain both ALK and EGFR positive clones.\textsuperscript{(76-78)} The mechanisms of resistance seen following ALK inhibitor therapy again demonstrate tumour evolution with secondary mutations in ALK, ALK copy number gain, secondary driver mutations in other genes and outgrowth of ALK fusion gene negative clones reported.\textsuperscript{(79-82)} Consequently contemporaneous sampling of progressive disease, by needle biopsy or analysis of cfDNA, may allow real time analysis of tumour evolution and guide therapy.

**Integration of multi-gene NGS testing in clinical practice**

Routine molecular profiling can be performed at scale on a national level. Large cooperative efforts in France and the USA used combinations of mutation specific PCRs, Sanger sequencing and FISH analysis to assay 6-10 oncogenic drivers in thousands of patients with NSCLC and survival was improved for those treated with gene directed targeted therapies.\textsuperscript{(83,84)} The use of next generation sequencing to simultaneously assay multiple oncogenic drivers is attractive because less DNA is required compared to multiple individual assays, there is a reduction in hands-on laboratory time, and complex FISH analysis for detection of fusion genes may be avoided. A recent NGS approach used an amplicon based approach to assay 14 genes used only 50ng of DNA from FFPE samples.\textsuperscript{(85)} This study provided a comprehensive assessment of the spectrum of mutations, and co-occurrence of mutations, in adenocarcinoma and squamous cell carcinomas with detection turn around times of less than two weeks. These studies and those of The Cancer Genome Atlas highlight the inter-patient molecular heterogeneity of NSCLC (Figure 2). Even within these molecular cohorts intra-tumour heterogeneity could have significant effects on outcome as exemplified by a recent study showing that the clonality of FGFR amplification is an important predictor of response to FGFR inhibition.\textsuperscript{(86)} A deeper understanding of the clonal or subclonal nature of driver events in NSCLC from sufficiently powered studies, is still awaited. Recruitment of patients with rare mutations to molecularly stratified trials is challenging,\textsuperscript{(87)} and some advocate that modifications to existing paradigms in drug
development are required in the era of genomic studies and precision medicine.\(^{(88)}\) Multi-

gene or WES NGS assays are likely to become standard practice in the years to come and the

ultimately automated provision of readable, applicable reports of complex genomic data is

another important challenge.

**Current challenges in molecular diagnostics for PDL1 testing in clinical practice**

Activation of inhibitory T cell checkpoint interactions in established tumours has been
demonstrated in a number of solid tumours, including NSCLC, and this suppresses the anti-
tumour immune response.\(^{(89,90)}\) The aim of immunotherapy using antagonists of these
inhibitory T cell checkpoint interactions is to reactivate anti-tumour immunity. PDL1 (B7-H1)
is a ligand present on antigen presenting cells (APCs), including tumour cells that interacts
with its receptor (PD-1) on T cells and inhibits T cell effector functions. PD-1 and CD8 positive
effector T cell population are thought to be the tumour reactive subset responsible for anti-
tumour immunity.\(^{(91)}\) There is limited knowledge of the spatial or functional heterogeneity of
tumour infiltrating lymphocyte (TIL) populations and the T cell checkpoint ligand-receptor
interactions within solid tumours.

Recent randomised trials have shown activity of PD1 and PDL1 targeting antibodies in
squamous and non-squamous NSCLC.\(^{(92-96)}\) In most instances these agents have shown
greater activity in patients whose tumour expresses PDL1 when tested using IHC. However
durable responses are seen in patients without PDL1 expression. This is unsurprising given
the technical and spatial heterogeneity of PDL1 expression in NSCLC, which hampers its use
as a predictive biomarker.\(^{(97-99)}\) Studies of the expression of PDL1 on APCs in NSCLC are also
contradictory with respect to any correlation with tumour infiltration of the effector CD8
positive T cells.\(^{(89,97,98)}\)

Regulation of PDL1 expression is complex and controlled by both cell intrinsic and cell
extrinsic factors.\(^{(100)}\) This means that oncogene driven expression of PDL1 can result in
increased expression in the absence of significant underlying immunogenicity.\(^{(101)}\) This
underlying immunogenicity is thought to be a result of non-synonymous SNVs which
generate neoantigens, mutated proteins, recognised by the TIL population.\(^{(102,103)}\) The
number of neoantigens harbour by a tumour could act as a potential biomarker for
immunotherapy although there are technical challenges inherent with such complex
assessments. Recent data also suggest that neo-antigen intratumour heterogeneity may also be associated with altered checkpoint inhibitor response, which may further complicate the use of such assays in a clinical setting.\textsuperscript{104}

The advent of immunotherapy presents additional challenges for molecular diagnostics in NSCLC. Although IHC for PDL1 can be performed on the small samples often used in lung cancer diagnostics there is the risk of significant sampling bias because of ITH. The dynamic nature of PDL1 gene expression,\textsuperscript{105,106} means a contemporaneous sample obtained by repeat biopsy may be the most accurate adding additional burden and expense to current clinical pathways. Characterisation of neo-antigens as a potential biomarker would require sufficient tumour DNA for WES and carries significant expense but given the cost of these therapies would be justified if the assay were sufficiently predictive. However neo-antigen prediction algorithms are still in their infancy and evidence suggests that there are a proportion of patients who derive no clinical benefit from checkpoint inhibitor therapy, yet have tumours with a neo-antigen burden above thresholds associated with sensitivity and conversely patients with low neo-antigen burden who benefit.

**Molecular diagnostics in practice: The United Kingdom National Lung Matrix Trial**

The Cancer Research UK Stratified Medicine Programme 2 (SMP2) screens samples from advanced NSCLC patients using NGS for known drivers that are considered clinically actionable. The aim of SMP2 is to establish high-throughput and quality genomic screening at a national level in the UK. Based on these results, patients are recruited to The National Lung Matrix Trial (NLMT) (NCT02664935) a phase II umbrella study with both targeted therapy and immunotherapy arms for patients who have progressed on first line therapy.\textsuperscript{107} In comparison, the Lung-MAP (NCT02154490) and SAFIR02 Lung trial (NCT02117167) are umbrella studies, outside of the UK, for patients with NSCLC where recruitment is preceded by molecular stratification (Table 1).

SMP2 molecular pathology workflow utilises DNA from excess diagnostic biopsy tissue. Sections are sent from the referring clinical site and extracted by one of three central technology hubs. Samples with sufficient amounts of DNA (>50ng) are then analysed using a custom 28-gene targeted NGS panel. Having successfully screened over 1000 patients, patterns of mutation and prevalence are emerging across the genomic and clinical data.
Preliminary analysis indicates prevalence and distribution of SNVs consistent to published reports, including 31.6% KRAS (of which 19.7% show concomitant STK11 mutation) and 15.1% EGFR mutations in patients with adenocarcinoma. Over the past year, SMP2 has led to the recruitment of over 60 patients to the NLMT. A number of detailed audits have identified areas of improvement along the SMP2 pathway; from patient recruitment, to sample preparation and result analysis.

Whilst utilising excess DNA from the FFPE diagnostic biopsy has significant advantage for patients and clinical workload (as repeat biopsies are not required), only 70% of samples sent have sufficient DNA to enter the sequencing pipeline. This is in part due to FFPE blocks being exhausted during the diagnostic process and a general reduction in the size of diagnostic cores over time. Consequently the minimum number of sections has since been increased to ensure enough DNA is obtained up front. Some recruiting centres quantify DNA upfront, which allows a faster feedback loop if insufficient DNA is present. Sites can then obtain additional sample from the diagnostic block or through re-biopsy, if appropriate. However, differences in quantification methodology between local clinical centres and central technology hubs have led to samples being sent with less than the required 50ng, resulting some of these samples failing quality control metrics prior to sequencing. As a result changes in extraction methods and a standardized DNA concentration have been introduced.

Unique to NLMT is the need to determine wild type status of some genes for eligibility to certain arms. Patients recruited to the CDK4/6 inhibitor palbociclib arm must have wildtype retinoblastoma 1 protein (RB1) in addition to deficiencies in cell cycle regulation. The determination of wildtype status requires a pre-sequencing assessment of tumour cellularity to determine appropriate sequencing depth. However there can be significant discordance between pathologist assessments of this. Clearer guidance and online training should ensure more concordance for visual assessment, whilst digital solutions may provide a useful alternative. A number of computational methods exist to assess tumour purity and control for both stromal cell admixture and cancer cell ploidy in DNA samples from next generation sequencing data.

Extremes of GC nucleotide content in certain genes (RB1 and FGFR3) can result in an increased number of sequencing failures. Additional probe coverage in the targeted panel  

11
and correction for GC content in the data processing stage will improve results for these
difficult to sequence regions. Following these incremental improvements at each step of the
molecular pathology workflow we have shown that the number of successfully sequenced
samples that would allow recruitment to the NLMT increased and there has also been an
increase in identification of potentially actionable mutations that would permit recruitment
to trials other than the NLMT.

Future solutions

Understanding tumour heterogeneity and cancer evolution

At present the technical limitation of the small, and potentially low tumour cellularity NSCLC
samples, obtained from bronchoscopic and EBUS-TBNA samples means that the main
challenge facing clinicians and pathologists is the need for ever greater amounts of
information from diminishing amounts of tissue. It is therefore imperative that the quality of
diagnostic samples in the advanced NSCLC setting is of the highest order. How best to
achieve this represents a challenge for health service providers that has received very little
attention thus far. However the spectre of ITH and cancer evolution means that sampling
bias and the presence of subclonal driver mutations, causing resistance to therapy, are likely
to hinder clinical benefit of targeted therapeutics. The UK Lung TRAcking Cancer
Evolution through Therapy trial (NCT01888601) is currently characterizing the extent of ITH
in early surgically resected NSCLC and with longitudinal follow-up aims to determine the
origins of tumour subclones contributing to relapse. There is evidence from other tumour
types of parallel evolution, acquisition of mutations in the same gene or signalling pathway
in distinct subclones, that may highlight an ‘evolutionary bottle neck’ that could be an
Achilles heel for subsequent cancer therapy. Clonal analyses of a drug target and
putative resistance events, whether they are present in all tumor cells or only a proportion,
may affect the response rate and progression free survival times on targeted therapy and
this is being addressed in clinical trials including the DARWIN studies (NCT02314481,
NCT02183883). Ultimately it may be that only through ‘warm’ autopsy studies, where
subclonal phylogenetic structures can be determined through sampling multiple sites of
disease, that effective strategies to forestall cancer evolution can be elucidated.

Circulating biomarkers

The use of minimally invasive methods to detect mutations in circulating cell-free DNA
(cfDNA) or ‘liquid biopsies’ offers the potential to obtain a mutation call in a patient where an invasive biopsy may not be feasible. As tumour DNA from all sites of disease has the potential to enter the blood stream it may also be a better reflection of tumour heterogeneity than a single biopsy.\(^{117,118}\) The use of cfDNA to detect resistance mechanisms in patients treated with EGFR TKIs, often prior to radiographic progression, has been demonstrated.\(^{45,119,120}\) This has resulted in the development and approval of a commercially available assay of cfDNA in plasma that can detect a spectrum of EGFR mutations in including the T790M mutation amenable to targeting with third generation TKIs.

Circulating tumour cells (CTCs) are tumour cells that can be isolated from the peripheral blood and are a complementary circulating biomarker to cfDNA. CTCs are a versatile tool, cell enumeration can be prognostic, immunohistochemistry permits further characterization, single cell DNA or RNA sequencing is possible and tumour xenografts can be generated to assess drug response.\(^{121-125}\) However at present the complexity of separation from other cells in the peripheral circulation and the need to process samples promptly for functional or genomic studies results in greater expense in comparison to cfDNA analysis. Circulating biomarkers will make a significant impact on cancer management in the near future and readers are directed to more extensive reviews focusing on CTCs, cfDNA and other circulating nucleotides.\(^{121,126-129}\)

**Conclusion**

The challenges for molecular diagnostics in NSCLC are largely paralleled across other tumour types. Resolving these issues will require technology improvements in addition to a greater understanding of tumour biology. The logistical challenges of implementing the next generation of molecular diagnostics into clinical practice are equally as challenging. Clinical governance, information technology infrastructure, data storage, pathways in sample processing and training and professional developments in histopathology, respiratory medicine and oncology will need investment. With these great challenges comes significant opportunity to improve the success rate and efficiency of drug development in NSCLC and ultimately patient outcomes.

Conflicts of interest: None
Figure 1: Summary of the key technical, logistical and biological challenges for molecular testing in NSCLC.

Figure 2: Pie charts showing the approximate distribution of clinically relevant driver mutations identified to date in individuals with NSCLC. The genomic variants shown are potentially clinically actionable variants. The proportions presented are based on estimates from the referenced studies and data sources, including the Stratified Medicine Programme 2 (unpublished data). These studies examine SNVs, copy number variants and gene fusion products using different sequencing technologies and sequencing depth resulting in inter-study variation and therefore the data is presented in aggregate form and represents an approximation. CCGA (Cell cycle genomic aberration); loss of Cyclin-Dependent Kinase Inhibitor 2A or amplification of Cyclin-Dependent Kinase 4 or Cyclin D1 in the presence of wildtype RB1. EGFR, Epidermal growth factor receptor; LKB1, Liver Kinase B1; ALK, anaplastic lymphoma kinase; MET, MET Proto-Oncogene, Receptor Tyrosine Kinase; FGFR, fibroblastic growth factor receptor; NRAS, neuroblastoma RAS viral (v-ras) oncogene homolog; DDR2 Discoidin Domain Receptor Tyrosine Kinase 2; AKT1, v-akt murine thymoma viral oncogene homolog 1; PTEN, Phosphatase And Tensin Homolog; PIK3CA, phosphoinositide-3-kinase, catalytic, α polypeptide; BRAF, v-raf murine sarcoma viral oncogene homolog B1; ERBB2, human epidermal growth factor receptor 2; KRAS, v-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog; TP53, Tumor Protein P53.

Table 1: Comparison of molecularly stratified umbrella studies in NSCLC. PFS = progression free survival, ORR = objective response rate, R = randomised, NR = non-randomised, SCC = squamous cell carcinoma, CCGA = cell cycle genomic aberration.
REFERENCES


70. IASLC atlas of alk testing in lung cancer: International Association for the Study of Lung Cancer Press; 2013.


75. Hou L WC. Comparative study of ALK rearrangement between primary tumor and paired lymphatic metastasis in NSCLC patients. ASCO 20152015.


### Technical
- Novel biopsy techniques (eg EBUS) generating smaller samples with diminished tumour cellularity
- Multiple tests with the potential for discordant results (eg IHC vs FISH for ALK mutation)
- Technology-specific failures due to differences in sensitivity/known artefacts (eg sequencing through repeats/high GC areas)
- Quality assurance of genomic medicine despite across multiple platforms and data analysis algorithms

### Logistical
- Turnaround time of assays in a clinically relevant timeframe
- Desirability of centralised vs distributed/local testing approaches
- Education and training of laboratory and clinical staff in new technologies
- Distillation of high volume data into useful standardised reports usable by clinicians
- Computational and data storage capacity for NGS within a healthcare system

### Tumour Biology
- Diversity of molecular subgroups within NSCLC
  - inter-patient heterogeneity
- Intra-tumour heterogeneity
  - sampling bias
  - differential responses
- Cancer evolution and resistance in response to treatments
  - need for longitudinal sampling
- Evolving treatment paradigms
  - immuno-oncology & new biomarkers (eg PDL-1, neoantigen load)
- Increasing complexity of detectable genomic changes in cancer
  - eg epigenetic changes & non-coding variants
Adenocarcinoma

- KRAS: 35%
- No variant
- EGFR
- LKB1
- ALK
- MET
- DDR2
- Other

Cell cycle variant

TP53 status

- mutant: 35%
- wildtype: 75%

Squamous Cell Carcinoma

- KRAS
- No variant
- FGFR
- PIK3CA
- PTEN
- MET
- DDR2
- Other

Cell cycle variant

TP53 status

- mutant: 57%
- wildtype: 43%
<table>
<thead>
<tr>
<th>Study</th>
<th>Line</th>
<th>Phase</th>
<th>POM</th>
<th>Molecular subgroups</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>National Lung Matrix Trial</td>
<td>2nd or later</td>
<td>NR Phase II</td>
<td>PFS ORR</td>
<td>AKT, PIK3CA/PTEN, TSC, LKB1, KRAS, NRAS, NF1, MET, ROS1, EGFR (T790M), CCGA, immunotherapy</td>
<td>UK</td>
</tr>
<tr>
<td>SAFIR_02 Lung Study</td>
<td>1st line maintenance</td>
<td>R Phase II</td>
<td>PFS</td>
<td>mTOR, AKT, FGFR, HER2, EGFR, MEK, immunotherapy</td>
<td>France</td>
</tr>
<tr>
<td>Lung MAP</td>
<td>2nd or later (SCC)</td>
<td>NR Phase II</td>
<td>PFS ORR</td>
<td>PIK3CA, FGFR, CCGA, immunotherapy</td>
<td>USA</td>
</tr>
<tr>
<td>Darwin I/II</td>
<td>1st or later</td>
<td>NR Phase II</td>
<td>PFS</td>
<td>EGFR, HER2, ALK, RET, BRAF, immunotherapy</td>
<td>UK</td>
</tr>
</tbody>
</table>
### Technical
- Novel biopsy techniques (eg EBUS) generating smaller samples with diminished tumour cellularity
- Multiple tests with the potential for discordant results (eg IHC vs FISH for ALK mutation)
- Technology-specific failures due to differences in sensitivity/known artefacts (eg sequencing through repeats/high GC areas)
- Quality assurance of genomic medicine despite across multiple platforms and data analysis algorithms

### Logistical
- Turnaround time of assays in a clinically relevant timeframe
- Desirability of centralised vs distributed/local testing approaches
- Education and training of laboratory and clinical staff in new technologies
- Distillation of high volume data into useful standardised reports usable by clinicians
- Computational and data storage capacity for NGS within a healthcare system

### Tumour Biology
- Diversity of molecular subgroups within NSCLC
  - inter-patient heterogeneity
- Intra-tumour heterogeneity
  - sampling bias
  - differential responses
- Cancer evolution and resistance in response to treatments
  - need for longitudinal sampling
- Evolving treatment paradigms
  - immuno-oncology & new biomarkers (eg PDL-1, neoantigen load)
- Increasing complexity of detectable genomic changes in cancer
  - eg epigenetic changes & non-coding variants
Adenocarcinoma

- KRAS: 35% mutant, 75% wildtype
- EGFR
- LKB1
- ALK
- MET
- DDR2
- Other: 75%

Squamous Cell Carcinoma

- KRAS: 57% mutant, 43% wildtype
- FGFR
- PIK3CA
- PTEN
- Other: 43%
- Cell cycle variant: 57%

TP53 status:

- Adenocarcinoma: 35% mutant, 75% wildtype
- Squamous Cell Carcinoma: 57% mutant, 43% wildtype
<table>
<thead>
<tr>
<th>Study</th>
<th>Line</th>
<th>Phase</th>
<th>POM</th>
<th>Molecular subgroups</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>National Lung Matrix Trial</td>
<td>2\textsuperscript{nd} or later</td>
<td>NR Phase II</td>
<td>PFS ORR</td>
<td>AKT, PIK3CA/PTEN, TSC, LKB1, KRAS, NRAS, NF1, MET, ROS1, EGFR (T790M), CCGA, immunotherapy</td>
<td>UK</td>
</tr>
<tr>
<td>SAFIR_02 Lung Study</td>
<td>1\textsuperscript{st} line maintenance</td>
<td>R Phase II</td>
<td>PFS</td>
<td>mTOR, AKT, FGFR, HER2, EGFR, MEK, immunotherapy</td>
<td>France</td>
</tr>
<tr>
<td>Lung MAP</td>
<td>2\textsuperscript{nd} or later (SCC)</td>
<td>NR Phase II</td>
<td>PFS ORR</td>
<td>PIK3CA, FGFR, CCGA, immunotherapy</td>
<td>USA</td>
</tr>
<tr>
<td>Darwin I/II</td>
<td>1\textsuperscript{st} or later</td>
<td>NR Phase II</td>
<td>PFS</td>
<td>EGFR, HER2, ALK, RET, BRAF, immunotherapy</td>
<td>UK</td>
</tr>
</tbody>
</table>